Cover Page

Universiteit Leiden

The handle http://hdl.handle.net/1887/77740 holds various files of this Leiden University dissertation.

Author: Kuo, C.L.
Title: Applications for activity-based probes in biomedical research on glycosidases
Issue Date: 2019-09-10

Applications for activity-based probes in biomedical research on glycosidases

PROEFSCHRIFT

ter verkrijging van

de graad van Doctor aan de Universiteit Leiden,

op gezag van Rector Magnificus prof. mr. C. J. J. M. Stolker,

volgens het besluit van het College voor Promoties

te verdedigen op dinsdag 10 september 2019

klokke 10:00 uur

door

Chi-Lin Kuo **郭季霖**

Geboren te Kaohsiung, Taiwan in 1988

Promotiecommissie

Promotores:	Prof. dr. J. M. F. G. Aerts	
	Prof. dr. H. S. Overkleeft	
Overige leden:	Prof. dr. J. Brouwer	
	Prof. dr. M. van der Stelt	
	Prof. dr. A. H. Meijer	
	Prof. dr. P. Saftig	(Kiel University)
	Prof. dr. D. Wachten	(The University of Bonn)

Applications for activity-based probes in biomedical research on glycosidases
Chi-Lin Kuo
Doctoral thesis, 324 pages, Leiden University, 2019
Front cover: Activity-based probe (ABP) ME569; structure model generated by MolView

© 2019, Chi-Lin Kuo, All Rights Reserved

ISBN: 978-94-6375-445-3
Printing: Ridderprint BV, the Netherlands
Cover design: Chi-Lin Kuo

To my family

Table of Contents

General Introduction	Lysosomal glycosidases: inherited deficiency, molecular aspects, and novel chemical probes	1
Scope of the Thesis		26
CHAPTER 1	Development of activity-based protein profiling methods for *in vitro* and *in situ* visualization of glucocerebrosidase	29
CHAPTER 2	Detecting *In vivo* target engagement of conduritol B epoxide and cyclophellitol by activity-based protein profiling	47
CHAPTER 3	Functionalized cyclophellitols are improved inhibitors for generating neuropathic Gaucher model in zebrafish	77
CHAPTER 4	Biological characterization of activity-based probes for α-glucosidase and β-glucuronidase	111
CHAPTER 5	Development and application of activity-based probes for α-L-iduronidase	157
CHAPTER 6	Activity-based probes for retaining exo-mannosidases	199
CHAPTER 7	β-Galactose configured cyclophellitol aziridine as activity-based probes for retaining exo-β-galactosidases	241

General Discussion and Future Prospects 281

Summary 300

Appendices 305

Summary in Dutch (Nederlandse Samenvatting) 306

Summary in Chinese (中文摘要) 310

List of Publications 314

Curriculum vitae 317

General Introduction

Lysosomal glycosidases: inherited deficiency, molecular aspects, and novel chemical probes

Carbohydrates are the most abundant biomolecules on earth and they are besides nucleic acids, proteins, and lipids one of the four major macromolecules essential for life.[1] They are structurally very diverse as the result of the occurrence of different monosaccharide building blocks, and different linkages among these. The most utilized monosaccharides are pentose and hexose (five- or six- carbon sugars), whose ketone or aldehyde can spontaneously react with one of its own hydroxyl, forming a stable five-membered ring (furanose, e.g. fructose and ribose) or a six-membered ring (pyranose, e.g. glucose). This is accompanied by two possible stereochemical outcomes (α- or a β-anomer, where the hemiacetal hydroxyl points to opposite directions from the ring plane). In addition, the monomers can have different hydroxyl configurations (e.g. glucose, galactose, and mannose), and their hydroxyl(s) may be substituted with other types of functional group (e.g. N-acetyl, carboxylic acid, sulfate). The ten commonly utilized monosaccharide building blocks in life are glucose, galactose, mannose, glucuronic acid, L-iduronic acid, neuraminic acid, N-acetylglucosamine (GlcNAc), N-acetylgalactosamine (GalNAc), L-fucose, and xylose (**Fig. 1A**).[2] When these are linked together through glycosidic linkages at different hydroxyl positions (on the reducing-end sugar), a large number of possible polymeric structures (oligosaccharides and polysaccharides) can be formed. For example, a simple hexasaccharide constituting of a single type of hexose can form 1.05 x 10^{12} different isomers.[3] Finally, carbohydrates can also linked to other biomolecules, such as protein and lipids, forming glycoconjugates that are also structurally diverse (**Fig 1B**). These entail specific cellular functions such as energy storage (i.e. glycogen and starch), structural integrity (i.e. cellulose and chitin), and a plethora of fundamental cellular processes in eukaryotes, ranging from glycoprotein folding and targeting, inter- and intra-cellular signaling, and cellular interaction/recognition events underlying infection, immune response, and cancer.[1] In vertebrates, glycan synthesis is carried out by ~200 glycosyltransferases, mostly are Golgi enzymes that transfer monosaccharides to the glycans' non-reducing end using high-energy nucleotide sugar donors.[2, 6]

In contrast, glycan degradation largely takes place in lysosomes. The lysosomes are ubiquitous cellular organelles with an acid interior; they contain a range of acid hydrolases fragmenting macromolecules that may enter lysosomes through endocytosis and autophagy. Degradation products are exported from lysosomes via specific transporters for re-use in the cytosol. Only ~30 glycosidases (glycoside hydrolases) orchestrate the step-wise degradation of

glycosidic bonds presented on the structurally diverse glycoconjugates. These lysosomal glycosidases are proficient catalytic machines that can accelerate the hydrolysis of glycosidic bonds by a factor of 10^{17}.[7] Most of them hydrolyze a specific type of reducing-end monosaccharide with either α- or β-anomeric linkage, from glycans presented on glycosphingolipids or liberated from other types of glycoconjugate. Paradoxically, the impressive substrate specificity and high catalytic efficiency of lysosomal glycosidases forms the basis of a number of inherited metabolic diseases in man (**Part I**). Deficiency of a single lysosomal glycosidase can cause major disturbances in metabolism, ultimately translating in clinical

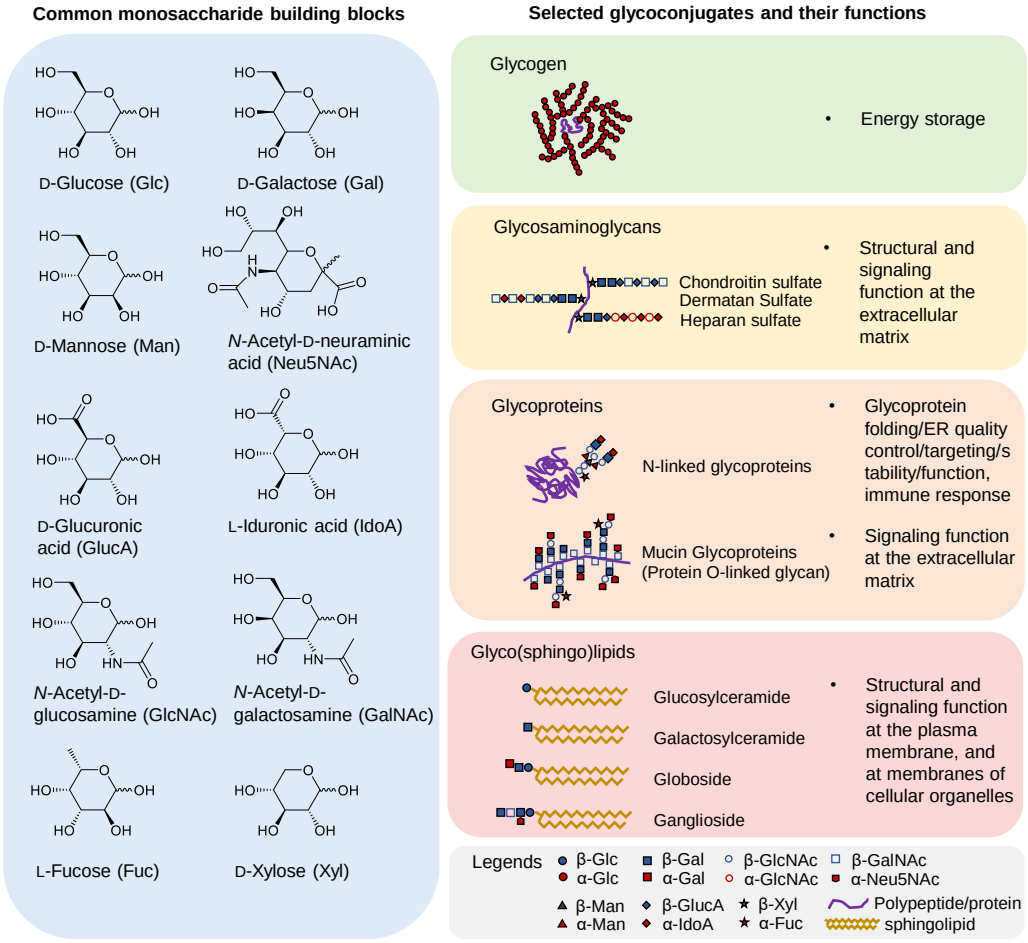


Figure 1. Structures of A) monosaccharide building blocks and B) a selected set of glycoconjugates with their main cellular functions.

symptoms. Research on lysosomal glycosidases and corresponding diseases have prompted the design of mechanism-based chemical probes for enzyme profiling to facilitate investigations on glycosidases in health and disease in an unprecedented manner (**Part II**).

Part I. Glycosidases and inherited metabolic diseases

Pioneering work in the 50's and 60's revealed that inborn deficiencies of particular lysosomal glycosidases are the molecular basis of a number of inherited metabolic disorders, collectively coined as lysosomal storage disorders (LSDs).[8–10] Individual LSDs are rare, and the combined incidence rate is about 1 in 5,000–8,000 newborns.[11, 12] The prevalence of some LSDs can be however high in specific populations: for example, the prevalence of Gaucher disease among Ashkenazi Jewish individuals is reported to be 1 in 450–850 newborns.[13, 14] To date, more than 50 discrete LSDs are recognized. Lysosome dysfunction may result from defects in formation of lysosomes, stability of the cell organelle, export of degradation products and failure of fragmentation processes. Most LSDs are caused by primary defects in enzymes: about one-third of the LSDs is caused by deficiencies of lysosomal glycosidases (**Table 1**).

The best studied LSD is Gaucher disease (GD), where deficiency of lysosomal β-glucosidase (glucocerebrosidase; GBA) leads to prominent accumulation of the substrate glucosylceramide in lysosomes of tissue macrophages that transform into lipid-laden Gaucher cells. These viable storage macrophages are thought to underlie characteristic symptoms of GD patients such as hepatosplenomegaly and pancytopenia (shortage of red blood cells and platelets).[15] In the case of GD, the progression and severity of disease, and occurrence of particular signs and symptoms varies greatly among patients, even homozygotic twins.[16, 17] The cause for this phenotypic variation of GD is presently largely unclear, hypothetically being attributed to modifier genes and other factors. [18–22] More severely affected GD patients (type 2 and type 3) develop marked skeletal disease, fatal neurological symptoms and impaired skin permeability. The pathophysiological mechanisms driving the latter symptoms have not been elucidated yet.[14] Alternative excessive metabolism of glucosylceramide by the cytosolic β-glucosidase GBA2 as well as the formation of toxic glucosylsphingosine from accumulating glucosylceramide in lysosomes is considered to contribute to particular symptoms of GD patients.[15, 21] Mutations in the *GBA* gene are at present also the highest known risk factor for developing Parkinson's disease[23] and Dementia of Lewy Bodies.[24, 25] The surprising link between partial lysosomal GBA impairment and risk for neurodegenerative disease remains enigmatic.

Table 1. Selected lysosomal glycosidases and underlying lysosomal storage diseases.

Enzyme	Gene	Major Substrates (Glycosidic Linkage)	Associated Diseases	Clinical Features
Glucocerebrosidase (GH45, E.C. 3.2.1.21)	*GBA*	Glucosylceramide (glucose β-1-ceramide))	Gaucher Disease (OMIM # 230800 (Type I), # 230900 (Type II), # 23100 (Type III))	Type I: Hepatosplenomegaly, anemia, thrombocytopenia, bone lesions Type II: Neurodegeneration and hepatosplenomegaly in infancy Type III: Intermediate between type I and II
Acid α-glucosidase (GH31, E.C. 3.2.1.20)	*GAA*	Glycogen (glucose α-1,4-glucose)	Pompe Disease (Glycogen Storage Disease Type II) (OMIM # 232300)	Infantile form: Cardiomegaly, hepatomegaly, hypotonia, severe muscle weakness, respiratory failure Late-onset form: Progressive muscle weakness
β-galactosidase (GH35, E.C. 3.2.1.23)	*GLB1*	Ganglioside GM1, Keratan sulfate (galactose β-1,4-saccharide)	GM1 Gangliosidosis (OMIM # 230500 (Type I), # 230600 (Type II), # 230650 (Type II)) Morquio Disease Type B (Mucopolysaccharidosis type IVB, MPS 4B) (OMIM # 253010)	Type I (infantile): Pyschomotor deterioration, central nervous system complication, hepatosplenomegaly, macular cherry-red spot, skeletal dysplasia, facial dysmorphism Type II (late infantile/juvenile): Pyschomotor deterioration, seizure, localized skeletal problem Type III (adult): Localized skeletal and central nervous system involvement Morquio Disease Type B: Bone abnormality, corneal clouding, hepatomegaly
Galactocerebrosidase (GH59, E.C. 3.2.1.46)	*GALC*	Galactosylceramide (galactose β-1-ceramide)	Krabbe Disease (OMIM # 245200)	Mental and motor degeneration, extreme irritability, spasticity
α-galactosidase A (GH27, E.C. 3.2.1.22)	*GLA*	Globotriaosylceramide, galabiosylceramide (galactose α-1,4-saccharide)	Fabry Disease (OMIM # 301500)	Extreme pain and paresthesia, renal failure, hypohydrosis, corneal opacity, angiokeratoma
β-mannosidase (GH2, E.C. 3.2.1.25)	*MANBA*	Oligosaccharides (mannose β-1,4-GlcNAc)	β-mannosidosis (OMIM # 248510)	Developmental delay, mental retardation, hearing loss, angiokeratoma
Lysosomal α-mannosidase (GH38, E.C. 3.2.1.24)	*MAN2B1*	Oligosaccharides (mannose α-1-2, -1,3, and 1,6-mannose)	α-mannosidosis (OMIM # 248500)	Facial dysmorphism, psychomotor impairment, hearing loss

Table 1. (Continued)

Enzyme	Gene	Major Substrates (Glycosidic Linkage)	Associated Diseases	Clinical Features
β-glucuronidase (GH2, E.C. 3.2.1.31)	*GUSB*	Heparan sulfate, Chondroitin sulfate (glucuronic acid β-1,4-saccharide)	Sly Syndrome (Mucopolysaccharidosis Type VII, MPS 7) (OMIM # 253220)	Hepatomegaly, skeletal abnormalities, facial dysmorphism, mental impairment, hydrops fetalis (in severely affected patients)
α-L-iduronidase (GH39, E.C. 3.2.1.76)	*IDUA*	Dermatan sulfate, heparin sulfate (iduronic acid α-1,4-saccharide)	Hurler Syndrome (Mucopolysaccharidosis Type I-H, MPS 1H) (OMIM # 607014) Scheie Syndrome (MPS 1S) (OMIM # 607016) Hurler-Scheie Syndrome (MPS 1H/S) (OMIM # 607015)	Hurler: Facial dysmorphism, corneal clouding, mental retardation, hernias, hepatosplenomegaly, dysostosis multiplex Scheie: Milder phenotype than Hurler Hurler-Scheie: Intermediate in phenotype
α-L-fucosidase (GH29, E.C. 3.2.1.51)	*FUCA1*	Protein N-linked glycan (fucose α-1,2-galactose, fucose α-1,3-, 1,4-, or 1,6-GlcNAc)	Fucosidosis (OMIM # 230000)	Angiokeratoma, facial dysmorphism, progressive psychomotor degeneration, dysostosis multiplex
Hexosaminidase A/B (heterodimer) (GH 20, E.C. 3.2.1.52)	*HEXA* *HEXB*	Ganglioside GM2, oligosaccharides	Tay-Sachs Disease (HEXA deficiency, GM2 Gangliosidosis Type I) (OMIM # 272800) Sandhoff Disease (HEXB deficiency, GM2 Gangliosidosis Type II) (OMIM # 268800)	Tay-Sachs: Mental retardation, paralysis, dementia, blindness, macular cherry-red spot Sandhoff: Clinically undistinguishable from Tay-Sachs Disease
α-galactosaminidase/ α-galactosidase B (GH27, E.C. 3.2.1.49 / 22)	*NAGA*	Mucin glycoprotein from human blood group A and AB (GalNac α-1,3-glycoprotein)	Schindler Disease (OMIM # 609241 (Type I & III), # 609242 (Type II, Kanzaki Disease)	Type I (infantile form): Neuroaxonal dystrophy, vision loss, hearing loss, seisure Type II (adult form): Mild cognitive impairment, hearing loss, angiokeratoma, Type III: Intermediate between Type I and II
α-glucosaminidase (GH89, E.C. 3.2.1.50)	*NAGLU*	Heparan sulfate (GlcNAC α-1,4-saccharide)	Sanfillipo Disease Type B (Mucopolysaccharidosis type IIIB, MPS 3B)	Progressive neurodegeneration, behavioral problems
Neuraminidase (GH 33, E.C. 3.2.1.18)	*NEU1*	Glycopeptide, oligosaccharide, glycoproteins (sialic acid α-2,3-, 2,6-, or 2,8-glycoconjugates)	Sialidosis (Mucolipidosi Type I, ML I) (OMIM # 256550)	Neurologic abnormalities, facial dysmorphism, spinal deformity, macular cherry-red spot

Pompe disease is another relatively common LSD caused by deficiency of a lysosomal glycosidase, in this case the acid α-glucosidase (GAA).[26] Affected patients are characterized by abnormal lysosomal accumulation of undegraded glycogen in muscle and bone cells, leading to progressive muscle weakness, respiratory failure, and in the severe cases death in early childhood (**Table 1**).[26] GM1 gangliosidosis is an LSD in the subgroup of (glyco)sphingolipidosis, in which the primary accumulating substrate is the (glyco)sphingolipid GM1 ganglioside. It is caused by deficiency in the lysosomal β-galactosidase (GLB1).[27] The lysosomal accumulation of GM1 ganglioside affects both the central and peripheral nervous system, resulting in neurodegeneration in affected patients (**Table 1**).[28] Hurler syndrome (Mucopolysaccharidosis Type I, MPS 1)[29] and Sly syndrome (MPS 7)[30] are mucopolysaccharidoses, in which enzyme deficiency (α- L-iduronidase and β-glucuronidase, respectively) causes accumulation of glycosaminoglycans (GAGs, or mucopolysaccharides). Patients typically exhibit abnormality in bone and connective tissues, as well as in the brain (**Table 1**), due to impaired GAG metabolism. Finally, the oligosaccharidoses α- and β-mannosidosis are caused by deficiency in lysosomal α- or β-mannosidases.[31, 32] Both diseases are characterized by the accumulation of undegraded oligosaccharides deriving from N-linked glycoproteins and lipid carriers of the N-linked glycan, which manifests in clinical features such as mental retardation and facial dysmorphism (**Table 1**).

Enzyme replacement therapy The realization that a single enzyme deficiency underlies specific LSD has prompted the design of therapeutic approaches that aim to either replace the defective enzyme with a normal one, to stabilize and activate the residual mutant enzyme, or to reduce the accumulating substrates for reducing their cytotoxic effects (**Fig. 2**).[12] Gaucher disease, in particular the non-neuropathic type 1 variant, has been the frontrunner in development of therapy approaches. Building on the finding that bone marrow transplantation of GD patients, replacing the white blood cell lineage with hematopoietic stem cells capable of producing GBA competent macrophages, renders major clinical improvements, correction of macrophages in type 1 GD patients by enzyme supplementation, so-called enzyme-replacement therapy (ERT) has been actively studied.[33] First conceived by Brady and co-workers in 1960s,[34] ERT with macrophage-targeted GBA results in remarkable reduction of hepatosplenomegaly, anemia, and thrombocytopenia in Gaucher patients.[35, 36] The originally approved therapeutic enzyme was isolated from human placenta, and its complex-type N-glycans were chemically modified into terminal high-mannose N-glycans,[35] resulting in efficient uptake by (affected) macrophages

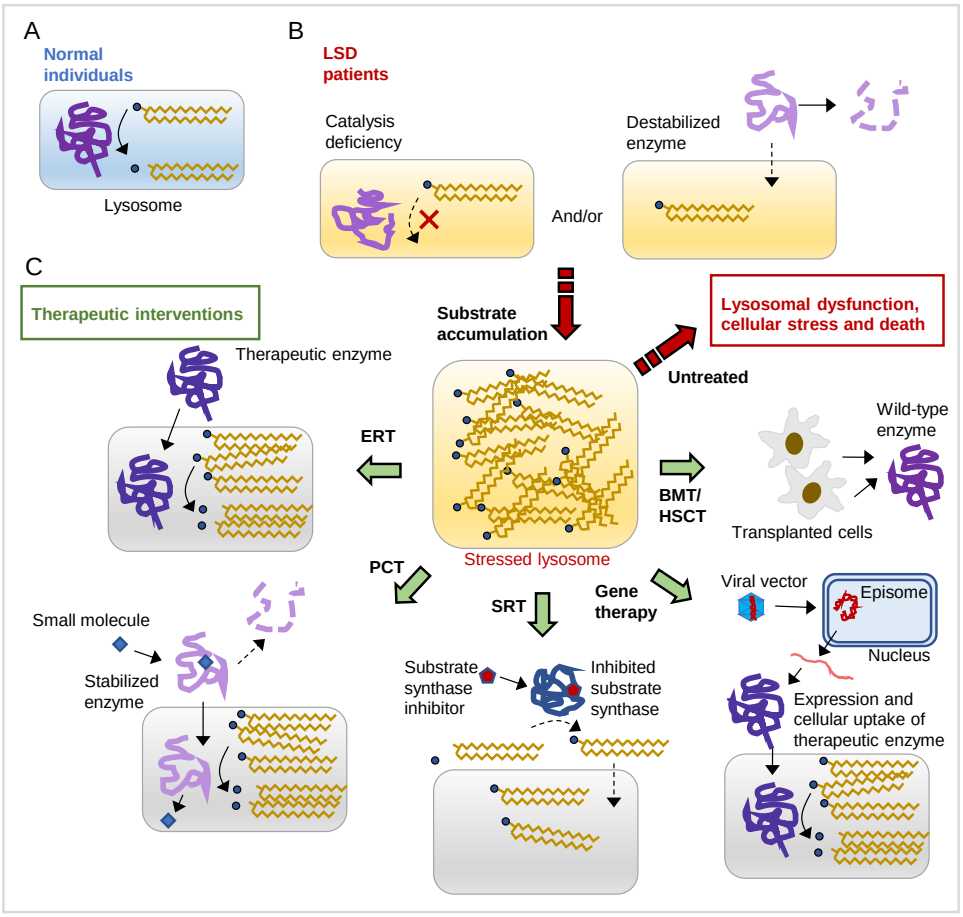


Figure 2. Lysosomal storage diseases and current therapeutic strategies. A) Substrate hydrolysis by glycosidases in normal individuals. B) Enzyme defects in lysosomal storage disease (LSD) patients lead to lysosomal dysfunction. C) Therapeutic strategies for LSDs. *ERT*, enzyme replacement therapy; *PCT*, pharmacological chaperon therapy; *SRT*, substrate reduction therapy; *BMT*, bone-marrow transplantation; *HSCT*, hematopoietic stem cell transplantation.

expressing the mannose receptors. Later therapeutic enzymes were produced in Chinese hamster ovary (CHO) cells (Imiglucerase, Cerezyme®). Now several therapeutic glucocerebrosidase preparations, produced in various cell types including carrot cells, are registered for therapeutic use in type 1 GD patients.[37, 38] The success of ERT for type 1 Gaucher disease, aided with the passing of Orphan Drug Act in the US in 1983, prompted the development of ERT for other diseases by the pharmacological industry. Approved by the Food and Drug Administration (FDA) in the USA or similar agencies in other parts of the world are therapeutic enzymes for Pompe disease (Aglucosidase alfa, Myozyme®/Lumizyme®), Fabry disease (Agalsidase alfa, Fabrazyme®

and Agalsidase beta, Replagal®), Hurler syndrome (Laronidase, Aldurazyme®),[32] and recently Sly syndrome (Vestronidase alfa, Mepsevii®)[39] and α-mannosidosis (Velmanase alfa, Lamzede®).[40]

Other therapeutic options A caveat of presently registered therapeutic enzymes is their poor penetration to certain types of tissue (including bone, muscle), but most particular the brain. At present, bone marrow transplantation or hematopoietic stem cells transplantation is investigated for Gaucher disease[41] and remain the only viable therapeutic options for some LSDs affecting the central nervous system, such as Krabbe disease and GM1 gangliosidosis.[27, 42, 43] Two other major therapeutic strategies for LSDs based on small molecules are now available on the market, namely pharmacological chaperon therapy (PCT) and substrate reduction therapy (SRT). The former makes use of small chemicals that interact with enzymes with folding defects, which stabilize the enzyme in during its maturation and promote their correct transportation to the lysosomes.[44] Such therapy is currently only available for Fabry disease,[45] while other types of chaperone therapy, such as induction of heat-shock proteins to stabilize glycosidase folding and function, are also under development.[46] For SRT, the aim is not to correct for the defective enzyme, but to reduce the amount of accumulating substrates by inhibiting the enzymes accounting for their synthesis.[47, 48] To date, Gaucher disease is the only LSD with lysosomal glycosidase deficiency with approved SRT options based on orally administered inhibitors of glucosylceramide synthase: the iminosugar Miglustat (N-butyl-deoxynojirimycin)[49] and the more potent and specific ceramide mimic Eliglustat.[50] Another actively pursued therapeutic approach for LSDs is gene therapy. Studied are several approaches, in modifying patients' hematopoietic stem cells *ex vivo* with lentiviral vectors encoding wild-type enzymes, or injecting the patient with engineered adeno-associated viral (AAV) vectors encoding the enzyme.[51, 52] The promotor and the site of injection for the later approach can be tailored according to the disease and affected tissue types, and the treatment would only require one-time injection—an obvious advantage compared to the life-long treatment regimen for ERT, PCT, and SRT. Safety remains the major concern, which has translated to strict manufacturing regulation and thus high pricing—the current major hurdle for developing such therapy.

Diagnosis Traditionally, the diagnosis of an LSD is made based on clinical symptoms, and later confirmed by genetic and biochemical analyses. However, due to the rarity and associated poor awareness of LSDs, as well as their considerable heterogeneity in clinical presentation, in most countries their correct diagnosis remains challenging.[22, 53] This prompts new-born screening

(NBS) programs in many parts of the world, aiming to detect patients before they show the first symptoms. Early treatment of many LSDs is increasingly considered to be essential to achieve maximum therapeutic outcome. The first-line screening sometimes builds on genetic testing. However, many LSDs show poor genotype-phenotype correlations, and in addition it remains challenging to interpret novel (mild) mutations as being a polymorphism or disease-causing trait.[54] Detection of biomarkers (surrogate biochemical markers of disease manifestation and/or progression) assist diagnosis, but for many LSD these are unavailable.[15, 55] In the case of LSDs caused by enzyme deficiencies, biochemical activity assays are viewed as the gold standard. In many cases these indeed allow reliable confirmation of diagnosis. Traditionally, enzyme activity assays involve measurement of residual activity in cultured patients' fibroblasts or collected white blood cells using artificial fluorogenic substrates, such as 4-methylumbelliferyl (4-MU) glycosides. Such assays may suffer from low dynamic range,[56] cross activity from other enzymes, and inability to multiplex with other enzyme assays. More recently, microfluidic-based 4-MU assay platform[57, 58] and an MS-based multiplex method using deuterium natural substrates[59, 60] have been developed and implemented in several NBS programs in the USA and other parts of the world, using dried-blood spots collected from patients as enzyme source. A major obstacle is the need for specialized laboratories employing validated measurements, still unavailable in most parts of the world. Furthermore, each method has its own limitations, as the microfluidic method does not solve the low-dynamic range issue intrinsic to the 4-MU assay, and the MS-based method requires laborious procedures and considerable technical know-how.[54]

Part II. Lysosomal glycosidases: from molecular understanding to novel chemical tools

Fundamental research in the last decades on lysosomal enzymes, including glycosidases, has revealed some commonalities among these enzymes. The hydrolases show optimal activity at acid pH, coinciding with the low lysosomal pH value. In general, lysosomal glycosidases are relatively resistant against proteolytic degradation and survive for more than one day in the lytic lysosome environment. For many of the lysosomal glycosidase the 3D-structure has been determined by means of crystallography and X-ray diffraction.

Life cycle of lysosomal glycosidases It is now well-established that lysosomal glycosidases are formed at the endoplasmic reticulum and are subsequently are routed to the lysosomes via the mannose-6-phosphate (M6P)-dependent pathway. Nascent lysosomal glycosidases are upon translation firstly targeted into the ER via their N-terminal signal peptide

(typically 20-25 amino acids), which is cleaved from the nascent preproproteins by the ER-resident signal peptidase.[61] The resulting proproteins are glycosylated at selective Asn residues (Asn-X-Ser/Thr, where X cannot be Pro or Asp) with the pre-synthesized $Glc_3Man_9GlcNac_2$ glycans transferred *en bloc* from the dolichol carrier (**Fig. 3**, step 1).[62] N-glycans prove to be crucial for the correct folding of many glycoproteins in the ER. Only after successful folding, successive trimming of terminal glucoses and of one terminal mannose in N-glycans takes place (by processing α-glucosidase I (MOGS), II (GANAB) and ER α-mannosidase (MAN1B1)). Next, the N-glycoproteins are transported to the Golgi apparatus for further modification and sorting (**Fig. 3**, step 2−4).[63] In the Golgi, the terminal α-1,2-linked mannoses in the N-glycans of lysosomal proteins are cleaved by the processing mannosidases (**Fig. 3**, step 5). This enables the terminal α-1,6-linked mannose to be modified at the 6 position with a N-acetylglucosamine (GlcNAc)-1-phosphate group, catalyzed by GlcNAc-1-phospotransferase (**Fig. 3**, step 6).[64] Next, the GlcNAc residue on the GlcNAc-1-phosphate group is released at the *trans*-Golgi network (TGN) by GlcNAc phosphodiesterase (i.e. uncovering enzyme), exposing the mannose-6-phosphate (M6P) group.[65] Two dedicated M6P receptor proteins (MPR300 in particular and MPR47 to far lesser extent) recognize the M6P groups presented on the glycosidases, and shuttle them via clathrin-coated vesicles to the early or late endosomes (**Fig. 3**, step 7−9).[61]

Some lysosomal glycosidases are co-transported to the lysosomes in a protein complex, such as the β-galactosidase (GLB1)-neuraminidase (Neu1)-cathepsin A complex.[66] Glucocerebrosidase (GBA) does not acquire M6P moieties in its glycans and contrary to other lysosomal glycosidases uses an M6P-independent lysosomal targeting pathway involving its hydrophobic association with the lysosomal integral protein 2 (LIMP2) (**Fig. 3**, step a−b).[67, 68, 69] As the early/late endosomes mature into lysosomes, gradual acidification of the lumen from pH 6.0 to below 5.0 facilitates (partial) dissociation of GBA from LIMP2 and dissociation of other lysosomal glycosidases from the MPRs (in particular MPR300) (**Fig. 3**, step c). LIMP2 stays in lysosomes, while MPRs are recycled back to the TGN via retrograde transport (**Fig. 3**, step 10).[61] An estimated 10 % of the total cellular MPRs reside at the plasma membrane, where they function in "fishing back" the secreted lysosomal glycosidases (both from the same cell or from other cells) to the lysosomes via endocytosis (**Fig. 3**, step 11−13).[61] In the lysosome, most of the glycosidases (e.g. α-galactosidase, α- L-fucosidase, α-glucosidase, β-glucuronidase, α- L-iduronidase, α-mannosidase, β-galactosidase, and galactocerebrosidase) undergo proteolytic

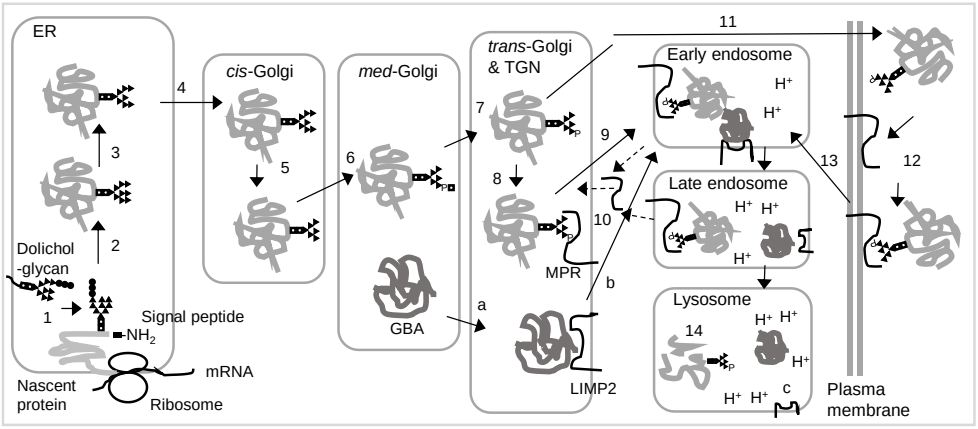


Figure 3. Lysosomal targeting of glycosidases. *TGN*, trans-Golgi network. *MPR*, mannose-6-phosphate receptor. *GBA*, glucocerebrosidase. *LIMP2*, lysosomal integral protein 2. *P*, phosphate group. Dashed lines, retrograde transport of MPR from early and late endosomes to TGN. Glycosylation on GBA is omitted for clarity.

cleavage (**Fig. 3**, step 14), producing the final mature enzymes.[70]

Catalytic mechanisms It has been observed in the first half of 20th century that most glycosidases can be categorized into two groups according to the stereochemistry outcome of their hydrolysis—either into inverting glycosidases or retaining glycosidases. These groups utilize distinct catalytic mechanisms, as firstly described by Daniel E. Koshland Jr. in 1953.[71] Inverting glycosidases typically possess a catalytic pocket 10.5 Å in width, which allow the accommodation of the glycoside substrate plus a water molecule.[72] When correctly orientated, the catalytic base deprotonate the water molecule, which attacks the electron-poor anomeric carbon by the assistance of the catalytic acid that attracts electrons from the glycosidic oxygen, distorting the glycosidic bond and disrupt the exo-anomeric effect that stabilizes the glycosidic linkage (**Fig. 4A**).[73] The reaction proceeds through a transient oxocarbenium-like transition state adopted by the substrate glycon (the sugar part), and completes in one-step where the hydroxyl replaces the aglycon (the rest of the reducing-end structures linked to the anomeric carbon) situated on the opposing side of the symmetric plane in a typical S_N2 mechanism, resulting a product with the inversion of stereochemistry (**Fig. 4A**). For retaining glycosidases, which include most lysosomal glycosidases, the S_N2 mechanism occurs twice (Koshland double-displacement mechanism), resulting in the net retention of stereochemistry on the product. The catalytic pocket is narrower (typically 5.5 Å apart), and the reaction occurs with protonation of the glycosidic oxygen by the

catalytic acid/base (typically aspartic acid/aspartate or glutamic acid/glutamate) and simultaneous nucleophilic attacks at the anomeric carbon by the catalytic nucleophile (aspartate or glutamate). This results in a transient oxocarbenium transition state, followed by a covalent substrate-enzyme intermediate (**Fig. 4B**).[72] Next, the catalytic acid/base deprotonates a water molecule, which displaces the nucleophile in a similar reaction sequence adopted by the inverting glycosidases, producing a product with retained stereochemistry (**Fig. 4B**).

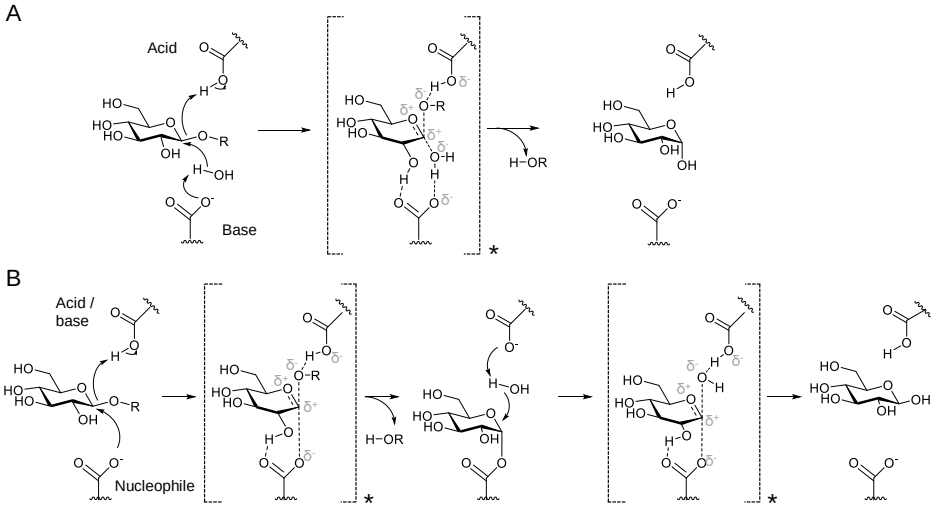


Figure 4. Reaction mechanism of exo-glycosidases towards a β-glucoside. A) Inverting mechanism. B) Retaining double-displacement mechanism. *, transition state. *R*, aglycon.

It is also known that glycosidases distort their sugar substrates during the course of their catalysis, and typical sets of conformational changes on the substrate's glycon during the reaction itinerary can be assigned for each glycosidase class (CAZy classification, discussed later).[74, 75] For example, the glucose on a β-glucoside, such as the β-glucose presented on a glucosylceramide, initially adopts a lowest-energy $^{1}C_{4}$ conformation, but as it accommodates in the active-site pocket of a retaining β-glucosidase, such as glucocerebrosidase, it is distorted to a skew-boat conformation ($^{1}S_{3}$) that facilitates both protonation by the acid/base and the nucleophilic attack by the nucleophile (**Fig. 5**). The glycon is further distorted to a half-chair conformation ($^{4}H_{3}$) during the transient transition state, and upon the formation of the new glycosidic bond with the nucleophile the glycon adopts a $^{4}C_{1}$ chair conformation, becoming an α-glucose (**Fig. 5**).[74, 75] The second half of the reaction distorts the glycon conformation in a reverse order, producing the product with identical conformation to the substrate (**Fig. 5**).

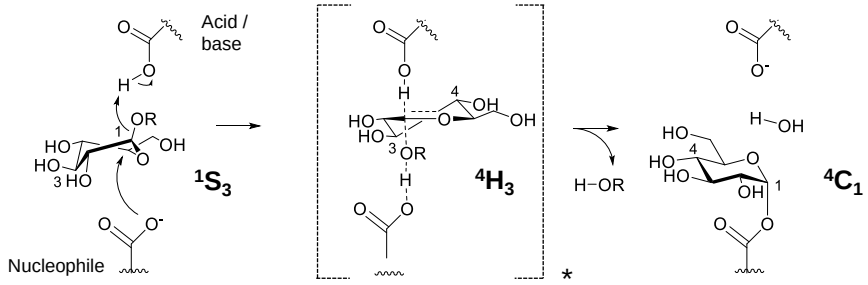


Figure 5. Distortion of substrate conformation by retaining β-glucosidase. Bold, conformation adopted by the glycon. *, transition state. *R*, aglycon.

Glycosidase can be additionally classified based on the site of hydrolysis in the substrate. Exo-glycosidase releases terminal monosaccharide from the reducing end of the substrate. Their catalytic sites is typically pocket-shaped, allowing interactions with only monosaccharide at the reducing end.[76] In contrast, endo-glycosidase have a wider cleft- or tunnel-shaped catalytic site, which can accommodate multiple sugar residues at the reducing end of the to-be-hydrolyzed glycosidic bond.[76]

Sequence-based classification In order to study glycosidases in a systematic approach, a framework named Carbohydrate-Active Enzyme (CAZy) database has been setup, which groups glycosidases into over 100 Glycoside Hydrolase (GH) families.[77] This classification system is based on amino acid sequence homology, instead of grouping glycosidases solely by their catalytic activity (the Enzyme Commission (E.C.) system). Using this system, it turns out that glycosidases classified into the same GH family usually possess similar structural features and reaction mechanisms. This in turn proves to be useful in predicting catalytic activities by proteins with known sequence, as well as the activity of a known active-site-interacting molecules towards enzymes in the same GH family. The second implication has greatly facilitated the design of glycosidase active-site interactors as well as inhibitors, and has assisted the development of activity-based probes (ABPs), as discussed in the next section.

Mechanism-based profiling Enzymes with known reaction mechanism invite the design of chemical probes that would react covalently with the enzyme's active site catalytic amino acids, allowing the detection of the enzymes through reporter moiety grafted on the probe. If the reaction depends on enzymatic activity rather than affinity alone, the probe is termed activity-

based probe (ABP). Crucial for an ABP is its ability to form covalent linkage to the target enzyme, for a period long enough to allow their detection. Key structural elements include a reactive "warhead"—typically an electrophile, a recognition element that confers specificity towards the target enzyme(s), and a reporter group for detection (or a chemical handle for subsequent installation of reporter group) (**Fig. 6A**).[78]

Known covalent inhibitors of GBA have been successfully used as scaffolds to design the desired ABPs for the enzyme. The suicide inhibitor conduritol B epoxide (CBE, **Fig. 6B**) irreversibly inactivates β-glucosidases and (with lower potency) other glycosidases in a mechanism-based manner.[79, 80] It has been exploited for the generation of a Gaucher-like mouse model,[81] and for the identification of active sites of both retaining β- and α-glucosidases.[80, 82] Cyclophellitol, a later discovered natural suicide inhibitor[83] and now synthetically available[84, 85], also reacts with β-glucosidases in a mechanism-based manner and with higher potency (**Fig. 6C**).[86, 87] Even superior potency is observed by the synthetic compound cyclophellitol aziridine, in which the epoxide on cyclophellitol is replaced by an aziridine group (**Fig. 6B**).[88]

Based on the cyclophellitol scaffold, the first true ABP for lysosomal glycosidase has been generated.[89] It was synthesized by installing an azide group at the methoxy carbon (C6 by glucopyranoside numbering; C8 by cyclophellitol numbering (**Fig. 6B**)) and the subsequent Cu(I)-catalyzed "click" reaction with a BODIPY-alkyne, which generates a BODIPY-substituted cyclophellitol (**Fig. 6D**, ABP **1**).[89] The ABPs (two BODIPY variants) label glucocerebrosidase (GBA) in a mechanism-based manner and with high potency and specificity, allowing detection of endogenous glucocerebrosidase by SDS-PAGE, fluorescence-activated cell sorting (FACS), and fluorescence microscopy, and were demonstrated to be useful in the study of GBA life cycle, inhibitor screening (for potential PCT), and diagnosis of Gaucher disease.[89] Later, a suite of cyclophellitol and cyclophellitol aziridine-based ABPs containing fluorophores and biotin have been developed for retaining β-glucosidases (GBA, non-lysosomal glucosylceramidase (GBA2), cytosolic β-glucosidase (GBA3), and lactase phlorizin hydrolase (LPH)) (**Fig. 6D**, ABP **2**, **3**),[90] GH27 α-galactosidases,[91] GH29 α- L-fucosidases,[92] and galactocerebrosidase (GALC).[93] The designing principle relies on tuning the cyclophellitol scaffold to match the configuration of the substrates from the target glycosidase, and this approach indeed has proved to be successful in the specific detection of enzymes in particular GH family.[94] Structural-activity-relationship (SAR)

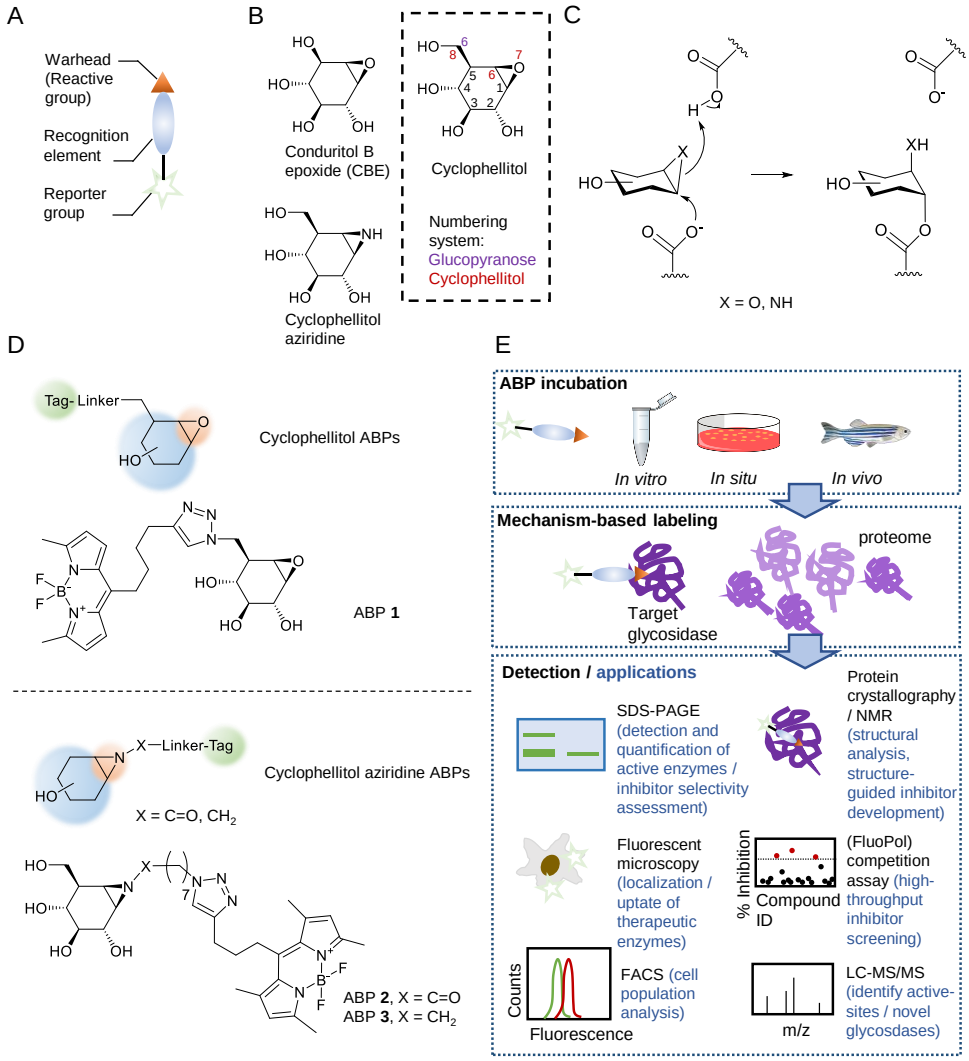


Figure 6. Mechanism-based activity profiling for retaining exo-glycosidases. A) Structural elements of activity-based probe (ABP). B) Structures of selected irreversible glycosidase inhibitors. C) Mechanism-based inactivation by cyclophellitol- and cyclophellitol aziridine-based structures. D) Structures of glycosidase activity-based probes (ABPs) based on cyclophellitol- (upper) and cyclophellitol aziridine (lower) scaffolds. ABP **1**, epoxide ABP for glucocerebrosidase[89]; ABP **2** and **3**, aziridine ABPs for β-glucosidases[90, 96]. Color scheme same as in A). E) Workflow for activity-based protein profiling (ABPP).

studies has confirmed this observation, as deleting the C2- or C4-hydroxyls on the aziridine ABP render the probes unselective towards specific glycosidase family.[95] SAR study also revealed that ABPs with the N-alkyl linker is equally potent than their N-acyl counterparts, while being more

synthetically tractable and more stable during their synthesis and purification[96] Meanwhile, this designing strategy has been extended for other (lysosomal) retaining glycosidases, for example GH31 α-glucosidases (Chapter 4, this thesis),[97] GH2 β-glucuronidases (Chapter 4, this thesis),[98] GH39 α-L-iduronidases (Chapter 5, this thesis),[99] GH38 α-mannosidases (Chapter 6, this thesis), GH2 β-mannosidases (Chapter 6, this thesis), and GH59 and GH35 β-galactosidases (Chapter 7, this thesis).

Applications of glycosidase activity-based probes Developed applications for the mechanism-based glycosidase inhibitors and ABPs include specific visualization of endogenous glycosidases in complex samples such as cell lysates and tissue homogenates based on their distinct pH range and molecular weight using SDS-PAGE and fluorescence scanning,[100] in intact cells using fluorescent microscopy[100] and fluorescence-activated cell sorting (FACS),[89] and in whole tissue sections by fluorescent microscopy (**Fig. 6E**).[93, 101, 102] These methods should prove to be useful in diagnostic conformation of LSDs, in investigations on outcome of lysosomal glycosidase mutations, and in the present development of LSD therapies such as PCT and gene therapy. Therapeutic enzymes for LSDs can also be pre-labeled by the ABPs, which allows their detection by these mentioned methods[99, 103]—crucial in monitoring ERT efficacy in pre-clinical and clinical settings. The ABPs can also be used to assess potency and specificity of known glycosidase inhibitors in intact cells or animals (Chapter 2, this thesis)[104] or in an *in vitro* high-throughput setup such as the fluorescence polarization (FluoPol) assay for discovering potential glycosidase interactors (inhibitors/stabilizers/activators) for PCT (**Fig. 6E**).[105] The covalent nature of labeling also allow structural analysis of (mutant) glycosidase (with or without ABP) by protein crystallography[93, 97-99] or NMR[106], which would offer structural insights crucial for the development of small-molecule chaperones for PCT (**Fig. 6E**). On the other hand, mechanism-based inhibitors, owing to their substrate specificity, can also be used to generate LSD cell and animal models (Chapter 3, this thesis).[107] Last but not least, ABPs appended with a biotin moiety would allow affinity-enrichment and identification of the labeled targets by LC-MS(/MS) based proteomics, which is useful in identification of glycosidase catalytic residues[108] or of unknown glycosidases (Chapter 7, this thesis) (**Fig. 6E**), with the later has already been exploited in plants.[109, 110] This thesis aims to illustrate a number of applications by the mechanism-based inhibitors and ABPs in the LSD field, as well as to present detailed biochemical characterization of novel ABPs targeting other (LSD-relevant) retaining exo-glycosidases (Chapter 4–7, this thesis).

References

1 Ghazarian H, Idoni B & Oppenheimer SB (2011) A glycobiology review: carbohydrates, lectins and implications in cancer therapeutics. *Acta Histochem*, **113**, 236–247.

2 Moremen KW, Tiemeyer M & Nairn AV (2012) Vertebrate protein glycosylation: diversity, synthesis and function. *Nat Rev Mol Cell Biol*, **13**, 448–62.

3 Laine RA (1994) A calculation of all possible oligosaccharide isomers both branched and linear yields 1.05 x 10(12) structures for a reducing hexasaccharide: the Isomer Barrier to development of single-method saccharide sequencing or synthesis systems. *Glycobiology* **4**, 759–767.

4 Cummings RD (2009) The repertoire of glycan determinants in the human glycome. *Mol Biosyst* **5**, 1087–104.

5 Davies GJ, Gloster TM & Henrissat B (2005) Recent structural insights into the expanding world of carbohydrate-active enzymes. *Curr Opin Struct Biol* **15**, 637–645.

6 Lairson LL, Henrissat B, Davies GJ & Withers SG (2008) Glycosyltransferases: structures, functions, and mechanisms. *Annu Rev Biochem* **77**, 521–555.

7 Wolfenden R, Lu X & Yang G (1998) Spontaneous Hydrolysis of Glycosides. *J Am Chem Soc* **120**, 6814–6815.

8 De Duve C, Pressman BC, Gianetto R, Wattlaux R & Appelmans F (1955) Tissue fractionation studies. 6. Intracellular distribution patterns of enzymes in rat-liver tissue. *Biochem J* **60**, 604–617.

9 Hers HG (1963) alpha-Glucosidase deficiency in generalized glycogenstorage disease (Pompe's disease). *Biochem J* **86**, 11–16.

10 Brady RO, Kanfer JN, Bradley RM & Shapiro D (1966) Demonstration of a deficiency of glucocerebroside-cleaving enzyme in Gaucher's disease. *J Clin Invest* **45**, 1112–1115.

11 Meikle PJ, Hopwood JJ, Clague AE & Carey WF (1999) Prevalence of lysosomal storage disorders. *JAMA* **281**, 249–254.

12 Platt FM, d'Azzo A, Davidson BL, Neufeld EF & Tifft CJ (2018) Lysosomal storage diseases. *Nat Rev Dis Primers* **4**, 27.

13 Zimran A, Gelbart T, Westwood B, Grabowski GA & Beutler E (1991) High frequency of the Gaucher disease mutation at nucleotide 1226 among Ashkenazi Jews. *Am J Hum Genet* **49**, 855–859.

14 Grabowski GA, Petsko GA & Kolodny EH (2013) Gaucher disease. In Valle D, Beaudet AL, Vogelstein B, Kinzler KW, Antonarakis SE, Ballabio A, Gibson KM & Mitchell G (Eds), *The Online Metabolic and Molecular Bases of Inherited Disease*. New York, NY: McGraw-Hill Medical.

15 Ferraz MJ, Kallemeijn WW, Mirzaian M, Herrera Moro D, Marques A, Wisse P, Boot RG, Willems LI, Overkleeft HS & Aerts JM (2014) Gaucher disease and Fabry disease: New markers and insights in pathophysiologyfor two distinct glycosphingolipidoses. *BBA Mol Cell Biol Lipids* **1841**, 811–825.

16 Lachmann, RH, Grant, IR, Halsall, D, & Cox, TM (2004) Twin pairs showing discordance of phenotype in adult Gaucher's disease. *QJM* **97**, 199–204.

17 Biegstraaten, M, van Schaik, IN, Aerts, JM, Langeveld, M, Mannens, MM, Bour, LJ, Sidransky, E, Tayebi, N, Fitzgibbon, E, & Hollak, CE (2011) A monozygotic twin pair with highly discordant Gaucher phenotypes. *Blood Cells Mol Dis* **46**, 39–41.

18 Ballabio A & Gieselmann V (2009) Lysosomal disorders: from storage to cellular damage. *Biochim Biophys Acta* **1793**, 684-96.

19 Cox TM & Cachón-González MB (2012) The cellular pathology of lysosomal diseases. *J Pathol* **226**, 241–254.

20 Parenti G, Andria G & Ballabio A (2015) Lysosomal storage diseases: from pathophysiology to therapy. *Annu Rev Med* **66**, 471–486.

21 Aerts JMFG, Guimaraes Da Lomba Ferraz MJ, Mirzaian M, Gaspar P, Oussoren SV, Wisse P, Kuo CL, Lelieveld LT, Kytidou K, Hazeu MD, Boer DEC, Meijer R, Lienden MJC van der, Herrera D, Gabriel TL, Aten J, Overkleeft HS, Eijk MC van, Boot RG & Marques ARA (2017) Lysosomal Storage Diseases. For Better or Worse: Adapting to Defective Lysosomal Glycosphingolipid Breakdown. In *eLS*. New York, NY: John Wiley & Sons, Ltd.

22 Platt FM (2018) Emptying the stores: lysosomal diseases and therapeutic strategies. *Nat Rev Drug Discov* **17**, 133–150.

23 Sidransky E, Nalls MA, Aasly JO, Aharon-Peretz J, Annesi G, Barbosa ER, Bar-Shira A, Berg D, Bras J, Brice A, Chen CM, Clark LN, Condroyer C, De Marco EV, Dürr A, Eblan MJ, Fahn S, Farrer MJ, Fung HC, Gan-Or Z, Gasser T, Gershoni-Baruch R, Giladi N, Griffith A, Gurevich T, Januario C, Kropp P, Lang AE, Lee-Chen GJ, Lesage S, Marder K, Mata IF, Mirelman A, Mitsui J, Mizuta I, Nicoletti G, Oliveira C, Ottman R, Orr-Urtreger A, Pereira LV, Quattrone A, Rogaeva E, Rolfs A, Rosenbaum H, Rozenberg R, Samii A, Samaddar T, Schulte C, Sharma M, Singleton A, Spitz M, Tan EK, Tayebi N, Toda T, Troiano AR, Tsuji S, Wittstock M, Wolfsberg TG, Wu YR, Zabetian CP, Zhao Y & Ziegler SG (2009) Multicenter analysis of glucocerebrosidase mutations in Parkinson's disease. *N Engl J Med* **361**, 1651–1661.

24 Nalls MA, Duran R, Lopez G, Kurzawa-Akanbi M, McKeith IG, Chinnery PF, Morris CM, Theuns J, Crosiers D, Cras P, Engelborghs S, De Deyn PP, Van Broeckhoven C, Mann DM, Snowden J, Pickering-Brown S, Halliwell N, Davidson Y, Gibbons L, Harris J, Sheerin UM, Bras J, Hardy J, Clark L, Marder K, Honig LS, Berg D, Maetzler W, Brockmann K, Gasser T, Novellino F, Quattrone A, Annesi G, De Marco EV, Rogaeva E, Masellis M, Black SE, Bilbao JM, Foroud T, Ghetti B, Nichols WC, Pankratz N, Halliday G, Lesage S, Klebe S, Durr A, Duyckaerts C, Brice A, Giasson BI, Trojanowski JQ, Hurtig HI, Tayebi N, Landazabal C, Knight MA, Keller M, Singleton AB, Wolfsberg TG & Sidransky E (2013) A multicenter study of glucocerebrosidase mutations in dementia with Lewy bodies. *JAMA Neurol* **70**, 727–735.

25 Aflaki E, Westbroek W & Sidransky E (2017) The Complicated Relationship between Gaucher Disease and Parkinsonism: Insights from a Rare Disease. *Neuron* **93**, 737–746.

26 Hirschhorn R, Reuser AJ (2001) Glycogen storage disease type II: acid alpha-glucosidase (acid maltase) deficiency. In Scriver CR, Beaudet A, Sly WS & Valle D (Eds) *The Metabolic and Molecular Basis of Inherited Disease*. (pp 3389–3420) New York, NY: McGraw-Hill.

27 Okada S & O'Brien JS (1968) Generalized gangliosidosis: beta-galactosidase deficiency. *Science* **160**, 1002–1004.

28 Brunetti-Pierri N & Scaglia F (2008) GM1 gangliosidosis: review of clinical, molecular, and therapeutic aspects. *Mol Genet Metab* **94**, 391–396.

29 Neufeld EF & Muenzer J (2001) The mucopolysaccharidosis. In Scriver CR, Beaudet AL, Sly WS & Valle D (Eds) *The Metabolic and Molecular Bases of Inherited Disease*. (pp 3421–3452) New York, NY: McGraw-Hill.

30 Sly WS, Quinton BA, McAlister WH & Rimoin DL (1973) Beta glucuronidase deficiency: report of clinical, radiologic, and biochemical features of a new mucopolysaccharidosis. *J Pediatr* **82**, 249–257.

31 Malm D & Nilssen Ø (2008) Alpha-mannosidosis. *Orphanet J Rare Dis* **3**, 21.

32 Percheron F, Foglietti MJ, Bernard M & Ricard B (1992) Mammalian beta-D-mannosidase and beta-mannosidosis. *Biochimie* **74**, 5–11.

33 Desnick RJ & Schuchman EH (2012) Enzyme replacement therapy for lysosomal diseases: lessons from 20 years of experience and remaining challenges. *Annu Rev Genomics Hum Genet* **13**, 307–335.

34 Brady RO (1966) The sphingolipidoses. *N Engl J Med* **275**, 312–318.

35 Brady, RO (2003) Enzyme replacement therapy: Conception, chaos and culmination. *Philos Trans R Soc Lond B Biol Sci* **358**, 915–919.

36 Barton NW, Brady RO, Dambrosia JM, Di Bisceglie AM, Doppelt SH, Hill SC, Mankin HJ, Murray GJ, Parker RI, Argoff CE, Grewal RP, & Yu K-T (1991) Replacement therapy for inherited enzyme deficiency--macrophage-targeted glucocerebrosidase for Gaucher's disease. *N Engl J Med* **324**, 1464–1470.

37 Burrow TA & Grabowski GA (2011) Velaglucerase alfa in the treatment of Gaucher disease type 1. *Clin Investig (Lond)* **1**, 285–293.

38 Grabowski GA, Golembo M & Shaaltiel Y (2014) Taliglucerase alfa: an enzyme replacement therapy using plant cell expression technology. *Mol Genet Metab* **112**, 1–8.

39 Harmatz P, Whitley CB, Wang RY, Bauer M, Song W, Haller C & Kakkis E (2018) A novel Blind Start study design to investigate vestronidase alfa for mucopolysaccharidosis VII, an ultra-rare genetic disease. *Mol Genet Metab* **123**, 488–494.

40 Harmatz P, Cattaneo F, Ardigò D, Geraci S, Hennermann JB, Guffon N, Lund A, Hendriksz CJ & Borgwardt L (2018) Enzyme replacement therapy with velmanase alfa (human recombinant alpha-mannosidase): Novel global treatment response model and outcomes in patients with alpha-mannosidosis. *Mol Genet Metab* **124**, 152–160.

41 Dahl M, Doyle A, Olsson K, Månsson JE, Marques ARA, Mirzaian M, Aerts JM, Ehinger M, Rothe M, Modlich U, Schambach A & Karlsson S (2015) Lentiviral gene therapy using cellular promoters cures type 1 Gaucher disease in mice. *Mol Ther* **23**, 835–844.

42 Krivit W, Peters C & Shapiro EG (1999) Bone marrow transplantation as effective treatment of central nervous system disease in globoid cell leukodystrophy, metachromatic leukodystrophy, adrenoleukodystrophy, mannosidosis, fucosidosis, aspartylglucosaminuria, Hurler, Maroteaux-Lamy, and Sly syndromes, and Gaucher disease type III. *Curr Opin Neurol* **12**, 167–176.

43 Won JS, Singh AK & Singh I (2016) Biochemical, cell biological, pathological, and therapeutic aspects of Krabbe's disease. *J Neurosci Res* **94**, 990–1006.

44 Boyd RE, Lee G, Rybczynski P, Benjamin ER, Khanna R, Wustman BA & Valenzano KJ (2013) Pharmacological chaperones as therapeutics for lysosomal storage diseases. *J Med Chem* **56**, 2705–25.

45 Hughes DA, Nicholls K, Shankar SP, Sunder-Plassmann G, Koeller D, Nedd K, Vockley G, Hamazaki T, Lachmann R, Ohashi T, Olivotto I, Sakai N, Deegan P, Dimmock D, Eyskens F, Germain DP, Goker-Alpan O, Hachulla E, Jovanovic A, Lourenco CM, Narita I, Thomas M, Wilcox WR, Bichet DG, Schiffmann R, Ludington E, Viereck C, Kirk J, Yu J, Johnson F, Boudes P, Benjamin ER, Lockhart DJ, Barlow C, Skuban N, Castelli JP, Barth J & Feldt-Rasmussen U (2017) Oral pharmacological chaperone migalastat compared with enzyme replacement therapy in Fabry disease: 18-month results from the randomised phase III ATTRACT study. *J Med Genet* 54, 288–296.

46 Kirkegaard T, Gray J, Priestman DA, Wallom KL, Atkins J, Olsen OD, Klein A, Drndarski S, Petersen NH, Ingemann L, Smith DA, Morris L, Bornæs C, Jørgensen SH, Williams I, Hinsby A, Arenz C, Begley D, Jäättelä M & Platt FM (2016) Heat shock protein-based therapy as a potential candidate for treating the sphingolipidoses. *Sci Transl Med* **8**, 355ra118.

47 Radin NS (1996) Treatment of Gaucher disease with an enzyme inhibitor. *Glycoconj J* **13**, 153–157.

48 Coutinho MF, Santos JI & Alves S (2016) Less Is More: Substrate Reduction Therapy for Lysosomal Storage Disorders. *Int J Mol Sci* **17**, pii: E1065.

49 Aerts JMFG, Hollak CEM, Boot RG, Groener JEM & Maas M (2006) Substrate reduction therapy of glycosphingolipid storage disorders. *J Inherit Metab Dis* **29**, 449–456.

50 McEachern KA, Fung J, Komarnitsky S, Siegel CS, Chuang WL, Hutto E, Shayman JA, Grabowski GA, Aerts JM, Cheng SH, Copeland DP, Marshall J (2007) A specific and potent inhibitor of glucosylceramide synthase for substrate inhibition therapy of Gaucher disease. *Mol Genet Metab* **91**, 259–267.

51 Cheng SH (2014) Gene therapy for the neurological manifestations in lysosomal storage disorders. *J Lipid Res* **55**, 1827–1838.

52 Biffi A (2016) Gene therapy for lysosomal storage disorders: a good start. *Hum Mol Genet* **25**, R65–75.

53 Matern D, Gavrilov D, Oglesbee D, Raymond K, Rinaldo P & Tortorelli S (2015) Newborn screening for lysosomal storage disorders. *Semin Perinatol* **39**, 206–216.

54 Schielen PCJI, Kemper EA & Gelb MH (2017) Newborn Screening for Lysosomal Storage Diseases: A Concise Review of the Literature on Screening Methods, Therapeutic Possibilities and Regional Programs. *Int J Neonatal Screen* **3**, pii: 6.

55 Aerts JM, Kallemeijn WW, Wegdam W, Joao Ferraz M, van Breemen MJ, Dekker N, Kramer G, Poorthuis BJ, Groener JE, Cox-Brinkman J, Rombach SM, Hollak CE, Linthorst GE, Witte MD, Gold H, van der Marel GA, Overkleeft HS & Boot RG (2011) Biomarkers in the diagnosis of lysosomal storage disorders: Proteins, lipids, and inhibodies. *J Inherit Metab Dis* **34**, 605–619.

56 Kumar AB, Masi S, Ghomashchi F, Chennamaneni NK, Ito M, Scott CR, Turecek F, Gelb MH & Spacil Z (2015) Tandem Mass Spectrometry Has a Larger Analytical Range than Fluorescence Assays of Lysosomal Enzymes: Application to Newborn Screening and Diagnosis of Mucopolysaccharidoses Types II, IVA, and VI. *Clin Chem* **61**, 1363−71.

57 Sista RS, Eckhardt AE, Wang T, Graham C, Rouse JL, Norton SM, Srinivasan V, Pollack MG, Tolun AA, Bali D, Millington DS & Pamula VK (2011) Digital microfluidic platform for multiplexing enzyme assays: implications for lysosomal storage disease screening in newborns. *Clin Chem* **57**, 444−451.

58 Hopkins PV, Klug T, Vermette L, Raburn-Miller J, Kiesling J & Rogers S (2018) Incidence of 4 Lysosomal Storage Disorders From 4 Years of Newborn Screening. *JAMA Pediatr* **172**, 696−697.

59 Li Y, Scott CR, Chamoles NA, Ghavami A, Pinto BM, Turecek F & Gelb MH (2004) Direct multiplex assay of lysosomal enzymes in dried blood spots for newborn screening. *Clin Chem* **50**, 1785−1796.

60 Spacil Z, Tatipaka H, Barcenas M, Scott CR, Turecek F & Gelb MH (2013) High-throughput assay of 9 lysosomal enzymes for newborn screening. *Clin Chem* **59**, 502−511.

61 Braulke T & Bonifacino JS (2009) Sorting of lysosomal proteins. *Biochim Biophys Acta* **1793**, 605−614.

62 Helenius A & Aebi M (2001) Intracellular functions of N-linked glycans. *Science* **291**, 2364−2369.

63 Ruddock LW & Molinari M (2006) N-glycan processing in ER quality control. *J Cell Sci* **119**, 4373−4380.

64 Reitman ML & Kornfeld S (1981) Lysosomal enzyme targeting. N-Acetylglucosaminylphosphotransferase selectively phosphorylates native lysosomal enzymes. *J Biol Chem* **256**, 11977−11980.

65 Rohrer J & Kornfeld R (2001) Lysosomal hydrolase mannose 6-phosphate uncovering enzyme resides in the trans-Golgi network. *Mol Biol Cell* **12**, 1623−1631.

66 Galjart NJ, Gillemans N, Harris A, van der Horst GT, Verheijen FW, Galjaard H & d'Azzo A (1988) Expression of cDNA encoding the human "protective protein" associated with lysosomal beta-galactosidase and neuraminidase: homology to yeast proteases. *Cell* **54**, 755−764.

67 Aerts JM, Schram AW, Strijland A, van Weely S, Jonsson LM, Tager JM, Sorrell SH, Ginns EI, Barranger JA & Murray GJ (1988) Glucocerebrosidase, a lysosomal enzyme that does not undergo oligosaccharide phosphorylation. *Biochim Biophys Acta* **964**, 303−208.

68 Reczek D, Schwake M, Schröder J, Hughes H, Blanz J, Jin X, Brondyk W, Van Patten S, Edmunds T & Saftig P (2007) LIMP-2 is a receptor for lysosomal mannose-6-phosphate-independent targeting of beta-glucocerebrosidase. *Cell* **131**, 770−783.

69 Zunke F, Andresen L, Wesseler S, Groth J, Arnold P, Rothaug M, Mazzulli JR, Krainc D, Blanz J & Saftig P, Schwake M (2016) Characterization of the complex formed byβ-glucocerebrosidase and the lysosomal integral membrane protein type-2. *Proc Natl Acad Sci USA* **113**, 3791−3796.

70 Hasilik A (1992) The early and late processing of lysosomal enzymes: proteolysis and compartmentation. *Experientia* **48**, 130−151.

71 Koshland, DE (1953). Stereochemistry and mechanism of enzymatic reactions. *Biol Rev Camb Philos Soc* **28**, 416–436.

72 Zechel DL & Withers SG (2000) Glycosidase mechanisms: anatomy of a finely tuned catalyst. *Acc Chem Res* **33**, 11−18.

73 Heightman TD & Vasella AT (1999) Recent Insights into Inhibition, Structure, and Mechanism of Configuration-Retaining Glycosidases. *Angew Chem Int Ed Engl* **38**, 750−770.

74 Davies, GJ, Ducros, VM, Varrot, A, & Zechel, DL (2003) Mapping the conformational itinerary of β-glycosidases by X-ray crystallography. *Biochem Soc Trans* **31**, 523–527.

75 Davies GJ, Planas A & Rovira C (2012) Conformational analyses of the reaction coordinate of glycosidases. *Acc Chem Res* **45**, 308−316.

76 Davies G & Henrissat B (1995) Structures and mechanisms of glycosyl hydrolases. *Structure* **3**, 853−859.

77 Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V & Henrissat B (2009) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res* **37**, D233−238.

78 Cravatt BF, Wright AT & Kozarich JW (2008) Activity-based protein profiling: from enzyme chemistry to proteomic chemistry. *Annu Rev Biochem* **77**, 383−414.

79 Legler G (1966) Studies on the action mechanism of glycoside splitting anzymes, I. Presentation and properties of specific inhibitors. *Hoppe Seylers Z Physiol Chem* **345**, 197−214.

80 Legler G (1990) Glycoside hydrolases: mechanistic information from studies with reversible and irreversible inhibitors. *Adv Carbohydr Chem Biochem* **48**, 319−384.

81 Kanfer JN, Legler G, Sullivan J, Raghavan SS & Mumford RA (1975) The Gaucher mouse. *Biochem Biophys Res Commun* **67**, 85−90.

82 Hermans MM, Kroos MA, van Beeumen J, Oostra BA & Reuser AJ (1991) Human lysosomal alpha-glucosidase. Characterization of the catalytic site. *J Biol Chem* **266**, 13507−13512.

83 Atsumi S, Umezawa K, Iinuma H, Naganawa H, Nakamura H, Iitaka Y & Takeuchi T (1990) Production, isolation and structure determination of a novel beta-glucosidase inhibitor, cyclophellitol, from Phellinus sp. J *Antibiot (Tokyo)* **43**, 49−53.

84 Hansen FG, Bundgaard E & Madsen R (2005) A short synthesis of (+)-cyclophellitol. *J Org Chem* **70**, 10139−10142.

85 Li K, Jiang J, Witte MD, Kallemeijn WW, van den Elst H, Wong C, Chander SD, Hoogendoorn S, Beenakker TJ, Codée JD, Aerts JM, van der Marel GA & Overkleeft HS (2014) Synthesis of Cyclophellitol, Cyclophellitol Aziridine, and Their Tagged Derivatives. *Eur J Org Chem* **2014**, 6030−6043.

86 Atsumi S, Iinuma H, Nosaka C & Umezawa K (1990) Biological activities of cyclophellitol. *J Antibiot (Tokyo)* **43**, 1579−1585.

87 Withers SG & Umezawa K (1991) Cyclophellitol: a naturally occurring mechanism-based inactivator of beta-glucosidases. *Biochem Biophys Res Commun* **177**, 532−537.

88 Nakata M, Chong C, Niwata Y, Toshima K & Tatsuta K (1993) A family of cyclophellitol analogs: synthesis and evaluation. *J Antibiot (Tokyo)* **46**, 1919−1922.

89 Witte MD, Kallemeijn WW, Aten J, Li KY, Strijland A, Donker-Koopman WE, van den Nieuwendijk AM, Bleijlevens B, Kramer G, Florea BI, Hooibrink B, Hollak CE, Ottenhoff R, Boot RG, van der Marel GA, Overkleeft HS, & Aerts JM (2010) Ultrasensitive in situ visualization of active glucocerebrosidase molecules. *Nat Chem Biol* **6**, 907−913.

90 Kallemeijn WW, Li KY, Witte MD, Marques ARA, Aten J, Scheij S, Jiang J, Willems LI, Voorn-Brouwer TM, Van Roomen CPAA, Ottenhoff R, Boot RG, Van Den Elst H, Walvoort MTC, Florea BI, Codée JDC, Van Der Marel GA, Aerts JMFG & Overkleeft HS (2012) Novel activity-based probes for broad-spectrum profiling of retaining β-exoglucosidases in situ and in vivo. *Angew Chemie Int Ed* **51**, 12529–12533.

91 Willems LI, Beenakker TJ, Murray B, Scheij S, Kallemeijn WW, Boot RG, Verhoek M, Donker-Koopman WE, Ferraz MJ, van Rijssel ER, Florea BI, Codée JD, van der Marel GA, Aerts JM & Overkleeft HS (2014) Potent and selective activity-based probes for GH27 human retaining α-galactosidases. *J Am Chem Soc* **136**, 11622−11625.

92 Jiang J, Kallemeijn WW, Wright DW, van den Nieuwendijk AMCH, Rohde VC, Folch EC, van den Elst H, Florea BI, Scheij S, Donker-Koopman WE, Verhoek M, Li N, Schürmann M, Mink D, Boot RG, Codée JDC, van der Marel GA, Davies GJ, Aerts JMFG & Overkleeft HS (2015) In vitro and in vivo comparative and competitive activity-based protein profiling of GH29 α-l-fucosidases. *Chem Sci* **6**, 2782−2789.

93 Marques AR, Willems LI, Herrera Moro D, Florea BI, Scheij S, Ottenhoff R, van Roomen CP, Verhoek M, Nelson JK, Kallemeijn WW, Biela-Banas A, Martin OR, Cachón-González MB, Kim NN, Cox TM, Boot RG, Overkleeft HS & Aerts JM (2017) A Specific Activity-Based Probe to Monitor Family GH59 Galactosylceramidase, the Enzyme Deficient in Krabbe Disease. *Chembiochem* **18**, 402–412.

94 Jiang J, Artola M, Beenakker TJ, Schröder SP, Petracca R, de Boer C, Aerts JM, van der Marel GA, Codée JD & Overkleeft HS (2016) The Synthesis of Cyclophellitol-Aziridine and Its Configurational and Functional Isomers. *Eur J Inorg Chem* **2016**, 3671−3678.

95 Schröder SP, van de Sande JW, Kallemeijn WW, Kuo C-L, Artola M, van Rooden EJ, Jiang J, Beenakker TJ, Florea BI, Offen WA, Davies G, Minnaard AJ, Aerts H, Codee JD, van der Marel G & Overkleeft HS (2017) Towards broad spectrum activity-based glycosidase probes: synthesis and evaluation of deoxygenated cyclophellitol aziridines. *Chem Commun* **53**, 12528–12531.

96 Jiang J, Beenakker TJ, Kallemeijn WW, van der Marel GA, van den Elst H, Codée JD, Aerts JM & Overkleeft HS (2015) Comparing Cyclophellitol N-Alkyl and N-Acyl Cyclophellitol Aziridines as Activity-Based Glycosidase Probes. *Chemistry* **21**, 10861−10869.

97 Jiang J, Kuo CL, Wu L, Franke C, Kallemeijn WW, Florea BI, Van Meel E, Van Der Marel GA, Codée JDC, Boot RG, Davies GJ, Overkleeft HS & Aerts JMFG (2016) Detection of active mammalian GH31 α-glucosidases in health and disease using in-class, broad-spectrum activity-based probes. *ACS Cent Sci* **2**, 351–358.

98 Wu L, Jiang J, Jin Y, Kallemeijn WW, Kuo C-L, Artola M, Dai W, van Elk C, van Eijk M, van der Marel GA, Codee JDC, Florea BI, Aerts JMFG, Overkleeft HS & Davies GJ (2017) Activity-based probes for functional interrogation of retaining β-glucuronidases. *Nat Chem Biol* **13**, 867–873.

99 Artola M, Kuo CL, McMahon SA, Oehler V, Hansen T, van der Lienden M, He X, van den Elst H, Florea BI, Kermode AR, van der Marel GA, Gloster TM, Codée JDC, Overkleeft HS & Aerts JMFG (2018) New Irreversible α-l-Iduronidase Inhibitors and Activity-Based Probes. *Chemistry* **24**, 19081−19088.

100 Kuo CL, van Meel E, Kytidou K, Kallemeijn WW, Witte M, Overkleeft HS, Artola ME & Aerts JM (2018) Activity-Based Probes for Glycosidases: Profiling and Other Applications. *Methods Enzymol* **598**, 217−235.

101 Kallemeijn WW, Witte MD, Wennekes T, Aerts JM (2014) Mechanism-based inhibitors of

glycosidases: design and applications. Adv. Carbohydr. *Chem Biochem* **71**, 297−338.

102 Herrera Moro Chao D, Kallemeijn WW, Marques AR, Orre M, Ottenhoff R, van Roomen C, Foppen E, Renner MC, Moeton M, van Eijk M, Boot RG, Kamphuis W, Hol EM, Aten J, Overkleeft HS, Kalsbeek A & Aerts JM (2015) Visualization of Active Glucocerebrosidase in Rodent Brain with High Spatial Resolution following In Situ Labeling with Fluorescent Activity Based Probes. *PLoS One* **10**, e0138107.

103 Kallemeijn WW, Scheij S, Hoogendoorn S, Witte MD, Herrera Moro Chao D, van Roomen CP, Ottenhoff R, Overkleeft HS, Boot RG & Aerts JM (2017) Investigations on therapeutic glucocerebrosidases through paired detection with fluorescent activity-based probes. *PLoS One* **12**, e0170268.

104 Kuo CL, Kallemeijn WW, Lelieveld LT, Mirzaian M, Zoutendijk I, Vardi A, Futerman AH, Meijer AH, Spaink HP, Overkleeft HS, Aerts JMFG & Artola M (2019) In vivo inactivation of glycosidases by conduritol B epoxide and cyclophellitol as revealed by activity-based protein profiling. *FEBS J* **286**, 584−600.

105 Lahav D, Liu B, Van Den Berg RJBHN, Van Den Nieuwendijk AMCH, Wennekes T, Ghisaidoobe AT, Breen I, Ferraz MJ, Kuo CL, Wu L, Geurink PP, Ovaa H, Van Der Marel GA, Van Der Stelt M, Boot RG, Davies GJ, Aerts JMFG & Overkleeft HS (2017) A Fluorescence Polarization Activity-Based Protein Profiling Assay in the Discovery of Potent, Selective Inhibitors for Human Nonlysosomal Glucosylceramidase. *J Am Chem Soc* **139**, 14192–14197.

106 Ben Bdira F, Kallemeijn WW, Oussoren SV, Scheij S, Bleijlevens B, van Roomen CPAA, Ottenhoff R, van Kooten, MJFM, Walvoort MTC, Witte MD, Boot RG, Ubbink M, Overkleeft HS, & Aerts JMFG (2017) Stabilization of glucocerebrosidase by active-site occupancy. *ACS Chem. Biol* **12**, 1830–1841.

107 Artola M, Kuo CL, Lelieveld LT, Rowland RJ, van der Marel GA, Codée JDC, Boot RG, Davies GJ, Aerts JMFG & Overkleeft HS (2019) Functionalized Cyclophellitols Are Selective Glucocerebrosidase Inhibitors and Induce a Bona Fide Neuropathic Gaucher Model in Zebrafish. *J Am Chem Soc* **141**, 4214–4218.

108 Kallemeijn WW, Witte MD, Voorn-Brouwer TM, Walvoort MT, Li KY, Codée JD, van der Marel GA, Boot RG, Overkleeft HS & Aerts JM (2014) A sensitive gel-based method combining distinct cyclophellitol-based probes for the identification of acid/base residues in human retaining β-glucosidases. *J Biol Chem* **289**, 35351–35362.

109 Kytidou K, Beekwilder J, Artola M, van Meel E, Wilbers RHP, Moolenaar GF, Goosen N, Ferraz MJ, Katzy R, Voskamp P, Florea BI, Hokke CH, Overkleeft HS, Schots A, Bosch D, Pannu N & Aerts JMFG (2018) Nicotiana benthamiana α-galactosidase A1.1 can functionally complement human α-galactosidase A deficiency associated with Fabry disease. *J Biol Chem* **293**, 10042–10058.

110 Husaini AM, Morimoto K, Chandrasekar B, Kelly S, Kaschani F, Palmero D, Jiang J, Kaiser M, Ahrazem O, Overkleeft HS & van der Hoorn RAL (2018) Multiplex Fluorescent, Activity-Based Protein Profiling Identifies Active α-Glycosidases and Other Hydrolases in Plants. *Plant Physiol* **177**, 24–37.

Scope of the thesis

The general goal of the thesis investigations was to apply newly designed and generated ABPs in fundamental research on lysosomal glycosidases. Described is the characterization of a large number of novel ABPs targeting various glycosidases as well as a broad scale of applications for ABPs in research.

Chapter 1 briefly reviews current research on Gaucher disease and applications of activity-based probes (ABPs) targeting lysosomal glucocerebrosidase (GBA). It provides detailed protocols for *in vitro* and *in situ* visualization of active GBA molecules.

Chapter 2 reports on the use of a suite of ABPs in a gel-based competitive activity-based protein profiling (cABPP) approach to investigate the *in vivo* glycosidase targets of the widely applied GBA inhibitor conduritol B epoxide (CBE), and its close structural analogue cyclophellitol (CP). Off-target glycosidases are identified for both cyclitols, and the selectivity windows for selective GBA inactivation by both inhibitors are assessed.

Chapter 3 demonstrates the use of novel modified CPs for selective GBA inactivation and generation of neuropathic Gaucher zebrafish models. These new compounds containing a biphenyl or adamantly group installed at the C8 position of CP, and when orally administered the adamantyl compound selectively and potently inactivates GBA in visceral organs and brain of adult zebrafish.

Chapter 4 reports on the biochemical and structural investigation of CP aziridine ABPs toward GH31 α-glucosidase and GH2 β-glucuronidase. These ABPs allow mechanism-based inactivation/labeling of their target glycosidases *in vitro* and in cells. This enables the profiling of the endogenous active enzymes across biological samples, including fibroblasts from Pompe patients for diagnostic purposes.

Chapter 5 documents the design and characterization of CP aziridine ABPs towards GH39 α-L-iduronidase, the enzyme deficient in Hurler syndrome (MPS 1). These compounds exhibit

mechanism-based labeling of iduronidase, allowing detailed structural investigation on the conformational itinerary of the cleavage reaction. In addition, the Cy5 ABP assists monitoring of cellular uptake of pre-labeled therapeutic iduronidase by fluorescent microscopy.

Chapter 6 presents the development and characterization of CP aziridine ABPs targeting GH38 α-mannosidases or GH2 β-mannosidase. For this purpose use was made of SDS-PAGE-based fluorescence detection, measurement of inhibitory potency and kinetic studies, structural analysis, and chemical proteomics. Obtained data on labeling potency and glycosidase selectivity for both sets of compounds, provide a basis for application in mannosidase ABPPs across a range of biological samples as well as their use in screening of compound libraries for inhibitors.

Chapter 7 investigates the labeling mechanism, potency, inhibition kinetics, and glycosidase targets of the β-galactose configured CP aziridine compounds and ABPs. This allows the activity-based profiling of both lysosomal β-galactosidase (GLB1) and galactocerebrosidase (GALC) in biological samples. Chemical proteomics also identified other GLB1-like proteins, whose biological functions and implication in GM1-gangliosidosis or Krabbe disease should be explored in the future.

In the **General discussion and future prospects** section, the design and applications of cyclophellitol and cyclophellitol aziridine ABPs reacting with lysosomal glycosidases are reviewed, with a focus on their application in LSD research. The present limitations of current ABPs and future prospects are discussed.

This thesis is concluded with **Appendices** that include a **List of publications** and ***Curriculum vitae*** of the author.

CHAPTER 1

Development of activity-based protein profiling methods for *in vitro* and *in situ* visualization of glucocerebrosidase

Manuscript published as:

Kuo CL, van Meel E, Kytidou K, Kallemeijn WW, Witte M, Overkleeft HS, Artola ME & Aerts JM (2018) Activity-Based Probes for Glycosidases: Profiling and Other Applications. *Methods Enzymol* **598**, 217–235.

ABSTRACT

Glycosidases mediate the fragmentation of glycoconjugates in the body, including the vital recycling of endogenous molecules. Several inherited diseases in men concern deficiencies in lysosomal glycosidases degrading glycosphingolipids. Prominent is Gaucher disease caused by an impaired lysosomal β-glucosidase (glucocerebrosidase; GBA) and resulting in pathological lysosomal storage of glucosylceramide (glucocerebroside) in tissue macrophages. GBA is a retaining glucosidase with a characteristic glycosyl–enzyme intermediate formed during catalysis. Using the natural suicide inhibitor cyclophellitol as lead, mechanism-based irreversible inhibitors of GBA equipped with a fluorescent reporter was developed. These covalently link to the catalytic nucleophile residue of GBA and permit specific and sensitive visualization of active enzyme molecules. The amphiphilic activity-based probes (ABPs) allow *in situ* detection of active glucocerebrosidase in cells and organisms. Furthermore, they may be used to biochemically confirm the diagnosis of Gaucher disease and they might assist screening for small compounds interacting with the catalytic pocket. While the focus of this chapter is ABPs for β-glucosidases and Gaucher disease, the described concept has meanwhile been extended to other retaining glycosidases and related disease conditions as well.

1.1 Introduction

The human body contains a huge number of glycoconjugates that are ongoingly synthesized and degraded by glycosidases. As first described by Daniel E. Koshland, the hydrolysis of the glycosidic bond by glycosidases results in inversion or retention of the anomeric stereochemistry in the glycon.[1] Inverting glycosidases employ two carboxylic acid residues, typically spaced 6–11 Å apart and positioned on opposing sides of the substrate glycoside, allowing joint entry of substrate and a water molecule.[2] The reaction proceeds through a single oxocarbenium ion-like transition state. In contrast, retaining glycosidases use a double-displacement mechanism that proceeds via a glycosyl–enzyme intermediate, flanked by two oxocarbenium ion-like transition states, preserving the initial configuration at the anomeric center. Two adjacent carboxylic acid residues, spaced 5.5 Å apart, function as the catalytic nucleophile and acid/base residue.[3] Variations on the classical Koshland-type mechanisms have been identified.[4]

Among the many cellular glycoconjugates are glycosphingolipids (GSLs), consisting of ceramide lipid backbones extended with oligosaccharides.[5] Their recycling involves fragmentation in lysosomes through sequential action of glycosidases and acid ceramidase. A genetic deficiency of a lysosomal glycosidase involved in GSL degradation causes accumulation of the corresponding lipid substrate.[6] Most prominent among these inherited glycosphingolipidoses are Gaucher disease (GD), Krabbe disease, Tay-Sachs disease and Fabry disease. Albeit monogenetic, the lysosomal disorders show remarkable variation in nature of symptoms, age of onset and progression of clinical presentations.

Over the years, GD has acted as frontrunner in fundamental research on lysosomal glycosidases and subsequent development of rational therapies for enzymopathies.[7] One and a half century ago Ernest Gaucher described a patient with a markedly enlarged spleen showing accumulation of lipid-laden phagocytes. It was soon recognized that this patient represented a distinct disease entity, subsequently referred to as Gaucher disease.[8] The accumulating lipid in the lysosomes of storage macrophages (Gaucher cells) of GD patients was identified as glucosylceramide (glucocerebroside). Next, Roscoe Brady and co-workers demonstrated that deficiency of glucocerebrosidase (GBA, acid β-glucosidase; EC. 3.2.1.45) is the molecular cause of GD.[9] Over 200 mutations in the GBA gene (locus 1q21) have meanwhile been linked to the disease.[10] The clinical manifestation of GD is remarkably heterogeneous, ranging from lethal

neonatal complications to an almost asymptomatic course.[8] The most prevalent manifestation, type 1 GD, does not involve pathology of the central nervous system. Major symptoms are enlargement of spleen and liver, infiltration of the bone marrow by storage cells, thrombocytopenia, anemia and bone disease. GD types 2 and 3 are accompanied by fatal neurological symptoms developing in the first years of life or at later age, respectively. The most severe manifestation is the so-called collodion baby, born with lethal skin permeability.[11] For several GBA genotypes considerable variability in disease severity has been documented, even among monozygotic twins.[12, 13] Deficiency of GBA has recently been recognized as a risk factor for multiple myeloma.[14] Intriguingly, carriers of a mutant GBA allele are also at markedly increased risk for developing Parkinsonism.[15]

The life cycle and structural features of GBA have been studied in great detail. The acquired fundamental knowledge has been translated to (experimental) therapeutic interventions. Given the success of bone marrow transplantation, presently gene therapy of GD is pursued, aiming at macrophage-specific expression of GBA in genetically modified hematopoietic stem cells.[16] Small compound glycomimetics (chemical chaperones) promoting folding and/or increasing the structural stability of GBA are searched and designed.[17] Substrate reduction therapy inhibiting biosynthesis of glucosylceramide through oral administration of inhibitors of glucosylceramide synthase has already been registered as treatment of GD (Miglustat, Actelion; Eliglustat, Sanofi-Genzyme).[18, 19] Moreover, intravenous supplementation of macrophage-targeted recombinant GBA results in major visceral improvements in type 1 GD patients, but fails to prevent neurological manifestations in type 2 and 3 GD patients.[20] Currently, several enzyme preparations have been registered for enzyme replacement therapy of type 1 GD.[7] The success of enzyme replacement therapy has led to similar approaches for other lysosomal enzymopathies in the lysosomal turnover of GSL, mucopolysaccharides and glycogen. Glycosidases, particularly those implicated in rare lysosomal storage diseases, receive consequently great attention. New tools to visualize active glycosidases in living cells and organisms are warranted.

GBA is a retaining β-glucosidase with E340 as nucleophile and E325 as acid/base residue (**Fig. 1.1**, upper panel).[21] Cyclophellitol **1** (**Fig. 1.1**, lower panel), isolated from the *Phellinus sp.* mushroom, is a mechanism-based inhibitor of GBA, reacting far more potently and specifically with GBA than cyclitol epoxide conduritol B-epoxide (CBE).[22, 23] Cyclophellitol has been

successfully used as a lead for the development of high affinity ABPs for retaining β-glucosidases, in particular for the human enzyme GBA.[24, 25] Grafting an azide group onto the C6 position of cyclophellitol **1** (glucopyranose numbering) yielded azido-cyclophellitol **2** (**Fig. 1.1**, lower panel). When **2** was covalently extended at C6 with BODIPY fluorophore moieties, the resulting compounds **3** and **4** were even 100-fold more potent inhibitors for GBA than parent cyclophellitol **1**.[24] The fluorescent ABPs, green-fluorescent β-epoxide **3** (MDW933) and red-fluorescent β-epoxide **4** (MDW941) allow immediate fluorescence scanning of SDS-PAGE gels with ABP-labeled proteins.[24] The attachment of the bulky BODIPY moiety at the C6-position in cyclophellitol enhanced the potency and specificity of the inhibitor for GBA.

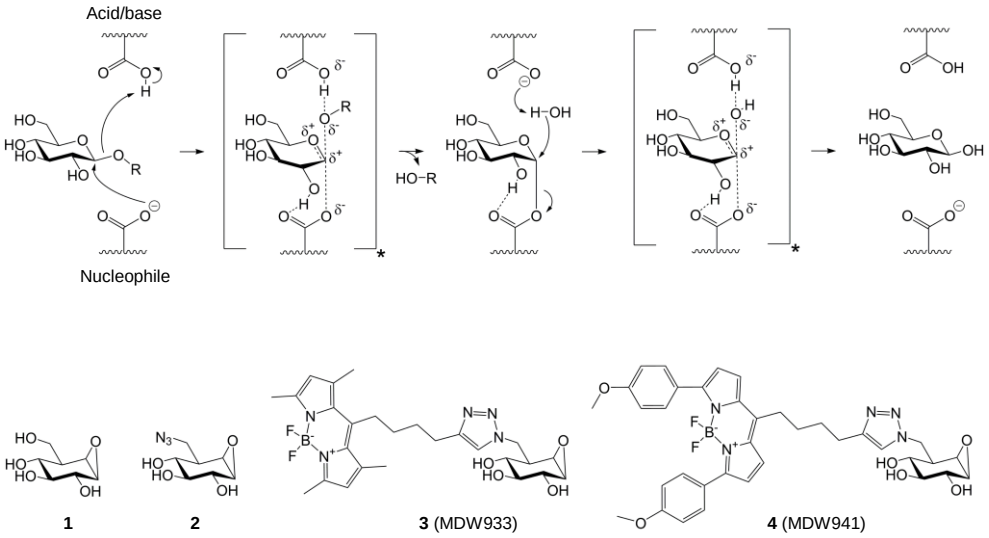


Figure 1.1. Catalytic mechanism of retaining glycosidases and ABPs. Upper panel: Catalytic mechanism of a retaining β-glycosidase, characterized by a covalent glycosyl–nucleophile adduct, flanked by two oxocarbenium ion-like transition states (*). The transition states are stabilized by hydrogen bonding (finely dotted line) between the C-2 hydroxyl group of the substrate and the nucleophile. Lower panel: Structures of cyclophellitol β-epoxide 1, azido-cyclophellitol **2** (KY170), fluorescent β-epoxide **3** (MDW933, green fluorescent) and β-epoxide **4** (MDW941, red fluorescent).

Active GBA can be labeled in lysates of cells with **3** or **4** and subsequently visualized by fluorescence scanning following SDS-PAGE electrophoresis. **Fig. 1.2** shows ABP-labeled GBA in lysates of human fibroblasts. Multiple molecular weight isoforms are detected, reflecting variations in the N-linked glycan composition of GBA.[26] Active GBA can also be labeled in living cells thanks to the cell permeability of the amphiphilic ABPs.[24] **Fig. 1.3** Shows the *in situ*

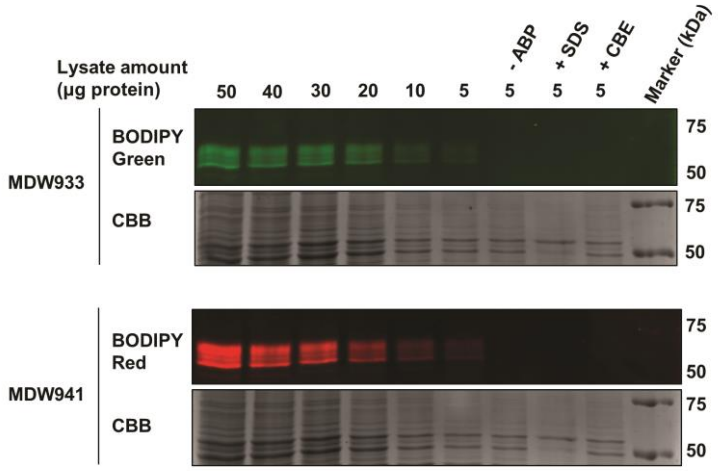


Figure 1.2. *In vitro* MDW933 (3) and MDW941 (4) labeling of GBA in human fibroblast lysates. Different amounts of samples were labeled with 100 nM MDW933 or MDW941 for 30 min at 37 °C, denatured, resolved using 7.5 % polyacrylamide SDS-PAGE, and scanned for BODIPY green or red fluorescence (upper and lower panel respectively). Coomassie Brilliant Blue G250 (CBB) staining was used to visualize total protein. No ABP was added to the negative control sample (- ABP). 2 % SDS was added to one sample (+SDS) and incubated for 5 min at 98°C prior to ABP addition. 1 mM CBE was added to one sample (+CBE) for 30 min at 37 °C prior to ABP addition.

labeling with **4** of GBA in human fibroblasts as visualized by confocal fluorescence microscopy.

The β-glucosyl configured cyclophellitols find broad applications as inspired by the seminal work by Cravatt and co-workers with ABPs.[27, 28] It has been earlier reported that they can be used in biochemical confirmation of the diagnosis GD through the demonstration of reduced ABP-labeled GBA in fibroblasts.[24, 29] The ABPs can be employed across species to visualize active GBA.[24, 30, 31] *In situ* ABP-labeling of active GBA in cultured cells and mice is feasible. In the case of cells, the ABP can be applied to the medium. In the case of mice, it needs to be intravenously infused.[32] Since BODIPY-tagged ABPs poorly penetrate the brain, *in situ* visualization of GBA in this tissue requires intrathecal administration of the ABP.[33] The use of the ABPs in correlative light and electron microscopy (CLEM) is presently pursued. Equipping the β-glucosyl

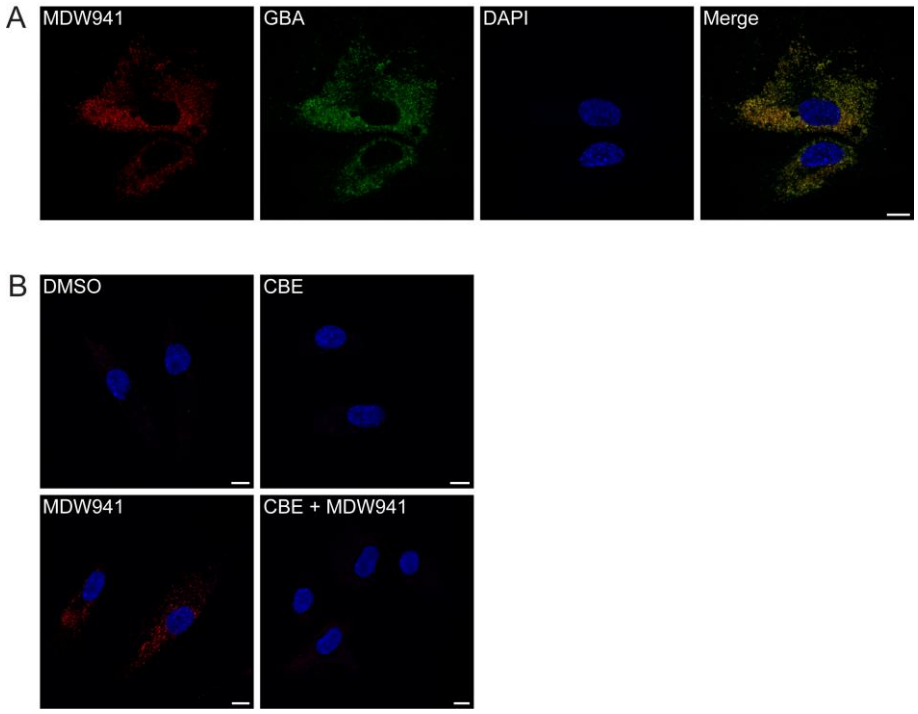


Figure 3. *In situ* labeling with MDW941 detects GBA in human fibroblasts by confocal fluorescence microscopy. A) In situ labeling with MDW941 for 2 h at 5 nM followed by immunostaining of GBA showing significant co-localization of active GBA (red) with total GBA (green). B) Active GBA is detected with MDW941, while pre-incubation with CBE followed by MDW941 results in loss of the fluorescent signal. Scale bars, 10 µm.

configured cyclophellitol with a biotin at C6 drastically decreased its activity against GBA[24] and was not suitable for streptavidin-bead pull down application, but novel biotin ABPs with longer linker are currently being developed for GBA (General Discussion and Future Prospects, this thesis). In principle, proteins strongly interacting with GBA could be enriched in this manner and identified by proteomics.

Another future application of the ABPs could be (ABPP) screening of compound libraries for structures showing affinity for the catalytic pocket of GBA. Such compounds may be of interest to be developed further as therapeutic chemical chaperones.[17] The β-glucosyl configured cyclophellitol ABP has been also used to study the interaction of isofagomine with GBA in intact cells.[24] Finally, the BODIPY-equipped ABPs, not penetrating the brain possibly due to active Pgp-mediated removal, may be used to pharmacologically generate a visceral model of

Gaucher disease without the fatal neuropathology in GBA-deficient mice.

Here, a detailed protocol for GBA detection using ABPs **3** and **4** is presented in this chapter. The first section lists the equipment and materials (**1.2**); the second section details the protocol for *in vitro* ABP labeling and SDS-PAGE-based fluorescence detection for GBA (**1.3**); the third section concerns the protocol for *in situ* ABP labeling of GBA in intact cells followed by SDS-PAGE-based fluorescence detection (**1.4**); and the fourth section reports on the protocol for *in situ* ABP labeling of GBA in intact cells followed by detection using fluorescent microscopy (**1.5**). Meanwhile, the concept of cyclophellitol-type ABPs has been extended to other retaining glycosidases (General introduction, this thesis). Thus, this protocol can also be used as a basis for activity-based protein profiling using other cyclophellitol-based or cyclophellitol-aziridine-based ABPs targeting different disease-relevant glycosidases.

1.2 List of equipment and materials

Equipment

Spectrophotometer
Heating block
Tabletop centrifuge
1.5-mL Microcentrifuge tubes
200-µL PCR tubes
15- and 50-mL Centrifuge tubes
Micropipettors
Micropipettor tips
SDS-PAGE equipment and glass plates
Fluorescence imager or scanner (with settings for BODIPY green or red fluorescence)
Sterile 6-well and 12-well plates
Coverslips
Incubator at 37 °C with 5 % CO2
Parafilm
Confocal fluorescence microscope
Microscope slides

Materials

a. 25 mM Potassium phosphate buffer (KPi) pH 6.5 (1 L)

Component	Final concentration	Stock concentration	Amount
K_2HPO_4	7.5 mM	1 M	7.5 mL
KH_2PO_4	17.5 mM	1 M	17.5 mL

*Add 950 mL of ultrapure-H_2O, adjust the pH to 6.5 and add H_2O to a total volume of 1 L

b. Cell lysis buffer (50 mL)

Component	Final concentration	Stock concentration	Amount
KPi buffer pH 6.5	25 mM	25 mM	49.5 mL
Protease inhibitor - cocktail tablet (Roche)	-		1 tablet
Triton X-100	0.1 % (v/v)	10 % (v/v)	0.5 mL

*Note: Store at -20 °C in 1 mL aliquots

c. 750 mM McIlvaine buffer pH 5.2 (40 mL)

Component	Final concentration	Stock concentration	Amount
Citric acid	0.27 M	0.5 M	21.4 mL
Na_2HPO_4	0.46 M	1 M	18.6 mL

*Note: 0.01% (w/v) NaN_3 can be added to both Na_2HPO_4 and citric acid stocks as bacteriostatic agent

d. 150 mM McIlvaine buffer pH 5.2 (40 mL)

Component	Final concentration	Stock concentration	Amount
Citric acid	0.054 M	0.1 M	21.4 mL
Na_2HPO_4	0.093 M	0.2 M	18.6 mL

*Note: 0.01% (w/v) NaN_3 can be added to both Na2HPO4 and citric acid stocks as bacteriostatic agent

e. 5x Laemmli sample buffer (10 mL)

Component	Final concentration	Stock concentration	Amount
Tris-HCl pH=6.8	0.3 M	1.25 M	2.5 mL
Glycerol	50% (v/v)	100% (v/v)	5 mL
Sodium dodecyl sulfate	10% (v/v)	-	1 g
Bromophenol blue	0.1% (v/v)	1% (w/v)	1 mL

*Add ultrapure-H_2O to a total volume of 10 mL

1.3 Protocol: Fluorescent detection of ABP-labeled GBA on wet slab gels

MDW933 (**3**) and MDW941 (**4**) are dissolved in DMSO and stored in small aliquots at -20 °C until use. The labeling can be performed on samples *in vitro* or directly in intact cells (*in situ*). For the *in vitro* labeling procedure, the samples and ABP are diluted separately in McIlvaine buffer pH 5.2 and combined for a 30 min incubation at 37 °C, and denatured for SDS-PAGE. For the *in situ* labeling procedure, MDW933 or MDW941 is diluted in culture medium and applied to the cell culture for a 2 h incubation at 37 °C. Subsequently, the cells are washed three times in PBS and lysed in lysis buffer. Proteins from both *in vitro* and *in situ* labeled samples are resolved by SDS-PAGE and wet slab gels are scanned for BODIPY green or red fluorescence using a fluorescence scanner.

Storage of ABPs

Aliquots of MDW933 and MDW941 are made in either 1 mM or 50 µM concentrations, in a volume of 1–2 µL.

1. To make the 1 mM stock, add 100 µL DMSO to the tube containing 100 nmol lyophilized MDW933 or MDW941, vortex to dissolve and transfer to a 1.5 mL Eppendorf tube.

2. Make 1 µL or 2 µL aliquots for the 1 mM stock in PCR tubes (200 µL volume).

3. To make the 50 µM stock, add 38 µL DMSO to one tube of the 2 µL 1 mM stock, vortex to dissolve and make aliquots of 1 µL in PCR tubes.

4. Place the PCR tubes into containers or 50 mL centrifuge tubes and store them at -20°C. MDW933 and MDW941 are stable at -20 °C for long-term storage (> 1 year).

*Tip

One tube of 1 µL 50 µM MDW933 or MDW941 is sufficient for the *in vitro* labeling of 66 samples at 100 nM labeling concentration (SDS-PAGE loading volume = 20 µL), or 1 x 6-well plate of culture cells at 5 nM labeling concentration (1 mL culture medium per well).

**Caution

1. Avoid freeze-thawing the ABP stocks (in DMSO) more than twice.

2. Do not store and re-use ABPs that have been diluted in H_2O or buffer.

3. Protect the ABPs from light.

In vitro labeling and in-gel detection

1. Prepare lysates and homogenates in lysis buffer and measure protein concentration by the BCA assay (Pierce, BCA protein assay kit).

2. On ice, prepare 2–5 µL samples in 1.5 mL tubes, containing 5–50 µg total protein, or 0.5–50 fmole GBA molecules of purified enzyme stock.

3. Add McIlvaine buffer (pH 5.2) to the samples, to a total volume of 10 µL, and incubate on ice for 5 min. Use 150 mM McIlvaine buffer for lysate volumes ≤ 2 uL, and 750 mM for lysate volumes of 2–5 µL.

4. Prepare 300 nM MDW933 or MDW941 diluted in McIlvaine buffer (pH 5.2, 150 mM or 750 mM) using the 50 µM MDW933 or MDW941 stock (e.g. 1 µL 50 µM stock + 166 µL McIlvaine buffer).

5. Add 5 µL 300 nM MDW933 or MDW941 to each sample (10 µL) and incubate at 37°C on a thermoshaker or in a waterbath for 30 min.

6. After the incubation, add 3.75 µL 5x Laemmli sample buffer to the sample, and incubate at 98 °C for 5 min.

7. Resolve the proteins by SDS-PAGE (7.5 % or 10 % polyacrylamide gels) and quantify the labeling on a Typhoon scanner (GE Healthcare) using the following fluorescence settings: green

BODIPY: λ_{ex} = 488 nm, λ_{em} = 500–550 nm; red BODIPY: λ_{ex} 594 nm, λ_{em} = 605–645 nm.

*Tip

1. For optimal labeling and fluorescence detection in lysates and homogenates, the protein concentration in the prepared cell lysates or tissue homogenates is best kept at 10-20 μg/μL.

2. When performing labeling on purified GBA, supplement the McIlvaine buffer with 0.1% (v/v) Triton X-100 and 0.2% (w/v) sodium taurocholate.

3. Optimal labeling visualization is achieved with 0.5–50 fmole GBA per sample.

4. Include one sample without MDW933 or MDW941 treatment as negative control for assessing autofluorescence from the samples.

1.4 Protocol: *In situ* labeling and in-gel detection

Cells are treated with 5 nM MDW933 or MDW941 at 37 °C for 2 h, washed three times with PBS, lysed in lysis buffer and detached by scraping. The homogenate is resolved by SDS-PAGE and the fluorescent signal on the gel is quantified by fluorescent scanning.

1. Maintain normal human dermal fibroblasts (Lonza) in DMEM containing 4.5 g/L glucose, sodium pyruvate and sodium bicarbonate (Sigma, D6546), or F-12 Ham/DMEM (Sigma, D8062), both supplemented with 10% (v/v) FBS, 100,000 units/L penicillin, 100 mg/L streptomycin and 2 mM glutamax.

Transfer the fibroblasts to 6-well plates and let them grow until at least 80 % confluency.

2. In a 15 mL centrifuge tube, combine 10 mL pre-warmed culture medium (37 °C) with 1 μL of 50 μM MDW933 or MDW941.

3. Aspirate the culture medium from the cell culture plate and add 1 mL of the prepared 5 nM MDW933 or MDW941 diluted in the pre-warmed culture medium.

4. Incubate for 2 h in a 37 °C incubator.

5. After the incubation, aspirate the MDW933 or MDW941-containing medium and wash the cells three times with 1 mL PBS.

6. Aspirate the PBS, place the culture plate on ice and add 50 µL pre-chilled 25 mM KPi buffer pH 6.5 (+0.1% (v/v) Triton X-100 and protease inhibitor cocktail) to each well of the plate.

7. Use a cell scraper or the back of a P200 pipette tip to detach the fibroblasts from the plate and collect the cells in 1.5 mL Eppendorf tubes.

8. Vortex, store the samples at -80 °C for 1 h, and thaw the samples on ice.

9. Determine the protein concentrations using the BCA assay (Pierce BCA protein assay kit) and a spectrophotometer (absorption set at 562 nm). The typical protein concentration for *in situ* treated fibroblast lysates is between 1–2 µg/µL.

10. For SDS-PAGE, samples containing the same amount of protein (typically 5-30 µg protein) are diluted in 25 mM KPi buffer (+0.1% (v/v) Triton X-100 and protease inhibitor cocktail), in a total volume of 15 µL, denatured by incubation with 3.75 µL 5x Laemmli sample buffer at 98 °C for 5 min, and resolved by SDS-PAGE.

11. Quantify fluorescence on wet slab gel using a Typhoon scanner (GE Healthcare).

*Tip

For ABP labeling experiment involving different ABP concentrations, or treatment at different time points, an alternative labeling method can be followed. For this method, first pre-dilute the ABP stock (1 mM or 50 µM) in DMSO into 1000x of the desired labeling concentrations in sterile 1.5 mL tubes, then directly pipette 1 µL of the prepared ABP dilutions per well to the 6-well plate (containing 1 mL refreshed medium per well). It is important to ensure proper mixing of the ABP in the culture medium.

1.5 Protocol: *In situ* labeling of GBA for fluorescence microscopy

To localize active GBA, the cells are *in situ* labeled with MDW941 by applying the ABP to the culture medium. It is important to include controls for autofluorescence, i.e. DMSO treated cells, and for the specificity of the ABP signal. For the latter, the cells are incubated with CBE prior to the in situ labeling with ABP. This compound will bind to the active GBA molecules, thereby blocking all binding by the ABP. Any remaining signal will represent unbound ABP.

1. The day before the experiment, transfer normal human dermal fibroblasts to four wells of a 12-well plate that contain autoclaved coverslips (Menzel, 15 mm diameter). Allow the cells to adhere to the coverslips by incubating them overnight in a humidified incubator at 37 °C with 5 % CO_2.

2. The cells should have reached ~70 % confluency on the day of the in situ labeling. Remove the culture medium and place 1 mL of fresh medium in each well.

3. Two wells are treated with CBE as a control for the specificity of the signal. Dissolve CBE (Enzo Life Sciences Inc.) in ultrapure-H_2O at a concentration of 150 mM, e.g. by adding 200 µL ultrapure-H_2O to 5 mg CBE. Add 2 µL of CBE per well (final concentration will be 0.3 mM) and incubate the cells at 37°C for 16 h.

4. After the 16 h incubation, remove the culture medium from the wells and wash twice with PBS. Subsequently, place 1 mL of culture medium in each well.

5. Dilute MDW941 in DMSO to a concentration of 5 µM, e.g. by adding 1 µL of the 1 mM stock to 200 µL DMSO. Subsequently, add 1 µL of the 5 µM dilution to one well that was pre-incubated with CBE and to one untreated well. This final concentration of 5 nM will label ~50 % of total endogenous GBA in these cells, but optimal concentrations should be determined for each cell line. Add 1 µL of DMSO to the second untreated well as a control for autofluorescence. From this step onwards it is important to protect the cells from light.

6. Incubate the cells for 2 h at 37 °C.

7. At the end of the incubation, remove the culture medium from the wells and wash 3 times with 1 mL PBS.

8. Fix the cells by placing 1 mL of 4 % (w/v) formaldehyde/PBS in each well and incubate for 25 min at RT. A stock of 16 % formaldehyde/H_2O can be made as described in Slot & Geuze, 2007.[34] Dilute this stock 1:4 in PBS, e.g. by mixing 4 mL of 16 % formaldehyde with 1.6 mL 10x concentrated PBS and 10.4 mL distilled H_2O.

9. Remove the fixative and wash the coverslips with PBS by placing 1 mL in each well. Remove the PBS and repeat this wash two times. To stain all GBA in the cells (active and inactive) for immunofluorescence microscopy continue with step 10. To mount the coverslips directly, skip

steps 10-14 and proceed with step 15.

10. Prepare a piece of parafilm with 200 µL drops of 2 % (w/v) bovine serum albumin/ 0.1% (w/v) saponin in PBS (= permeabilization buffer). Incubate the coverslips onto these drops, with the cells facing the drop, for 10 min. Cover the coverslips with a petri dish wrapped in aluminum foil to protect from light and evaporation.

11. Incubate the coverslips for 1 h at RT onto 100 µL drops of 1:500 diluted mouse anti-GBA monoclonal antibody 8E4 (generated in the Aerts lab) in permeabilization buffer.

12. Transfer the coverslips to 200 µL drops of permeabilization buffer and incubate for 10 min. Repeat this washing step twice.

13. Incubate the coverslips for 1 h at RT onto 100 µL drops of 1:500 diluted Alexa Fluor 488-coupled donkey anti-mouse IgG (H+L) (Invitrogen).

14. Wash the coverslips three times on 200 µL drops of permeabilization buffer.

15. Wash the coverslips quickly by dipping them in distilled H_2O.

16. Drain off the H_2O from the coverslips and mount them on a microscope slide (VWR) with ProLong Diamond antifade reagent containing DAPI (Molecular Probes). Allow the coverslips to dry overnight, seal with nail polish and store at 4°C.

17. Image the cells under a Leica SP8 confocal microscope with a 63x/1.40 NA HC Plan Apo CS2 oil immersion objective and equipped with a hybrid detector (HyD). Image MDW941 with excitation at 552 nm, emission 590–650 nm, Alexa Fluor 488 with excitation at 488 nm, emission 500–540 nm and DAPI with excitation at 405 nm, emission 420–480 nm.

*Tip

The concentration of *in situ* applied MDW933 or MDW941 can be adjusted from 5 nM to 100 nM, according to cell types and confluency, and it is advisable to validate the percentage of GBA inhibition by *in vitro* 4-methylumbelliferyl substrate assay.[35]

1.6 Conclusion

Deficiency of GBA leads to Gaucher disease and constitutes a risk factor for multiple myeloma and Parkinsonism. The biochemical and clinical study on Gaucher disease has enabled tremendous therapeutic development in the past decades, which has encouraged the research and therapeutic development on other lysosomal storage diseases involving glycosidases. The availability of the here-described ABPs specifically labeling GBA in a mechanism-based manner offer new valuable tools for research on the enzyme, diagnosis and therapy development. Detailed protocols for *in vitro* and *in situ* detection of GBA using these ABPs were described in this chapter, which should also serve as a basis for ABP detection of other glycosidases relevant in lysosomal storage diseases.

1.7 References

1 Koshland DE (1953) Stereochemistry and mechanism of enzymatic reactions. *Biol Rev Camb Philos Soc* **28**, 416–436.

2 Vasella A, Davies GJ & Bohm M (2002) Glycosidase mechanisms. *Curr Opin Chem Biol*, **6**, 619–629.

3 Davies GJ, Ducros VM, Varrot A & Zechel DL (2003) Mapping the conformational itinerary of β-glycosidases by X-ray crystallography. *Biochem Soc Trans* **31**, 523–527.

4 Kallemeijn WW, Witt, MD, Wennekes T & Aerts JM (2014) Mechanism-based inhibitors of glycosidases: design and applications. *Adv Carbohydr Chem Biochem* **71**, 297–338.

5 Wennekes T, van den Berg RJ, Boot RG, van der Marel GA, Overkleeft HS & Aerts JM (2009) Glycosphingolipids—nature, function, and pharmacological modulation. *Angew Chem Int Ed Engl* **48**, 8848–8869.

6 Ballabio A & Gieselmann V (2009) Lysosomal disorders: from storage to cellular damage. *Biochim Biophys Acta* **1793**, 684–696.

7 Ferraz MJ, Kallemeijn WW, Mirzaian M, Herrera Moro D, Marques A, Wisse P, Boot RG, Willems LI, Overkleeft HS & Aerts JM (2014) Gaucher disease and Fabry disease: new markers and insights in pathophysiology for two distinct glycosphingolipidoses. *Biochim Biophys Acta* **1841**, 811–825.

8 Beutler E & Grabowski GA (2001) Gaucher disease. In Scriver CR, Sly WS, Childs B, Beaudet AL, Valle D, Kinzler KW, & Vogelstein B (Eds) *The Metabolic and Molecular Bases of Inherited Disease*, *8th Ed* (pp 3635–3668). New York, NY: McGraw-Hill.

9 Brady RO, Kanfer JN, Bradley RM & D Shapiro, D (1966) Demonstration of a deficiency of glucocerebroside-cleaving enzyme in Gaucher's disease. *J Clin Invest* **45**, 1112–1115.

10 Hruska KS, LaMarca ME, Scott CR & Sidransky E (2008) Gaucher disease: mutation and polymorphism spectrum in the glucocerebrosidase gene (GBA). *Hum Mutat* **29**, 567–583.

11 Eblan MJ, Goker-Alpan O & Sidransky E (2005) Perinatal lethal Gaucher disease: a distinct phenotype along the neuronopathic continuum. *Fetal Pediatr Pathol* **24**, 205–222

12 Lachmann RH, Grant IR, Halsall D & Cox TM (2004) Twin pairs showing discordance of phenotype in adult Gaucher's disease. *QJM* **97**, 199–204.

13 Biegstraaten M, van Schaik IN, Aerts JM, Langeveld M, Mannens MM, Bour LJ, Sidransky E, Tayebi N, Fitzgibbon E & Hollak, CE (2011) A monozygotic twin pair with highly discordant Gaucher phenotypes. *Blood Cells Mol Dis* **46**, 39–41.

14 Nair S, Branagan AR, Liu J, Boddupalli CS, Mistry PK & Dhodapkar, MV (2016) Clonal Immunoglobulin against Lysolipids in the Origin of Myeloma. *N Engl J Med* **374**, 555–561.

15 Aflaki E, Westbroek W & Sidransky, E (2017) The Complicated Relationship between Gaucher Disease and Parkinsonism: Insights from a Rare Disease. Neuron **93**, 737–746.

16 Dahl M, Doyle A, Olsson K, Månsson JE, Marques AR, Mirzaian M, Aerts JM, Ehinger M, Rothe M, Modlich U, Schambach A & Karlsson S (2015) Lentiviral gene therapy using cellular promoters cures type 1 Gaucher disease in mice. *Mol Ther* **23**, 835–844.

17 Sánchez-Fernández EM, García Fernández JM & Mellet CO (2016) Glycomimetic-based pharmacological chaperones for lysosomal storage disorders: lessons from Gaucher, GM1-gangliosidosis and Fabry diseases. *Chem Commun (Camb)* **52**, 5497–5515.

18 Cox T, Lachmann R, Hollak C, Aerts J, van Weely S, Hrebícek M, Platt F, Butters T, Dwek R, Moyses C, Gow I, Elstein D & Zimran A (2000) Novel oral treatment of Gaucher's disease with N-butyldeoxynojirimycin (OGT 918) to decrease substrate biosynthesis. *Lancet* **355**, 1481–1485.

19 Cox TM, Drelichman G, Cravo R, Balwani M, Burrow TA, Martins AM, Lukina E, Rosenbloom B, Ross L, Angell J & Puga AC (2015) Eliglustat compared with imiglucerase in patients with Gaucher's disease type 1 stabilised on enzyme replacement therapy: a phase 3, randomised, open-label, non-inferiority trial *Lancet* **385**, 2355–2362.

20 Barton NW, Furbish FS, Murray GJ, Garfield M & Brady RO (1990) Therapeutic response to intravenous infusions of glucocerebrosidase in a patient with Gaucher disease. *Proc Natl Acad Sci USA* **87**, 1913–1916.

21 Kallemeijn WW, Witte MD, Voorn-Brouwer TM, Walvoort MT, Li KY, Codée JD, van der Marel GA, Boot RG, Overkleeft HS & Aerts JM (2014) A sensitive gel-based method combining distinct cyclophellitol-based probes

for the identification of acid/base residues in human retaining β-glucosidases. *J Biol Chem* **289**, 35351–35362.

22 Atsumi S, Nosaka C, Iinuma H & Umezawa K (1992) Inhibition of glucocerebrosidase and induction of neural abnormality by cyclophellitol in mice. *Arch Biochem Biophys* **297**, 362–367.

23 Witte MD, van der Marel GA, Aerts JM & Overkleeft HS (2011) Irreversible inhibitors and activity-based probes as research tools in chemical glycobiology. *Org Biomol Chem* **9**, 5908–5926.

24 Witte MD, Kallemeijn WW, Aten J, Li KY, Strijland A, Donker-Koopman WE, van den Nieuwendijk AM, Bleijlevens B, Kramer G, Florea BI, Hooibrink B, Hollak CE, Ottenhoff R, Boot RG, van der Marel GA, Overkleeft HS & Aerts JM (2010) Ultrasensitive in situ visualization of active glucocerebrosidase molecules. *Nat Chem Biol* **6**, 907–913.

25 Li KY, Jiang J, Witte MD, Kallemeijn WW, Donker-Koopman WE, Boot RG, Aerts JM, Codée JD, van der Marel GA & Overkleeft HS (2014) Exploring functional cyclophellitol analogues as human retaining beta-glucosidase inhibitors. *Org Biomol Chem* **12**, 7786–7791.

26 Van Weely S, Aerts JM, Van Leeuwen,MB, Heikoop JC, Donker-Koopman WE, Barranger JA, Tager JM & Schram AW (1990) Function of oligosaccharide modification in glucocerebrosidase, a membrane-associated lysosomal hydrolase. *Eur J Biochem* **191**, 669–677.

27 Cravatt BF, Wright AT & Kozarich JW (2008) Activity-based protein profiling: from enzyme chemistry to proteomic chemistry. *Annu Rev Biochem* **77**, 383–414.

28 Nomura DK, Dix MM & Cravatt BF (2010) Activity-based protein profiling for biochemical pathway discovery in cancer. *Nat Rev Cancer* **10**, 630–638.

29 Aerts JM, Kallemeijn WW, Wegdam W, Joao Ferraz M, van Breemen MJ, Dekker N, Kramer G, Poorthuis BJ, Groener JE, Cox-Brinkman J, Rombach SM, Hollak CE, Linthorst GE, Witte MD, Gold H, van der Marel GA, Overkleeft HS & Boot RG (2011) Biomarkers in the diagnosis of lysosomal storage disorders: proteins, lipids, and inhibodies. *J Inherit Metab Dis* **3**4, 605–619.

30 Gaspar P, Kallemeijn WW, Strijland A, Scheij S, Van Eijk M, Aten J, Overkleeft HS, Balreira A, Zunke F, Schwake M, Sá Miranda C & Aerts JM (2014) Action myoclonus-renal failure syndrome: diagnostic applications of activity-based probes and lipid analysis. *J Lipid Res* **55**, 138–145.

31 Marques AR, Aten J, Ottenhoff R, van Roomen CP, Herrera Moro D, Claessen N, Vinueza Veloz MF, Zhou K, Lin Z, Mirzaian M, Boot RG, De Zeeuw CI, Overkleeft HS, Yildiz Y & Aerts JM (2015) Reducing GBA2 Activity Ameliorates Neuropathology in Niemann-Pick Type C Mice. *PLoS One* **10**, e0135889.

32 Ben Bdira F, Kallemeijn WW, Oussoren SV, Scheij S, Bleijlevens B, van Roomen CPAA, Ottenhoff R, van Kooten MJFM, Walvoort MTC, Witte MD, Boot RG, Ubbink M, Overkleeft HS & Aerts JMFG (2017) Stabilization of glucocerebrosidase by active-site occupancy. *ACS Chem Biol* **12**, 1830–1841.

33 Herrera Moro Chao D, Kallemeijn WW, Marques AR, Orre M, Ottenhoff R, van Roomen C, Foppen E, Renner MC, Moeton M, van Eijk M, Boot RG, Kamphuis W, Hol EM, Aten J, Overkleeft HS, Kalsbeek A & Aerts JM (2015) Visualization of Active Glucocerebrosidase in Rodent Brain with High Spatial Resolution following In Situ Labeling with Fluorescent Activity Based Probes. *PLoS One* **10**, e0138107.

34 Slot JW & Geuze HJ (2007) Cryosectioning and immunolabeling. *Nat Protoc* **2**, 2480–2491.

35 Aerts JM, Donker-Koopman WE, van der Vliet MK, Jonsson LM, Ginns EI, Murray GJ, Barranger JA, Tager JM & Schram, AW (1985) The occurrence of two immunologically distinguishable beta-glucocerebrosidases in human spleen. *Eur J Biochem* **150**, 565–574.

CHAPTER 2

Detecting *In vivo* target engagement of conduritol B epoxide and cyclophellitol by activity-based protein profiling

Manuscript published as:

Kuo CL, Kallemeijn WW, Lelieveld LT, Mirzaian M, Zoutendijk I, Vardi A, Futerman AH, Meijer AH, Spaink HP, Overkleeft HS, Aerts JMFG & Artola ME (2019) In vivo inactivation of glycosidases by conduritol B epoxide and cyclophellitol as revealed by activity-based protein profiling. *FEBS J* **286**, 584–600.

ABSTRACT

Glucocerebrosidase (GBA) is a lysosomal β-glucosidase catalyzing the hydrolysis of glucosylceramide into glucose and ceramide. Its deficiency causes Gaucher disease (GD), the most common lysosomal storage disorder (LSD). Carrying a genetic abnormality in GBA also constitutes at present the largest genetic risk factor for Parkinson's disease (PD). Conduritol B epoxide (CBE) is a commonly used pharmacological agent to generate cell and animal models for investigations on GD and PD, due to its perceived specificity towards GBA. It's inhibition towards GBA is mechanism-based and irreversible, but it may have additional *in vivo* glycosidase targets besides GBA, depending on the applied dose and treatment duration. In this chapter, the first *in vivo* target engagement study for CBE is presented. A suite of activity-based probes and gel-based fluorescence detection was employed, allowing visualization of the covalent catalytic pocket occupancy by CBE on candidate off-target glycosidases in cultured cells, zebrafish larvae, and brain of mice, at various CBE concentrations and fixed incubation time. Comparatively studied is the *in vivo* target engagement of cyclophellitol (CP)—another structurally related GBA inhibitor that has been used for generate GD animal model. The results revealed that only at significantly higher CBE concentrations, non-lysosomal glucosylceramidase (GBA2) and lysosomal α-glucosidase (GAA) were identified as major off-targets in cells and zebrafish larvae. A tight, but acceptable window for selective inhibition of GBA in the brain of mice was observed. On the other hand, CP, a closer glucose mimic, was found to inactivate with equal affinity GBA and GBA2 and therefore is not suitable to generate genuine GD-like models.

2.1 Introduction

The lysosomal enzyme glucocerebrosidase (GBA, EC 3.2.1.45) is a retaining β-glucosidase that degrades the glycosphingolipid, glucosylceramide. Inherited deficiency of GBA is the cause of autosomal recessive Gaucher disease (GD).[1] Most GD patients display heterogeneous symptoms including spleen and liver enlargement, bone deterioration, anaemia, leukopenia, and thrombocytopenia. Some patients also develop fatal neurological symptoms.[2] The GBA genotype poorly predicts the onset and severity of disease in individual GD patients, even in monozygotic twins.[3, 4] Carriers of a GBA defect do not develop GD but show a markedly increased risk for Parkinson's disease (PD) and Lewy-body dementia.[5, 6] The molecular basis for this risk is unknown and a subject of research.

Because complete genetic abrogation of GBA results in premature death in mice, research models of GBA deficiency are often generated with conduritol B epoxide (CBE) (**Figure 2.1A**).[7–9] CBE is a cyclitol epoxide that covalently and irreversibly reacts with the catalytic nucleophile of GBA and thus inactivates irreversibly the enzyme (**Fig. 2.1B**). The crystal structure of GBA with bound CBE confirmed the covalent linkage of the compound to the catalytic nucleophile Glu340.[10, 11] Building on the initial work by Kanfer and co-workers, a regimen using different doses of CBE has been established to generate a phenotypic copy of neuronopathic GD in mice.[9–12] This pharmacological model is now widely used to study the nature of neuropathology resulting from GBA deficiency, including Parkinsonism.[13–15]

A major advantage of CBE's pharmacological use in cultured cells and mice is its tunability: the extent of GBA inactivation can be adjusted by variation of the inhibitor concentration and/or exposure time.[9] However, this has led to use of distinct treatment regimens across studies: exposure of cells ranging from 50 μM to 100 mM CBE for 2 hours up to 60 days[16–22] and daily exposure of mice from 25 to 300 mg kg^{-1} body weight during 2 hours up to 36 days.[9] The use of a high CBE concentration raises concerns about specificity since the compound has been reported to inhibit at high concentration other glycosidases than GBA. Examples are *in vitro* inhibition of retaining α-glucosidases (EC 3.2.1.20)[23–26], *in vitro*[27] and *in situ*[28] cell inhibition of the non-lysosomal glucosylceramidase (GBA2, EC 3.2.1.45), and inhibition of the lysosomal β-glucuronidase (GUSB, EC 3.2.1.31) in mice[29]. The reactivity of CBE towards both β- and α-glucosidases can be explained by the C2-symmetry found in its structure[26] (**Fig. 2.1B**), which

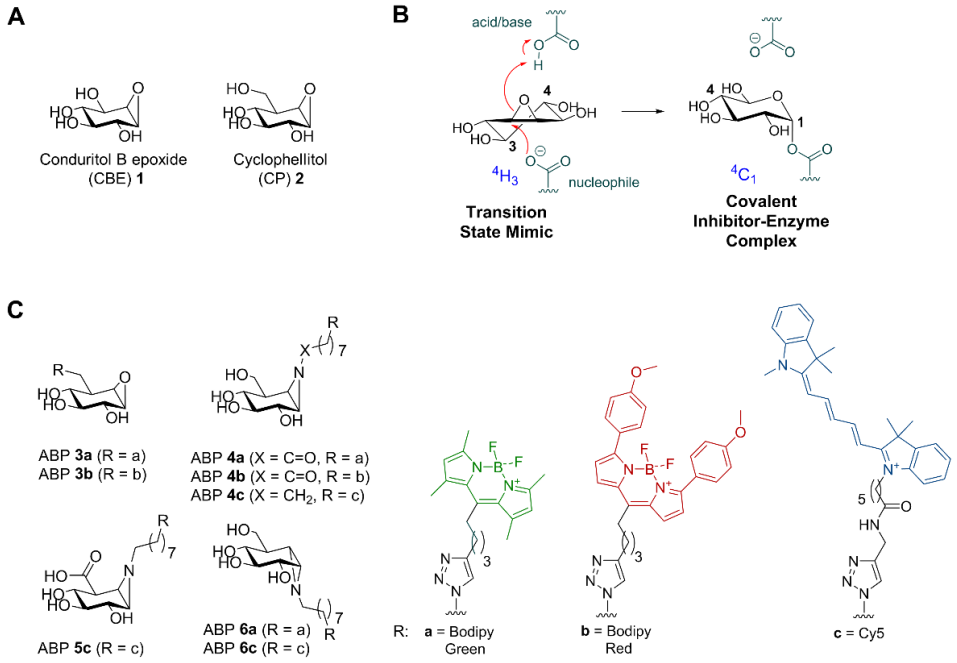


Figure 2.1. Structures of compounds used in this study and inactivation of β-glucosidases by CBE. A) Chemical structure of CBE **1** and cyclophellitol (CP) **2**. B) Reaction coordinates of CBE during inhibition of β-glucosidases. C) Activity-based probes (ABPs) used in this study: GBA ABPs **3a** and **3b**, GBA and GBA2 ABPs **4a-c**, GUSB ABP **5c**, and GAA and GANAB ABPs **6a** and **6c**.

allows reaction with the catalytic nucleophile of both classes of enzymes. Another structurally related cyclitol epoxide, cyclophellitol (CP, **Fig. 2.1A**), is a structurally closer β-glucose mimic and inhibits GBA with far higher affinity than CBE.[30, 31] It exhibits selectivity over α-glucosidases due to the C5-hydroxymethyl group[30–32], and was also shown to induce Gaucher phenotypes in mice.[30] Its reactivity *in vivo* towards GBA2 and other glycosidases is unknown.

The aim of this chapter is to systematically study the *in vivo* selectivity of CBE and CP in cells and animal models. It is envisioned that besides the traditional enzymatic assays employing fluorogenic substrates, activity-based probes (ABPs) could be superior tools for this study. Unlike enzymatic substrate assays, which do not easily distinguish similar enzymatic activities such as GBA vs GBA2, ABPs would allow direct and unambiguous visualization of respective target glycosidases that are not occupied/inactivated by CBE or CP at the active site pocket. Cravatt and co-workers and van der Stelt and colleagues earlier used ABPs directed towards

proteases and lipases in a competitive activity-based protein profiling (ABPP) approach to identify *in vivo* target engagement of small compounds.[33–35] Used ABPs in this study are cyclophellitol-epoxide tagged with a fluorophore that labels specifically GBA[32], and appropriately configured cyclophellitol-aziridines tagged with a fluorophore that label multiple β-glucosidases, (GBA and GBA2)[36], β-glucuronidase (GUSB) (descried in **Chapter 4**, this thesis)[37] and α-glucosidases (GAA and GANAB) (descried in **Chapter 4**, this thesis)[38] (**Fig. 2.1C**, ABPs **3**−**6**). Through parallel application of both the competitive ABPP method and enzymatic assay in lysates of cultured cells, zebrafish (*Danio rerio*) larvae, and brains of mice treated *in vivo* with a relevant range of concentrations of CBE or CP, the *in vivo* off-targets and selectivity windows for GBA by both CBR and CP i*n vivo* was systematically studied and reported.

2.2 Results

2.2.1 A competitive ABPP method to determine GBA target engagement of CBE and CP

ABPs specifically label GBA have been reported, in which their use in profiling GBA in cells and identifying active-site pocket interactors have been demonstrated (**Fig. 2.1C**, ABP **3** and **4**).[32, 39] Here, they were employed to visualize GBA target engagement of CBE **1** and CP **2** using a competitive ABPP method, by which the irreversible occupancy of the catalytic nucleophile of GBA by the inhibitors during pre-incubation is assessed. As a validation, competition of ABP labeling by CBE and CP was compared to the loss of GBA activity measured using 3.75 mM 4-methylumbelliferyl-β-glucoside as substrate.[40] For this, recombinant human GBA (rGBA) was pre-incubated with CBE across a range of concentrations at 37 °C for 0, 30 or 180 minutes in McIlvaine buffer (pH 5.2) containing 0.2 % taurocholate and 0.1 % Triton X-100, a condition optimal for enzymatic activity.[41] Subsequent labeling of GBA by ABP **4c** was quantified by SDS-PAGE and fluorescence scanning. IC_{50} values (concentrations of inhibitor yielding a 50 % reduction of ABP **4c** labeling) were determined and found to be 26.6 μM at 30 min CBE pre-incubation, and 2.30 μM at 180 min pre-incubation (**Fig. 2.2A**, **Table 2.1**). These values match the ones determined by measurement of residual enzymatic activity of GBA assay (**Fig. 2.2A** lower right panel, **Table 2.1**), validating the competitive ABPP methodology.

Next, CP was comparably studied and its IC_{50} values determined by ABPP were 0.15 μM at 30 min pre-incubation and 0.03 μM at 180 min pre-incubation, comparable to values

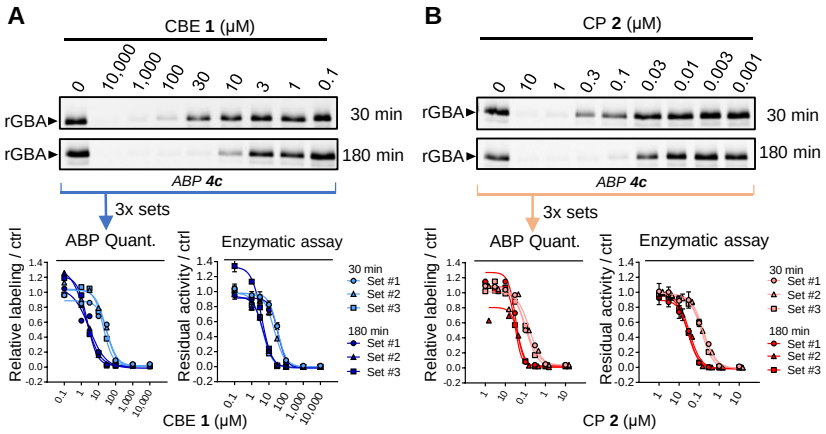


Figure 2.2. Effect of pre-incubation with CBE or CP on ABP labeling of recombinant GBA. A) Representative gel images (1 set from *n* = 3 biological replicates) of rGBA pre-incubated with CBE for 30 min or 180 min and fluorescently labeled with ABP **4c**. Fluorescent signals were quantified and normalized to the untreated sample (ctrl, 0 μM CBE) (bottom left), and compared to the inhibition curves obtained with enzymatic assay (bottom right). B) Same as A), with CP at 10–0.001 μM. Error ranges in enzymatic assay = ± SD, *n* = 3 (technical replicates).

Table 2.1. *In vitro* inhibition of CBE or CP towards human recombinant GBA. Apparent IC_{50} values (μM) were derived from the average of 3 individual experiments as measured by either enzymatic assays (*Enz. assay*) or quantification of ABP labeling on residual active enzymes (*ABP*). Error ranges = ± SD, *n* = 3 biological replicates.

Incubation time	CBE 1		CP 2	
	Enz. Assay	ABP	Enz. Assay	ABP
30 min	28.8 ± 6.90	26.6 ± 7.03	0.150 ± 0.013	0.118 ± 0.034
180 min	4.28 ± 0.50	2.30 ± 0.58	0.030 ± 0.002	0.036 ± 0.008

determined by enzymatic assay (**Fig. 2.2B**, **Table 2.1**).

2.2.2 *In vivo* targets of CBE

Next analysed were the targets of CBE in intact human cells, zebrafish larvae, and brain of treated mice. These biological materials contain besides GBA the candidate off-target glycosidases: GBA2, α-glucosidases (GAA and GANAB), and lysosomal β-glucuronidase GUSB. For each of these enzymes ABPs have been designed, and enzymatic activity assays with fluorogenic substrates established.[36–38] Of note, ABP **4** allows simultaneous visualization of

active GBA (58-66 kDa) and GBA2 (110 kDa) following SDS-PAGE analysis. HEK293T cells expressing GBA2 were exposed for 24 h to different concentrations of CBE (0.1 μM–10 mM) after which the residual amount of GBA, GBA2, GAA, GANAB and GUSB in cell lysates that can still be labeled with the appropriate ABPs (**Fig. 2.1C**) was determined (**Fig. 2.3A**). Competitive ABPP showed that besides GBA, all other candidate off-target enzymes are inactivated by CBE but with marked lower affinity: GBA (IC_{50} = 0.59 μM), GBA2 (IC_{50} = 315 μM), GAA (IC_{50} = 249 μM), GANAB (IC_{50} = 2900 μM) and GUSB (IC_{50} = 857 μM). (**Fig. 2.3B**, **Table 2.2**). Comparable results were obtained by determination of residual enzyme activities: GBA (IC_{50} = 0.33 μM), GBA2 (IC_{50} =272 μM), GAA (IC_{50} = 309 μM), GANAB (IC_{50} = 1580 μM) and GUSB (IC_{50} = 607 μM) (**Fig. 2.3C**, **Table 2.2**). Next, cultured human fibroblasts, rich in lysosomal enzymes, were used to analyse a panel of additional glycosidases (α- and β-mannosidase, *N*-acetyl α-galactosidase, β-hexosaminidase, α-fucosidase, α-iduronidase, α- and β-galactosidase) for their possible inactivation by CBE. At the highest concentration of CBE (10 mM) tested, no significant loss of activity was observed for any of these additional lysosomal enzymes (**Fig. 2.S1**). In addition, *in vivo* competitive ABPP experiments were also performed. For this, the appropriate ABPs were added to the culture medium of HEK293T cells during 4 hours prior to their harvesting. The outcome of this *in vivo* ABP labeling was comparable to that of *in vitro* ABP labeling in the lysates of harvested cells (**Fig. 2.S2**, **Table 2.S1**).

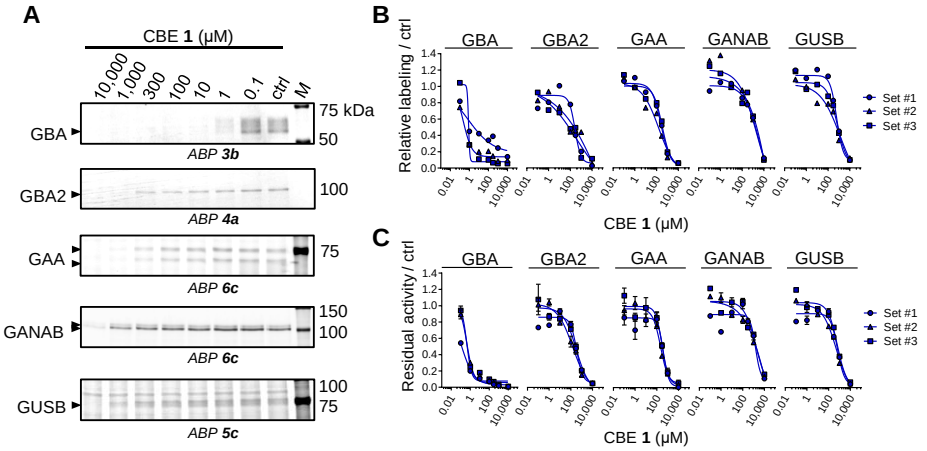


Figure 2.3. *In vivo* glycosidase targets of CBE in cultured cells. A) Representative gel images (1 set from *n* = 3 biological replicates) showing fluorescent ABP labeling of GBA, GBA2, GAA, GANAB, and GUSB in lysates of cells treated *in vivo* for 24 h with CBE. B) Quantification of relative ABP labeling. C) Residual activities of glycosidases in cell lysates treated *in vivo* with CBE. Error ranges = ± SD, *n* = 3 (technical replicates).

Table 2.2. *In vivo* inhibition of CBE or CP towards glycosidases in cultured cells. Apparent IC_{50} values (μM) were derived from the average of biological triplicates as measured by either *in vitro* enzymatic assays (*Enz. assay*) or quantification of *in vitro* ABP labeling on residual active enzymes (*ABP*). -, value cannot be calculated (no inhibition at the tested concentrations). Error ranges = ± SD, *n* = 3 biological replicates.

	CBE 1		CP 2	
Enzyme	Enz. Assay	ABP	Enz. assay	ABP
GBA	0.331 ± 0.189	0.594 ± 0.316	0.086 ± 0.003	0.063 ± 0.035
GBA2	272 ± 101	315 ± 62.8	0.198 ± 0.008	0.154 ± 0.070
GAA	309 ± 88.2	249 ± 83.9	-	-
GANAB	1580 ± 116	2900 ± 1120	-	-
GUSB	607 ± 70.0	857 ± 341	-	-

Next, CBE target engagement was investigated in intact zebrafish larvae. Fertilized zebrafish eggs (0 dpf) were exposed to CBE (1 μM–100 mM) in the egg water for five days. The larvae were collected, lysed, and analysed by the competitive ABPP method. Exposure to 100 mM CBE was found to reduce ABP labeling of all five glycosidases (**Fig. 2.4A**, **Fig. 2.S3A**). The IC_{50} values determined by the competitive ABPP method were: GBA (IC_{50} = 44.1 μM), GBA2 (IC_{50} = 890 μM), GAA (IC_{50} = 9550 μM), GANAB (IC_{50} = 4700 μM) and GUSB (IC_{50} = 6470 μM) (**Table 2.3**). Thus, inactivation of GBA in zebrafish larvae takes place 20-fold more avidly than that of GBA2 and 100- to 200-fold more potently than that of GAA, GANAB and GUSB. Analysis of residual enzymatic activity of the various enzymes gave similar results (**Table 2.3**, **Fig. 2.S3B**). Analysis by enzymatic assay revealed that exposure of the animals to 10 mM CBE did not lead to significant inactivation of other glycosidases (α- and β-mannosidase, *N*-acetyl α-galactosidase, β-hexosaminidase, α-fucosidase and α-iduronidase) except for β-galactosidase (IC_{50} = 11.2 mM) and α-galactosidase (IC_{50} = 22.5 mM) (**Fig. 2.S3B**).

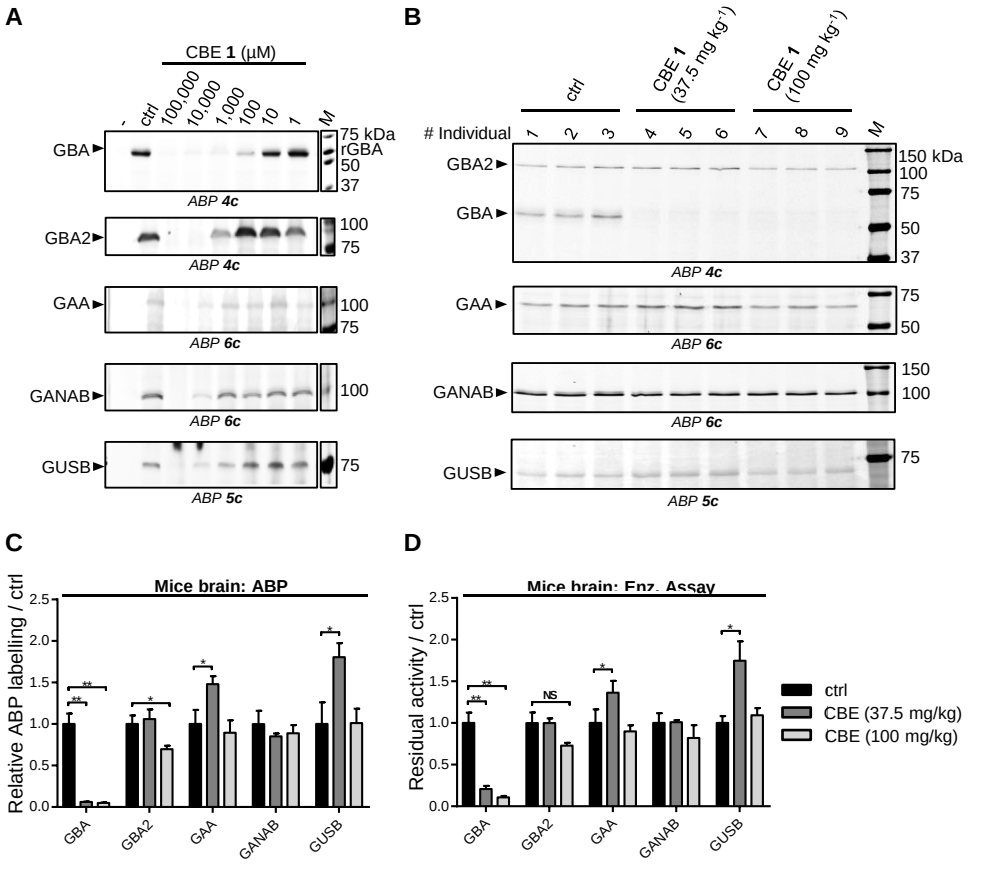


Figure 2.4. ***In vivo*** **glycosidase targets of CBE in animal models.** A) Effect of CBE incubation in zebrafish to ABP labeling of glycosidases in zebrafish lysates. B) Effect of CBE injection in mice to ABP labeling of glycosidases in mice brain homogenates. C) Quantification of labeled bands in B). D) Residual activity of glycosidases in brain homogenates of CBE-injected mice. Error ranges = ± SD, *n* = 3 (biological replicates). Two-tailed unpaired t-test was used to derive statistical significance (* $p < 0.05$, ** $p < 0.01$, NS = not significant).

Finally, *in vivo* targets of CBE was also investigated the in brain of mice treated daily from day 8 to either day 25 with 37.5 mg CBE kg^{-1} body weight or to day 14 with 100 mg CBE kg^{-1} body weight.[9] Brain homogenates were prepared from three individuals of each treatment group and the untreated group. As above, both competitive ABPP and measurement of residual enzymatic activity was performed. Interestingly, GBA was found to be the only enzyme significantly targeted by CBE in the brain of mice treated with both CBE dosages (**Fig. 2.4B**). Quantification of the gels revealed only a slight reduction of ABP labeling of GBA2 in brain of

Table 2.3. *In vivo* inhibition of CBE or CP towards glycosidases in zebrafish larvae. Apparent IC_{50} values (μM) were derived from measurements from lysates made with n = 48 individuals. *Enz. assay*, enzymatic assay. *ABP*, activity-based probes detection. -, value cannot be calculated (no inhibition at the tested concentrations).

	CBE 1		CP 2	
Enzyme	**Enz. Assay**	**ABP**	**Enz. assay**	**ABP**
GBA	58.50	44.1	0.130	0.083
GBA2	1160	890	0.176	0.059
GAA	5010	9550	-	-
GANAB	4550	4700	-	-
GUSB	6380	6470	-	-

mice treated with 100 mg CBE kg^{-1} body weight (**Fig. 2.4C**), but this reduction (30 %) is far less pronounced than the complete reduction of GBA labeling in the same mice. Measurement of residual enzymatic activities in these samples yielded comparable results (**Fig. 2.4D**), showing that besides GBA the other enzymes GBA2, GAA, GANAB, and GUSB were not affected in the brains of CBE-treated mice. Taking together these results, it can be confirmed that there is a therapeutic window for selective inactivation of GBA with CBE. However, when using higher CBE concentration, looking whether concomitant inhibition of GBA2 takes place is warranted.

2.2.3 *In vivo* targets of CP

CP has been previously used to generate a GD mouse model.[30] Compared to CBE, CP is a much more potent inhibitor of GBA[30] and reported to inhibit poorly α-glucosidases *in vitro*.[31] However, its *in vivo* reactivity towards other glycosidases has not been thoroughly investigated. The *in vivo* targets of CP were therefore comparatively studied in cultured cells and zebrafish larvae using the competitive ABPP methodology and measurement of the residual enzymatic activities. To compare the selectivity windows of CP to the ones of CBE, the concentration range of CP in living models was chosen at 0.001–10 μM for cultured cells and 0.001–100 μM for zebrafish, so that it would match the extent of GBA inhibition by CBE *in vitro* (**Fig. 2.2A**, **B**). As determined by the competitive ABPP method, CP was found to inhibit GBA2 on a par with GBA in HEK293T cells upon incubation with varying inhibitor concentrations (0.1−10 μM) for 24 hours. IC_{50} values of CP for blocking ABP labeling were 0.063 μM for GBA and 0.154 μM for GBA2 (**Fig. 2.5A**, **B**, **Table 2.2**). No reduction of ABP labeling of GAA, GANAB

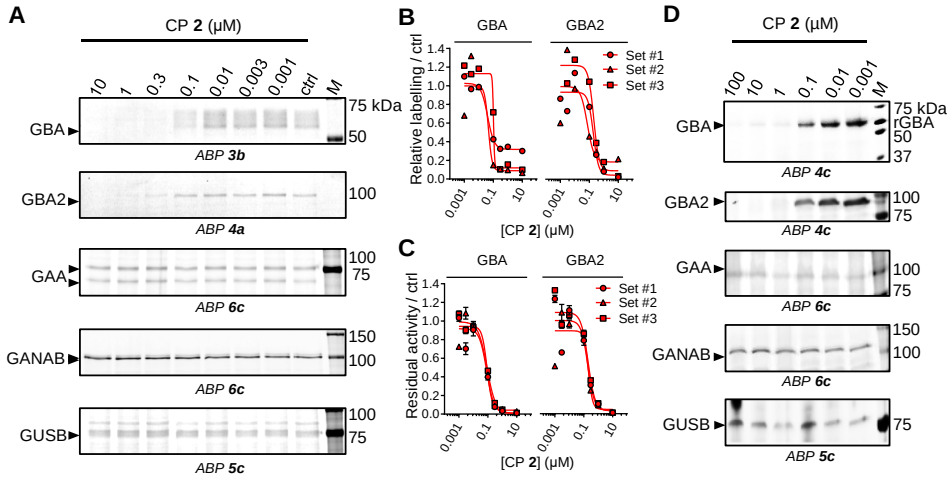


Figure 2.5. *In vivo* glycosidase targets of CP. A) Representative gel images (1 of 3 biological replicates) showing fluorescent ABP labeling of GBA, GBA2, GAA, GANAB, and GUSB in lysates of cells treated *in vivo* for 24 h with CP. B) Quantification of relative ABP labeling of GBA and GBA2. C) Residual activities of GBA and GBA2 in cell lysates treated *in vivo* with CP. Error ranges = ± SD, *n* = 3 (technical replicates). D) Effect of CP incubation in zebrafish to ABP labeling of glycosidases in zebrafish lysates.

and GUSB was observed in lysates of cells incubated for 24 h with the highest concentration of CP (10 μM) (**Fig. 2.5A**). As determined by enzymatic assay, the apparent IC_{50} values for inactivation were quite comparable, being 0.086 μM for GBA and 0.198 μM for GBA2 (**Fig. 2.5C**, **Table 2.2**). For the other enzymes the IC_{50} values exceeded at least 10 μM (**Fig. 2.S1**, **Table 2.2**).

Exposure of zebrafish to CP (5 days at 1–100 μM) also comparably competed GBA and GBA2 labeling, but not that of GUSB, GAA and GANAB (**Fig. 2.5D**). This finding was again supported by results obtained from measurement of residual enzymatic activities (**Table 2.3**, **Fig. 2.S3**). From the noted lack of selectivity of CP with respect to inactivation of GBA and GBA2, it is obvious that CP does not allow generation of specific GBA-deficiency in cell and animal models.

2.2.4 Inhibitor sensitivity of GBA3

Besides GBA and GBA2, there exists in humans and zebrafish another β-glucosidase named GBA3 (**EC 3.2.1.21**).[42] This is a cytosolic enzyme able to degrade a variety of glycoside substrates. GBA3 is not sensitive towards inhibition by CBE. [42] The enzyme is selectively

expressed in cells and tissues, for example being particularly high in kidney and absent in fibroblasts, and its activity can be visualized by labeling with nanomolar concentrations of ABP **4**.[36] In lysates of mice brain (**Fig. 2.4B**) and 5 dpf zebrafish larvae (**Fig. 2.4A**, **2.5D**), the presence of GBA3 could not be detected (expected m.w. = 45–53 kDa[42]) by labeling with 200 nM ABP **4c**. Therefore, HEK293T cells over-expressing GBA3 were generated to study the inhibitor sensitivity of the enzyme. In these cells the *in vivo* interaction of CBE and CP with GBA3 could be measured by ABP detection (ABP **4c**) (**Fig. 2.6A**, **B**). The apparent IC_{50} values towards GBA3 for CBE and CP were 485 µM and > 10 µM, respectively (**Table 2.4**). The inhibition of GBA3 by the two compounds was also measured by enzymatic assays, and similar results were obtained (apparent IC_{50} values for CBE and CP were 474 µM and > 10 µM) (**Fig.**

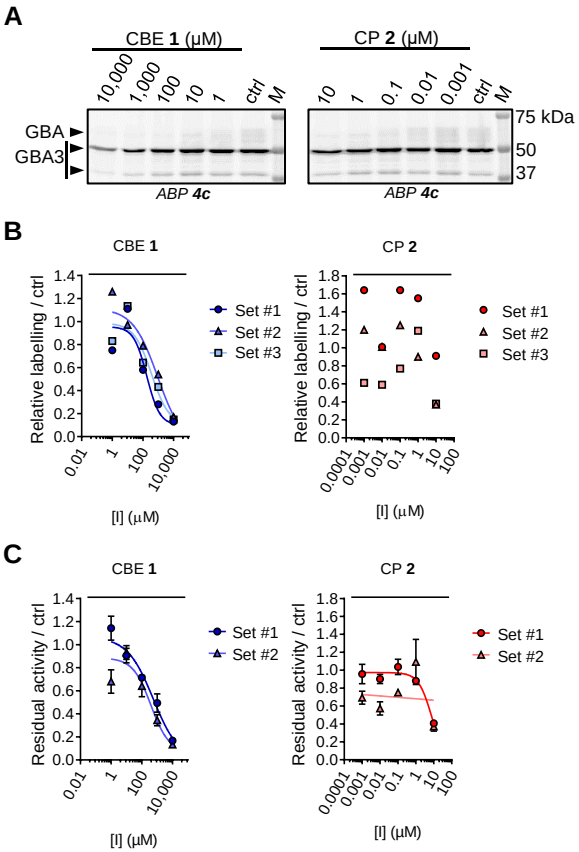


Figure 2.6. *In vivo* GBA3 inhibition by CBE and CP in cultured cells. A) Representative fluorescent gel images (1 of 3 biological replicates) of ABP labeling of GBA3 in lysates of HEK293T cells treated with CBE or CP for 24 h. B) Quantification of ABP gels . C) Residual activities of GBA3 by enzymatic assay. Error ranges = ± SD, *n* = 3 (technical replicates).

Table 2.4. Overview of apparent IC_{50} values (µM) determined in this study. *Enz. assay*, enzymatic assay. *ABP*, activity-based probes (ABP) detection. Error ranges = ± SD, *n* =3 biological replicates.

		CBE 1 (µM)		CP 2 (µM)	
	Enzyme	**Enz. Assay**	**ABP**	**Enz. Assay**	**ABP**
rGBA (*in vitro* 3h)	GBA	**4.28** ± 0.50	**2.30** ± 0.58	**0.030** ± 0.002	**0.036** ± 0.008
Cultured cells (*in vivo* 24h)	GBA	0.331 ± 0.189	0.594 ± 0.316	0.086 ± 0.003	0.063 ± 0.035
	GBA2	272 ± 101	315 ± 62.8	0.198 ± 0.008	0.154 ± 0.070
	GAA	309 ± 88.2	249 ± 83.9	>10	>10
	GBA3	474 ± 124	485 ± 386	>10	>10
Ratio	***GBA2 /GBA***	***883***	***530***	***2.30***	***2.44***
	GAA /GBA	***934***	***418***	***>116***	***>159***
	GBA3 /GBA	***1430***	***816***	***>116***	***>159***
***Danio rerio* larvae** (*in vivo* 5d)	GBA	58.5	44.1	0.130	0.083
	GBA2	116	890	0.176	0.059
	GAA	5010	9550	>100	>100
Ratio	***GBA2 /GBA***	***19.8***	***20.2***	***1.35***	***0.710***
	GAA /GBA	***85.6***	***226***	***>769***	***>1200***

2.6C, **Table 2.4**). Hence, both compounds show selectivity towards GBA over GBA3; the ratio of IC_{50} values for GBA3/GBA by ABP detection were 816 for CBE and > 159 for CP (**Table 2.4**).

2.2.5 Impact of CBE and CP on glycosphingolipids in exposed zebrafish larvae

Next studied were the functional impact of exposing fish embryos for 5 days to CBE or CP. It is known that *in vivo* inactivation of GBA leads to compensatory formation of glucosylsphingosine (GlcSph) by acid ceramidase- mediated conversion of accumulating GlcCer

in lysosomes.[40, 43] In other words, formation of GlcSph is a biomarker of the inactivation of GBA. The exposure of zebrafish larvae to CBE and CP led to pronounced accumulation of GlcSph (4-fold increase in the case of 1000 µM CBE; 6-fold increase at 10 µM CP) (**Fig. 2.7**). Thus, the observed *in vivo* GBA inhibition by both compounds according to ABP detection was confirmed by the accumulation of GlcSph at comparable inhibitor concentrations.

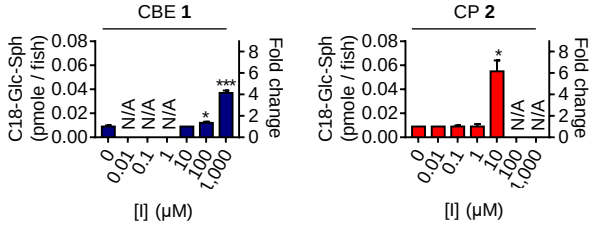


Figure 2.7. C18-Glc-Sphingosine levels in zebrafish larvae treated for five days with CBE or CP. Error ranges ± SD, *n* = 3 (biological replicates). Two-tailed unpaired t-test was used to derive statistical significance (* p < 0.05, ** p < 0.01, *** p < 0.001). *N/A*, not determined

2.3 Discussion

Our present study demonstrates that gel-based fluorescence ABPP can be appropriately used to determine *in vivo* target engagement of irreversible glycosidase inhibitors such as CBE and CP. This method provides important insights on dose-dependent off-targets of the tested compounds. The off-targets may vary among cells and organismal models. From our findings, it is concluded that it is wise to test pharmacologically induced Gaucher disease models with respect to selectivity of GBA inactivation.

In this chapter, the *in vivo* target engagement of CBE is assessed in cells and zebrafish model, and compared to that of CP. **Table 2.4** provides an overview of the IC_{50} values observed for CBE and CP across rGBA, cultured cells, and zebrafish larvae. The table highlights that although CBE has *in vivo* off-targets such as GBA2 and the lysosomal α-glucosidase GAA, it is still selective towards GBA in cells (GBA2/GBA inhibition ratio = 530; GAA/GBA inhibition ratio = 418; ABP detection) and in zebrafish larvae (GBA2/GBA inhibition ratio = 20.2; GAA/GBA inhibition ratio = 226; ABP detection). Thus, a selective window for *in vivo* GBA inactivation by CBE exists (in cells: 0.6−315 µM; in zebrafish larvae: 44−890 µM; ABP detection) (**Fig. 2.8**). Importantly, such selectivity has also been observed by us in brain of mice treated

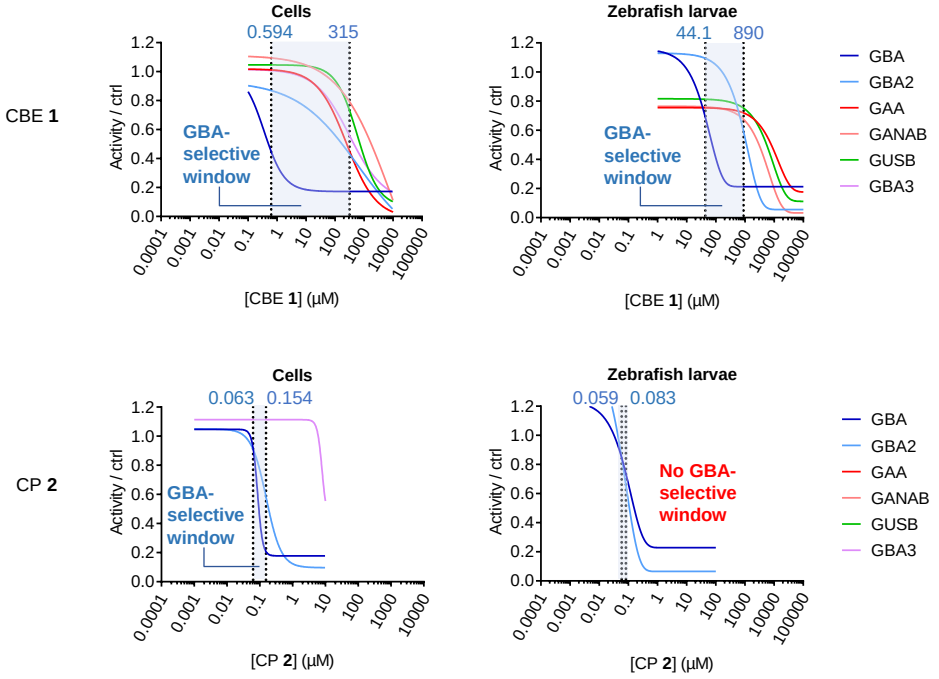


Figure 2.8. Windows for selective GBA inhibition by CBE and CP in cultured cells and zebrafish larvae. Inhibition curves for CBE or CP towards GBA and other glycosidases were derived from the results of ABP detection (average of *n* = 3 biological replicates). GBA selective window is shown in blue area, defined as the concentration range for CBE or CP between its IC_{50} values towards GBA and the next major glycosidase target (in all cases, GBA2).

with 37.5 or 100 mg CBE kg^{-1} body weight (**Fig. 2.4B**). It therefore can be concluded that the CBE treatment generates a valuable neuronopathic Gaucher disease model. It should however be stressed that the use of higher CBE concentrations or longer incubation periods may cause undesired inhibition of other glycosidases.

The presented investigation also reveals that CP is a more potent GBA inactivator. However, CP inhibits GBA2 on a par with GBA in both cultured cells and zebrafish larvae (**Table 2.4**), and therefore does not offer a window for selective GBA inactivation (**Fig. 2.8**). In other words, CP seems of little use to generate a Gaucher disease model. Previous work showed that a CP functionalized at C6 with a BODIPY moiety is an ABP that inactivates potently and selectively GBA.[32] This compound was found not to penetrate well into the brain, likely by active removal via Pgp proteins[39, 44]. It will be of interest to design CP analogues that inactivate GBA

with higher selectivity and concomitantly are brain-permeable. Such compounds might create the desired genuine Gaucher disease models solely deficient in GBA.

2.4 Experimental procedures

2.4.1 General materials and methods

Cyclophellitol (CP) and the ABPs were synthesized as described earlier.[32, 36–38, 45] Chemicals were obtained from Sigma-Aldrich (St. Louis, MO, USA) if not otherwise indicated. Conduritol B-epoxide (CBE) was purchased from Enzo Life Sciences (Farmingdale, NY, USA). Recombinant GBA (rGBA, imiglucerase) was obtained from Sanofi Genzyme (Cambridge, MA, USA). Human fibroblasts (CC-2511) were obtained from Cambrex-Lonza (East Rutherford, NJ, USA). HEK293T (CRL-3216) and RAW-264.7 (TIB-71) cell lines were purchased from ATCC (Manassas, VA, USA). HEK293T cells overexpressing human GBA2 and GBA3 were generated as previously described[46]. Cell lines were cultured in HAMF12-DMEM medium (Sigma-Aldrich) for fibroblasts or DMEM medium (Sigma-Aldrich) for HEK293T cell line, supplied with 10 % (v/v) FCS, 0.1% (w/v) penicillin/streptomycin, and 1 % (v/v) Glutamax, under 5 % CO_2 (fibroblasts) or 7 % CO_2 (HEK293T). Zebrafish (*Danio rerio*) were handled and maintained according to standard protocols (zfin.org). Adult zebrafish were housed at a density of 40 per tank, with a cycle of 14 h of light and 10 h of dark. Adults, embryos and larvae were kept at a constant temperature of 28.5 °C. Embryos and larvae were raised in egg water (60 μg L^{-1} sea salt, Sera marin). Synchronized wild-type ABTL zebrafish embryos were acquired after mating of single male and female couples (both > 3 months old). Frozen brain samples from wild-type and CBE-treated mice were obtained from a previous study.[9] Protein concentration was measured using Pierce BCA assay kit (Thermo Fisher Scientific, Waltham, MA, USA).

2.4.2 Enzymatic assays

All assays were performed in 96-well plates at 37 °C for human, zebrafish, and mice material. Samples were diluted with McIlvaine buffer (150 mM citric acid—Na_2HPO_4) to a final volume of 25 μL, at pH appropriate for each enzyme. Assays were performed by incubating the samples with 100 μL 4-MU- (4-methylumbelliferyl-) substrates diluted in McIlvaine buffer (with 0.1% (w/v) bovine serum albumin (BSA)) for a period of 30 min to 2 h. After stopping the substrate reaction with 200 μL 1M NaOH-glycine (pH 10.3), 4-MU-emitted fluorescence was measured with a fluorimeter LS55 (Perkin Elmer, Waltham, MA, USA) using λ_{EX} 366 nm and λ_{EM} 445 nm.[32] Measured activities were subtracted with background values (from samples without enzyme), normalized with the average values from the control samples (no inhibitor), and curve-

fitted to inhibitor concentrations using Prism 7.0 (GraphPad Software, San Diego, CA, USA) by the [inhibitor] vs response—various slope (four parameters) method to obtain IC_{50} values. The substrate mixtures used for each enzyme are listed as follows: GBA, 3.75 mM 4-MU-β-D-glucopyranoside (Glycosynth, Warrington Cheshire, UK) at pH 5.2, supplemented with 0.2% (w/v) sodium taurocholate and 0.1% (v/v) Triton X-100, and 25 nM *N*-(5-adamantane-1-yl-methoxy-pentyl)-deoxynojirimycin (AMP-DNM), a GBA2-specific inhibitor[47]; GBA2, 3.75 mM 4-MU-β-D-glucopyranoside at pH 5.8, with pre-incubation with 1 μM ABP **3a** for 30 min to specifically inhibit GBA activity; α-glucosidases, 3 mM 4-MU-α-D-glucopyranoside at pH 4.0 (GAA) or at 7.0 (GANAB), GUSB, 2 mM 4-MU-β-D-glucuronide at pH 5.0; α-galactosidases, 2 mM 4-MU-α-D-galactopyranoside at pH 4.6; β-galactosidases, 1 mM 4-MU-β-D-galactopyranoside at pH 4.3 with 0.2 M NaCl; α-mannosidases, 10 mM 4-MU-α-D-mannopyranoside at pH 4.0; β-mannosidases, 2 mM 4-MU-β-D-mannopyranoside (Glycosynth) at pH 4.2; β-hexosaminidase HexA, 5 mM 4-MU-β-D-6-sulpho-2-acetamido-2-deoxy-glucopyranoside at pH 4.4; β-hexosaminidases HexA/B, 5 mM 4-MU-β-*N*-acetyl-glucosaminide at pH 4.5; α-*N*-acetyl-galactosaminidase, 1 mM 4-MU-α-*N*-acetyl-galactosaminide at pH 4.5; α-L-fucosidase, 1 mM 4-MU-α-L-fucopyranoside at pH 5.0, α-L-iduronidase, 2 mM 4-MU-α-L-iduronide (Glycosynth) at pH 4.0; GBA3, 3.75 mM 4-MU-β-D-glucopyranoside at pH 6.0.

2.4.3 Fluorescent ABP labeling and detection

Residual active, not irreversibly inhibited glycosidases were labeled with excess fluorescent ABPs in the optimum McIlvaine buffer, if not otherwise stated (see above). ABP labeling was performed at 37 °C for 30 min for all materials, in a total sample volume of 20–40 μL and 0.5–1 % DMSO concentration. GBA was labeled with 200 nM ABP **3b** (pH 5.2, 0.1 % (v/v) Triton-100, 0.2 % (w/v) sodium taurocholate), or labeled together with GBA2 using 200 nM β-aziridine ABP **4c** at pH 5.5. GBA2 was labeled with 200 nM β-aziridine ABP **4a**, **4b** or **4c**. The α-glucosidases GAA and GANAB were first pre-incubated with 200 nM ABP **4a** for 30 min (pH 4.0 for GAA and pH 7.0 for GANAB), followed by labeling with 500 nM ABP **6a** or **6c** at pH 4.0 or 7.0**.** The β-glucuronidase GUSB was pre-incubated with 200 nM ABP **4a** for 30 min, followed by labeling with 200 nM β-aziridine ABP **5c**. After ABP incubation, proteins were denatured by boiling the samples with 5× Laemmli buffer (50 % (v/v) 1 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) for 5 min at 98 °C, and separated by electrophoresis on 7.5 % or 10 % (w/v) SDS-PAGE gels

running continuously at 90V.[32, 36–38] Wet slab-gels were scanned on fluorescence using the Typhoon FLA 9500 (GE Healthcare) at λ_{EX} 473 nm and $\lambda_{EM} \geq 510$ nm for green fluorescent ABP **4a** and **6a**; at λ_{EX} 532 nm and $\lambda_{EM} \geq 575$ nm for ABP **3b** and **4b**; and at λ_{EX} 635 nm and $\lambda_{EM} \geq 665$ nm for ABP **4c**, **5c** and **6c**. ABP-emitted fluorescence was quantified using ImageQuant software (GE Healthcare, Chicago, IL, USA) and curve-fitted using Prism 7.0 (GraphPad Software). After fluorescence scanning, SDS-PAGE gels were stained for total protein with Coomassie G250 and scanned on a ChemiDoc MP imager (Bio-Rad, Hercules, CA, USA).

2.4.4 *In vitro* effects of inhibitors on rGBA

For IC_{50} measurements using enzymatic assay, 3.16 ng (53 fmole) of rGBA was prepared in 12.5 µL McIlvaine buffer (150 mM, pH 5.2) supplemented with 0.1 % (v/v) Triton X-100, and 0.2 % (w/v) sodium taurocholate, and incubated with 12.5 µL of inhibitors (CBE or CP) diluted in McIlvaine buffer at 37 °C for various time periods. Residual activity of rGBA was measured as described in previous section. For assessing the occupancy of active site pocket by the inhibitors using ABP labeling, same amount of rGBA was prepared in 10 µL of the same McIlvaine buffer, incubated firstly with 2.5 µL inhibitor dilutions prepared in McIlvaine buffer at 37 °C for various time periods, then with 2.5 µL ABP dilutions prepared in McIlvaine buffer; detection of ABP-labeled rGBA follows the procedures described in the previous section.

2.4.5 *In vivo* effects of CBE and CP in intact cultured cells

Confluent HEK293T stably expressing human GBA2 were cultured in 12-well plates in triplicates with(out) CBE (0.01−10,000 µM) or CP (0.001−10 µM) for 24 h at 37 °C. In addition, confluent human fibroblasts were similarly treated in 15 cm dishes for 2, 24, and 72 h. For lysis, cells were washed three times with PBS, subsequently lysed by scraping in potassium phosphate buffer (K_2HPO_4−KH_2PO_4, 25 mM, pH 6.5, supplemented with 0.1 % (v/v) Triton X-100 and protease inhibitor cocktail (Roche, Basel, Switzerland)), aliquoted, and stored at –80 °C. After determination of the protein concentration, lysates containing equal protein amount (5−20 µg total protein per measurement) were adjusted to 12 µL with potassium phosphate buffer and subjected to residual activity measurements using enzymatic assay (n = 3 technical replicates for each biological triplicate at each treatment condition) or ABP detection (n = 3 biological replicates). For *in vivo* ABP labeling, HEK293T cells expressing GBA2 were incubated with

culture medium containing CBE (0.015–15,000 μM) or CP (0.001−10 μM) for 24 hours. Medium was removed and cells were incubated for 4 hours with culture medium containing a mixture of 200 nM ABP **4b**, 1 μM ABP **5c**, and 500 nM ABP **6c**, or with DMSO only (negative control). Lysates were prepared and measured for protein concentration as described above, and samples containing 60 μg total protein (diluted with potassium phosphate buffer to 15 μL total volume) were subjected to ABP detection (n = 3 biological replicates).

2.4.6 *In vivo* effects of inhibitors in living zebrafish larvae

Adult zebrafish were not sacrificed for this study; all experiments were performed on embryos/larvae before the free-feeding stage (120 h, i.e. 5 days post-fertilization) and did not fall under animal experimentation law according to the EU Animal Protection Directive 2010/63/EU. For *in vivo* inhibitor treatment, a single fertilized embryo was seeded in each well of a 96-wells plate, and exposed to 200 μL CBE (1−100,000 μM), or CP (0.001−100 μM) for 120 hours at 28.5 °C. Per condition, n = 48 embryos were used. At 120 hours (5 dpf), larvae were collected, rinsed three times with egg water, fully aspirated, snap-frozen in liquid nitrogen and stored at –80 °C until homogenization in 200 μL 25 mM potassium phosphate buffer per 48 individuals. Lysis was conducted by sonication with a Polytron PT 1300D sonicator (Kinematica, Luzern, Switzerland) on ice at 20 % power for three seconds, and repeated three times. Samples containing 20–45 μg total protein were subjected to ABP detection or enzymatic assay.

2.4.7 *In vivo* effects of CBE in brain of mice

Brain hemispheres were obtained from mice injected daily with CBE at either 37.5 mg or 100 mg kg^{-1} body weight, or PBS, from day 8 until day of sacrifice (day 24 for the 37.5 mg CBE kg^{-1} group; day 14 for the 100 mg CBE kg^{-1} group) as previously described.[9] Brain hemispheres were homogenized in 4x volumes of tissue wet weight in 25 mM potassium-phosphate buffer (4x volume/wet tissue weight) with 1.0 mm glass beads using a Fastprep-24 instrument (MP Biomedicals, Santa Ana, CA, USA) set at 6 m s^{-1} for 20 seconds, repeated three times, while chilling samples on ice for 2 min between separate runs. Crude lysates were isolated from the glass beads by pipetting into sterile Safe-Lock Eppendorf tubes. Homogenates were measured for protein concentration, aliquoted, and snap-frozen in liquid nitrogen. Samples containing 50 μg total protein were subjected to ABP detection or enzymatic assay. Two-tailed unpaired t-test

was performed in Prism 7.0 software (GraphPad) to derive statistical significance, where $p < 0.05$ was considered significant.

2.4.8 *In vivo* effects of inhibitors towards GBA3 in cells

HEK293T cells expressing human GBA3 were cultured, treated (triplicate sets), and lysed in an identical setup as described with GBA2-expressing HEK293T cells. Lysates (12 μg protein) were subjected to ABP detection using 200 nM ABP **4c** at pH 6.0. For enzymatic assay, lysates were separated from GBA2 and GBA by centrifugation (16,000 g for 10 min at 4°C) and incubation of the resulting supernatant with 20 μL concanavalin A sepharose beads (Sigma-Aldrich) in 200 μL of binding buffer (0.1 M Sodium acetate, 0.1 M NaCl, 1 mM $MgCL_2$, 1 mM $MnCl_2$, 1 mM $CaCl_2$, pH 6.0) for 1 h at 4°C on a rotor. Next, beads were removed from the supernatant by centrifugation. 5 μL of the supernatant was added with 20 μL McIlvaine buffer (pH 6.0), and subjected to enzymatic assays for GBA3 activity. Measured activity was normalized with the corresponding protein concentration of each sample, and data were processed as earlier described.

2.4.9 Sphingolipid extraction and analysis by mass spectrometry in zebrafish larvae

Zebrafish embryos at 8 hours post fertilization were seeded in 12-well plates (15 fish/well, 3 mL egg water/well) and treated with CBE (10–1,000 μM) or CP (0.01–10 μM)) for 112 hours at 28 °C. Thereafter, zebrafish larvae were washed three times with egg water, and collected in clean screw-cap Eppendorf tubes (three tubes of three larvae per inhibitor concentration). Lipids were extracted and measured according to methods described previously.[48] Briefly, after removing of the egg water, 20 μL of ^{13}C-GlcSph[49] from concentration 0.1 pmol μL^{-1} in MeOH, 480 μL MeOH, and 250 μL $CHCl_3$ were added to the sample, stirred, incubated for 30 min at RT, sonicated (5 x 1 min in sonication water bath), and centrifuged for 10 min at 15,700 g. Supernatant was collected in a clean tube, 250 μL $CHCl_3$ and 450 μL 100 mM formate buffer (pH 3.2) were added. The sample was stirred and centrifuged, the upper phase was transferred to a clean tube. The lower phase was extracted with 500 μL MeOH and 500 μL formate buffer. The upper phases were pooled and taken to dryness in a vacuum concentrator at 45 °C. The residue was extracted with 700 μL butanol and 700 μL water, stirred and centrifuged. The upper phase (butanol phase) was dried and the residue was dissolved in 100 μL MeOH. 10 μL of this sample was injected to the LC-MS for lipid measurement. Two-tailed unpaired t-test was

performed in Prism 7.0 software (GraphPad) to derive statistical significance, where $p < 0.05$ was considered significant.

2.5 References

1 Brady RO, Kanfer JN, Bradley RM & Shapiro D (1966) Demonstration of a deficiency of glucocerebroside-cleaving enzyme in Gaucher's disease. *J Clin Invest* **45**, 1112–1115.

2 Beutler E & Grabowski GA (2001) Gaucher disease. In Scriver CR, Sly WS, Childs B, Beaudet AL, Valle D, Kinzler KW, & Vogelstein B (Eds) *The Metabolic and Molecular Bases of Inherited Disease*, *8th Ed* (pp 3635–3668). New York, NY: McGraw-Hill.

3 Lachmann RH, Grant IR, Halsall D & Cox TM (2004) Twin pairs showing discordance of phenotype in adult Gaucher's disease. *QJM - Mon J Assoc Physicians* **97**, 199–204.

4 Biegstraaten M, van Schaik IN, Aerts JMFG, Langeveld M, Mannens MMAM, Bour LJ, Sidransky E, Tayebi N, Fitzgibbon E & Hollak CEM (2011) A monozygotic twin pair with highly discordant Gaucher phenotypes. *Blood Cells Mol Dis* **46**, 39–41.

5 Sidransky E, Nalls MA, Aasly JO, Aharon-Peretz J, Annesi G, Barbosa ER, Bar-Shira A, Berg D, Bras J, Brice A, Chen C-M, Clark LN, Condroyer C, De Marco EV, Dürr A, Eblan MJ, Fahn S, Farrer MJ, Fung H-C, Gan-Or Z, Gasser T, Gershoni-Baruch R, Giladi N, Griffith A, Gurevich T, Januario C, Kropp P, Lang AE, Lee-Chen G-J, Lesage S, Marder K, Mata IF, Mirelman A, Mitsui J, Mizuta I, Nicoletti G, Oliveira C, Ottman R, Orr-Urtreger A, Pereira LV, Quattrone A, Rogaeva E, Rolfs A, Rosenbaum H, Rozenberg R, Samii A, Samaddar T, Schulte C, Sharma M, Singleton A, Spitz M, Tan E-K, Tayebi N, Toda T, Troiano AR, Tsuji S, Wittstock M, Wolfsberg TG, Wu Y-R, Zabetian CP, Zhao Y & Ziegler SG (2009) Multicenter Analysis of Glucocerebrosidase Mutations in Parkinson's Disease. *N Engl J Med* **361**, 1651–1661.

6 Tsuang D, Leverenz JB, Lopez OL, Hamilton RL, Bennett DA, Schneider JA, Buchman AS, Larson EB, Crane PK, Kaye JA, Kramer P, Woltjer R, Kukull W, Nelson PT, Jicha GA, Neltner JH, Galasko D, Masliah E, Trojanowski JQ, Schellenberg GD, Yearout D, Huston H, Fritts-Penniman A, Mata IF, Wan JY, Edwards KL, Montine TJ & Zabetian CP (2012) GBA mutations increase risk for Lewy body disease with and without Alzheimer disease pathology. *Neurology* **79**, 1944–1950.

7 Legler G (1966) Studies on the action mechanism of glycoside splitting anzymes, I. Presentation and properties of specific inhibitors. *Hoppe Seylers Z Physiol Chem* **345**, 197–214.

8 Legler G (1990) Glycoside Hydrolases: Mechanistic Information from Studies with Reversible and Irreversible Inhibitors. *Adv Carbohydr Chem Biochem* **48**, 319–384.

9 Vardi A, Zigdon H, Meshcheriakova A, Klein AD, Yaacobi C, Eilam R, Kenwood BM, Rahim AA, Massaro G, Merrill AH, Vitner EB & Futerman AH (2016) Delineating pathological pathways in a chemically induced mouse model of Gaucher disease. *J Pathol* **239**, 496–509.

10 Premkumar L, Sawkar AR, Boldin-Adamsky S, Toker L, Silman I, Kelly JW, Futerman AH & Sussman JL (2005) X-ray structure of human acid-β-glucosidase covalently bound to conduritol-B-epoxide: Implications for Gaucher disease. *J Biol Chem* **280**, 23815–23819.

11 Kacher Y, Brumshtein B, Boldin-Adamsky S, Toker L, Shainskaya A, Silman I, Sussman JL & Futerman AH (2008) Acid beta-glucosidase: insights from structural analysis and relevance to Gaucher disease therapy. *Biol Chem* **389**, 1361–1369.

12 Kanfer JN, Legler G, Sullivan J, Raghavan SS & Mumford RA (1975) The Gaucher mouse. *Biochem Biophys Res Commun* **67**, 85–90.

13 Manning-Boğ AB, Schüle B & Langston JW (2009) Alpha-synuclein-glucocerebrosidase interactions in pharmacological Gaucher models: A biological link between Gaucher disease and parkinsonism. *Neurotoxicology* **30**, 1127–1132.

14 Xu YH, Sun Y, Ran H, Quinn B, Witte D & Grabowski GA (2011) Accumulation and distribution of β-synuclein and ubiquitin in the CNS of Gaucher disease mouse models. *Mol Genet Metab* **102**, 436–447.

15 Rocha EM, Smith GA, Park E, Cao H, Graham A-R, Brown E, McLean JR, Hayes MA, Beagan J, Izen SC, Perez-Torres E, Hallett PJ & Isacson O (2015) Sustained Systemic Glucocerebrosidase Inhibition Induces Brain α-Synuclein Aggregation, Microglia and Complement C1q Activation in Mice. *Antioxid Redox Signal* **23**, 550–564.

16 Newburg DS, Yatziv S, McCluer RH & Raghavan S (1986) beta-Glucosidase inhibition in murine peritoneal macrophages by conduritol-B-epoxide: an in vitro model of the Gaucher cell. *Biochim Biophys Acta* **877**, 121–126.

17 Bieberich E, Freischütz B, Suzuki M & Yu RK (1999) Differential effects of glycolipid biosynthesis inhibitors on ceramide-induced cell death in neuroblastoma cells. *J Neurochem* **72**, 1040–1049.

18 Prence EM, Chaturvedi P & Newburg DS (1996) In vitro accumulation of glucocerebroside in neuroblastoma cells: A model for study of Gaucher disease pathobiology. *J Neurosci Res* **43**, 365–371.

19 Newburg DS, Shea TB, Yatziv S, Raghavan SS & McCluer RH (1988) Macrophages exposed in vitro to conduritol B epoxide resemble Gaucher cells. *Exp Mol Pathol* **48**, 317–323.

20 Berger J, Lecourt S, Vanneaux V, Rapatel C, Boisgard S, Caillaud C, Boiret-Dupré N, Chomienne C, Marolleau JP, Larghero J & Berger MG (2010) Glucocerebrosidase deficiency dramatically impairs human bone marrow haematopoiesis in an in vitro model of Gaucher disease: Research paper. *Br J Haematol* **150**, 93–101.

21 Schueler UH, Kolter T, Kaneski CR, Zirzow GC, Sandhoff K & Brady RO (2004) Correlation between enzyme activity and substrate storage in a cell culture model system for Gaucher disease. *J Inherit Metab Dis* **27**, 649–658.

22 Lecourt S, Vanneaux V, Cras A, Freida D, Heraoui D, Herbi L, Caillaud C, Chomienne C, Marolleau J-P, Belmatoug N & Larghero J (2012) Bone Marrow Microenvironment in an In Vitro Model of Gaucher Disease: Consequences of Glucocerebrosidase Deficiency. *Stem Cells Dev* **21**, 239–248.

23 Quaroni A, Gershon E & Semenza G (1974) Affinity labeling of the active sites in the sucrase-isomaltase complex from small intestine. *J Biol Chem* **249**, 6424–6433.

24 Shou-jun Y, Su-guo G, Yu-cheng Z & Shu-zheng Z (1985) Inactivation of α-glucosidase by the active-site-directed inhibitor, conduritol B epoxide. *BBA - Protein Struct Mol* **828**, 236–240.

25 Hermans MM, Krooss MA, van Beeurnens J, Oostras BA & Reuser AJ (1991) Human lysosomal alpha-glucosidase. Characterization of the catalytic site. *Biochemistry* **266**, 13507–13512.

26 Braun H, Legler G, Deshusses J & Semenza G (1977) Stereospecific ring opening of conduritol-B-epoxide by an active site aspartate residue of sucrase-isomaltase. *BBA - Enzymol* **483**, 135–140.

27 van Weely S, Brandsma M, Strijland A, Tager JM & Aerts JMFG (1993) Demonstration of the existence of a second, non-lysosomal glucocerebrosidase that is not deficient in Gaucher disease. *BBA - Mol Basis Dis* **1181**, 55–62.

28 Ridley CM, Thur KE, Shanahan J, Thillaiappan NB, Shen A, Uhl K, Walden CM, Rahim AA, Waddington SN, Platt FM & Van Der Spoel AC (2013) β-Glucosidase 2 (GBA2) activity and imino sugar pharmacology. *J Biol Chem* **288**, 26052–26066.

29 Hara A & Radin NS (1979) Enzymic effects of β-glucosidase destruction in mice changes in glucuronidase levels. *BBA - Gen Subj* **582**, 423–433.

30 Atsumi S, Iinuma H, Nosaka C & Umezawa K (1990) Biological activities of cyclophellitol. *J Antibiot (Tokyo).* **43**, 1579–1585.

31 Withers SG & Umezawa K (1991) Cyclophellitol: A naturally occurring mechanism-based inactivator of beta-glucosidases. *Biochem Biophys Res Commun* **177**, 532–537.

32 Witte MD, Kallemeijn WW, Aten J, Li KY, Strijland A, Donker-Koopman WE, Van Den Nieuwendijk AMCH, Bleijlevens B, Kramer G, Florea BI, Hooibrink B, Hollak CEM, Ottenhoff R, Boot RG, Van Der Marel GA, Overkleeft HS & Aerts JMFG (2010) Ultrasensitive in situ visualization of active glucocerebrosidase molecules. *Nat Chem Biol* **6**, 907–913.

33 Liu Y, Patricelli MP & Cravatt BF (1999) Activity-based protein profiling: The serine hydrolases. *Proc Natl Acad Sci* **96**, 14694–14699.

34 Niphakis MJ & Cravatt BF (2014) Enzyme Inhibitor Discovery by Activity-Based Protein Profiling. *Annu Rev Biochem* **83**, 341–377.

35 Van Esbroeck ACM, Janssen APA, Cognetta AB, Ogasawara D, Shpak G, Van Der Kroeg M, Kantae V, Baggelaar MP, De Vrij FMS, Deng H, Allarà M, Fezza F, Lin Z, Van Der Wel T, Soethoudt M, Mock ED, Den Dulk H, Baak IL, Florea BI, Hendriks G, De Petrocellis L, Overkleeft HS, Hankemeier T, De Zeeuw CI, Di Marzo V, Maccarrone M, Cravatt BF, Kushner SA & Van Der Stelt M (2017) Activity-based protein profiling reveals off-target proteins of the FAAH inhibitor BIA 10-2474. *Science* **356**, 1084–1087.

36 Kallemeijn WW, Li KY, Witte MD, Marques ARA, Aten J, Scheij S, Jiang J, Willems LI, Voorn-Brouwer TM, Van Roomen CPAA, Ottenhoff R, Boot RG, Van Den Elst H, Walvoort MTC, Florea BI, Codée JDC, Van Der Marel GA, Aerts JMFG & Overkleeft HS (2012) Novel activity-based probes for broad-spectrum profiling of retaining β-exoglucosidases in situ and in vivo. *Angew Chemie - Int Ed* **51**, 12529–12533.

37 Wu L, Jiang J, Jin Y, Kallemeijn WW, Kuo C-L, Artola M, Dai W, van Elk C, van Eijk M, van der Marel GA, Codee JDC, Florea BI, Aerts JMFG, Overkleeft HS & Davies GJ (2017) Activity-based probes for functional interrogation of retaining β-glucuronidases. *Nat Chem Biol* **13**, 867–873.

38 Jiang J, Kuo CL, Wu L, Franke C, Kallemeijn WW, Florea BI, Van Meel E, Van Der Marel GA, Codée JDC, Boot RG, Davies GJ, Overkleeft HS & Aerts JMFG (2016) Detection of active mammalian GH31 α-glucosidases in health and disease using in-class, broad-spectrum activity-based probes. *ACS Cent Sci* **2**, 351–358.

39 Chao DHM, Kallemeijn WW, Marques ARA, Orre M, Ottenhoff R, Van Roomen C, Foppen E, Renner MC, Moeton M, Van Eijk M, Boot RG, Kamphuis W, Hol EM, Aten J, Overkleeft HS, Kalsbeek A & Aerts JMFG (2015) Visualization of active glucocerebrosidase in rodent brain with high spatial resolution following in situ labeling with fluorescent activity based probes. *PLoS One* **10**, 1–20.

40 Dekker N, Van Dussen L, Hollak CEM, Overkleeft H, Scheij S, Ghauharali K, Van Breemen MJ, Ferraz MJ, Groener JEM, Maas M, Wijburg FA, Speijer D, Tylki-Szymanska A, Mistry PK, Boot RG & Aerts JM (2011) Elevated plasma glucosylsphingosine in Gaucher disease: Relation to phenotype, storage cell markers, and therapeutic response. *Blood* **118**, e118–e127.

41 Aerts JMFG, Donker-Koopman WE, van Laar C, Brul S, Murray GJ, Wenger DA, Barranger JA, Tager JM & Schram AW (1987) Relationship between the two immunologically distinguishable forms of glucocerebrosidase in tissue extracts. *Eur J Biochem* **163**, 583–589.

42 Dekker N, Voorn-Brouwer T, Verhoek M, Wennekes T, Narayan RS, Speijer D, Hollak CEM, Overkleeft HS, Boot RG & Aerts JMFG (2011) The cytosolic β-glucosidase GBA3 does not influence type 1 Gaucher disease manifestation. *Blood Cells Mol Dis* **46**, 19–26.

43 Ferraz MJ, Marques ARA, Appelman MD, Verhoek M, Strijland A, Mirzaian M, Scheij S, Ouairy CM, Lahav D, Wisse P, Overkleeft HS, Boot RG & Aerts JM (2016) Lysosomal glycosphingolipid catabolism by acid ceramidase: Formation of glycosphingoid bases during deficiency of glycosidases. *FEBS Lett* **590**, 716–725.

44 Fischer S, Klüver N, Burkhardt-Medicke K, Pietsch M, Schmidt AM, Wellner P, Schirmer K & Luckenbach T (2013) Abcb4 acts as multixenobiotic transporter and active barrier against chemical uptake in zebrafish (Danio rerio) embryos. *BMC Biol* **11**, 1–16.

45 Schröder SP, van de Sande JW, Kallemeijn WW, Kuo C-L, Artola M, van Rooden EJ, Jiang J, Beenakker TJ, Florea BI, Offen WA, Davies G, Minnaard AJ, Aerts H, Codee JD, van der Marel G & Overkleeft HS (2017) Towards broad spectrum activity-based glycosidase probes: synthesis and evaluation of deoxygenated cyclophellitol aziridines. *Chem Commun* **53**, 12528–12531.

46 Lahav D, Liu B, Van Den Berg RJBHN, Van Den Nieuwendijk AMCH, Wennekes T, Ghisaidoobe AT, Breen I, Ferraz MJ, Kuo CL, Wu L, Geurink PP, Ovaa H, Van Der Marel GA, Van Der Stelt M, Boot RG, Davies GJ, Aerts JMFG & Overkleeft HS (2017) A Fluorescence Polarization Activity-Based Protein Profiling Assay in the Discovery of Potent, Selective Inhibitors for Human Nonlysosomal Glucosylceramidase. *J Am Chem Soc* **139**, 14192–14197.

47 Overkleeft HS, Renkema GH, Neele J, Vianello P, Hung IO, Strijland A, Van Der Burg AM, Koomen GJ, Pandit UK & Aerts JMFG (1998) Generation of specific deoxynojirimycin-type inhibitors of the non- lysosomal glucosylceramidase. *J Biol Chem* **273**, 26522–26527.

48 Mirzaian M, Wisse P, Ferraz MJ, Marques ARA, Gaspar P, Oussoren SV, Kytidou K, Codée JDC, van der Marel G, Overkleeft HS & Aerts JM (2017) Simultaneous quantitation of sphingoid bases by UPLC-ESI-MS/MS with identical 13 C-encoded internal standards. *Clin Chim Acta* **466**, 178–184.

49 Wisse P, Gold H, Mirzaian M, Ferraz MJ, Lutteke G, Van Den Berg RJBHN, Van Den Elst H, Lugtenburg J, Van Der Marel GA, Aerts JMFG, Codée JDC & Overkleeft HS (2015) Synthesis of a panel of carbon-13-labeled (glyco)sphingolipids. *European J Org Chem* **2015**, 2661–2677.

APPENDIX

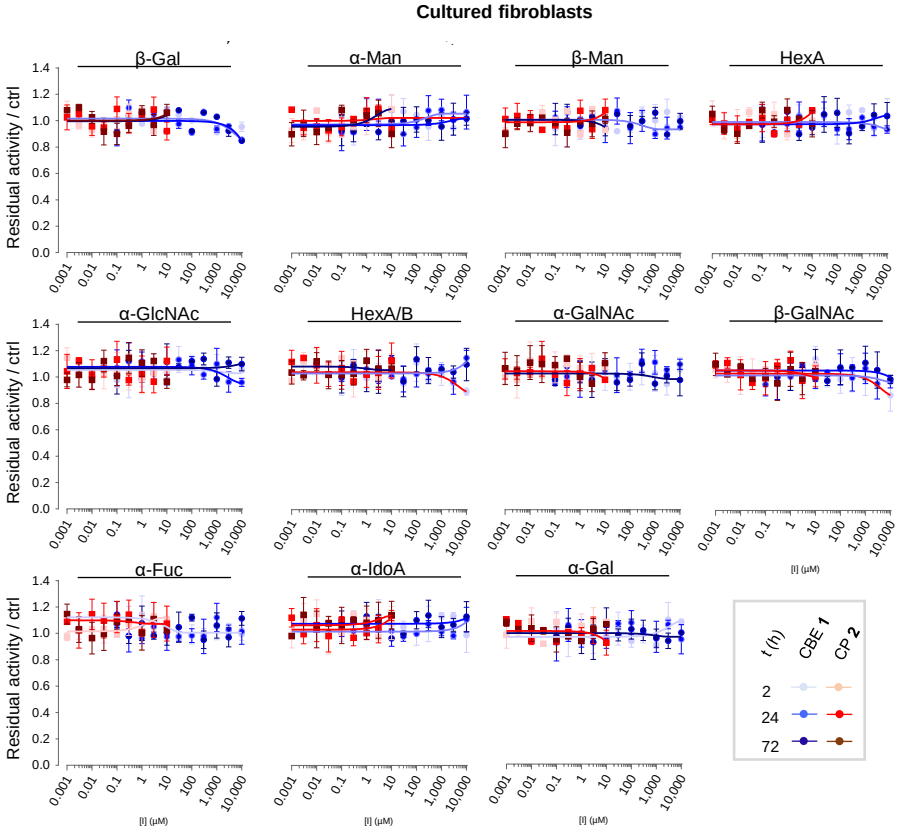


Figure 2.S1. Effect of *in vivo* incubation with CBE and CP on enzymatic activity of glycosidases in cultured cells. Confluent cultured human fibroblasts were treated *in vivo* with different concentrations of CBE (blue) and CP (red) for 2 h, 24 h, or 72 h, before harvested and lysed in potassium phosphate buffer. Residual activities of glycosidases were measured by *in vitro* enzymatic assay using appropriate glycoside substrates, and normalized to the activity from the control samples. Error range = ± SD, *n* = 3 technical replicates.

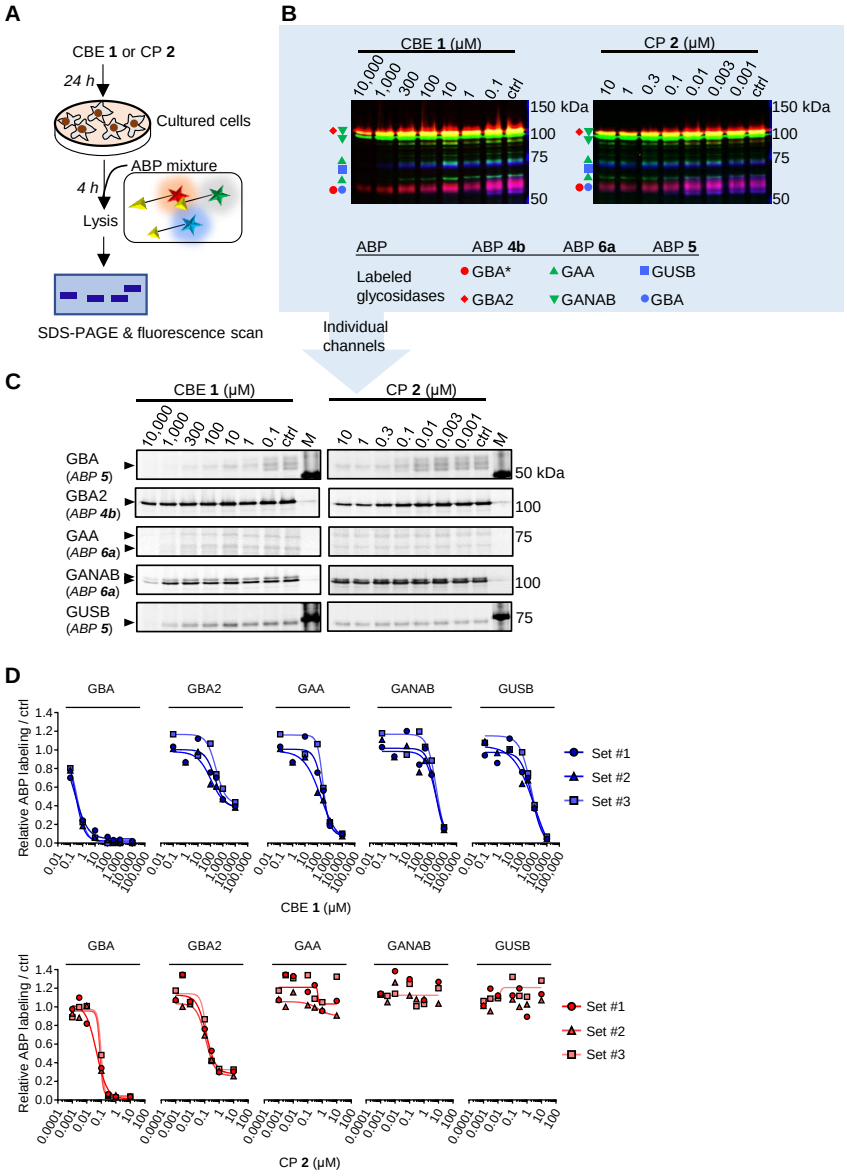


Figure 2.S2. Effect of *in vivo* incubation with CBE and CP on *in vivo* ABP labeling of glycosidases in cultured cells. A) Schematic representation of the experimental procedure. B) Overlay of fluorescence channels showing *in vivo* ABP labeling of glycosidases that have not been inactivated by CBE or CP; one representative image is selected from $n = 3$ biological replicates. C) Individual fluorescence channels of B); one representative image is selected from $n = 3$ biological replicates. *GBA labeling by ABP **4b** was not used for analysis, due to the presence of an unidentified off-target by ABP 4b in living cells. ABP **5** labels both GUSB and GBA.[37] D) Quantification of ABP labeling. Error ranges = ± SD, $n = 3$ (biological replicates).

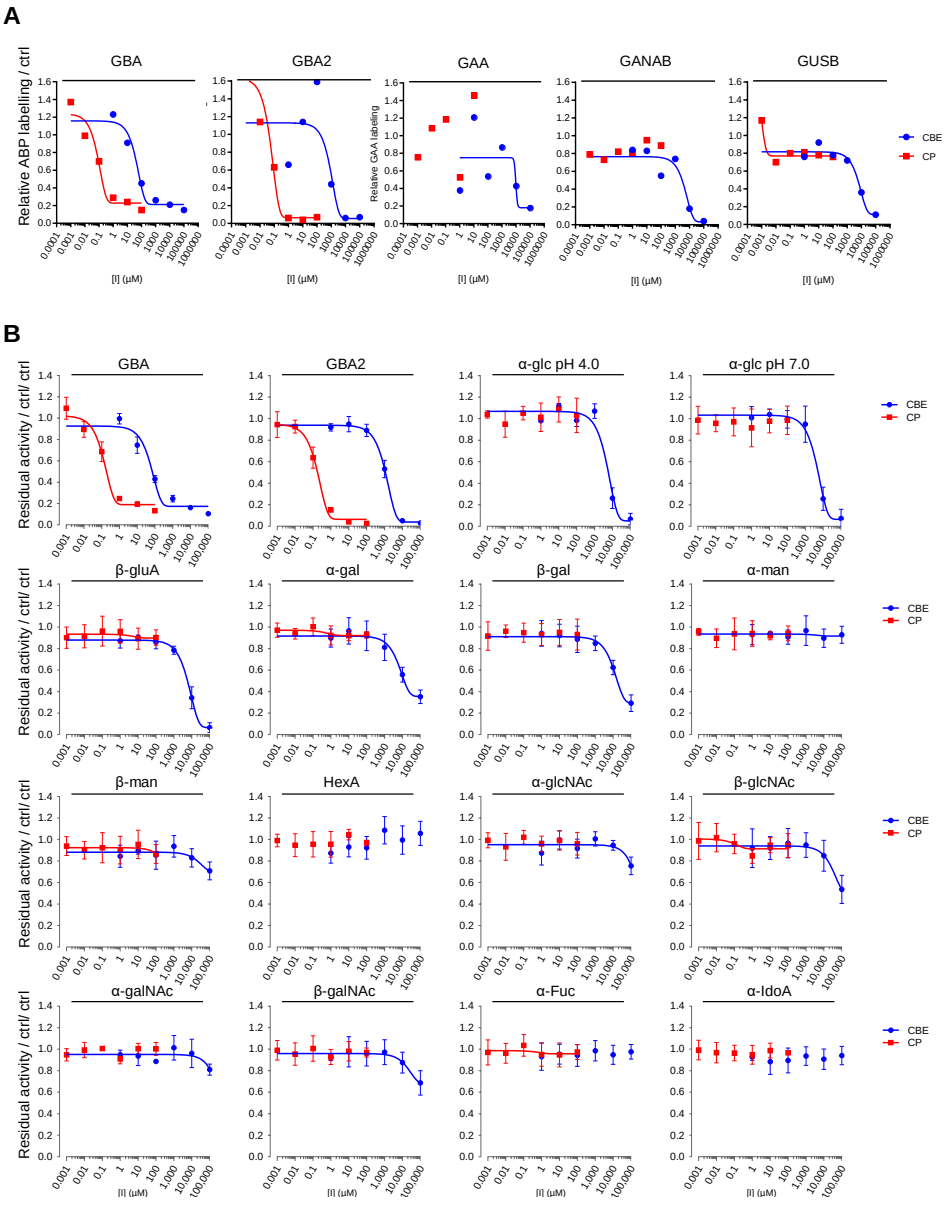


Figure 2.S3. Effect of *in vivo* incubation with CBE and CP on glycosidases in zebrafish larvae. Zebrafish embryos (n = 48) were incubated with CBE or CP from 8–120 hpf before lysed in potassium phosphate buffer. A) Quantification of gels from **Fig. 2.4A** and **2.5D**. B) Residual activity of glycosidases by enzymatic assay. Error ranges = ± SD, *n* = 2 technical replicates.

Table 2.S1 Effect of *in vivo* incubation with CBE and CP on *in vivo* ABP labeling of glycosidases in cultured cells. Apparent IC_{50} values (μM) were derived from the average of biological triplicates as measured by quantification of *in vivo* ABP labeling on residual active enzymes (*ABP*). -, value cannot be calculated (no inhibition at the tested concentrations). Error ranges = ± SD, *n* = 3 biological replicates.

Enzyme	CBE	CP
GBA	0.282 (± 0.05)	0.075 (± 0.026)
GBA2	210 (± 87.8)	0.128 (± 0.006)
GAA	249 (± 53.8)	-
GANAB	2280 (± 163)	-
GUSB	748 (± 241)	-

in vivo target engagement of CBE and CP

CHAPTER 3

Functionalized cyclophellitols are improved glucocerebrosidase inhibitors for generating neuropathic Gaucher model in zebrafish

Manuscript published as:

Artola M, **Kuo CL**, Lelieveld LT, Rowland RJ, van der Marel GA, Codée JDC, Boot RG, Davies DJ, Aerts JMFG, Overkleeft HS (2019) Functionalized Cyclophellitols Are Selective Glucocerebrosidase Inhibitors and Induce a Bona Fide Neuropathic Gaucher Model in Zebrafish. *J Am Chem Soc* **141**, 4214–4218.

ABSTRACT

Gaucher disease is caused by inherited deficiency in glucocerebrosidase (GBA, a retaining β-glucosidase), and deficiency in GBA constitutes the largest known genetic risk factor for Parkinson disease. In the past, animal models of Gaucher disease have been generated by treatment with the mechanism-based GBA inhibitors, conduritol B epoxide (CBE) and cyclophellitol. Both compounds however also target other retaining glycosidases, rendering generation and interpretation of such chemical knockout models complicated. Here it is demonstrated that cyclophellitol derivatives carrying a bulky hydrophobic substituent at C8 (cyclophellitol numbering) are potent and selective GBA inhibitors and that an unambiguous Gaucher animal model can be readily generated by treatment of zebrafish with these.

3.1 Introduction

Glucocerebrosidase (acid glucosylceramidase, GBA, E.C. 3.2.1.45,) is a lysosomal retaining β-glucosidase that belongs to the glycoside hydrolase (GH) 30 (www.cazy.org)[1] family and degrades the glycosphingolipid, glucosylceramide through a two-step Koshland double displacement mechanism (**Fig. 3.1A**). Inherited deficiency in GBA causes the most common autosomal recessive lysosomal storage disorder, Gaucher disease.[2] Individuals carrying heterozygous mutations in the gene coding for GBA do not develop Gaucher disease but have a remarkable increased risk for developing Parkinson disease (PD) and Lewy-body dementia.[3–5] Appropriate animal models linking impaired GBA functioning to Gaucher disease and Parkinson's disease are imperative both for understanding the pathophysiology of these diseases and for the development of effective treatments for these. Because complete genetic abrogation of GBA hampers animal viability due to skin permeability problems,[6] research models have been generated in the past in a chemical knockdown strategy by making use of the mechanism-based, covalent and irreversible retaining β-glucosidase inhibitor, conduritol B epoxide (CBE, **1**, **Fig. 3.1B**) or its close structural analogue, cyclophellitol (**2**, **Fig. 3.1B**).[7, 8] One complication in the use of these compounds is their relative lack of selectivity.[9] It has been found that cyclophellitol **2** is unsuited for creating a reliable Gaucher animal model because it targets GBA and GBA2 with about equal efficiency.[9] On the other hand, CBE **1** exhibits some GBA selectivity but it also inhibits lysosomal α-glucosidase (GAA),[10–13] non-lysosomal glucosylceramidase (GBA2),[14, 15] and lysosomal β-glucuronidase (GUSB)[16]. Effective mouse models can be generated with CBE **1**, but the therapeutic window is rather narrow and varies in cellular and animal models.

Recent research from the Aerts and Overkleeft group has revealed that functionalized cyclophellitol derivatives carrying a BODIPY substituent at C8 (cyclophellitol numbering, the primary carbon corresponding to C6 in glucose) are very potent and very selective activity-based probes (ABPs) for monitoring GBA activity *in vitro, in situ,* and *in vivo*.[17,18] The presence of a bulky and hydrophobic substituent at this position at once proved beneficial for GBA inactivation (ABPs **3** and **4**, (**Fig. 3.1C, D**)) proved to inhibit GBA in the nanomolar range, whereas cyclophellitol **2** is a high nanomolar to micromolar GBA inactivator and detrimental to inhibition of other retaining β-glucosidases. Following these studies, Vocadlo and co-workers designed a set of fluorogenic substrates featuring a fluorophore at C6 of a β-glucoside, the

aglycon of which carried a fluorescence quencher, compounds that proved to be very selective GBA substrates *in situ*.[19] These results altogether evoked the question whether cyclophellitols bearing a simple, hydrophobic moiety at C8, such as compounds **6** and **7** (**Fig. 3.1D**), would be suitable compounds for generating chemical knockdown Gaucher animal models. The validity of this reasoning is shown in the generation of a GBA-deficient *Dario rerio* zebrafish model, as revealed by the accumulation of elevated levels of the Gaucher harbinger lysolipid, glucosylsphingosine, using cyclophellitol derivatives **6** and **7**.

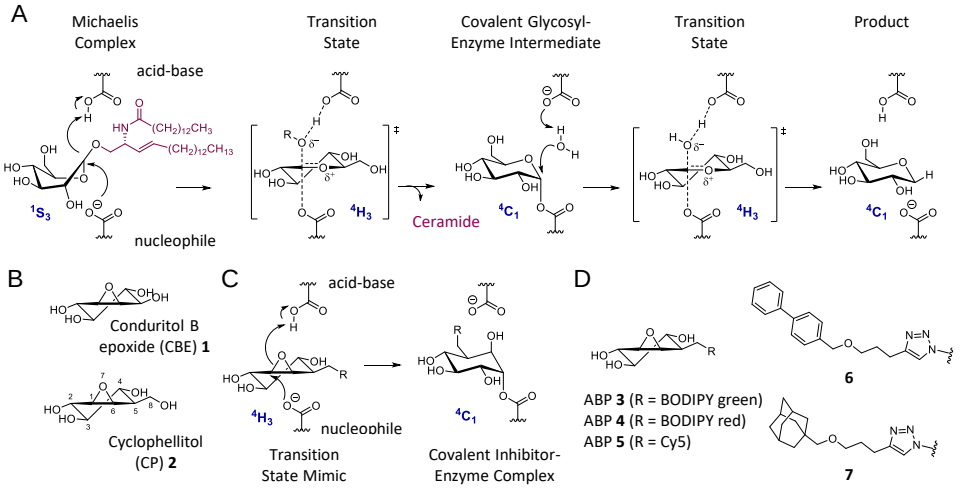


Figure 3.1. Structures and reaction mechanism of compounds in this chapter. A) Glucocerebrosidase (GBA) hydrolyses glucosylceramide in a two-step double displacement mechanism to yield glucose and ceramide. B) Chemical structure of CBE **1** and cyclophellitol **2**. C) Mechanism-based inactivation of GBA by glucopyranoside-configured cyclitol epoxides (shown for cyclophellitol). D) Structures of C6-extended cyclophellitol derivatives used in the here-presented studies: GBA1 activity-based probes ABPs **3**-**5** and selective inhibitors **6** and **7** (see for the full chemical structures of ABPs **3**-**5** and **8**-**14** in **Fig. 3.S1**).

3.2 Results and Discussion

3.2.1 Structural support for inhibitor design

At the onset of the studies, structural support was sought for the design of compounds **6** and **7**. The Cy5-functionalized cyclophellitol **5** was synthesized (Department of Bio-organic Chemistry, Leiden University) and a crystal structure of human recombinant GBA soaked with this ABP was obtained (University of York). As expected (**Fig. 3.2A**), the active site nucleophile (in both molecules of the asymmetric unit) had reacted with the epoxide to yield the covalently bound cyclitol in $^4C_1$-conformation, with the Cy5 moiety, via its flexible linker, clearly bound in

one molecule of the asymmetric unit (the differences may reflect crystal packing constraints in a soaking experiment) accommodated by a hydrophobic pocket in GBA. Previous studies on the bacterial glycoside hydrolase, *Thermoanaerobacterium xylanolyticum* TxGH116 β-glucosidase, a close homologue of human GBA2 with a conserved active site, instead showed an "inwards" position of O6 (**Fig. 3.S2A**) and a narrower and less hydrophobic pocket) (**Fig. 3.S2B**), which may partially mitigate against the binding of O6-functionaised probes, thus allowing sufficient discrimination for GBA over GBA2.[20, 21]

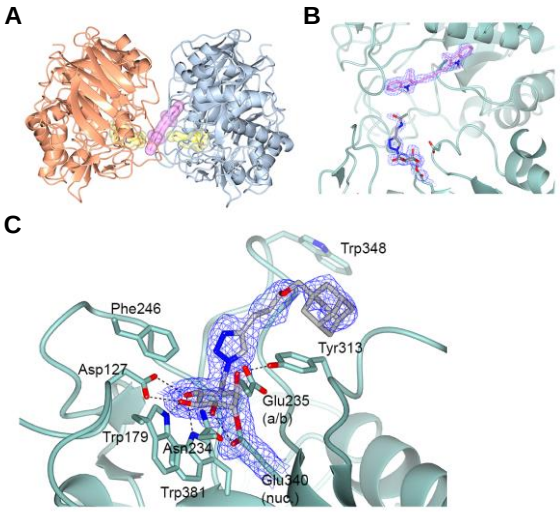


Figure 3.2. Structure of GBA reacted with ABP 5 and adamantyl-cyclophellitol 7. A) GBA dimer, with the cyclophellitol and linker moiety of **5** shaded in yellow, and a single observed Cy5 in pink. B) Zoomed view of a GBA monomer reacted with ABP **5**. C) Structure of GBA with adamantyl-cyclophellitol **7**. The linker-adamantyl moiety of **7** is observed in slightly different positions in the two molecules of the asymmetric unit (PDB Code 6Q6L, **Fig. 3.S2**) reflecting its binding through predominantly hydrophobic interactions.

Biphenyl-cyclophellitol **6** and adamantyl-cyclophellitol **7** were synthesized following adaptations of literature cyclophellitol syntheses (see **Appendix section 3.S2** for synthesis details) to generate superior selective GBA inhibitors for the generation of a Gaucher model zebrafish.[22, 23]

Although soaking of GBA crystals with **6** did not yield suitable structures for structural analysis (**Fig. 3.S2)**, soaking with **7** did (**Fig. 3.2C**), and again revealed binding of the

hydrophobic moiety (here, the adamantane) to the same hydrophobic cavity and pocket occupied by the O6 linker on Cy5 ABP **5**. Several hydrophobic residues, including Tyr313, Phe246, and Trp348 provide the wide cavity that is able to accommodate different hydrophobic O6 substituents which is absent in other human β-glucosidases and which provides the structural basis for the inhibitory (and substrate) preferences of GBA.

3.2.2 *In vitro* and *in vivo* activity of inhibitors

The *in vitro* activity and selectivity of compounds **6** and **7** towards GBA and the two major off-target glycosidases of CBE **1** (GBA2 and GAA)[9] were evaluated, by pre- incubating the inhibitors with recombinant human GBA (rGBA, Cerezyme), human GBA2 (from lysates of GBA2 overexpressed cells), and recombinant human GAA (rGAA, Myozyme) for 3 h, followed by enzymatic activity measurement. Both compound **6** and **7** were nanomolar inhibitors of rGBA (apparent IC_{50} values = 1.0 nM), which were 4,000-fold more potent than CBE **1** (apparent IC_{50} values = 4.28 μM) (**Fig. 3.3A**, **3.S3**) with improved lipophilic ligand efficiencies (LipE) (**Table 3.S2**). Both compounds **6** and **7** were rather inactive towards GBA2 and GAA (apparent IC_{50} values > 100 μM), similar to ABP **3** and **5** (**Fig. 3.3A**, **3.S4**). When comparing their selectivity towards GBA, both compounds **6** and **7** exhibited IC_{50} ratio **(**GBA2/GBA and GAA/GBA**)** of > 100,000, thus making them 4,000-times and 200-times more selective than CBE **1** (IC_{50} ratio **=** 23.6 for GBA2/GBA and 444 for GAA/GBA) (**Fig. 3.3A)**.

To evaluate the *in vivo* activity of compound **6** and **7**, compounds were added to the egg-water containing zebrafish (*Danio rerio*) embryos, and incubated for 5 days at 28 °C before subsequent homogenization and enzyme selectivity analysis by appropriate ABP labeling.[9, 24] Quantification of ABP-labeled bands revealed that compounds **6** and **7** had *in vivo* apparent IC_{50} values towards GBA of 4 to 6 nM, and that they were 5- to 70-fold more potent than ABP **3** or **5** and 7,500 fold more potent than CBE **1** (**Fig. 3.3B**, **3.S5**) in the zebrafish larvae. More importantly, an improved selective inactivation of GBA was achieved with both compounds **6** and **7**. At a concentration of 0.1-10 μM of compound **6** or **7**, ABP labeling of GBA with broad-spectrum retaining β-glucosidase ABP **8** (**Fig. 3.S1**) and GBA-specific ABP **5** was abrogated (**Fig. 3.3C**), while other enzymes such as GBA2 and LPH (**Figure 3.3C**), GAA, ER α-glucosidase GANAB, and lysosomal β-glucuronidase GUSB (**Fig. 3.S6A**, **3.S6B**) were not affected.

A

	CBE **1**	CP **2**	ABP **3**	ABP **5**	**6**	**7**
***In vitro* (nM)**						
rGBA	4.28 x 10^3 ± 500 [9]	29.6 ± 2.40 [9]	1.20 ± 0.30	3.20 ± 0.17	1.06 ± 0.19	0.96 ± 0.17
GBA2 (HEK2 93T lysate)	101 x 10^3 ± 20.1 x 10^3	29.7 ± 3.13	> 10^5	412 x 10^3 ± 10.1 x 10^3	> 10^5	> 10^5
rGAA	1.90 x 10^6 ± 192 x 10^3	> 10^5	> 10^5	> 10^5	> 10^5	> 10^5
Ratio (*In vitro*)						
GBA2/ GBA	24	1	> 10^5	> 10^4	> 10^5	> 10^5
GAA/ GBA	444	> 10^3	> 10^6	> 10^6	> 10^6	> 10^6

B

	CBE **1**	CP **2**	ABP **3**	ABP **5**	**6**	**7**
***In vivo* (*Danio Rerio* larvae) (nM)**						
GBA	4.41 x 10^4 [9]	83 [9]	31.6 ± 8.88	284 ± 31.5	5.85 ± 2.44	3.94 ± 1.21
GBA 2	8.90 x 10^5 [9]	59 [9]	> 10^4	> 10^4	> 10^4	> 10^4
GAA	9.55 x 10^6 [9]	> 10^5 [9]	> 10^4	> 10^4	> 10^4	> 10^4
Ratio (*in vivo*)						
GBA 2/ GBA	22.0	0.714	> 316	> 35.2	> 1,710	> 2,540
GAA / GBA	223	> 120	> 316	> 35.2	> 1,710	> 2,540

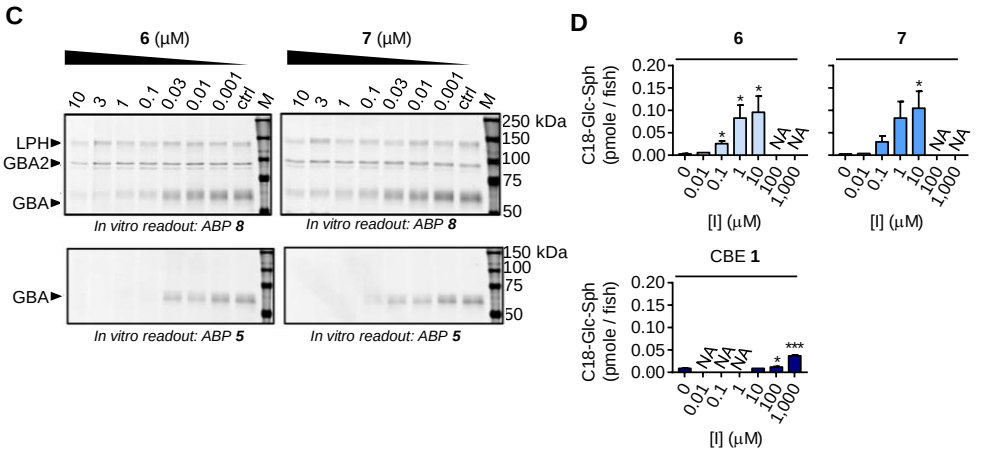


Figure 3.3. ***In vitro* and *in vivo* activity of compounds used in this study.** A) Apparent IC_{50} values for *in vitro* inhibition of GBA, GBA2 and GAA in recombinant enzymes (rGBA and rGAA) or overexpressed cell lysates (GBA2) by compounds **1**, **2**, **3**, **5**, **6** and **7**. Error ranges depict standard deviations from biological duplicates. B) Apparent IC_{50} values for *in vivo* inhibition in 5-day treated zebrafish embryo with compounds **1**, **2**, **3**, **5**, **6**, and **7**. Error ranges depict standard deviations from n = 12−24 individuals. C) Competitive ABPP in lysates of zebrafish treated in vivo with compounds **6** and **7** using broad-spectrum retaining β-glucosidase ABP **8** and selective GBA ABP **5** as readout. D) Glucosylsphingosine levels produced in zebrafish embryos treated for 5 days with inhibitors **6**, **7** or CBE **1**.[9] Error ranges depict standard deviations from n = 3 individuals. *N/A*, not analyzed; *, $p < 0.5$; ***, $p < 0.001$.

At 0.1-10 μM of inhibitor **6** or **7**, a 10- to 30-fold elevation in the level of glucosylsphingosine (GlcSph) was observed, which is known to be formed by acid ceramidase-mediated conversion of accumulating GlcCer in lysosomes.[25, 26] Therefore, this observation also strongly points to *in vivo* inactivation of lysosomal GBA. For comparison, reaching similar GlcSph levels in the zebrafish with CBE required 1,000 to 10,000-fold higher concentration in contrast with compounds **6** or **7** (**Fig. 3.3D**), concentrations at which GBA2 and GAA may also be targeted.

3.2.3 Brain permeability of inhibitors

Finally, the brain permeability of these new inhibitors was investigated, a crucial feature for their future application in the study of neuropathic Gaucher disease and Parkinson disease. Adult zebrafish of 3 months' age were treated with DMSO, ABP **3** or compound **7** (1,6 nmol/fish, approximately 4 μmol/kg) administered via food intake, and after 16 h brains and other organs were isolated, homogenized, and analyzed by ABP labeling using ABP **5** (GBA), ABP **8** (GBA2 + GBA), ABP **11** (GAA at pH 4.0 and ER α-glucosidase GANAB at pH 7.0), and ABP **13** (lysosomal β-glucuronidase GUSB) (**Fig. 3.S1**). Labeling of brain homogenate of adult zebrafish with ABP **5** resulted in considerable GBA labeling in control and ABP **3**-treated fish, but no labeling in brain homogenates from fish treated with compound **7** (**Fig. 3.4**). Labeling by the broad-spectrum β-glucosidase ABP **8** showed that GBA2 was not a target of compound **7**, nor was the lower running band (48 kDa), which is hypothesized to be the cytosolic β-glucosidase, GBA3 (E.C. 3.2.1.21).[27] It could be labeled (in lysates of zebrafish larvae) with the broad spectrum β-glucosidase ABP **8** at 0.1 μM[28] optimally at pH of 6.0 (**Fig. 3.S8**) and could be competed away by pre-incubating the lysates with another broad spectrum β-glucosidase ABP—ABP **10**, while not showing reactivity with the GBA-specific epoxide ABP **5** (**Fig. 3.S8**). It is noted that the expression level of this protein is likely variable among individual fish, as four out of six fish in the control group lacked this band (**Fig. 3.4**). In the visceral organs (both liver and spleen), both ABP **3** and compound **7** selectively abrogated GBA, while not affecting the labeling on other tested glycosidases (**Fig. 3.S7**).

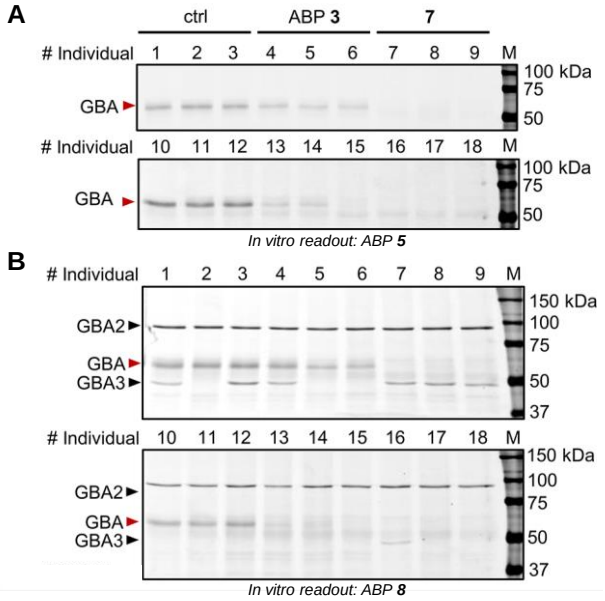


Figure 3.4. *In vivo* targets of ABP 3 and 7 in brains of adult zebrafish. Competitive ABPP in adult zebrafish homogenates with selective GBA ABP **5** or broad-spectrum retaining β-glucosidase ABP **8** as read-out.

3.3 Conclusion

To summarize, crystallographic studies aided the rational design of novel cyclophellitol analogues **6** and **7**, that turned out to be very potent and selective GBA inhibitors, also in zebrafish embryos and adult zebrafish (GBA2/GBA inhibition ratio > 1,000). Compound **7**, which also completely block GBA activity in the brain, should be superior to CBE **1** and CP **2** for generating GBA deficiency on demand in zebrafish, thus to create zebrafish models for neuropathic Gaucher disease, to assist research in the context of neuropathic GD and PD.

3.4 Experimental procedures

3.4.1 General materials and methods

Cyclophellitol (CP) **1** and the ABPs **3**, **4** and **8**–**14** were synthesized as described earlier[17, 27–30]. Conduritol B-epoxide (CBE) was purchased from Enzo Life Sciences (Farmingdale, NY, USA). Recombinant human GBA (rGBA, imiglucerase, Cerezyme®) and recombinant human GAA (rGAA, alglucosidase alfa, Myozyme) were obtained from Sanofi Genzyme (Cambridge, MA, USA). HEK293T (CRL-3216) cell lines were purchased from ATCC (Manassas, VA, USA). Cell lines were cultured in DMEM medium (Sigma-Aldrich, St. Lois, MO, USA), supplied with 10% (v/v) FCS, 0.1 % (w/v) penicillin/streptomycin, and 1 % (v/v) Glutamax, under 7 % CO_2 at 37°C. Zebrafish (*Danio rerio*) were handled and maintained as previously described.[9] Protein concentration was measured using Pierce BCA assay kit (Thermo Fisher Scientific, Waltham, MA, USA).

3.4.2 Preparation of cell lysates containing human GBA2

HEK293T cells overexpressing human GBA2 were generated as previously described[20]. For preparing lysates, cell pellets from two 15-cm culture dishes were resuspended in 1200 µL of lysis buffer (25 mM KH_2PO_4–K_2HPO_4, pH 6.5, protease inhibitor cocktail (EDTA-free, Roche, Basel, Switzerland), 2.5 U/mL benzonase) and incubated for 30 min on ice. The suspension was passed through a 30-gauge needle 10 times using a 1 mL syringe, and stored at –80°C.

3.4.3 *In vitro* activity of inhibitors on glycosidases measured by 4-MU substrates

In vitro apparent IC_{50} measurements with compounds (ABP **3**, ABP **5**, **6**, and **7**) in GBA (rGBA) followed the fluorogenic substrate methods described[9] previously at 3 h incubation time and 2 nM enzyme. For *in vitro* apparent IC_{50} measurements with compounds (CBE **1**, CP **2**, ABP **3**, ABP **5**, **6**, and **7**) in GBA2, 8 volumes of cell lysates (7 µg total protein/µL) containing overexpressed human GBA2 were firstly pre-incubated with 1 volume of ABP **4** (100 nM final concentration, 0.5% (v/v DMSO)) for 30 min at 37°C to selectively inhibit GBA activity. After which, lysates were incubated with 1 volume of compounds at various concentrations for 3 h at 37°C, before subsequent enzymatic assay for GBA2 activity described earlier.[9] For measurement in rGAA, 2.1 ng rGAA was prepared in 12.5 µL assay buffer (150 mM McIlvaine buffer, pH 4.0, 0.1% (w/v) bovine serum albumin (Sigma)), and incubated with compounds (CBE **1**, CP **2**, ABP

3, ABP **5**, **6**, and **7**) for 3 h at 37°C and subjected to GAA activity readout as described earlier ([enzyme] =19 nM during incubation).[9] All assays were performed in duplicate sets, each set with 3 technical replicates at each inhibitor concentration. DMSO concentration was kept at 1 % (v/v) in all assays during incubation with compounds. *In vitro* apparent IC_{50} values were calculated by fitting data with [inhibitor] vs response—various slope (four parameters) function using GraphPad Prism 7.0 software, and the average values and standard deviations were calculated from the two sets for each compound.

3.4.4 Zebrafish housing and breeding

Zebrafish were housed and maintained at the University of Leiden, the Netherlands, at a density of 40-50 adults per tank, on a cycle of 14 hours of light, 10 hours of darkness and at constant temperature of 28°C. The breeding of fish lines was approved by the local animal welfare committee (Instantie voor dierwelzijn, IvD) of the University of Leiden (license number: 10612). Experiments with embryos and larvae were performed before the free-feeding stage, not falling under animal experimentation law according the EU animal Protection Directive 2010/63/EU. Embryos and larvae were grown and incubated in egg water (60 µg/mL Instant Ocean Sera marin TM aquarium salts (Sera; Heinsberg, Germany) at constant temperature of 28 °C.

3.4.5 *In vivo* activity of inhibitors and fluorescent ABP labeling and detection in zebrafish larvae

For *in vivo* treatment of compounds (ABP **3**, ABP **5**, **6**, and **7**), a single fertilized embryo was seeded in each well of a 96-wells plate, and exposed to ABP **3** (0.001−10 µM), ABP **5** (0.0001−10 µM), **6** (0.001−10 µM) and **7** (0.001−10 µM) diluted in 200 µL egg water (0.5% (v/v) DMSO) for 120 hours at 28.5 °C. Per condition, n = 24 embryos were used. At 120 hours (5 dpf), larvae were collected, rinsed three times with egg water, fully aspirated, snap-frozen in liquid nitrogen and stored at –80 °C until homogenization in lysis buffer (without benzonase) at 4 µL/fish. Lysis was conducted by sonication with a Polytron PT 1300D sonicator (Kinematica, Luzern, Switzerland) on ice at 20% power for three seconds, and repeated three times. Samples containing 20–45 µg total protein were diluted in 14 µL lysis buffer, added with McIlvaine buffer at various pHs, and subjected to ABP detection at a final volume of 32 µL for 30 min at 37 °C using the following conditions: GBA with McIlvaine buffer pH 5.2 (with 0.1 % (v/v) Triton X-100 and 0.2 % (w/v) sodium taurocholate (Sigma)), and 200 nM ABP **4** or ABP **5**; β-glucosidases (GBA, GBA2, GBA3, LPH) with McIlvaine buffer pH 5.5 and 100 nM ABP **8** or ABP **9**; GAA

with McIlvaine buffer pH 4.0 and pre-incubation with 200 nM ABP **10**, before incubation with 500 nM ABP **11** or 3 μM ABP **12**; GANAB with McIlvaine buffer pH 7.0 and pre-incubation with 200 nM ABP **10**, before incubation with 500 nM ABP **11** or ABP **12**; GUSB with McIlvaine buffer pH 5.0 and pre-incubation with 200 nM ABP **10**, before incubation with 200 nM ABP **13** or ABP **14**. After ABP incubation, proteins were denatured and separated by SDS-PAGE, and analysed according to the previously described method.[9] Coomassie staining was used to assess total protein loading.

3.4.6 *In vivo* activity of inhibitors in adult zebrafish

Surplus wild-type adult zebrafish of 3 months of age were administrated with a single dose of food grain mixed with DMSO, ABP **3** or **7** (1.6 nmol/fish, approximately 4 μmol/kg, $n = 3$ for each treatment) in $n = 2$ sets, according to project license AVD1060020184725,1-04 held by Dr. R.G. Boot. An initial experiment was performed with 3 adult zebrafish and the effect was confirmed by additional 3 individuals per experimental condition. Zebrafish were sacrificed after 24 h using Tricaine (250 mg/L), organs were harvested, snap-frozen in liquid nitrogen and stored at -80°C until use. Food grain consisted of Gemma micro mixed with Gemma diamond (Skretting, Stavanger, Norway). Lysis was performed with 50 μL lysis buffer (without benzonase) per sample, and lysates containing 20–60 μg total protein were analyzed by ABP method for GBA, GBA2, GAA, GANAB, and GUSB, as described in the previous section.

3.4.7 Sphingolipid extraction and analysis by mass spectrometry in treated zebrafish larvae

Zebrafish embryos at 8 hours post fertilization were seeded in 12-well plates (15 fish/well, 3 mL egg water/well) and treated with CBE **1** (10–1,000 μM)[9], **6** (0.001−10 μM) or **7** (0.001−10 μM) for 112 hours at 28 °C. Thereafter, zebrafish larvae were washed three times with egg water, and collected in clean screw-cap Eppendorf tubes (three tubes of three larvae per inhibitor concentration). Lipids were extracted and measured according to methods described previously.[31] Briefly, after removing of the egg water, 20 μL of ^{13}C-GlcSph[32] from concentration of 0.1 pmol μL^{-1} in MeOH, 480 μL MeOH, and 250 μL $CHCl_3$ were added to the sample, stirred, incubated for 30 min at RT, sonicated (5 x 1 min in sonication water bath), and centrifuged for 10 min at 15,700 g. Supernatant was collected in a clean tube, where 250 μL $CHCl_3$ and 450 μL 100 mM formate buffer (pH 3.2) were added. The sample was stirred and centrifuged, the upper phase was transferred to a clean tube. The lower phase was extracted with 500 μL MeOH and

450 μL formate buffer. The upper phases were pooled and taken to dryness in a vacuum concentrator at 45 °C. The residue was extracted with 700 μL butanol and 700 μL water, stirred and centrifuged. The upper phase (butanol phase) was dried and the residue was dissolved in 100 μL MeOH. 10 μL of this sample was injected to the LC-MS for lipid measurement. Two-tailed unpaired t-test was performed in Prism 7.0 software (GraphPad) to derive statistical significance, where $p < 0.05$ was considered significant.

3.5 References

1 Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, & Henrissat B (2014) The Carbohydrate-Active Enzymes Database (CAZy) in 2013. *Nucleic Acids Res* **42**, 490–495.

2 Brady RO, Kanfer JN, Bradley RM & Shapiro D (1966) Demonstration of a Deficiency of Glucocerebroside-Cleaving Enzyme in Gaucher's Disease. *J Clin Invest* **45**, 1112–1115.

3 Schapira AHV (2015) Glucocerebrosidase and Parkinson Disease: Recent Advances. *Mol Cell Neurosci* **66**, 37–42.

4 Tsuang D, Leverenz JB, Lopez OL, Hamilton RL, Bennett DA, Schneider JA, Buchman AS, Larson EB, Crane PK, Kaye JA Kramer P, Woltjer R, Kukull W, Nelson PT, Jicha GA, Neltner JH, Galasko D, Masliah E, Trojanowski JQ, Schellenberg GD, Yearout D, Huston H, Fritts-Penniman A, Mata IF, Wan JY, Edwards KL, Montine TJ & Zabetian CP (2012) GBA Mutations Increase Risk for Lewy Body Disease with and without Alzheimer Disease Pathology. *Neurology* **79**, 1944–1950.

5 Sidransky E, Nalls MA, Aasly JO, Aharon-Peretz J, Annesi G, Barbosa ER, Bar-Shira A, Berg D, Bras J, Brice A, Chen CM, Clark LN, Condroyer C, De Marco EV, Dürr A, Eblan MJ, Fahn S, Farrer MJ, Fung HC, Gan-Or Z, Gasser T, Gershoni-Baruch R, Giladi N, Griffith A, Gurevich T, Januario C, Kropp P, Lang AE, Lee-Chen GJ, Lesage S, Marder K, Mata IF, Mirelman A, Mitsui J, Mizuta I, Nicoletti G, Oliveira C, Ottman R, Orr-Urtreger A, Pereira LV, Quattrone A, Rogaeva E, Rolfs A, Rosenbaum H, Rozenberg R, Samii A, Samaddar T, Schulte C, Sharma M, Singleton A, Spitz M, Tan EK, Tayebi N, Toda T, Troiano AR, Tsuji S, Wittstock M, Wolfsberg TG, Wu YR, Zabetian CP, Zhao Y & Ziegler SG (2009) Multicenter Analysis of Glucocerebrosidase Mutations in Parkinson's Disease. *N Engl J Med* **361**, 1651–1661.

6 Holleran WM, Ginns EI, Menon GK, Grundmann JU, Fartasch M, McKinney CE, Elias PM & Sidransky E (1994) Consequences of β-Glucocerebrosidase Deficiency in Epidermis. Ultrastructure and Permeability Barrier Alterations in Gaucher Disease. *J Clin Invest* **93**, 1756–1764.

7 Farfel-Becker T, Vitner EB & Futerman AH (2011) Animal Models for Gaucher Disease Research. *Dis Model Mech* **4**, 746–52.

8 Vardi A, Zigdon H, Meshcheriakova A, Klein AD, Yaacobi C, Eilam R, Kenwood BM, Rahim AA, Massaro G, Merrill AH Jr, Vitner EB & Futerman AH (2016) Delineating Pathological Pathways in a Chemically Induced Mouse Model of Gaucher Disease. *J Pathol* **239**, 496–509.

9 Kuo CL, Kallemeijn WW, Lelieveld LT, Mirzaian M, Zoutendijk I, Vardi A, Futerman AH, Meijer AH, Spaink HP, Overkleeft HS, Aerts JMFG, Artola M (2019) In Vivo Inactivation of Glycosidases by Conduritol B Epoxide and Cyclophellitol as Revealed by Activity-Based Protein Profiling. *FEBS J* **286**, 584– 600.

10 Quaroni A, Gershon E & Semenza G (1974) Affinity Labeling of the Active Sites in the Sucrase-Isomaltase Complex from Small Intestine. *J Biol Chem* **249**, 6424–6433.

11 Yang SJ, Ge SG, Zeng YC & Zhang SZ (1985) Inactivation of α-Glucosidase by the Active-Site-Directed Inhibitor, Conduritol B Epoxide. *BBA - Protein Struct Mol Enzymol* **828**, 236–240.

12 Hermans MM, Krooss MA, van Beeurnens J, Oostras BA, Reuser AJ (1991) Human Lysosomal alpha-Glucosidase. Characterization of the Catalytic Site. *Biochemistry* **266**, 13507–13512.

13 Braun H, Legler G, Deshusses J & Semenza G (1977) Stereospecific Ring Opening of Conduritol-B-Epoxide by an Active Site Aspartate Residue of Sucrase-Isomaltase. *BBA - Enzymol* **483**, 135–140.

14 van Weely S, Brandsma M, Strijland A, Tager JM & Aerts JMFG (1993) Demonstration of the Existence of a Second, Non-Lysosomal Glucocerebrosidase That Is Not Deficient in Gaucher Disease. *BBA - Mol Basis Dis* **1181**, 55–62.

15 Ridley CM, Thur KE, Shanahan J, Thillaiappan NB, Shen A, Uhl K, Walden CM, Rahim AA, Waddington SN, Platt FM & van der Spoel AC (2013) β-Glucosidase 2 (GBA2) Activity and Imino Sugar Pharmacology. *J Biol Chem* **288**, 26052–26066.

16 Hara A & Radin NS (1979) Enzymic Effects of β-Glucosidase Destruction in Mice Changes in Glucuronidase Levels. *BBA - Gen Subj* ***582***, 423–433.

17 Witte MD, Kallemeijn WW, Aten J, Li KY, Strijland A, Donker-Koopman WE, van den Nieuwendijk AM, Bleijlevens B, Kramer G, Florea BI, Hooibrink B, Hollak CE, Ottenhoff R, Boot RG, van der Marel GA, Overkleeft HS & Aerts JM (2010) Ultrasensitive in Situ Visualization of Active Glucocerebrosidase Molecules. *Nat Chem Biol* **6**, 907–913.

18 Herrera Moro Chao D, Kallemeijn WW, Marques AR, Orre M, Ottenhoff R, van Roomen C, Foppen E, Renner

MC, Moeton M, van Eijk M, Boot RG, Kamphuis W, Hol EM, Aten J, Overkleeft HS, Kalsbeek A & Aerts JM (2015) Visualization of Active Glucocerebrosidase in Rodent Brain with High Spatial Resolution Following in Situ Labeling with Fluorescent Activity Based Probes. *PLoS One* **10**, e0138107.

19 Yadav AK, Shen DL, Shan X, He X, Kermode AR, Vocadlo DJ (2015) Fluorescence-Quenched Substrates for Live Cell Imaging of Human Glucocerebrosidase Activity. *J Am Chem Soc* **137**, 1181–1189.

20 Lahav D, Liu B, van den Berg RJBHN, van den Nieuwendijk AMCH, Wennekes T, Ghisaidoobe AT, Breen I, Ferraz MJ, Kuo CL, Wu L, Geurink PP, Ovaa H, van der Marel GA, van der Stelt M, Boot RG, Davies GJ, Aerts JMFG & Overkleeft HS (2017) A Fluorescence Polarization Activity-Based Protein Profiling Assay in the Discovery of Potent, Selective Inhibitors for Human Nonlysosomal Glucosylceramidase. *J Am Chem Soc* **139**, 14192–14197.

21 Charoenwattanasatien R, Pengthaisong S, Breen I, Mutoh R, Sansenya S, Hua Y, Tankrathok A, Wu L, Songsiriritthigul C, Tanaka H, Williams SJ, Davies GJ, Kurisu G & Cairns JR (2016) Bacterial β-Glucosidase Reveals the Structural and Functional Basis of Genetic Defects in Human Glucocerebrosidase 2 (GBA2). *ACS Chem Biol* ***11***, 1891–1900.

22 Berger J, Lecourt S, Vanneaux V, Rapatel C, Boisgard S, Caillaud C, Boiret-Dupré N, Chomienne C, Marolleau JP, Larghero J & Berger MG (2010) Glucocerebrosidase Deficiency Dramatically Impairs Human Bone Marrow Haematopoiesis in an in Vitro Model of Gaucher Disease. *Br J Haematol* **150**, 93–101.

23 Schueler UH, Kolter T, Kaneski CR, Zirzow GC, Sandhoff K & Brady RO (2004) Correlation between Enzyme Activity and Substrate Storage in a Cell Culture Model System for Gaucher Disease. *J Inherit Metab Dis* **27**, 649–658.

24 Kuo CL, van Meel E, Kytidou K, Kallemeijn WW, Witte M, Overkleeft HS, Artola ME & Aerts JM (2018) Activity-Based Probes for Glycosidases: Profiling and Other Applications. *Methods Enzymol* **598**, 217– 235.

25 Dekker N, van Dussen L, Hollak CE, Overkleeft H, Scheij S, Ghauharali K, van Breemen MJ, Ferraz MJ, Groener JE, Maas M, Wijburg FA, Speijer D, Tylki-Szymanska A, Mistry PK, Boot RG & Aerts JM (2011) Elevated Plasma Glucosylsphingosine in Gaucher Disease: Relation to Phenotype, Storage Cell Markers, and Therapeutic Response. *Blood* **118**, e118–e127.

26 Ferraz MJ, Marques AR, Appelman MD, Verhoek M, Strijland A, Mirzaian M, Scheij S, Ouairy CM, Lahav D, Wisse P, Overkleeft HS, Boot RG & Aerts JM (2016) Lysosomal Glycosphingolipid Catabolism by Acid Ceramidase: Formation of Glycosphingoid Bases during Deficiency of Glycosidases. *FEBS Lett* **590**, 716–725.

27 Schröder SP, van de Sande JW, Kallemeijn WW, Kuo CL, Artola M, van Rooden EJ, Jiang J, Beenakker TJM, Florea BI, Offen WA, Davies GJ, Minnaard AJ, Aerts JMFG, Codée JDC, van der Marel GA & Overkleeft HS (2017) Towards Broad Spectrum Activity-Based Glycosidase Probes: Synthesis and Evaluation of Deoxygenated Cyclophellitol Aziridines. *Chem Commun* **53**, 12528–12531.

28 Jiang J, Beenakker TJ, Kallemeijn WW, van der Marel GA, van den Elst H, Codée JD, Aerts JM & Overkleeft HS (2015) Comparing Cyclophellitol N-Alkyl and N-Acyl Cyclophellitol Aziridines as Activity-Based Glycosidase Probes. *Chemistry* **21**, 10861–10869.

29 Jiang J, Kuo CL, Wu L, Franke C, Kallemeijn WW, Florea BI, van Meel E, van der Marel GA, Codée JD, Boot RG, Davies GJ, Overkleeft HS & Aerts JM (2016) Detection of Active Mammalian GH31 α-Glucosidases in Health and Disease Using in-Class, Broad-Spectrum Activity-Based Probes. *ACS Cent Sci* **2**, 351–358.

30 Wu L, Jiang J, Jin Y, Kallemeijn WW, Kuo CL, Artola M, Dai W, van Elk C, van Eijk M, van der Marel GA, Codée JDC, Florea BI, Aerts JMFG, Overkleeft HS & Davies GJ. (2017) Activity-Based Probes for Functional Interrogation of Retaining β-Glucuronidases. *Nat Chem Biol* **13**, 867–873.

31 Mirzaian M, Wisse P, Ferraz MJ, Marques ARA, Gabriel TL, van Roomen CPAA, Ottenhoff R, van Eijk M, Codée JDC, van der Marel GA, Overkleeft HS & Aerts JM (2016) Accurate Quantification of Sphingosine-1-Phosphate in Normal and Fabry Disease Plasma, Cells and Tissues by LC-MS/MS with13C-Encoded Natural S1P as Internal Standard. *Clin Chim Acta* **459**, 36–44.

32 Wisse P, Gold H, Mirzaian M, Ferraz MJ, Lutteke G, Van Den Berg RJB H N, Van Den Elst H, Lugtenburg J, Van Der Marel GA, Aerts JMFG, Codée JDC & Overkleeft HS (2015) Synthesis of a Panel of Carbon-13-Labelled (Glyco)sphingolipids. *European J Org Chem* **2015**, 2661-2677.

APPENDIX

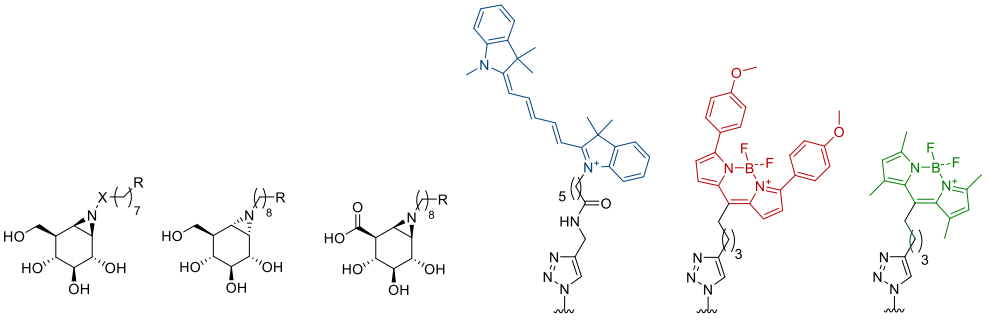


Figure 3.S1. Chemical structures of ABPs 8-14 used in this work. Code for each ABP: ABP **8** (JJB367), ABP **9** (JJB75), ABP **10** (JJB70), ABP **11** (JJB383), ABP **12** (JJB347), ABP **13** (JJB392), and ABP **14** (JJB391).

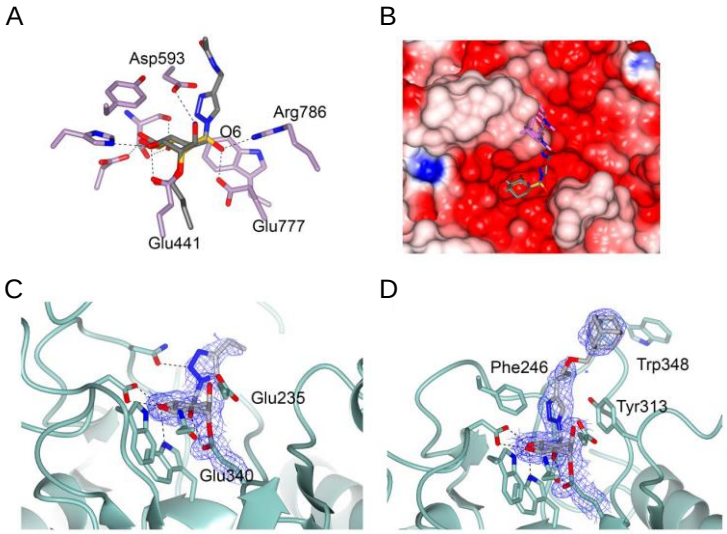


Figure 3.S2. Accommodation of O6-derivatized ligands in GBA1 and GBA2. A) Overlay of the bacterial GBA2 homology TxGH116 (PDB 5NCX, carbon-atoms colored lavender, aziridine complex bound to Glu441 in yellow) with the GBA1 complex with **5** reported here (grey carbons). Of note is the "inwards" orientation of O6, via interactions with Arg786 and Glu777 and the constricted space by virtue of the position of Asp593. B) Partial electrostatic surface of TxGH116 overlaid with the ligand coordinates for **5** bound to GBA1 (this work) overlaid (indicating potential steric clashes were the ligand bound exactly as in GBA1). Together these overlays suggest that TXGH116 is less able to accommodate O6 substituted ligands, at least to the extent that O6 substituted ligands gain sufficient discriminatory power for GBA1/GBA2. C) Structure of GBA reacted with biphenyl-cyclophellitol **6** and adamantyl-cyclophellitol. The biphenyl moiety is not visible and just the cyclophellitol-GBA complex structure is observed (PDB Code 6Q6N). D) The linker-adamantyl moiety of **6** is observed in slightly different positions in the two molecules of the asymmetric unit (PDB Code 6Q6L) reflecting its binding through predominantly hydrophobic interactions and the flexibility of its accommodation in crystal.

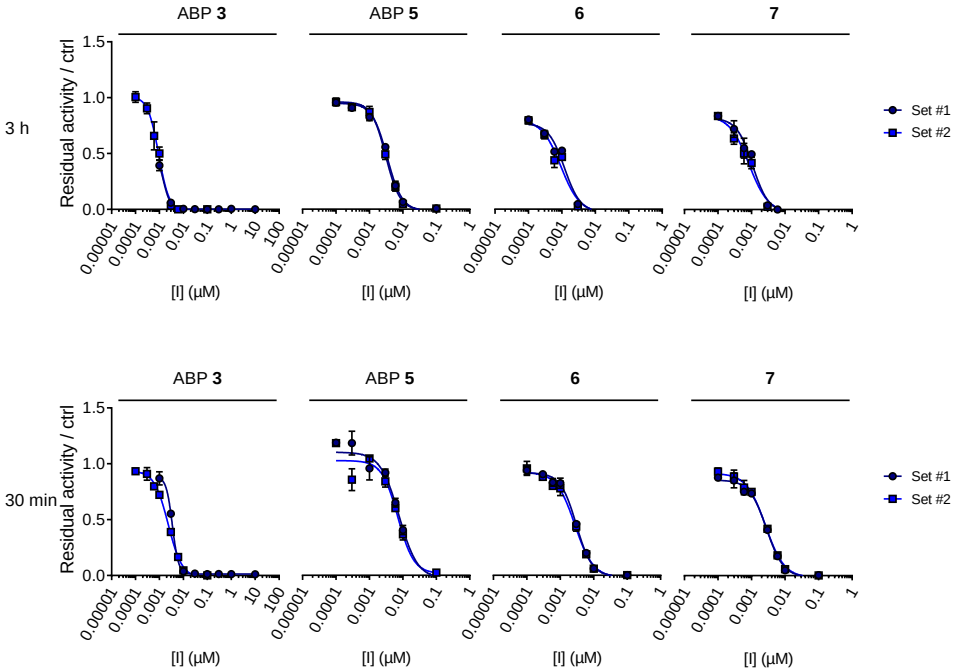


Figure 3.S3. *In vitro* inhibition curves of ABPs 3 and 5 and inhibitors 6 and 7 on rGBA: 3 h *vs* 30 min incubation.

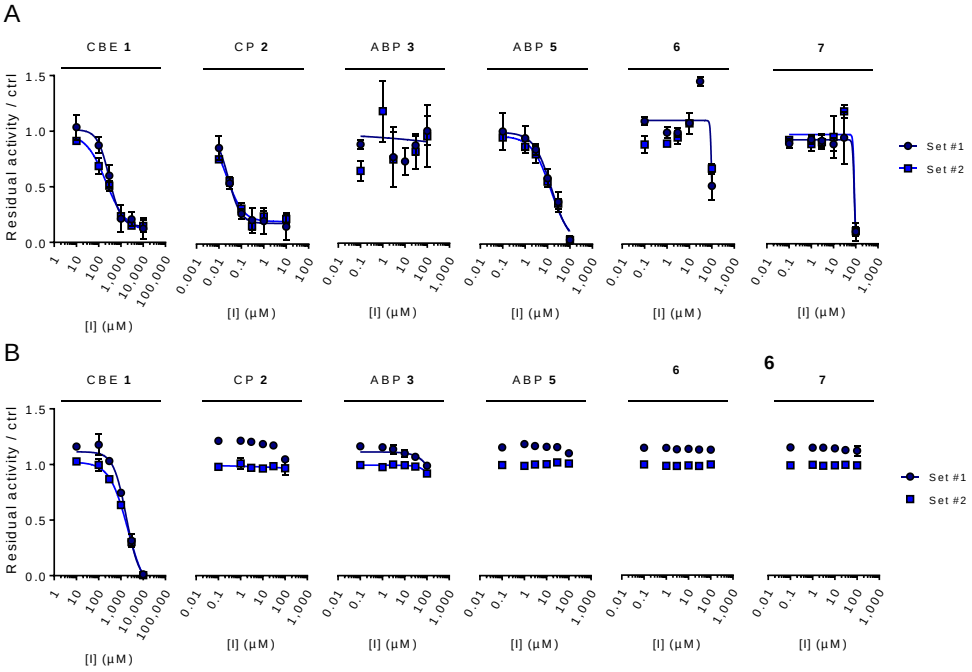


Figure 3.S4. *In vitro* inhibition curves of CBE 1, CP 2, ABPs 3 and 5, and inhibitors 6 and 7 on GBA2 from A) overexpressed cell lysates and on B) rGAA.

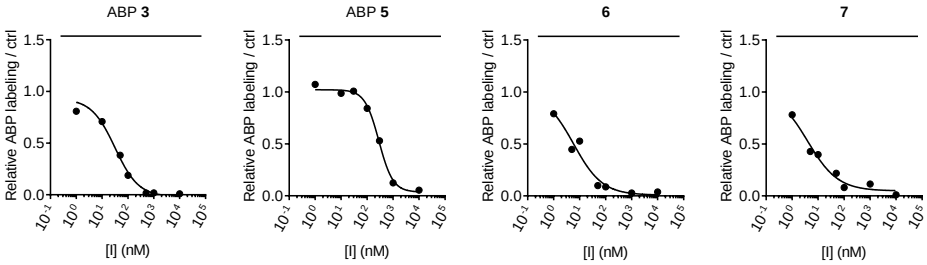


Figure 3.S5. Inhibition curves for i*n vivo* GBA inactivation in zebrafish larvae by ABPs 3 and 5, and inhibitors 6 and 7 quantified by ABP labeling.

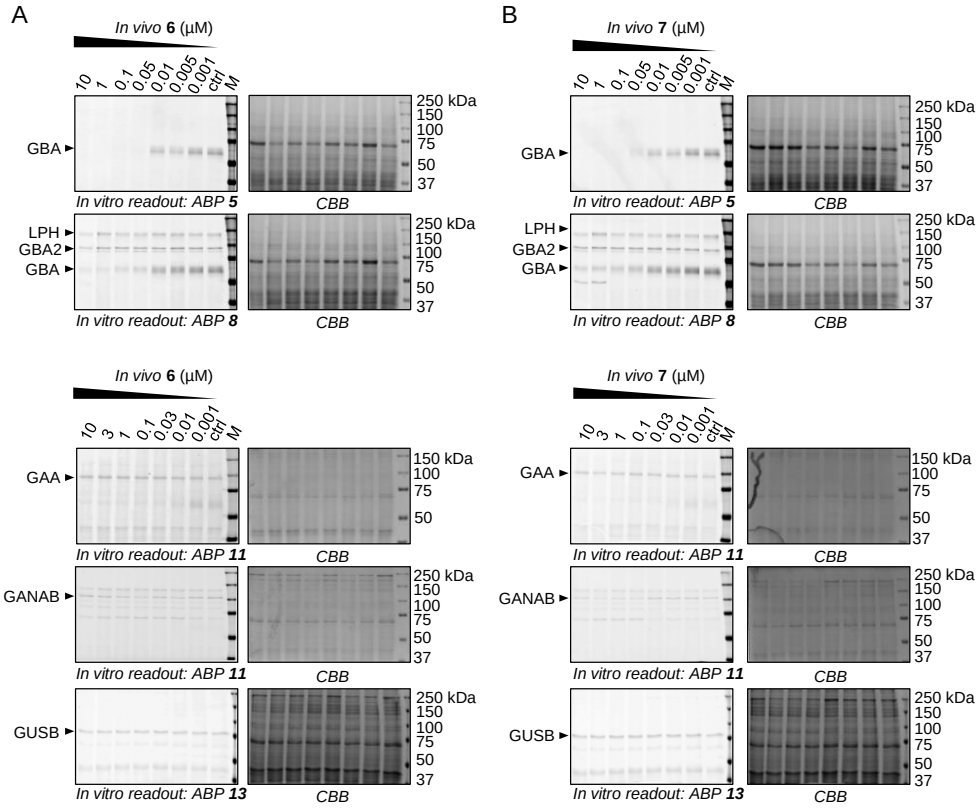


Figure 3.S6 (continued, 1 of 2).

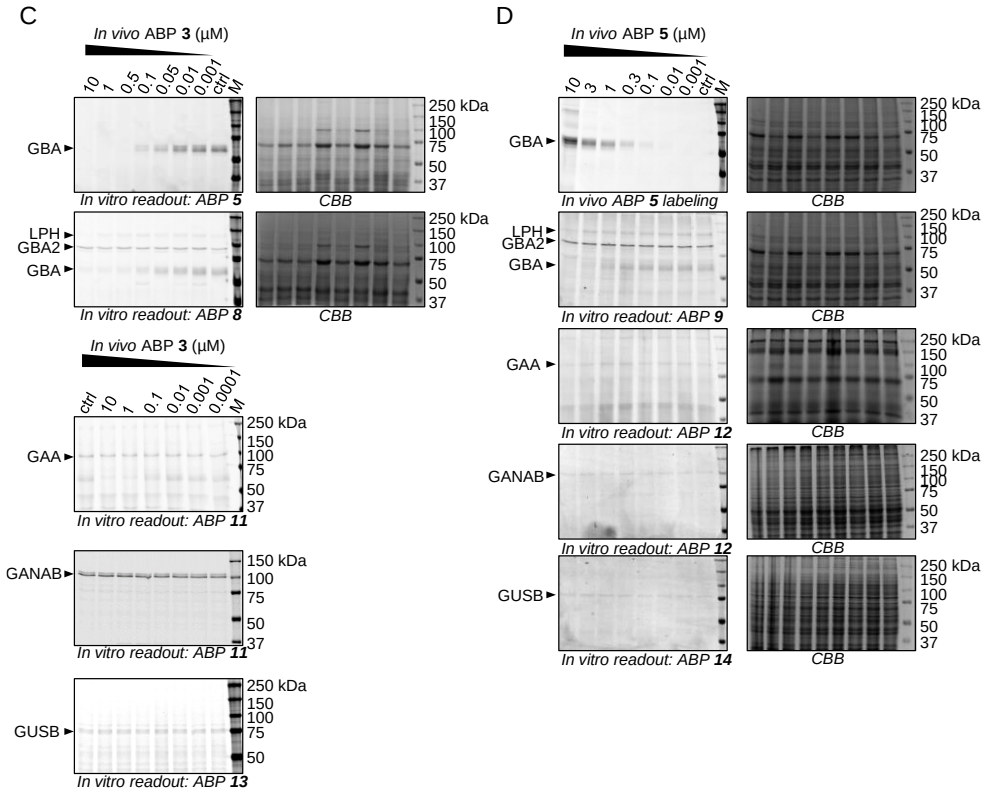


Figure 3.S6. ***In vivo*** **targets of A) inhibitor 6, B) inhibitor 7, C) ABP 3, and D) ABP 5 in zebrafish larvae revealed by competitive ABPP.** ***CBB*****, coomassie brilliant blue stain.**

A

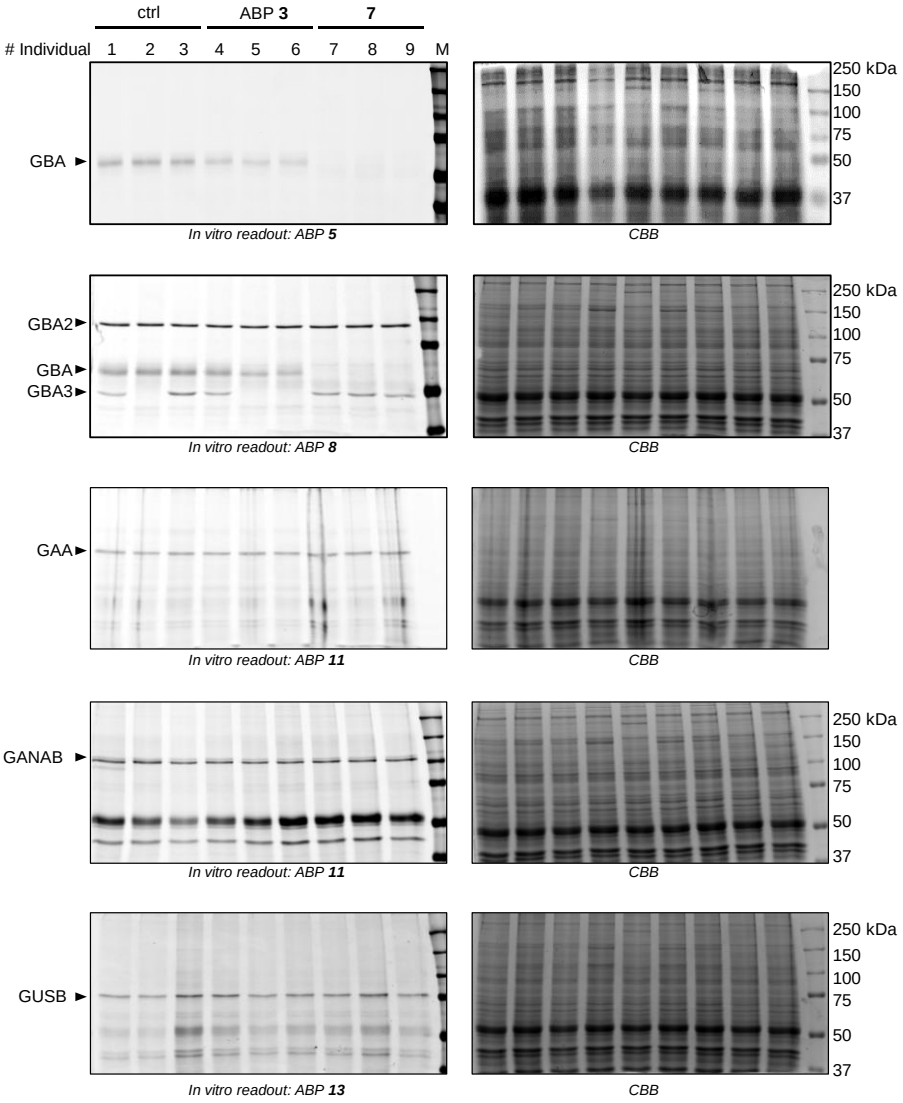


Figure 3.S7 (continued, 1 of 4).

B

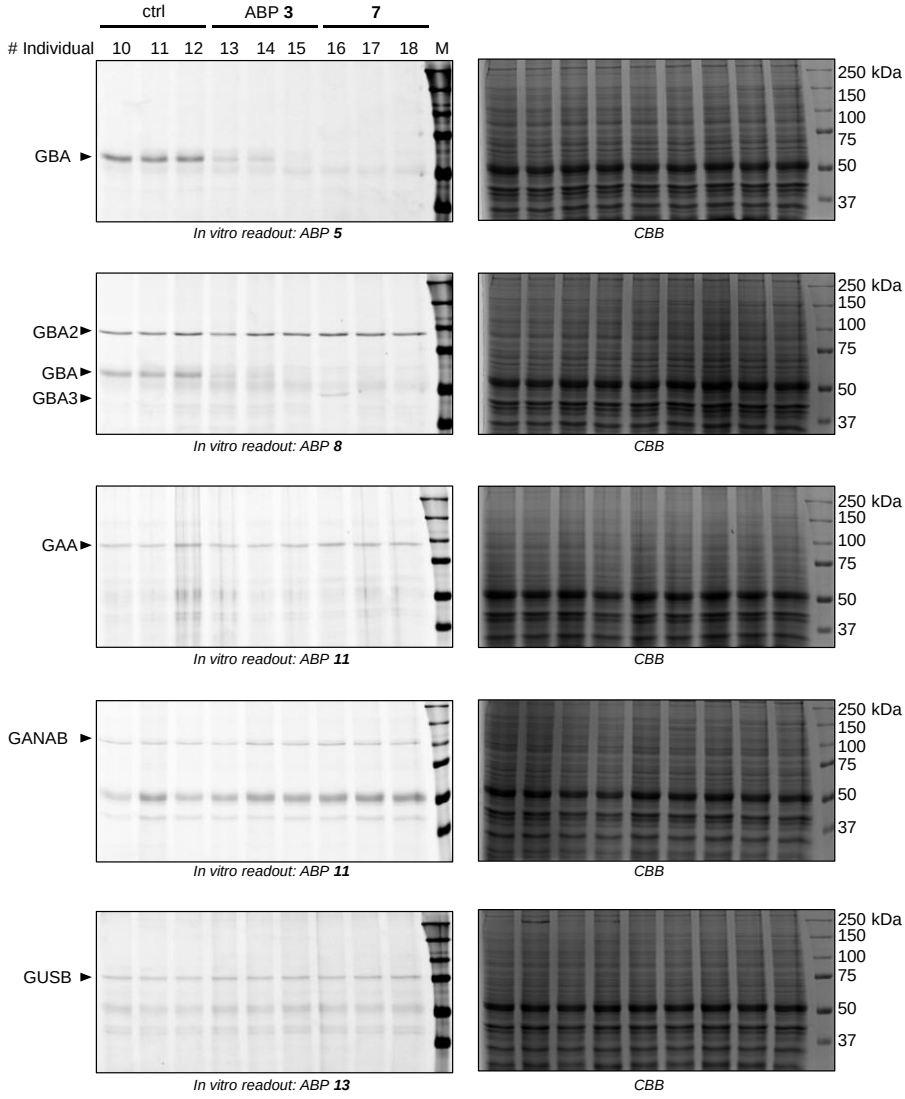


Figure 3.S7 (continued, 2 of 4).

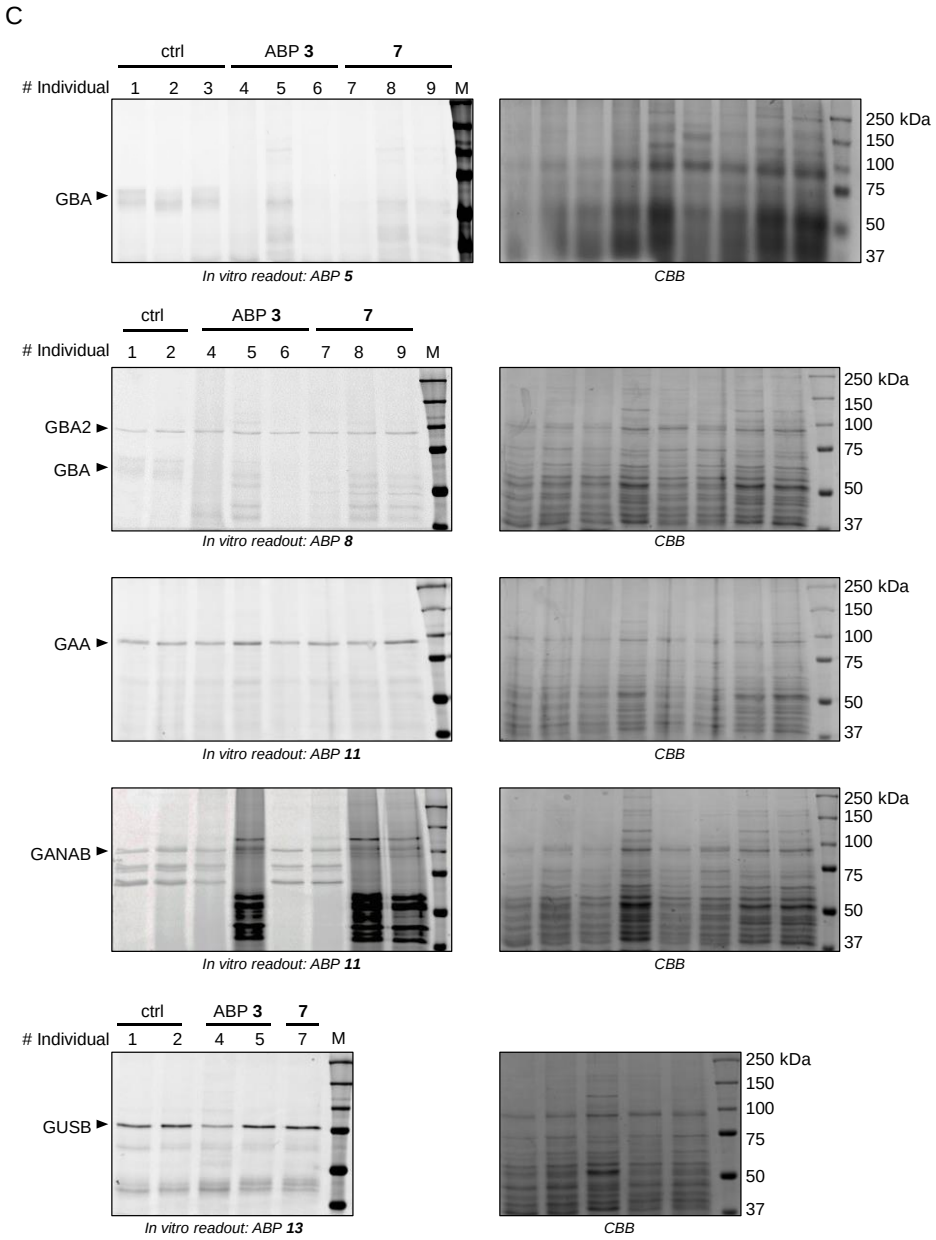


Figure 3.S7 (continued, 3 of 4).

D

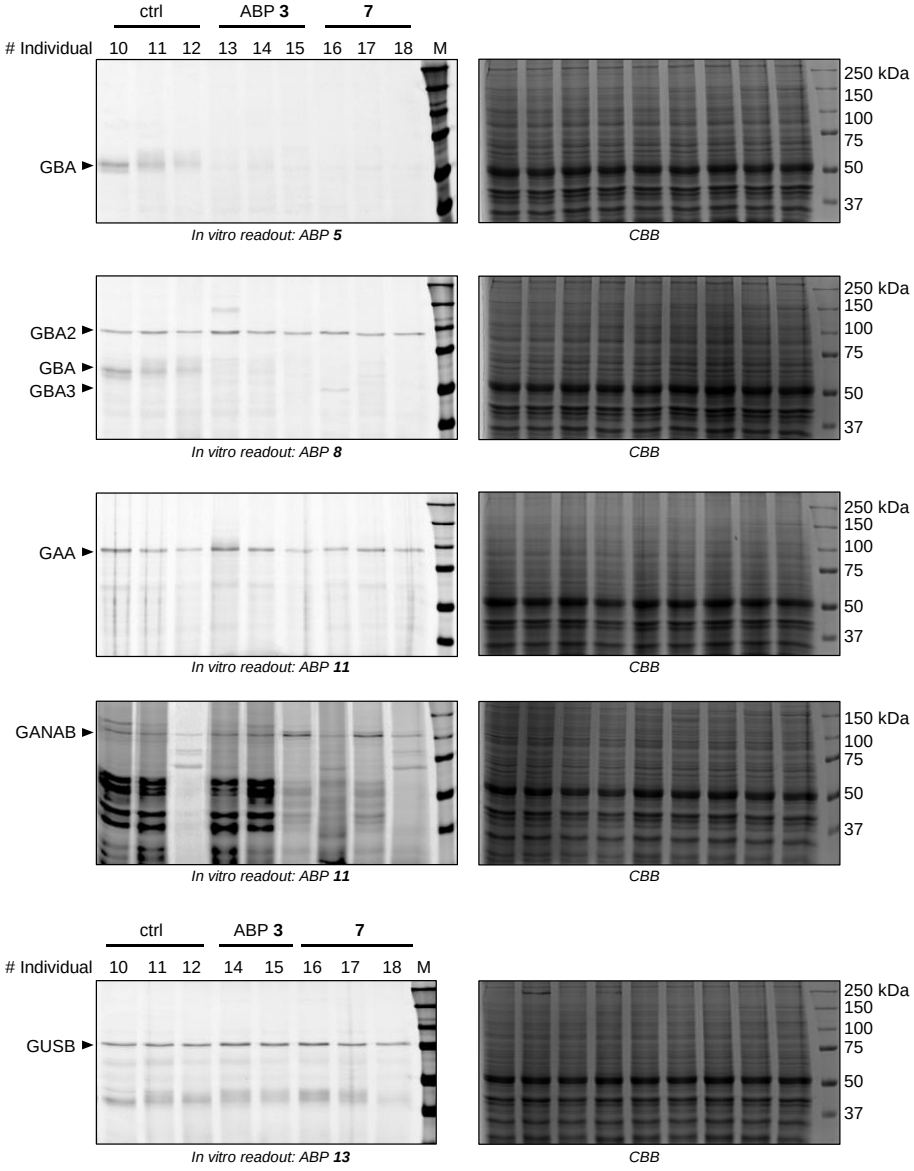


Figure 3.S7. *In vivo* targets of ABPs 3 and inhibitor 7 of treated adult zebrafish revealed by competitive ABPP. A), B) labeling in brain homogenates C), D) Labeling in liver and spleen homogenates. *CBB*, Coomassie brilliant blue stain.

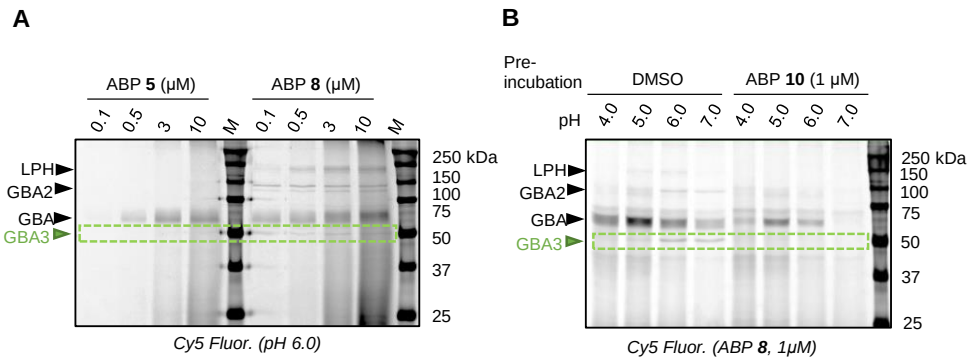


Figure 3.S8. Detection of GBA3 in lysates of zebrafish larvae. A) Labeling using ABP **5** (ME569) or ABP **8** (JJB367) at various ABP concentrations. B) Labeling using ABP **8** (JJB367) in lysates at various pH, with or without pre-incubation of ABP **10** (JJB70).

Table 3.S1. Data collection and refinement statistics for rGBA in complex with ABP **5** and inhibitors **6** and **7**.

	5	**6**	**7**
Data collection			
Space group	*C* 2 2 2_1	*C* 2 2 2_1	*C* 2 2 2_1
Cell dimensions			
a, *b*, *c* (Å)	110.6, 285.9, 92.3	110.4, 285.2, 91.9	110.2, 285.1, 92.0
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å)	77.52-1.92 (1.95-1.92)*	72.13-1.63 (1.66-1.63)	71.98-1.81 (1.84-1.81)
R_{merge}	0.161 (1.623)	0.101 (1.609)	0.132 (1.661)
R_{pim}	0.064 (0.647)	0.047 (0.735)	0.061 (0.757)
I / σI	6.5 (1.1)	8.8 (1.1)	7.7 (1.0)
Completeness (%)	99.9 (99.8)	99.9 (98..8)	100.0 (98.7)
Redundancy	8.3 (8.1)	6.49 (6.50)	6.53 (6.58)
Refinement			
Resolution (Å)	77.52-1.92	72.13-1.63	71.98-1.81
No. reflections	925691	1167262	859243
R_{work} / R_{free}	0.18/0.22	0.18/0.21	0.18/0.21
No. atoms			
Protein	7939	7870	7960
Ligand/ion	354	273	334
Water	699	966	936
B-factors (Å^2)			
Protein	39	26	28
Ligand/ion	65	52	55
Water	47	39	39
R.m.s. deviations			
Bond lengths (Å)	0.009	0.012	0.010
Bond angles (°)	1.60	1.67	1.63
Ramachandran Plot Residues			
In most favourable regions (%)	94.5	94.7	94.6
In allowed regions (%)	4.3	4.2	4.3
PDB code	6Q6K	6Q6N	6Q6L

*Values in parentheses are for highest-resolution shell.

Table 3.S2. Lipophilic ligand efficiency (LipE) values of CBE **1**, CP **2**, ABP **3**, ABP **5** and new functionalized cyclophellitol **6** and **7** for rGBA activity. LipE = pIC_{50} – cLogP. cLogP values were calculated using the ChemDraw version 16 software.

	CBE 1	**CP** 2	**ABP** 3	**ABP** 5	6	7
cLogP	-1.76	-2.77	2.47	1.66	1.16	1.60
pIC_{50} GBA	5.37	7.52	9.00	8.52	9.00	9.00
LipE	7.13	10.29	6.53	6.86	7.84	7.40

3.S1 Crystallographic data collection and refinement statistics (University of York)

3.S1.1 Deglycosylation, Purification and Crystallization

Prior to crystallization, Cerezyme (3.4 mg mL^{-1}, NEB) was deglycosylated with PNGase F (20 μL, NEB) for 5 days at room temperature. The digested material was purified by size exclusion chromatography on an S75 16/600 column. Crystals were obtained using hanging-drop vapor diffusion, based on parent conditions outlined by Dvir et al. (2003).[1] Drops contained 1 μL purified Cerezyme (9.1 mg mL^{-1}) and 1 μL mother liquor (1.1 M $(NH_2)_2SO_4$, 0.19 M guanidine HCl, 0.04 M KCl, 0.1 M Na acetate, pH 4.6). Crystals were grown for 1 week at 18 °C.

3.S1.2 Inhibitor Complexes

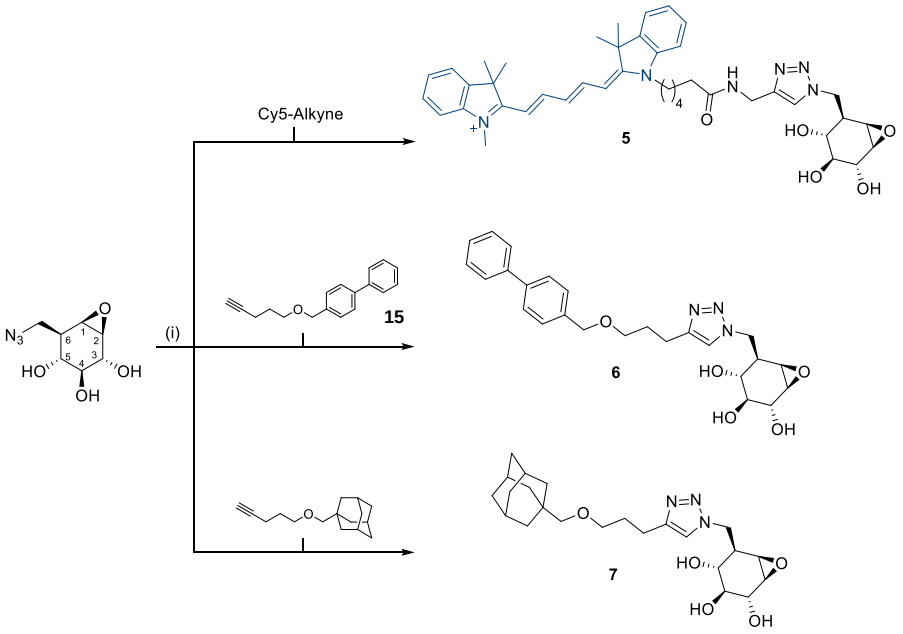
Inhibitors were prepared at 20 mM in HEPES buffer (20 mM, pH 7.0) and diluted to 2 mM in mother liquor comprising 1.1 M $(NH_2)_2SO_4$, 0.19 M guanidine HCl, 0.04 M KCl and 0.1 M Na acetate (pH 4.6). Crystals were soaked in 5 μL of inhibitor-mother liquor solution for 4 hours at 18 °C. Crystals were transferred to a lithium sulfate cryoprotectant (0.2 M Li_2SO_4, 0.17 M guanidine HCl, 0.04 M KCl, 0.1 M Na acetate, pH 4.6) before freezing in liquid nitrogen.

3.S1.3 Structure Solution and Refinement

For the ABP **5** complex, data were collected at the i04_1 beamline of the Diamond Light Source Facility to 1.79 Å. Data were collected at the i24 beamline to 1.81 Å and 1.63 Å for the inhibitors **6** and **7** complexes respectively. For all complexes, data were processed using the XIA2[2, 3] and AIMLESS data reduction pipelines through the CCP4i2 software.[4] The structures were solved by molecular replacement using MOLREP[5] with the previous GBA1 PDB (2NT0)[6] as the homologous search model. Refinement was performed using REFMAC[7] followed by several rounds of manual model building with COOT.[8] Idealised coordinate sets and refinement dictionaries for the ligands were generated using JLIGAND[9] and sugar conformations were validated using Privateer.[10] Crystal structure figures were generated using ccp4mg.[11] For ABP **5**, the structure could be traced from residues 1–497 in chain A and chain B. In chain A, residue 344 is disordered in density and has been omitted. In chain B residues 29–31 and 346–348 are disordered in density and have been omitted. For inhibitor **6**, the structure could be traced from residues 1–497 in chain A and chain B. In chain B, residues 344–345 are disordered in density and have been omitted.

For inhibitor **7** the structure could be traced from residues 1–497 in chain A and chain B. In chain A, residue 345 is disordered in density and has been omitted. In chain B, residues 316–318 and residues 344–347 are disordered in density and have been omitted.

3.S2 Synthesis of ABP 5 and Compounds 6 and 7 (Department of Bio-organic Synthesis, Leiden University)



Scheme 3.S1. Synthesis of C-8 functionalized cyclophellitol ABP 5 and inhibitors 6 and 7. Reagents and conditions: (i) $CuSO_4 \cdot H_2O$, sodium ascorbate, DMF, RT, 18 h, **5**: 38%, **6**: 78%, **7**: 89%.

3.S2.1 4-((pent-4-yn-1-yloxy)methyl)-1,1'-biphenyl (**15**)

[1,1'-biphenyl]-4-ylmethanol (689 mg, 3.74 mmol), tetrabutyl ammonium iodide (1.38 g, 3.74 mmol) and NaH (204 mg, 5.10 mmol) were added to a cooled (0 °C) solution of 5-bromopent-1-yne (500 mg, 3.74 mmol) in DMF (15 mL). The reaction mixture was allowed to warm to room temperature and stirred 24 h and subsequently poured into ice, extracted with EtOAc (20 mL), and washed with water (20 mL) and saturated aqueous NaCl (10 mL). The organic phase was dried ($MgSO_4$), filtered and concentrated. The residue was purified by silica gel column chromatography with Pentane/EtOAc (from 10:0 to 8:2, v/v) to give the desired product as a yellow oil (102 mg, mmol, 12% yield). ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.62 – 7.49 (m, 4H, 4 x

CH Ar), 7.49 – 7.28 (m, 5H, 5 x CH Ar), 4.54 (s, 2H, CH_2O), 3.59 (t, J = 6.1 Hz, 2H, CH_2O), 2.33 (td, J = 7.1, 2.7 Hz, 2H, CH_2), 1.94 (t, J = 2.7 Hz, 1H, CH), 1.91 – 1.78 (m, 2H, CH_2); ^{13}C-NMR (101 MHz, $CDCl_3$) δ 141.0, 140.6, 137.6, 128.9, 128.7, 128.4, 128.2, 127.4, 127.3, 127.2, 84.1, 72.8, 68.8, 68.6, 28.8, 15.4. LC/MS: (linear gradient 10%→90% B in 12.5 min), m/z 251.1 $[M+H]^+$.

3.S2.2 (1*S*,2*R*,3*S*,4*R*,5*R*,6*R*)-5-((4-(3-([1,1'-biphenyl]-4-ylmethoxy)propyl)-1H-1,2,3-triazol-1-yl)methyl)-7-oxabicyclo[4.1.0]heptane-2,3,4-triol (**6**)

Azido cyclophellitol[12] (4.4 mg, 0.022 mmol) was dissolved in DMF (1 mL) and $CuSO_4 \cdot 5H_2O$ (0.1 M, 50 µL, 4.4 µmol, 0.2 eq.) and sodium ascorbate (0.1 M, 50 µL, 4.4 µmol, 0.2 eq.) were added to the solution under argon atmosphere. Then, a solution of biphenyl alkyne (4.4 mg, 26 µmol, 1.2 eq.) was added and the reaction mixture was stirred at room temperature for 18 h. The crude was purified by column chromatography (DCM to DCM/MeOH 9:1) to afford the desired compound **6** (7.7 mg, 0.017 mmol, 78%). Rf: 0.20 (DCM/MeOH 9:1). ^{1}H-NMR (400 MHz, CD_3OD) δ 7.65 – 7.55 (m, 4H, 4 x CH Ar), 7.43 (t, J = 7.7 Hz, 4H, 4 x CH Ar), 7.32 (t, J = 7.4 Hz, 1H, CH Ar), 5.49 (s, 1H, CH=N), 4.81 (d, J = 14.3 Hz, 1H, C*H*H-7), 4.69 – 4.56 (m, 1H, CH*H*-7), 4.55 (s, 2H, CH_2O), 3.65 – 3.48 (m, 3H, CH_2O, CH-2), 3.21 (t, J = 8.8 Hz, 1H, CH-3), 3.17 – 3.06 (m, 1H, CH-4), 3.02 – 2.97 (s, 2H, CH-1, CH-6), 2.85 (br s, 2H, CH_2), 2.40 – 2.35 (t, J = 7.3 Hz, 1H, CH-5), 2.03 (br s, 2H, CH_2); ^{13}C-NMR (101 MHz, CD_3OD) δ 142.1, 141.8, 138.93 (3 x C Ar), 131.1, 129.9, 129.4, 128.4, 127.9 (9 x CH Ar), 78.2 (CH-4), 73.6 (CH_2O), 72.5 (CH-2), 70.1 (CH_2O), 68.7 (CH-3), 57.5, 55.5 (CH-1/CH-6), 51.1 (CH_2), 44.5 (CH-5), 30.2 (CH_2). HRMS: calculated for $C_{25}H_{30}N_3O_5^+$ $[M+H]^+$ 452.21855, found: 452.21759.

3.S2.3 (1*S*,2*R*,3*S*,4*R*,5*R*,6*R*)-5-((4-(3-(((1*s*,3*S*)-adamantan-1-yl)methoxy)propyl)-1*H*-1,2,3-triazol-1-yl)methyl)-7-oxabicyclo[4.1.0]heptane-2,3,4-triol (**7**)

Azido-cyclophellitol[12] (4.4 mg, 0.022 mmol) was dissolved in DMF (1 mL) and $CuSO_4 \cdot 5H_2O$ (0.1 M, 50 µL, 4.4 µmol, 0.2 eq.) and sodium ascorbate (0.1 M, 50 µL, 4.4 µmol, 0.2 eq.) were added to the solution under argon atmosphere. Then, a solution of adamantane alkyne (6.10 mg, 26 µmol, 1.2 eq.) was added and the reaction mixture was stirred at room temperature for 18 h. The crude was purified by column chromatography (DCM to DCM/MeOH 9:1) to afford the desired compound **7** (8.4 mg, 0.019 mmol, 89%). Rf: 0.20 (DCM/MeOH 9:1). ^{1}H-NMR (400 MHz, CD_3OD) δ 4.82 (dd, J = 13.7, 3.9 Hz, 1H, CH*H*), 4.62 (dd, J = 13.8, 8.5 Hz, 1H, C*H*H),

3.62 (d, *J* = 8.1 Hz, 1H, CH-2), 3.42 (t, *J* = 5.4 Hz, 2H, CH_2), 3.23 (dd, *J* = 10.0, 8.1 Hz, 1H, CH-3), 3.13 (t, *J* = 9.8 Hz, 1H, CH-4), 3.03 (s, 2H, CH-1, CH-6), 2.97 (s, 2H, CH_2O), 2.80 (br s, 2H, CH_2), 2.39 (td, *J* = 9.0, 3.5 Hz, 1H, CH-5), 1.95 (br s, 4CH, 4 x CH adamantane), 1.85 – 1.63 (m, 6H, 3 x CH_2 adamantane), 1.57 (br s, 6H, 3 x CH_2 adamantane); ^{13}C-NMR (101 MHz, CD_3OD) *δ* 83.0 (CH_2O), 78.2 (CH-3), 72.5 (CH-2), 71.4 (CH_2), 68.7 (CH-4), 57.6, 55.5 (CH-1/CH-6), 44.6 (CH-5), 40.8 (3 x CH_2), 38.3 (3 x CH_2), 35.1 (C adamantane), 30.3 (CH_2), 29.8 (4 x CH adamantane), 23.2 (CH_2); HRMS: calculated for $C_{25}H_{30}N_3O_5^+$ $[M+H]^+$ 434.26550, found: 434.26441.

3.S2.4 Cyclophellitol Cy5 ABP (**5**)

Azido-cyclophellitol[12] (24.2 mg, 0.12 mmol) and the desired Cy5-alkyne (84 mg, 0.15 mmol) were dissolved in BuOH/toluene/H_2O (6 mL, 1:1:1, v/v/v). $CuSO_4$ (0.024 mL, 1 M in H_2O) and sodium ascorbate (0.024 mL, 1 M in H_2O) were added and the reaction mixture was heated at 80 ºC for 18 h. Then, the solution was diluted with CH_2Cl_2, washed with H_2O, dried over $MgSO_4$ and concentrated under reduced pressure. The crude was purified by silica gel column chromatography (CH_2Cl_2 to CH_2Cl_2/MeOH 9:1), subsequently purified by semipreparative reversed-phase HPLC (linear gradient: 45 to 48% B in A, 12 min, solutions used A: 50mM NH_4HCO_3 in H_2O, B: MeCN) and lyophilized to yield ABP **5** as a blue powder (33.5 mg, 44 μmol, 37%). ^{1}H NMR (400 MHz, CD_3OD): *δ* 8.25 (t, *J* = 13.1 Hz, 2H), 7.91 (d, *J* = 3.3 Hz, 1H), 7.49 (d, *J* = 7.4 Hz, 2H), 7.44 – 7.39 (m, 2H), 7.31 – 7.24 (m, 4H), 6.63 (t, *J* = 12.4 Hz, 1H), 6.29 (d, *J* = 13.8 Hz, 2H), 4.81 (dd, *J* = 13.9, 3.8 Hz, 1H), 4.61 (dd, *J* = 13.9, 8.6 Hz, 1H), 4.42 (s, 2H), 4.09 (t, *J* = 7.4 Hz, 2H), 3.63 (s, 3H), 3.60 (d, *J* = 8.0 Hz, 1H), 3.24 – 3.20 (m, 1H), 3.12 (t, *J* = 14.9 Hz, 1H), 3.02 – 3.00 (m, 1H), 2.41 – 2.35 (m, 1H), 2.25 (t, *J* = 7.3 Hz, 2H), 1.91 (s, 6H), 1.85 – 1.78 (m, 2H), 1.72 (s, 12H), 1.50 – 1.43 (m, 2H) ppm; ^{13}C NMR (101 MHz, CD_3OD): *δ* 175.8, 175.4, 174.6, 155.5, 155.5, 146.2, 144.2, 143.5, 142.6, 142.5, 129.8, 129.7, 126.7, 126.2, 126.2, 125.2, 123.4, 123.3, 112.0, 111.8, 104.5, 104.3, 78.2, 72.5, 68.6, 57.6, 55.5, 50.8, 50.5, 50.5, 44.8, 44.6, 36.5, 35.6, 31.6, 28.1, 28.0, 27.3, 26.4 ppm; HRMS: calculated for $C_{42}H_{53}N_6O_5$ $[M]^+$ 721,4072, found: 721.4070.

3.S3 Supplementary References

1 Dvir H, Harel M, McCarthy AA, Toker L, Silman I, Futerman AH & Sussman JL (2003) X-Ray Structure of Human Acid-β-Glucosidase, the Defective Enzyme in Gaucher Disease. *EMBO Rep* **4**, 704–709.

2 Kabsch W (2010) XDS. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 125–132.

3 Winter G. (2010) Xia2: An Expert System for Macromolecular Crystallography Data Reduction. *J. Appl. Crystallogr.* **43**, 186–190.

4 Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AG, McCoy A, McNicholas SJ, Murshudov GN, Pannu NS, Potterton EA, Powell HR, Read RJ, Vagin A & Wilson KS (2011) Overview of the CCP4 Suite and Current Developments. *Acta Crystallogr Sect D Biol Crystallogr* **67**, 235–242.

5 Vagin A, Teplyakov A (2010) Molecular Replacement with MOLREP. *Acta Crystallogr Sect D Biol Crystallogr* **66**, 22–25.

6 Lieberman RL, Wustman BA, Huertas P, Powe A C, Pine CW, Khanna R, Schlossmacher MG, Ringe D & Petsko GA (2007) Structure of Acid β-Glucosidase with Pharmacological Chaperone Provides Insight into Gaucher Disease. *Nat Chem Biol* **3**, 101–107.

7 Murshudov GN, Vagin, AA & Dodson EJ (1997) Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallogr Sect D Biol Crystallogr* **53**, 240–255.

8 Emsley P & Cowtan K (2004) Coot: Model-Building Tools for Molecular Graphics. *Acta Crystallogr Sect D Biol Crystallogr* **60**, 2126–2132.

9 Lebedev AA, Young P, Isupov MN, Moroz OV, Vagin AA, & Murshudov GN (2012) JLigand: A Graphical Tool for the CCP4 Template-Restraint Library. *Acta Crystallogr Sect D Biol Crystallogr* **68**, 431–440.

10 Agirre J, Iglesias-Fernández J, Rovira C, Davies GJ, Wilson KS & Cowtan KD (2015) Privateer: Software for the Conformational Validation of Carbohydrate Structures. *Nat Struct Mol Biol* **22**, 833–834.

11 McNicholas S, Potterton E, Wilson KS & Noble MEM (2011) Presenting Your Structures: The CCP4mg Molecular-Graphics Software. *Acta Crystallogr Sect D Biol Crystallogr* **67**, 386–394.

12 Li KY, Jiang J, Witte MD, Kallemeijn WW, Van Den Elst H, Wong CS, Chander SD, Hoogendoorn S, Beenakker TJM, Codée JDC, Aerts JM, van der Marel GA & Overkleeft HS (2014) Synthesis of Cyclophellitol, Cyclophellitol Aziridine, and Their Tagged Derivatives. *European J Org Chem* **2014**, 6030–6043.

CHAPTER 4

Biological characterization of activity-based probes for α-glucosidase and β-glucuronidase

Manuscript published as:

(1) Jiang J, **Kuo CL**, Wu L, Franke C, Kallemeijn WW, Florea BI, van Meel E, van der Marel GA, Codée JDC, Boot RG, Davies GJ, Overkleeft HS & Aerts JMFG (2016) Detection of active mammalian GH31 α-glucosidases in health and disease using in-class, broad-spectrum activity-based probes. *ACS Central Science* **2**, 351–358.

(2) Wu L, Jiang J, Jin Y, Kallemeijn WW, **Kuo CL**, Artola M, Dai W, van Elk C, van Eijk M, van der Marel GA, Codée JDC, Florea BI, Aerts JMFG, Overkleeft HS & Davies GJ (2017) Activity-based probes for functional interrogation of retainingβ-glucuronidases. *Nat Chem Biol* **8**, 867–873.

ABSTRACT

The lysosomal glycosidases α-glucosidase (GAA) and β-glucuronidase (GUSB) catalyze fragmentation of glycogen and glycosaminoglycans, respectively.. Their deficiency due to genetic aberrations leads to the lysosomal storage disorder (LSD) Pompe disease (GAA deficiency) and mucopolysaccharidosis type VII (MPSVII; Sly Syndrome), both severe disease conditions with still unmet clinical needs. To generate tools to facilitate biochemical and biomedical research on these two enzymes, activity-based probes targeting either enzyme (GH31 α-glucosidases or GH2 β-glucuronidase) were developed and evaluated for their mechanism-of-action, inhibition potency, and target selectivity. Both sets of ABPs successfully label their respective target enzyme class in activity-based manner. They are nanomolar irreversible inhibitors with fast inhibition kinetics. At acidic labeling pH and in cells, the lysosomal acid α-glucosidase GAA or β-glucuronidase GUSB are labeled by their respective ABPs, as revealed by both in-gel fluorescence and chemical proteomics. In the case of both GAA and GUSB ABPs, some minor out-of-class targets are observed. By altering the labeling pH or pre-incubating samples with previously generated inhibitors for the off-targets, GAA or GUSB can be visualized with their corresponding ABPs simultaneously with the off-target glycosidases. The new ABPs for GH31 α-glucosidases and GH2 β-glucuronidase successfully expand the chemical toolbox for lysosomal glycosidases and should find future applications in studies of fundamental and applied LSD research.

4.1 Introduction

Deficiency in lysosomal proteins due to genetic mutations underlies lysosomal storage disorders (LSDs) in which undegraded substrate accumulate in the lysosomes of particular cell types in the affected patients, and lead to a complex cascade of downstream pathological events[1]. An important class of deficient proteins in LSDs are the lysosomal glycosidases (E.C. 3.2.1), accounting for a third of known LSD enzymopathies. Glycosidases can be broadly classified by their substrate preference on the glycon part (e.g. sugar configuration, exo-acting *vs* endo-acting), mechanism of hydrolysis (e.g. retaining, inverting, substrate assisted), or evolutionary similarity by amino acid sequence and folding by the glycoside hydrolase (GH) family system.[1]

Man expresses at least four exo-acting retaining α-glucosidases that belong to the glycoside hydrolase family 31 (GH31): the α-subunit of ER α-glucosidase II (GANAB) (E.C. 3.2.1.20 and 3.2.1.84) which plays a role in quality control of N-glycoprotein synthesis, the two intestinal enzymes maltase-glucoamylase (MGAM) (E.C. 3.2.1.20) and sucrase-isomaltase (SI) (E.C. 3.2.1.20 and 3.2.1.48) degrading dietary glycosides, and the lysosomal acid α-glucosidase (GAA) (E.C. 3.2.1.20) degrading glycogen.[2] The human GAA is formed as a 110 kDa proenzyme, that upon mannose-6-phosphate dependent transport to the lysosome is processed into the 76 kDa and 70 kDa active forms.[3, 4] Mutations in GAA gene may lead to Glycogen storage disease type II, also known as Pompe disease.[5] This relatively common LSD (1:40,000 live birth worldwide)[6] is characterized by the intra-lysosomal accumulation of undegraded glycogen, which causes progressive muscle weakness in the heart and skeletal muscles and eventually affects the liver and central nervous system.[7] Classification of the disease is based on the age of onset, and patients classified into the infantile form (4 to 8 months of age) usually do not survive before one year of age.[8] Later onset forms (juvenile and adult forms) are characterized by a slower progressive decrease in muscle strength, firstly in the legs and then the trunks and the arms, and finally to fatality through respiratory failure.[9] Severity of Pompe disease correlates with the extent of lost GAA activity.[10] The disease is currently treated by chronic intravenous injection of recombinant GAA (rGAA, alglucosidase alpha, Myozyme®), which impressively delays the fatal

symptoms in infantile patients.[11] However, due to poor correction in muscle cell pathology, it only has limited response in adult patients.[12]

Another LSD-relevant lysosomal glycosidase, the retaining exo-β-glucuronidase (GUSB), belongs to the GH2 family.[2] It catalyzes the hydrolysis of terminal β-glucuronic acid from heparan sulfates, dermatan sulfates, and chondroitin-4,6-sulfate—polysaccharides presented in glycosaminoglycans (GAGs) which are found in the extracellular matrix that serve structural and signaling functions. Deficiency of GUSB activity due to hereditary genetic mutations forms the basis of another LSD known as mucopolysaccharidosis type VII (MPSVII), or Sly syndrome.[13] Affected individual are characterized by lysosomal accumulation of GAGs in many tissues, which ultimately affects the central nervous system and causes hepatosplenomegaly and bone dysplasia.[14] No targeted therapy is currently available for this rare disease. Notably, there exists another lysosomal enzyme in human that has exhibits β-glucuronidase activity—heparanase (HPSE). This is an endo-acting retaining β-glucuronidase, and belongs to the GH79 family, that functions in the turnover of heparan sulfate.[2] Interestingly, the bacterial GH79 homologue from *Acidobacterium capsulatum,* that exhibits similar protein fold, is an exo-acting retaining β-glucuronidase.[15] HPSE activity is implicated in disease-related processes such as inflammation, tumor metastasis and angiogenesis.[16, 17]

In the past, the mechanism-based irreversible β-glucosidase inhibitor cyclophellitol (CP) has been used as a scaffold to generate potent and specific inhibitors and ABPs towards the lysosomal glucocerebrosidase (GBA).[1] The configurational and functional isomers of CP-aziridine have been found to inactivate and label other classes of retaining exo-glycosidases that follow the formal Koshland double-displacement mechanism.[1] Both GAA and GUSB are retaining exo-glycosidases (**Fig. 4.1A**),[18] which makes them possible candidates for activity-based probe labeling (**Fig. 4.1B**). It is therefore hypothesized that the α-glucose and β-glucuronic acid configured cyclophellitol epoxides and aziridines may render inhibitors and ABPs targeting GH31 retaining exo-α-glucosidases (GAA and GANAB) and GH2 retaining exo-β-glucuronidase (GUSB), respectively. In this chapter the characterization of these newly generated compounds is presented in detail, which provides a basis for future laboratory, pre-clinical, and

clinical applications.

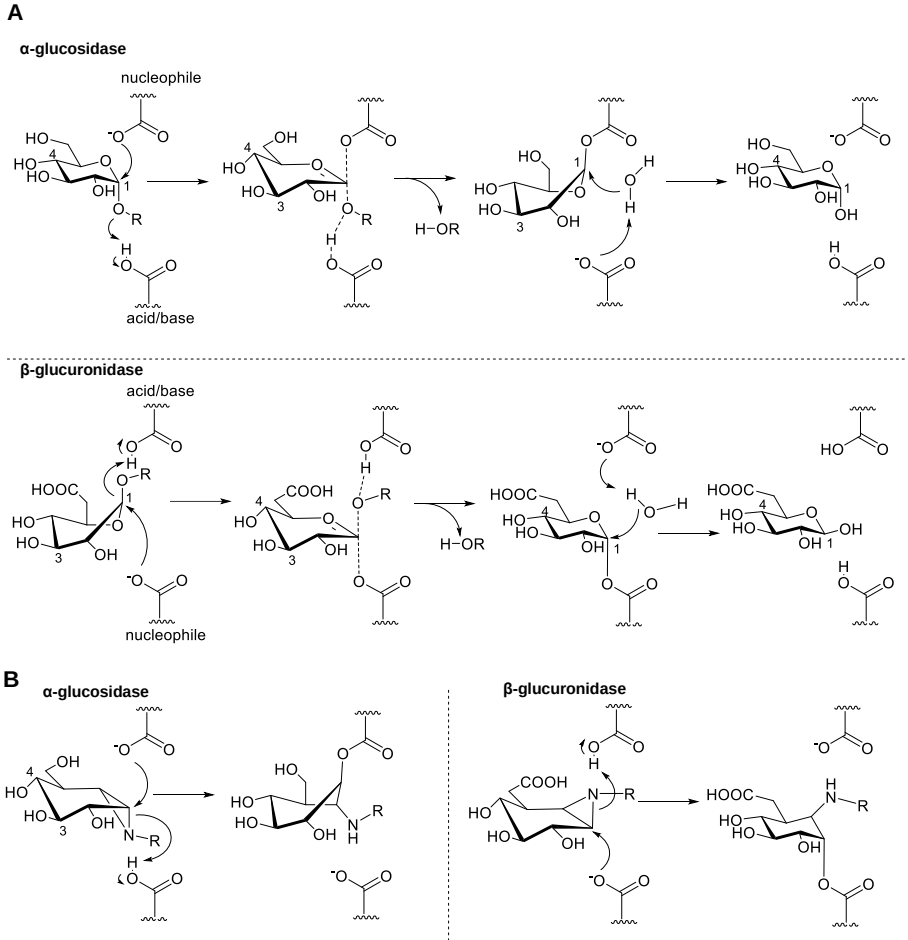


Figure 4.1. Reaction mechanisms of compounds used in this chapter. A) Koshland double-displacement mechanism employed by retaining α-glucosidase and β-glucuronidase. B) Proposed inhibition mechanism of cyclophellitol aziridine-based α-glucosidase inhibitor (left) and β-glucuronidase inhibitor. The corresponding ABPs may be obtained by grafting a detection group (fluorophore or biotin) at the R position.

4.2 Results and Discussion

4.2.1 Synthesis of α-glucose- and β-glucuronic acid-configured inhibitors and ABPs

Compounds were synthesized at the at the Department of Bio-organic Synthesis, Leiden University. The α-glucose-configured-configured cyclophellitol (**1**) and cyclophellitol aziridines (**2**) were synthesized (**Fig. 4.2**) from a cyclohexane diol (**17**) (see **Scheme 4.S1** for synthetic strategy, and reference Jiang et al.[19] for synthetic and characterization details of compounds). Compound **2** was alkylated to give compound **3**, which was further reacted with different substituted alkynes by copper (I)-catalyzed [2+3] azide-alkyne click reaction to afford ABPs **4**–**7** (**Fig. 4.2**, **Scheme 4.S1**).

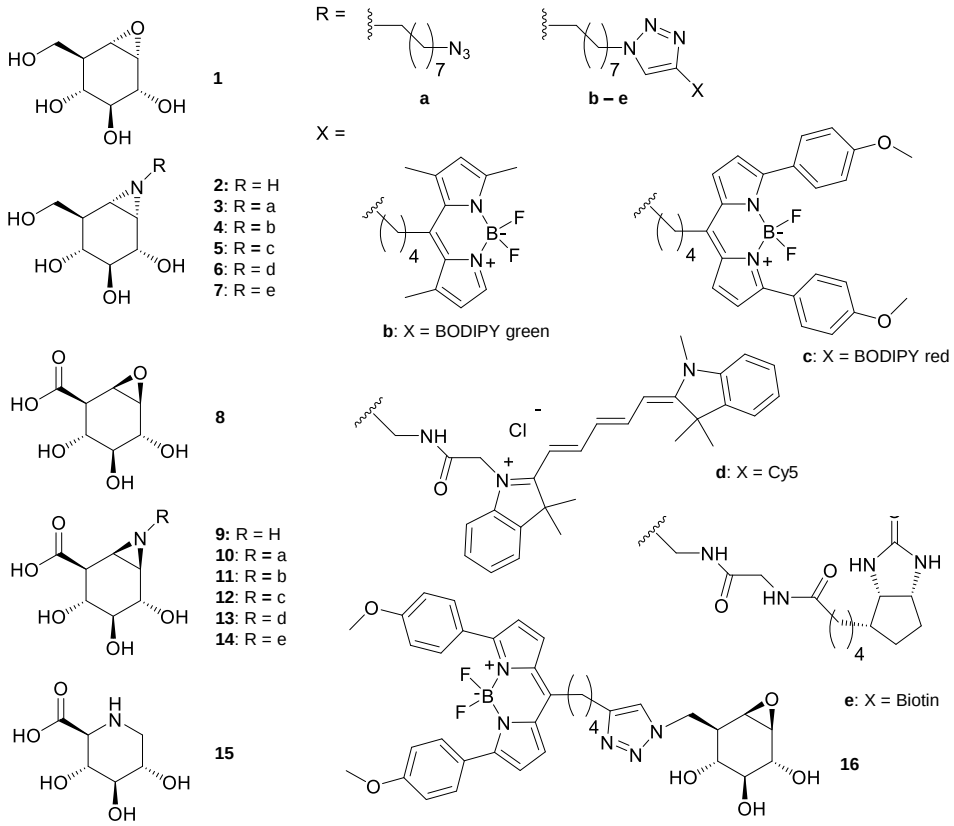


Figure 4.2. Structures of compounds used in this chapter.

The β-glucuronic acid-configured cyclophellitol **8** and cyclophellitol aziridine **9** (**Fig. 4.2**) were also synthesized from **17** (**Scheme 4.S2**). The synthesis of *N*-alkyl-substituted β-glucuronic acid-configured cyclophellitol aziridines utilized the *N*-alkyl-substituted β-glucose-configured cyclophellitol aziridine[20] as a precursor, and upon selective oxidation at the primary alcohol, the β-glucuronic acid-configured compound **10** was obtained (**Fig. 4.2**, **Scheme 4.S3**). Further click reaction afforded the *N*-alkyl-substituted β-glucuronic acid-configured ABPs **11–14** (**Fig. 4.2**). The iminosugar **15** (**Fig. 4.2**) was synthesized as a putative GUSB inhibitor, according to the strategy depicted in **Scheme 4.S4**.

4.2.2 *In vitro* inhibition and labeling of enzyme targets by compounds **1–14**

At the start of the evaluation of the newly generated compounds, the therapeutic recombinant human GAA (rGAA) and recombinant GH79 retaining exo-β-glucuronidase (AcGH79) from the bacteria *Acidobacterium capsulatum* were used as enzyme sources. The inhibitory activity of compounds **1–7** towards rGAA, and compounds **8–14** on AcGH79, were determined by measuring the apparent IC_{50} value of each compound at 30 min incubation time using fluorogenic 4-methylumblliferyl (4-MU)-glycoside substrate assays. As can be seen in **Fig. 4.3A** (top panel), the value for the α-glucose configured cyclophellitol epoxide **1** was at a low-micromolar apparent range (14.6 ± 1.6 μM) towards rGAA; the values for cyclophellitol aziridine **2** and its alkyl derivative **3** were at mid-nanomolar range (apparent IC_{50} = 30–50 nM), and that the values for ABPs **4–7** were at high-nanomolar range (apparent IC_{50} = 200–800 nM). For the β-glucuronic-acid-configured compounds **8–14**, all exhibited nanomolar apparent IC_{50} values towards AcGH79. The *N*-alkyl-substituted compound **10** and ABPs **11–14** were the most potent, reaching low- to sub-nanomolar potency (**Fig. 4.3A**, lower panel). The aziridine compound **9** and the epoxide compound **8** were also nanomolar inhibitors towards AcGH79 (**Fig. 4.3A**, lower panel), but slightly less reactive compared to the N-alkylated compounds.

Next, the formation of covalent enzyme-ABP complex was examined by fluorescent gel-based ABPP assay. ABP **4** or ABP **11**, both containing a green BODIPY fluorophore, were incubated with their respective target enzyme (rGAA or AcGH79) at 1 μM concentration for 30

minutes. Afterwards, samples were denatured and subjected to SDS-PAGE. The subsequent fluorescent scanning on the wet slab-gels containing the rGAA samples (**Fig. 4.3B**, left panel) revealed a clear green band corresponds to the expected molecular weight for rGAA (~100 kDa) in the ABP **4**-treated lane (second lane of the gel). This band is absent in samples without ABP treatment or pre-incubated with compounds **1–3** and **7**, the GAA substrate 4-MU-α-glc and maltose, and the protein-denaturing agent SDS. In samples pre-incubated with ABP **5** or **6**, red

A

Compounds	IC_{50} (nM) (Myozyme)
1 (JJB307)	14,600 ± 1,600
2 (CF21)	37.8 ± 3.40
3 (CF22)	51.7 ± 2.30
4 (JJB382)	459 ± 38.4
5 (JJB347)	383 ± 22.10
6 (JJB383)	269 ± 7.10
7 (JJB384)	767 ± 3.50

Compounds	IC_{50} (nM) (AcGH79)
8 (JJB400)	33.4 ± 3.1
9 DW222	4.6 ± 0.03
10 (JJB355)	0.1 ± 0.01
11 (JJB395)	0.6 ± 0.2
12 (JJB391)	0.8 ± 0.2
13 (JJB392)	1.1 ± 0.1
14 (JJB397)	0.4 ± 0.02

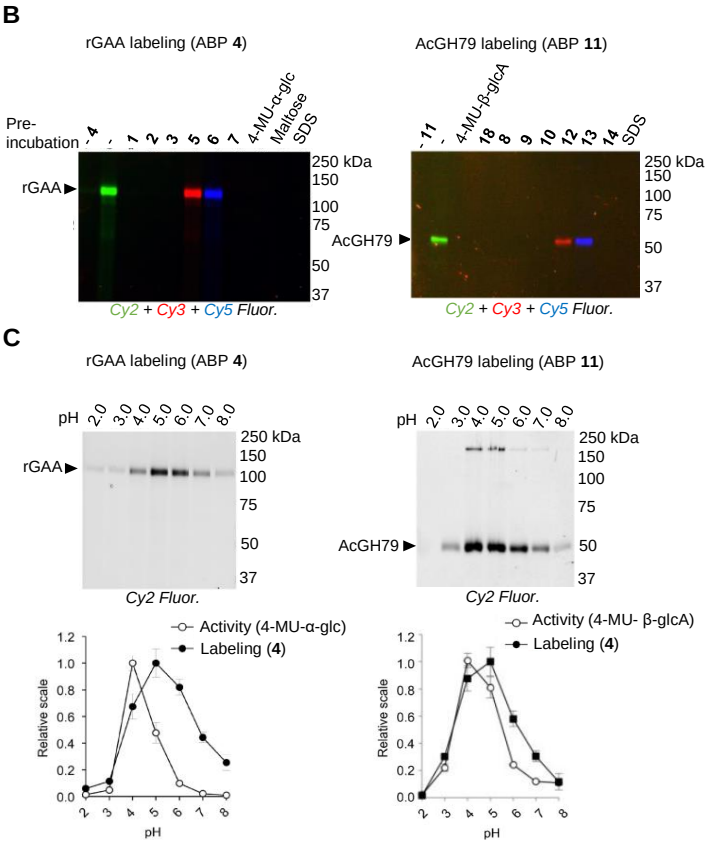


Figure 4.3. In vitro inhibition and labeling on recombinant enzymes by compounds 1–14. A) Apparent IC_{50} values for compounds **1–7** towards rGAA (top) and for compounds **8–14** towards AcGH79 (bottom). B) Fluorescent gel-based cABPP by compound **1–7** towards rGAA (left) and by compounds **8–14** towards AcGH79 (right). Enzymes were pre-incubated with inhibitors shown on the top of gels, and labeled with ABP **4** or **11**. C) pH-dependent labeling of ABP **4** (left) or ABP **11** (right) towards corresponding enzymes. The intensity of the labeled bands was quantified, and compared with the activity measured by fluorogenic substrates at each concentration (lower panels).

or Cy5 (shown in blue) fluorescence replaced the green fluorescence, suggesting the labeling of these ABP on rGAA competed away for the one from ABP **4**. Similarly, the gel containing the AcGH79 samples (**Fig. 4.3B**, right panel) showed a comparable labeling pattern. A green fluorescent band at the expected molecular weight of AcGH79 (~50 kDa) was observed in the ABP **11**-treated sample (second lane), and absent (or replaced by fluorescence of other colors) in samples pre-incubated with compounds **8–10** and **12–14** for AcGH79), known substrate (4-MU-β-glcA), putative competitive enzyme inhibitor (compound **15**), or 2 % (w/v) SDS. These results suggested that the labeling of ABP **4** or **11** towards their respective enzyme target occurred at the active site. To further evaluate whether the labeling by these ABPs occurred in an activity-based manner, ABP labeling on both enzymes was examined for the pH-dependence. For this, enzymes were labeled by ABPs across a range of pH (**Fig. 4.3C**, top panels), and the intensity of ABP labeling were compared to the measured enzymatic activity across the same pH range. The comparison revealed that the pH optimum of 5.0 for ABP labeling differed only slightly with that for substrate hydrolysis at pH 4.0 (**Fig. 4.3C**, lower panels).

Next, the kinetics for ABP labeling was examined. Both ABP **4** and **11** showed time-dependent labeling on their respective enzyme target, and reached saturation during 10–15 minutes of incubation time (**Fig. 4.S1**). The inhibition kinetic parameters were measured for the β-glucuronic acid-configured compound **8–14** on AcGH79 by fluorogenic substrate assay in a continuous setup (see **4.S2 Supporting note**), and the results revealed that beside the epoxide **8**, all compounds had pseudo-first order rate constant (k_i/K_I) in the rage of 3.5–25 μM^{-1} min^{-1} (**Table 4.1**; see **Fig. 4.S2** for processing curves and k_{obs} vs [I] data, **Fig 4.S3** for Michaelis-Menton plot). These values were comparably high (i.e. fast in labeling) to the reported values for the previously generated cyclophellitol ABPs towards rGBA (0.8–25 μM^{-1} min^{-1}).[21]

Table 4.1. Kinetic parameters for compounds 8–14 towards AcGH79. Error ranges = SD from triplicate sets of experiment.

Compounds	8 JJB400	9 DW222	10 JJB355	11 JJB395	12 JJB391	13 JJB392	14 JJB397
k_i/K_I (μM^{-1} min^{-1}) (AcGH79)	0.49 ± 0.05	3.5 ± 0.2	18.8 ± 0.7	25.0 ± 0.7	18.2 ± 0.9	14.0 ± 0.8	5.5 ± 0.2

4.2.3 Structural analysis of compounds in complex with recombinant enzymes

The mechanistic aspect of inhibition/labeling by the compounds were further examined by using protein crystallography studies (University of York), which could reveal structural insights how compounds bind in the active-site pocket of an enzyme, such as the formation of covalent bond, or the conformation that compounds adopt. For this purpose, crystals of the bacterial GAA homologue CjAgd31B from *Cellvibrio japonicus* or the bacterial GUSB homologue AcGH79 (wild-type or the E287Q nucleophile mutant) were soaked with either compound **3** or compound **10**, before subjecting to X-ray crystallography and structural determination. The results revealed that compound **3** adopted a $^{1}S_{3}$ conformation in the active site of CjAgd31B (**Fig. 4.4A**), identical with the known conformation adopted by other GAA substrates[22, 23] (**Fig. 4.1A**, third structure in the top panel). It also showed a covalent glycosidic bond between anomeric carbon (C1) of **3** and the nucleophile Asp412 of the enzyme, confirming the predicted mechanism-based inhibition/labeling (**Fig. 4.1**).

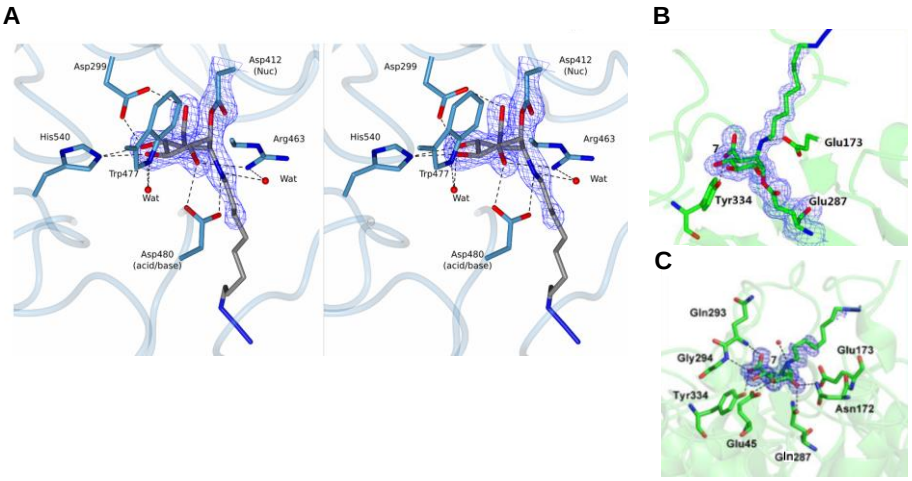


Figure 4.4. Structures of inhibitor-enzyme complex determined by protein crystallography. A) Spectroscopic view of compound **3** in the active site of CjAgd31B. B) Structure of compound **10** in the active site of AcGH79. C) Structure of compound **10** in the active site of the E287Q catalytic nucleophile mutant of AcGH79.

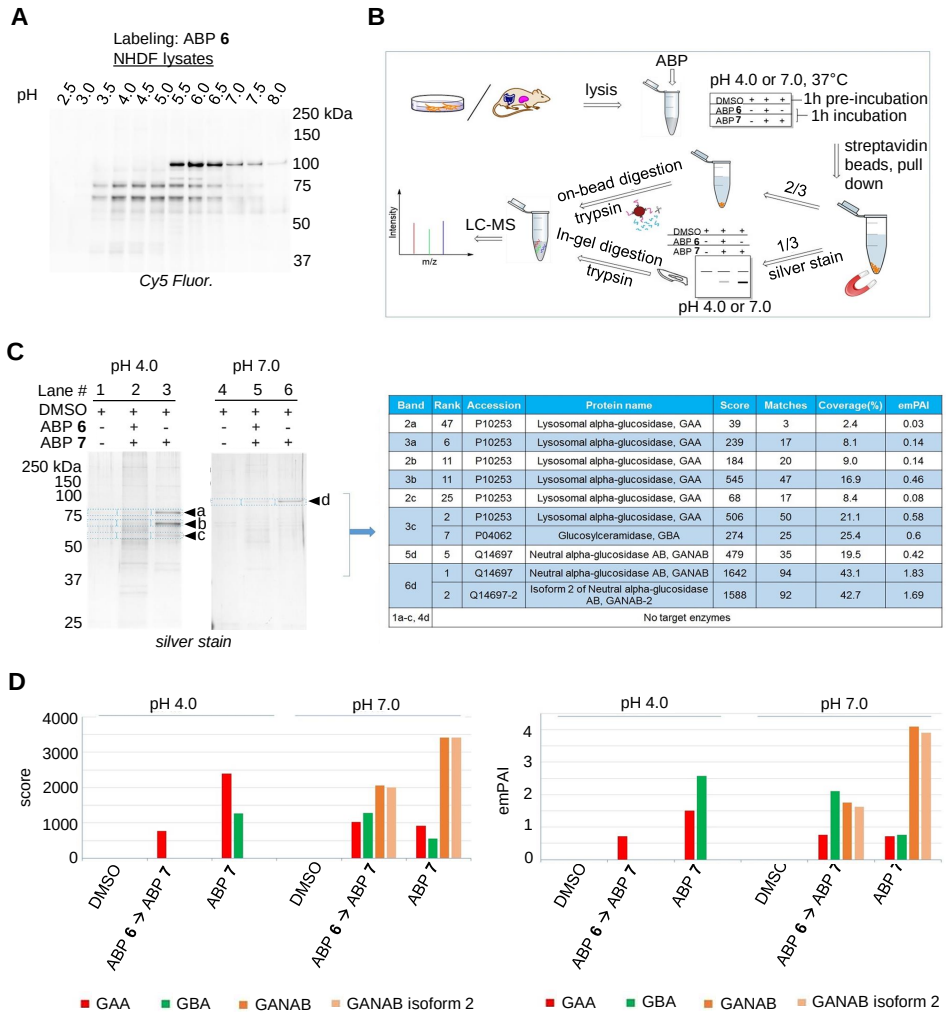
The structure of compound **10** coupled to AcGH79 revealed a $^{4}C_{1}$ conformation adopted by the compound, as well as the presence of the trans-diaxial C1-nucleophile glycosidic linkage.

On the other hand, the structure of **10** in complex with the E287Q catalytic nucleophile mutant of AcGH79 revealed a $^4H_3$ conformation adopted by the compound, and the absence of the glycosidic linkage between the compound and the enzyme (**Fig. 4.4C**). This demonstrated that the compound is a transition state-mimic of the substrates of β-glucuronidase, and that the inhibition/labeling followed the expected $^4H_3 \rightarrow {}^4C_1$ itinerary from the transition-state to the covalent substrate-enzyme intermediate (**Fig. 4.1**).[15]

4.2.4 Glycosidase targets of ABPs **4**–**7** in complex biological samples

Whether the generated ABPs are applicable to detecting GAA or GUSB in complex biological samples was tested next. For GAA, lysates of human fibroblasts were incubated with 1 μM of the Cy5 ABP **6** for 30 minutes at 37°C at pH range from 2.5 to 8.0, and subjected to gel-based fluorescence detection. The results clearly showed two lower bands (around 75 kDa) prominently presented at pH from 3.5 to 6.5, and a higher band (around 100 kDa) seen at pH from 5.5 to 7.5 (**Fig. 4.5A**). These two lower bands had acidic pH optimum and molecular weights characteristic of the two mature lysosomal GAA forms (70 and 76 kDa), while the higher band has a neutral pH optimum despite havening similar molecular weight of pro-GAA (110 kDa). The amounts of pro-form of GAA should be low in cells, and it is—as well as the mature forms—active at acidic pH. Hence, this higher running band is likely to be another target, likely the α-subunit of the ER-residing α-glucosidase II (GANAB), which possesses a similar molecular weight and a more neutral pH optimum. To identify these ABP targets, chemical proteomic method was employed. In this method, the biotin-containing ABP **7** was incubated with fibroblasts lysates at 5 μM concentration at either pH 4.0 or 7.0, with or without pre-incubation of ABP **4** (5 μM , as a negative control). Samples were subsequently denatured, affinity-enriched for biotin-containing molecules using streptavidin beads, subjected to either on-bead tryptic digestion or in-gel tryptic digestion, and peptides were analyzed by nano-LC-MS-based protein identification (**Fig. 4.5B**). For the in-gel digestion protocol, biotin-enriched samples were subjected to SDS-PAGE and silver stained. The resulting bands were excised out and subjected to in-gel tryptic digestion and proteomic identification. The silver stain results (**Fig. 4.5C**) showed similar band patterns with those by the labeling of Cy5 ABP **6** (**Fig. 4.5A**): two prominent bands at around 75 kDa were observed in the sample labeled with the biotin ABP **7** at pH 4.0, and one 100 kDa band were detected in the sample labeled with the biotin

ABP **7** at pH 7.0 (**Fig. 4.5C**). These band were absent from samples not treated with ABP **7**, and from samples pre-incubated with the green BODIPY ABP **4** (**Fig. 4.5C**), suggesting they were truly protein targets that were labeled with the biotin-containing ABP **7** and affinity-enriched. Subsequent proteomic identification revealed that GAA was the



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 4.5. Identification of ABP targets in human fibroblast lysates by chemical proteomics. A) Gel-based fluorescent labeling of ABP **6** in lysates of normal human dermal fibroblasts (NHDF) at different pHs. B) Schematic representation of the chemical proteomic workflow. C) Silver stain of gels containing samples enriched for biotin-containing molecules (left), and the resulting list of identified glycosidases at the indicated positions (lane 1–6, band a–d). D) Comparison of relative abundance of glycosidases by Mascot search engine peptide score (left) and exponentially modified Protein Abundance Index (emPAI) (right).

prominent glycosidase in position a–c (around 75–50 kDa) in the gel containing samples labeled at pH 4.0, while GANAB was exclusively identified in position d (around 100 kDa) in the gel containing samples labeled at pH 7.0 (**Fig. 4.5C**, right). Interestingly, the lysosomal glucocerebrosidase (GBA) was also identified as a target at position c of the pH 4.0 gel (**Fig. 4.5C**). To verify the activity of compound **1–7** towards GBA, compounds were subjected to enzymatic assay with the therapeutic recombinant human GBA (rGBA, Imiglucerase, Cerezyme®). The results demonstrated that, besides the epoxide compound **2** and the aziridine compound **3**, all other N-alkylated compounds inhibited GBA at high nanomolar to low micromolar range. They were, however, less reactive towards rGBA compared to rGAA (**Table. 4.S1**).

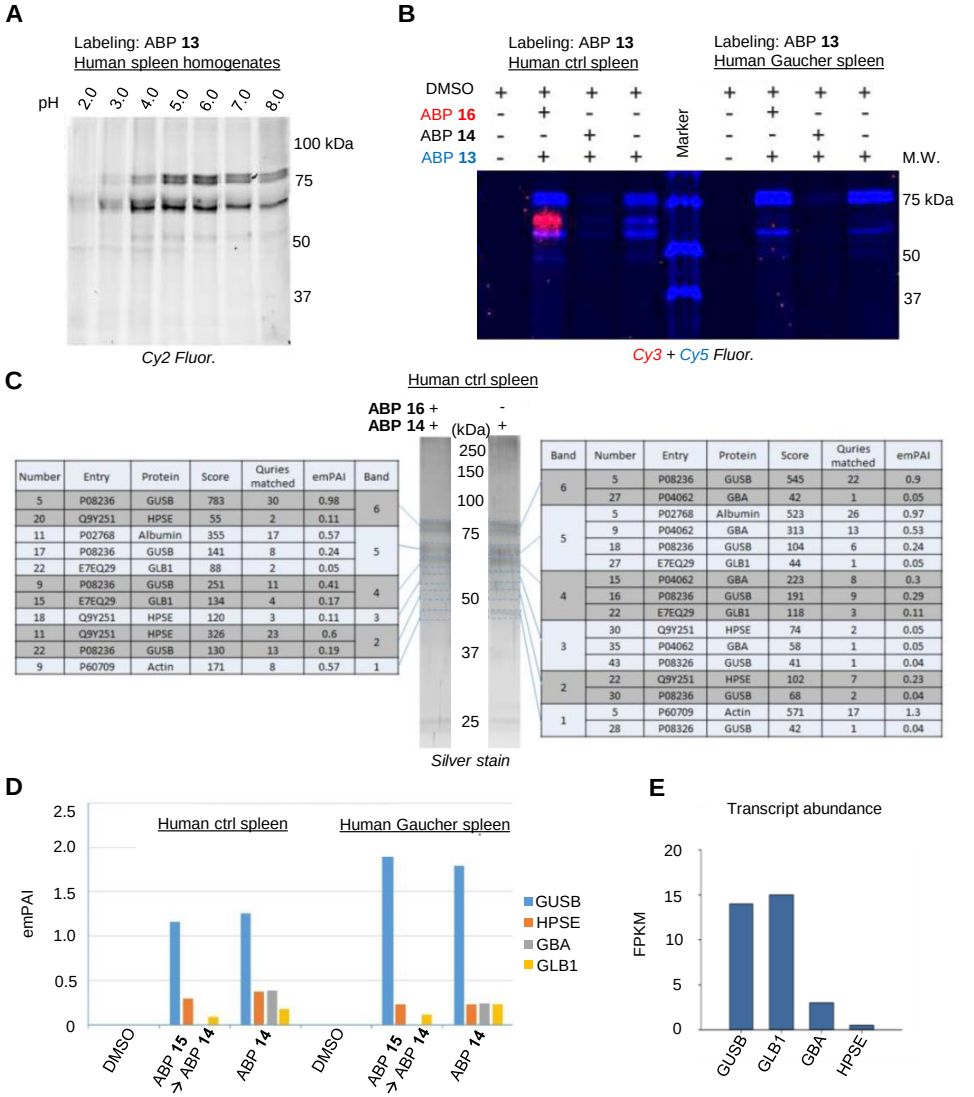
In a parallel experiment, the affinity enriched samples were trypsin-digested in the solution (i.e. on-bead digest), and the resulting peptides were comparably analyzed by nano-LC-MS. The identified glycosidases in each sample were compared for relative abundance using the Mascot search score (**Fig. 4.5D**, left) and the exponentially modified Protein Abundance Index (emPAI) (**Fig. 4.5D**, right), which is based on comparing the coverage of matched peptides for each protein against the search database. The results were in excellent agreement with the previous finding, in which GAA and GBA were the only two ABP glycosidase targets identified at pH 4.0, and that GANAB (and its isoform 2) were additionally identified at pH 7.0.

The chemical proteomic protocol was also applied to tissue extracts from mouse. Labeling by ABP **7** in mouse liver homogenates also enabled identification of GAA and GBA at pH 4.0, and GANAB at pH 7.0 (**Fig. 4.S4**), while labeling by ABP **7** in mouse intestine homogenates identified three additional enzymes: maltase-glucoamylase (MGAM, predominantly at pH 4.0), sucrase-isomaltase (SI, at both pH 4.0 and 7.0), and lactase-phlorizin hydrolase (LPH, or protein Lct, at pH 7.0) (**Fig. 4.S5**). Thus, compound **7** labeled all four mammalian GH31 α-glucosidases, while also reactive towards the retaining exo-β-glucosidases GBA and LPH.

4.2.5 Glycosidase targets of ABPs **11–14** in complex human proteome

Identification of targets by the β-glucuronic acid configured ABPs was performed in

homogenates of human spleen, a material which is predicted to express high GUSB level,



Number	Entry	Protein	Score	Quries matched	emPAI	Band
5	P08236	GUSB	783	30	0.98	6
20	Q9Y251	HPSE	55	2	0.11	
11	P02768	Albumin	355	17	0.57	5
17	P08236	GUSB	141	8	0.24	
22	E7EQ29	GLB1	88	2	0.05	
9	P08236	GUSB	251	11	0.41	4
15	E7EQ29	GLB1	134	4	0.17	
18	Q9Y251	HPSE	120	3	0.11	3
11	Q9Y251	HPSE	326	23	0.6	2
22	P08236	GUSB	130	13	0.19	
9	P60709	Actin	171	8	0.57	1

Band	Number	Entry	Protein	Score	Quries matched	emPAI
6	5	P08236	GUSB	545	22	0.9
	27	P04062	GBA	42	1	0.05
5	5	P02768	Albumin	523	26	0.97
	9	P04062	GBA	313	13	0.53
	18	P08236	GUSB	104	6	0.24
	27	E7EQ29	GLB1	44	1	0.05
4	15	P04062	GBA	223	8	0.3
	16	P08236	GUSB	191	9	0.29
	22	E7EQ29	GLB1	118	3	0.11
3	30	Q9Y251	HPSE	74	2	0.05
	35	P04062	GBA	58	1	0.05
	43	P08326	GUSB	41	1	0.04
2	22	Q9Y251	HPSE	102	7	0.23
	30	P08236	GUSB	68	2	0.04
1	5	P60709	Actin	571	17	1.3
	28	P08326	GUSB	42	1	0.04

Figure 4.6. Identification of ABP targets in human spleen homogenates. A) Gel-based fluorescent labeling of ABP **6** at different pHs. B) Competitive ABPP of GBA-specific ABP **14** or biotin ABP **16** with the Cy5 ABP **13** in control (left) or Gaucher (right) spleen homogenates. C) List of glycosidases identified at each gel position by chemical proteomics. Albumin and actin were included based on their absence in the control samples. D) emPAI values of identified glycosidases. E) Transcriptomic abundance of the identified glycosidases in human spleen, based on data from FANTOM5 consortium experiment E-MTAB-3358. FPKM, Fragments sequenced per kilobase transcript per million reads.

according to the Expression Atlas transcriptome data base.[24] Gel-based fluorescence detection

in samples treated with 100 nM Cy5 ABP **13** at various pH showed four prominent bands between around 80–60 kDa, all of which have labeling optimum at around pH 5.0 (**Fig. 4.6A**). Human GUSB is processed from the 78 kDa pro-form into a 60 kDa and a 18 kDa fragment in the lysosomes[25, 26], while two other isoforms of 70 (isoform 2) and 58 kDa (isoform 3) were known. It is likely that three of the bands were GUSB isoforms, and one was another protein. Thus, the two higher bands were likely the pro-form of GUSB and the isoform 2, and the lowest band could be a mixture of mature GUSB and isoform 3. The identity of the second-lowest band was hypothesized as the mature form of GUSB, or GBA, which has a similar molecular weight as appeared on SDS-PAGE.[21] To verify the identity of this band, gel-based competitive ABPP was employed, in which the biotin ABP **14** or the GBA-specific red BODIPY ABP **16**[21] (**Fig. 4.2**) were pre-incubated with the sample at pH 5.0, prior to the addition of the Cy5 ABP **13**. Indeed, in sample pre-treated with ABP **16**, a red band was observed, in place of the second-lowest band labeled by ABP **13**, suggesting that the second-lowest band was GBA (**Fig. 4.6B**, left). In support of this, ABP **13** labeling was also performed in homogenates of human Gaucher spleen that is deficient in GBA, where this band was absent from samples labeled with either ABP **13** or ABP **16** (**Fig. 4.6B**, right). In samples pre-treated with ABP **14,** all labeling were abrogated, which indicated that both ABP **13** and **14** could label at least GBA, and possibly on GUSB.

Similar to the target identification for the α-glucose configured cyclophellitol aziridine compound **7**, chemical proteomics were applied to further identify the protein targets of the biotin ABP **14**. Human spleen homogenates were pre-incubated with or without the GBA-specific ABP **16**, and labeled with ABP **14** at 10 μM concentration. Upon denaturation and affinity-purification of biotinylated molecules, samples were again divided for in-gel tryptic digestion and on-bead tryptic digestion. As expected, in-gel digestion protocol resulted in the identification of GUSB and GBA at the indicated positions (**Fig. 4.6C**) in samples treated with ABP **14**, while pre-incubation of ABP **16** abrogated the identification of GBA by both in-gel (**Fig. 4.6C**, left) and on-bead digestion protocol (**Fig. 4.6D**). Also identified were the human GH79 retaining endo-β-glucuronidase heparanase (HPSE, at around 50 kDa) and the lysosomal

acid β-galactosidase (GLB1, at around 60 kDa). However, the abundance of GBA, HPSE, and GLB1 were all much lower compared to the level of GUSB, based on emPAI value (**Fig. 4.6D**). Analysis of transcript abundance of the identified proteins in human spleen (based on data from the FANTOM5 consortium experiment E-MTAB-3358) showed that GUSB and GLB1 transcripts are much more abundant compared to the GBA and HPSE transcripts. In combination, these results suggested that GUSB is the major target by ABP **14**, and that GLB1 is a minor target. GBA and HPSE could also be major targets of ABP **14**, given sufficient expression levels in the labeled tissue.

4.2.5 GAA and GUSB labeling in intact cells

The applicability of the ABPs to inhibit and label their target glycosidases in intact cells were next investigated. For inhibition, confluent human fibroblasts were treated with the ABPs in a range of different concentrations for 2 hours, washed and lysed, and the lysates were measured for the relative loss of GAA or GUSB activity. For most of the α-glucose-configured compounds, it was observed that their *in situ* apparent IC_{50} values towards the endogenous GAA in intact cells were similar to their *in vitro* apparent IC_{50} values towards rGAA (mid- to high nanomolar) (**Fig. 4.7A**, left). The epoxide compound **1** and the biotin ABP **7** were the only two exceptions, showing no inhibition at the highest applied concentration (50 or 10 μM). This was in contrast to the β-glucuronic acid configured compounds, in which only the red BODIPY ABP **12** and the Cy5 ABP **13** showed moderate activity towards GUSB (apparent *in situ* IC_{50} value = 1.7 and 1.8 μM, respectively), and other compounds did not show inhibition up to the highest applied concentration (15 μM) (**Fig. 4.7A**, right). This suggested that installation of the biotin moiety on the α-glucose-configured cyclophellitol aziridine, as well as the presence of carboxylic acid at the C8 position of the cyclophellitol aziridine, are detrimental for *in situ* inhibition of the two lysosomal glycosidases.

To visualize the *in situ* labeling, the α-glucose-configured Cy5 ABP **6** was treated to cells at 100 nM concentration for various time periods (10 minutes to 4 hours), and the fluorescence in cell lysates was visualized by gel-based ABPP. Another set of cells were identically treated, but was lysed in lysis buffer containing an excess of green BODIPY ABP **4** (10 μM), to exclude the

possible labeling of GAA by ABP **6** during cell lysis. The gel-based ABPP analysis showed three prominent bands corresponding to the molecular weight of GANAB (~100 kDa) and GAA (76 and 70 kDa), as identified previously, and that the labeling intensity increased with incubation time (**Fig. 4.7B**, left). Addition of the green BODIPY ABP **4** resulted in a band with overlapping molecular weight as the putative GANAB, and another lower band at around 55 kDa that was not labeled by the Cy5 ABP **6**—likely to be GBA according to its appeared molecular weight.[21] Because the addition of ABP **4** in the lysis buffer did not change the labeling pattern of the *in situ* treated Cy5 ABP **6**, it can be concluded that ABP **6** labeling on GANAB and GAA was most likely taken place in intact cells. The higher band labeled by ABP **4** in the lysis buffer was also

A

Compounds	IC_{50} (nM) GAA in NHDF	Compounds	IC_{50} (nM) GUSB in NHDF
1 (JJB307)	> 50,000	8 (JJB400)	> 15,000
2 (CF21)	71.2 ± 9.0	9 DW222	> 15,000
3 (CF22)	75.8 ± 10.9	10 (JJB355)	> 15,000
4 (JJB382)	502 ± 4.1	11 (JJB395)	> 15,000
5 (JJB347)	562 ± 352	12 (JJB391)	1,700 ± 600
6 (JJB383)	42.2 ± 4.9	13 (JJB392)	1,800 ± 400
7 (JJB384)	> 10,000	14 (JJB397)	> 15,000

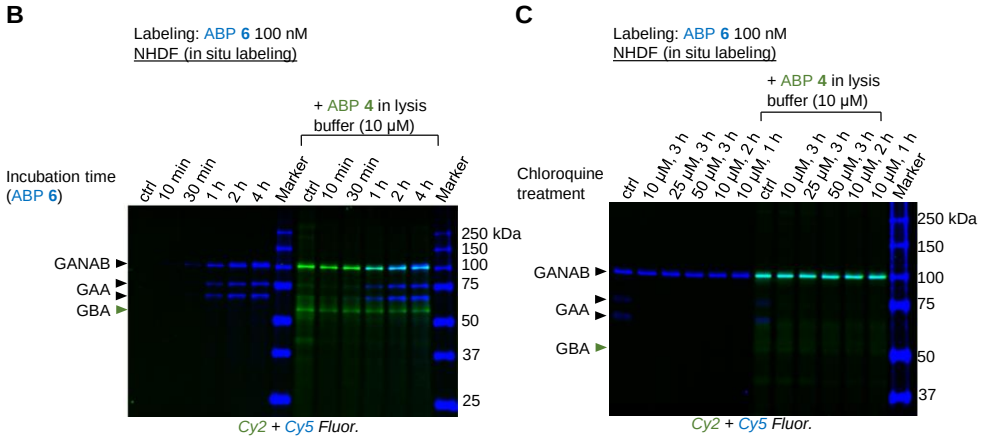


Figure 4.7. Glycosidase inhibition and labeling in intact cells by compounds 1–14. A) *In situ* IC_{50} values towards GAA by compounds **1**–**7** (left) or GUSB by compounds **8**–**14** (right). B) *In situ* labeling by ABP **6**, with or without an excess of ABP 4 in the lysis buffer. C) Effect of chloroquine pre-incubation on

ABP **6** labeling in cells.

putatively assigned to GANAB, as the *in situ* treated ABP **6** at 100 nM seemed not to reach saturate labeling of that protein at 4 h incubation time. This is in contrast to the labeling on the GAA bands, which were saturated at around 2 h incubation. In another experimental setup, chloroquine was added at various dosage and incubation time to the cells, prior to ABP **6** incubation and the subsequent lysis with or without an excess of ABP **4**. Chloroquine is a chemical agent that is known to raise lysosomal pH and thereby disrupt lysosomal functions[27], and its presence could inhibit GAA activity in the lysosome and therefore abrogate ABP labeling. As predicted, it was clearly observed that the labeling by ABP **6** on the putative GAA bands was abrogated in the presence of chloroquine, while that the labeling on GANAB was not affected (**Fig 4.7C**). This result confirmed the ability of ABP **6** to inhibit and label the lysosomal GAA in intact cells in an activity-based manner, and demonstrated the applicability of this ABP in monitoring GAA activity in cells.

4.2.6 Visualizing GAA deficiency in patient fibroblasts

As ABP **6** was shown to allow specific detection of active GAA in gel-based assays, it's applicability for GAA labeling in a diagnostic setup was examined. Fibroblasts from healthy donors or Pompe disease patients were cultured, and cell lysates were labeled with 1 μM ABP **6** for 30 minutes, at either pH 4.0 or 7.0. The results showed the marked loss of GAA labeling by ABP **6** at pH 4.0 in samples of Pompe patients (**Fig. 4.8A**), but not on those of GANAB at pH 7.0 (**Fig. 4.8B**). This loss of active GAA in patient samples correlated with the loss of GAA proteins, as determined by the subsequent Western blot detection of GAA from the same gel (**Fig. 4.8A** and **B**, lower panels). This result therefore demonstrated the potential application of ABP **6** in laboratory diagnosis of Pompe disease.

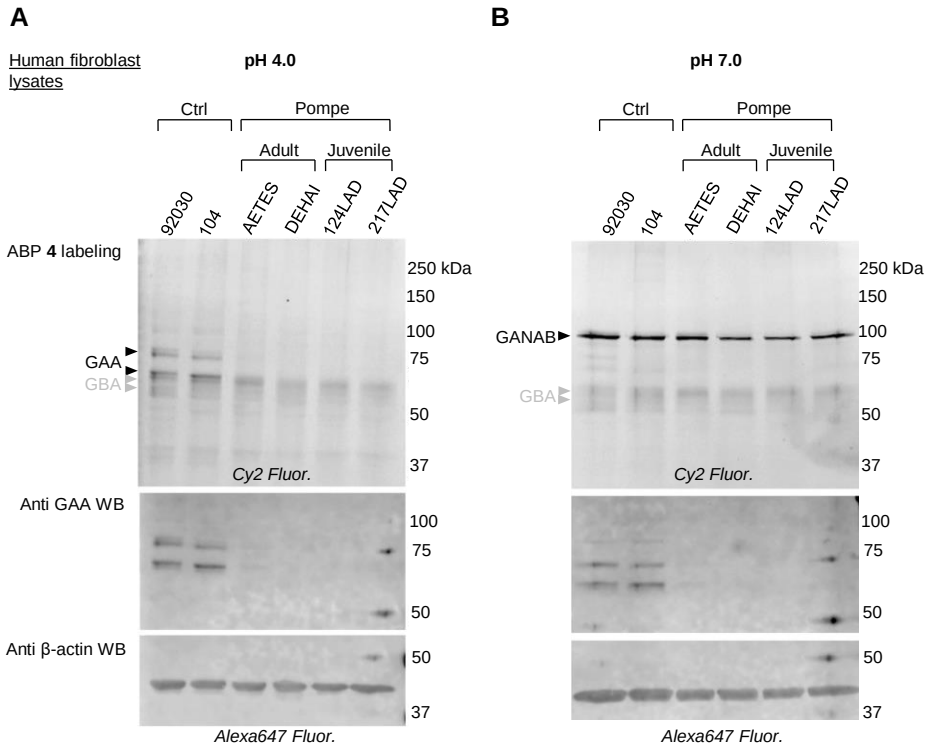


Figure 4.8. ABP 6 labeling in fibroblasts lysates of control or Pompe patients. A) Labeling at pH 4.0. B) Labeling at pH 7.0. The same gels were subjected to subsequent Western blot detection of GAA and β-actin (protein loading control).

4.3 Discussion

Activity-based protein profiling has emerged in the past decade allowing unprecedented studies on active enzymes such as quantitative visualization and enrichment. Glycosidases possessing the Koshland double-displacement mechanism have been shown to be applicable for activity-based protein profiling through the appropriate designing of small molecules grafted with an electrophilic trap and exhibit sufficient substrate mimicry. Functionalization and tuning the configuration of cyclophellitol epoxide and cyclophellitol aziridine was found to offer a viable strategy for generating ABPs against a number of retaining glycosidases.[20, 21, 28–30] In this chapter, the α-glucose- and β-glucuronic acid-configured cyclophellitol and (*N*-tagged) cyclophellitol aziridines are examined for their biological activities, including mechanism of

action in recombinant enzymes and selectivity in complex proteomes and in intact cells.

Both sets of compounds inhibit recombinant enzymes in a pH- and time-dependent manner, and generally with high (nanomolar) potency. Protein crystallographic studies reveals that they react with the nucleophile of their respective target enzyme, and adopting identical conformation as the true substrates. Targets of these ABPs have been identified qualitatively by gel-based fluorescent detection, and quantitatively by chemical proteomics. For α-glucose-configured cyclophellitol aziridine ABPs in human and mouse samples these encompass four GH31 retaining exo-α-glucosidases, including GAA, GANAB, and the intestinal dietary enzymes MGAM and SI. Intriguingly, the retaining β-glucosidases GBA and LPH are also identified as targets by chemical proteomics. As these two enzymes exhibit no α-glucosidase activity towards artificial 4-MU substrates, the labeling mechanism of ABP **14** towards the retaining β-glucosidases remains to be explored in the future. The β-glucuronic acid-configured cyclophellitol aziridine ABP labels the intended target GUSB with exceptionally high potency. It nevertheless shows additional minor targets in GBA, HPSE, and GLB1. Labeling of the GUSB ABPs of HPSE, as well as pro-HPSE, has been further demonstrated by chemical proteomics and protein crystallography in the published article based on the results from this chapter.[31] Because GUSB and HPSE have distinct molecular weights, the β-glucuronic acid configured cyclophellitol aziridine ABPs should also be viable tools to specifically detect HPSE activity in complex biological samples using an SDS-PAGE setup. Altogether, the *in vitro* labeling results suggest that tuning the glycon configuration of the cyclophellitol aziridine ABP is a viable strategy to target the intended glycosidase class, but labeling of other out-of-class glycosidases can in some cases still occur. Nevertheless, by identifying the off-target glycosidases, specific visualization of target glycosidases, such as GAA or GUSB, can be achieved by varying the labeling pH, ABP concentration, and/or pre-incubation of known ABPs/inhibitors for the off-target glycosidases, as demonstrated in this chapter.

It is shown that the α-glucose-configured cyclophellitol aziridine compounds are cell-permeable, and that the Cy5 ABP **6** at 100 nM specifically labeled GAA over GBA in cells. The β-glucuronic acid configured cyclophellitol aziridines are less effective in intact cells, as they

exhibit at least 1000-fold reduction in apparent potency in cells when compared to *in vitro* assay. Nevertheless, the red BODIPY ABP **12** and Cy5 ABP **13** should be applicable for *in situ* GUSB labeling, given enough incubation time and prior incubation with a GBA inhibitor or ABP. In the future, the ability of the ABPs to label active lysosomal enzymes in lysates/homogenates and in intact cells should assist important studies such as those on processing and distribution of therapeutic enzymes, and the efficacy of other LSD therapeutic approaches that aim to promote the lysosomal activity of deficient enzymes at target tissues/organs like gene therapy and chaperone therapy.

Finally, in this chapter the applicability of the GAA ABP in the laboratory diagnosis of Pompe disease is demonstrated. Common biochemical diagnosis methods for Pompe disease rely on GAA ativity measurement using the 4-MU-α-glc substrate assay from lysates of cultured fibroblasts, muscle biopsy, urine, or blood samples.[32] The biggest drawbacks for the substrate assay is the low detection dynamic range (typically in the order of two)[33], due to background activity from 4-MU, and the presence of MGAM in blood samples that requires the addition of acarbose to relatively block MGAM activity in the assay[34]. Mass spectrometry-based activity assays have improved the detection sensitivity[33, 35], but suffer from the associated laborious procedures. Therefore, it is envisioned that the gel-based ABPP described in this chapter could circumvent these shortcomings, offering rapid and sensitive visualization of active GAA on a simple and rapid SDS-PAGE-based assay. Other possible GAA activity readouts using ABPs are fluorescent microscopy (Chapter 1, this thesis) and fluorescence-assisted cell sorting (FACS).[21] However, in both cases, the concomitant labeling on GANAB in intact cells would introduce undesirable background fluorescence, which could complicate data interpretation. This issue can be overcome by pre-incubation with a GANAB-selective or GAA-selective active-site blocker. Although no such inhibitors are currently known, using the ABPs in a gel-based competitive ABPP assay (cABPP) or in a fluorescent polarization assay[36] would assist identification of such inhibitors. Studies are currently pursued in this direction (Daniel Lahav, Department of Bio-organic Synthesis, Leiden University).

In conclusion, true mechanism-based inhibitors and ABPs for GAA and GUSB have been

successfully developed and extensively characterized. They enable specific detection of their target enzymes in complex biological samples. The applicability of the ABP in GAA detection has been demonstrated in intact cells, and its value for diagnosis of Pompe disease has been demonstrated. The novel ABPs should find future applications in the study of lysosomal glycosidases in both fundamental and applied research in the context of Pompe disease, Sly Syndrome, and even pathologies involving heparanase activity.

4.4 Experimental procedures

4.4.1 Materials

Chemicals were obtained from Sigma-Aldrich (St. Louis, MO, USA), if not otherwise indicated. Trypsin and was commercially available from Promega (Madison, WI, USA). Recombinant GAA was obtained from Sanofi Genzyme (Cambridge, MA, USA). Fibroblasts were obtained with consent from donors. Pompe patients were diagnosed on the basis of reduced GAA activity. Fibroblast cell lines were cultured in HAMF12-DMEM medium (ThermoFisher Invitrogen™, Waltham, MA, USA) 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 (KPi) lysis buffer (25 mM K_2HPO_4/KH_2PO_4, pH 6.5, supplemented with protease inhibitor cocktail (EDTA-free, Roche, Basel, Switzerland) and 0.1 % (v/v) Triton X-100) via homogenization with silent crusher S equipped with Typ 7 F/S head (30 rpm x 1000, 3 × 7 sec) on ice. Protein concentration in lysates was determined with BCA Protein Assay Kit (ThermoFisher Pierce™). Lysates were stored in small aliquots at –80 °C until use.

4.4.2 Cloning, expression, and purification of bacterial enzymes (University of York)

CjAgd31B expression and purification was carried out as previously described.[37] For AcGH79, the coding sequence of the AcGH79 gene was cloned into the pET28a (Novagen, Madison, WI, USA) expression vector with an N-terminal His6 tag. The E287N mutant was produced by polymerase chain reaction (PCR) using the primer 5'-CCTGACCCAAACGAATTC-3' (forward primer) and 5'-GAATTCGTTTGGGTCAGG-3' (reverse primer). Both the wild-type and mutant proteins were overexpressed in *Escherichia coli* strain BL21 (DE3) GOLD using LB medium. The transformed cells were grown at 37 °C in LB media containing 50 μg mL^{-1} kanamycin until the A600 nm reached 0.8. Expression of the recombinant proteins was induced by the addition of 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) for 12 h at 25 °C. The cells were harvested by centrifugation at 8000 g for 30 min and resuspended in 50 mL HEPES lysis buffer (20 mM HEPES, NaCl 200 mM, imidazole 5 mM, pH 7.0). After 20 min of

sonication and 30 min of centrifugation at 12000 g, the filtered supernatant containing His6-AcGH79 was loaded onto a His Trap column (GE Healthcare, Chicago, IL, USA), equilibrated with the lysis buffer. The column was washed with HEPES lysis buffer and the His6-AcGH79 protein was eluted with the same buffer with supplement of 400 mM imidazole over a gradient of 100 mL. The fractions containing the His6-AcGH79 were then loaded onto a Hiload 16/60 Superdex 75 column (GE Healthcare). The fractions containing the His6-AcGH79 were pooled and concentrated to the final concentration of 14.5 mg mL^{-1}.

4.4.3 Enzyme activity assays and *in vitro* IC_{50} measurements

The α-D-glucosidase activity of lysosomal α-D-glucosidase GAA was assayed in individual wells of medium-binding flat-bottomed black 96-well plates (Greiner, Kremsmünster, Austria) at 37 °C by incubating samples with 3.0 mM 4- methylumbelliferyl-α-D-glucopyranoside (4-MU-α-glc) as substrate in McIlvaine buffer (150 mM sodium citrate-Na_2HPO_4, supplemented with 0.01 % (w/v) NaN_3 as bacteriostatic)[38], supplemented with 0.1 % (w/v) BSA, at pH 5.0 for rGAA and at pH 4.0 for GAA in cell lysates or tissue homogenates. Activity of rGBA was measured using similar conditions but with 3.75 mM 4-methylumbelliferyl-β-D-glucopyranoside (4-MU-β-glc) as substrate at pH 5.2, supplemented with 0.1 % (v/v) Triton X-100 and 0.2 % (w/v) sodium taurocholate. Activity of AcGH79 and GUSB was performed with 2.5 mM 4-methylumbelliferyl-β-D-glucuronic acid (4-MU-β-glc) as substrate at pH 5.0, supplemented with 0.1 % (w/v) BSA. To determine the apparent *in vitro* IC_{50} value, recombinant enzymes (12.5 µL) were firstly pre-incubated with inhibitor dilutions (12.5 µL) for 30 min at 37 °C, prior to incubation with substrates (100 µL) for a further 30 min at 37 °C. The enzymatic reaction was quenched by adding 200 µL of Glycine-NaOH (1 M, pH 10.3), after which fluorescence of liberated 4-methylumbelliferyl was measured with a fluorimeter LS55 (Perkin Elmer, Waltham, MA, USA) at λ_{EX} = 366 nm and λ_{EM} = 445 nm. For experiments with varying pH, enzymes, inhibitors, and substrates were prepared in McIlvaine buffer at the indicated pH. All apparent IC_{50} values were determined from biological duplicates. Data was corrected for background fluorescence, then normalized to the untreated control condition and finally curve-fitted via one phase exponential decay function using Prism 5.0 (GraphPad Software, San Diego, CA, USA).

4.4.4 *In vitro* ABP labeling for recombinant enzymes

All enzymes were pre-incubated in McIlvaine buffer at pH 5.0 (or other indicated pH) on ice for 5 min. *In vitro* labeling on rGAA was performed by incubating 100 fmol rGAA with(out) prior incubation with inhibitors in McIlvaine buffer (pH 5.0 if not otherwise indicated) for 30 min at 37 °C, and ABP **4** labeling at 1 μM probe concentration for 30 min at 37 °C. For *In vitro* labeling on AcGH79, 200 fmol of the enzyme was incubated with(out) inhibitors with the identical conditions as those for rGAA, and ABP incubation was performed with 1 μM ABP **11** for 30 min at 37 °C. Inhibitor concentrations were: 10 μM for compounds **3**, **5**–**7**, **10**, and **12**–**14**; 100 μM for compound **1**–**2**, **8**, **9**, and **15**; 2.5 mM for maltose, 10 mM for 4-MU substrates, and 2 % (w/v) for sodium dodecyl sulfate (SDS). After ABP incubation, samples were denatured with 5x 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) by boiling for 5 min at 98 °C. Denatured samples were separated by SDS-PAGE using 7.5 % or 10% polyacrylamide gels, at 90 V for 30 min and 200 V for 50–70 min. Wet slab-gels were scanned for ABP-emitted fluorescence using a Bio-Rad ChemiDoc MP imager (Bio-Rad, Hercules, CA, USA) using Cy2 (λ_{EX} = 470 nm, bandpass 30 nm; λ_{EM} = 530 nm, bandpass 28), Cy3 (λ_{EX} = 530 nm, bandpass 28 nm; λ_{EM} = 605 nm, bandpass 50), and Cy5 (λ_{EX} = 625 nm, bandpass 30 nm; λ_{EM} = 695 nm, bandpass 55) channels. Detected fluorescence was quantified by ImageLab Software (Bio-Rad) and plotted by Prism 5.0 Software (GraphPad). For labeling kinetics, enzymes and ABPs were prepared as earlier described, and incubated at either 4°C or 37 °C for 2 to 60 min. Samples without ABP addition were used as the 0 min controls.

4.4.5 Measurement of inhibition kinetic parameters

Kinetic parameters for inhibition of AcGH79 were determined using a continuous method involving simultaneous exposure of the enzyme to substrates and inhibitor (see **Supporting Note 4.S2**; ref. [39]), using fluorogenic substrate assay. During the enzymatic reaction, the concentration of enzyme (AcGH79) and substrates were 0.26 nM and 2 mM, respectively. For each inhibitor, triplicate sets of eight 2 mL Eppendorf tubes were prepared. These were added

firstly with 8.13 µL inhibitor at various concentrations (diluted in DMSO, 200x of reaction concentration, control = DMSO) and secondly with 154.4 µL McIlvaine buffer (150 mM, pH 5.0) and 1300 µL substrate mixture (2.5 mM 4-MU-β-GlcA, pH 5.0, 0.1 % (w/v) BSA). The reaction tubes were pre-warmed on a thermoshaker at 37 °C for 10 min. For each inhibitor, the enzyme (AcGH79) was prepared by diluting with McIlvaine buffer in an Eppendorf tube to a volume of 1250 µL and 2.6 nM enzyme concentration. A negative control (enzyme blank) was prepared in a separate Eppendorf tube by substituting the enzyme with denatured enzyme (boiled at 98 ºC for 5 min). Both were pre-warmed on a thermoshaker at 37 °C for 10 min. Prior to the reaction, a black 96-well plate was loaded with 200 µL stop buffer (1 M glycine-NaOH, pH 10.3) in each well. The t = 0 samples were prepared by filling the first two columns of the plate with 12.5 µL enzyme (or enzyme blank, in the last row) and 112.6 µL of each from the substrate-inhibitor mixture (1 concentration per row, in duplicate wells). To start the reaction, 137.6 µL aliquots of the pre-warmed 26 nM enzyme was transferred to each but the last (containing DMSO dilution and substrate mixture) of the pre-warmed 2 mL tube containing the substrate-inhibitor mixture, with a time interval of 15 seconds between samples. 137.6 µL aliquots from the pre-warmed enzyme blank were similarly transferred to the last 2 mL tube. To ensure proper mixing of the components, during the addition of enzyme aliquots the 2 mL tubes were continuously incubated in the thermal shaker at 37 °C under constant shaking at 850 rpm. At t = 5, 10, 15, 30 and 60 min, duplicates of 125 µL aliquots from each of the reaction tubes were transferred to the 96-well plate that contained stop buffer, with a 15-second interval between each tube. The plate was vortexed shortly and 4-MU fluorescence was measured immediately. The observed pseudo-first order inactivation rate (k_{obs}) was calculated for each concentration of each inhibitor by fitting the data with the one-phase exponential association function: 4-MU fluorescence = A*(1–e^(−k_{obs}*t)) using GraphPad Prism. The obtained k_{obs} values for each concentration of each inhibitor were then plotted against the inhibitor concentration, and the resulting plots were fitted using a linear function, which gives the combined apparent inhibition parameter k_{inact}/K'_I as the slope. k_{inact}/K_I was derived from k_{inact}/K'_I by correcting for the presence of competing 4-MU-β-GlcA substrate, using the relationship $K'_I = K_I (1+[S]/K_M)$, where [S] = 2 mM and K_M = 18.2 µM (**Fig. 4.S3**). The k_{inact}/K_I

values from three independent sets measured for each inhibitor were used to calculate the average and SD of the k_{inact}/K_I value of each inhibitor.

4.4.6 Protein crystallography (University of York)

Protein crystals of CjAgd31B 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 compound **3** was obtained by soaking in mother liquor containing 5.0 mM compound for 2 h, before cryoprotecting in 2.0 M lithium sulfate, 0.10 M HEPES (pH = 7.0), 2 % (w/v) PEG 400, and flash freezing in liquid N_2 for data collection. AcGH79 was tested against a range of commercial crystallization screens. Well diffracting crystals of wild-type AcGH79 were obtained by the sitting-drop vapor-diffusion method at 20 °C using 0.8–1.2 M of 0.5:9.5 NaH_2PO_4:K_2HPO_4 (v/v) at a protein:well ratio of 700:500 nL. Crystals of AcGH79 E287Q mutant were obtained using 1.2–1.5 M of 1.0:9.0 NaH_2PO_4:K_2HPO_4 (v/v) at a protein:well ratio of 500:500 nL. Crystals typically appeared after 1 day. Protein complexes with compound **10** were prepared by supplying the drops containing crystals with 1.0 µL of 5.0 mM compound, in the precipitant solution freshly prepared, before soaking procedure for 1 h at 20 °C before fishing. Crystals of AcGH79 were cryoprotected using 2 M lithium sulfate before flash freezing in liquid N_2 for data collection. All data were collected at 100 K on beamline I04 of the Diamond Light Source UK. Data were processed manually using XDS31[40] and reduced using Aimless[41] (CjAgd31B), or with the xia2 pipeline[42] of the CCP4 software suite (AcGH79). Complex structures were solved by molecular replacement using MolRep,[43] before subsequent rounds of manual model building and refinement using Coot[44] and REFMAC5[45] respectively. Refinements were carried out using TLS determination of molecular motions.[46] Ligand coordinates were built using jLigand.[47] Ribbon and protein surface diagrams were generated using PyMOL. Crystal structure figures were generated using CCP4mg.[48] The data processing statistics and structure refinement are listed in **Table 4.S2**.

4.4.7 *In vitro* ABP labeling in cell lysates and tissue homogenates

All ABP incubation were performed at 37 °C for 30 min. For ABP **6** labeling in human fibroblast lysates, 10 µg total protein from the lysates was incubated with McIlvaine buffer at various pH in 10 µL of volume for 5 min on ice, followed by incubation with 5 µL ABP **6** (diluted with McIlvaine buffer at the matching pH) at a final ABP concentration of 1 µM. For ABP **13** labeling in human spleen homogenates, 20 µg total protein from the samples was incubated firstly with McIlvaine buffer at various pH and secondly with 5 µL ABP **13** (diluted with McIlvaine buffer at the matching pH) at a final ABP concentration of 100 nM, in the same volume and incubation time described for ABP **6**. For competitive ABPP in human spleen homogenates, samples (40 µg total protein) were diluted in 10 µL McIlvaine buffer (pH 5.0), pre-incubated with 2.5 µL ABP **14**, ABP **16,** or DMSO (diluted in McIlvaine buffer) at 100 nM ABP concentration during reaction. These were followed by incubation with 2.5 µL ABP **13** or DMSO (diluted in McIlvaine buffer), at 100 nM ABP **13** during incubation.

4.4.8 Chemical proteomics

For target identification of ABP **7**, human fibroblast lysates (3 mg total protein), mouse liver homogenates (10 mg total protein) or mouse intestine homogenates (6 mg total protein) were incubated at pH 4.0 or 7.0 in McIlvaine buffer, at 37 °C in 0.5 mL volume, with either (1) 0.1 % (v/v) DMSO for 2 h, (2) firstly with 5.0 µM ABP **4** for 1 h, followed by ABP **7** (5 µM for fibroblasts lysates and mouse intestine homogenates; 10 µM for mouse liver homogenates) for 1 h, or (3) ABP **7** (5 µM for fibroblasts lysates and mouse intestine homogenates; 10 µM for mouse liver homogenates) for 1 h. After incubation, samples were denatured by the addition of 125 µL 10 % (w/v) SDS, and boiling for 5 min at 100 °C. From here on, samples were prepared for pull-down with streptavidin beads as published earlier.[49] After pull-down procedure, the samples were divided, two-thirds for on-bead digestion and one-thirds for in-gel digestion. On-bead digestion samples were incubated with 400 µL of trypsin digestion buffer (100 mM Tris-HCl pH 7.8, 100 mM NaCl, 1.0 mM $CaCl_2$, 2% (v/v) acetonitrile (ACN) and 10 ng/µL trypsin) at 37 °C overnight under constant shaking. The supernatant was desalted using stage tips, followed by evaporation of ACN and dilution in 70 µL sample solution (H_2O/ACN/TFA,

95:3:0.1, v:v:v) for LC-MS analysis following previously described procedures.[49] For the in-gel digestion samples, peptides were eluted from the streptavidin resin by boiling at 100 °C with 30 µL of 1 x Laemmli buffer (supplemented with 10 µL biotin). The eluted proteins were separated on 0.75 mm 10 % protein gels by SDS-PAGE at 200 V for 1 h, and silver-stained using the SilverQuest kit (Thermo Fisher). The resulting protein bands were excised and cut into 1 mm^3 cubes with a sterile surgery knife, transferred into 1.5 mL Eppendorf tubes and treated with 30 µL of in-gel digestion buffer (10 mM NH_4HCO_3, 5 % (v/v) ACN, 1 mM $CaCl_2$, 10 ng/µL trypsin). The supernatant containing the trypsin-digested peptides was extracted, and desalted for LC-MS measurement following the published protocol.[49] For LC-MS measurement, all 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.[49] 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.[49] For target identification of ABP **14**, 3 mg total protein from human spleen or Gaucher spleen homogenates was incubated with either (1) 0.1% (v/v) DMSO, (2) firstly with 10 µM ABP **16** followed by 10 µM **14**, or (3) 10 µM **14**, each step taking 30 min at 37 °C, in a total volume of 0.5 mL McIlvaine buffer at pH 5.0. Subsequent steps were identical to the method describes for target identification of ABP **7**.

4.4.9 Determination of *in situ* apparent IC_{50} values

The *in situ* IC_{50} value was determined by incubating fibroblast cell lines expressing wild-type GAA and GUSB, 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 KPi lysis buffer. *In situ* apparent IC_{50} values for each compound were determined by using enzymatic assay method described earlier, with GAA activity measured at pH 4.0 to avoid possible readout of GANAB activity.

4.4.10 *In situ* ABP labeling

Cultured fibroblasts were cultured in 6-well plates to confluency, and treated with 100 nM ABP **6** diluted in culture medium, at 1 mL medium per well, for 10 to 240 min. For chloroquine treatment experiment, cultured fibroblasts were incubated with chloroquine (10–50 μM, for 1 h up to 3 h) and subsequently with ABP **6** (100 nM, for 3 h) in medium. After washing 3 times with PBS, cells were scraped from culture dishes in the presence of KPi lysis buffer (50 μL per well, with or without 10 μM ABP **4**), collected in Eppendorf tubes, incubated on ice for 30 min for lysis, and stored at –80 °C until use. Lysates containing 5–10 μg total protein were subjected to SDS-PAGE and fluorescence scanning by methods described earlier. Labeling were visualized by fluorescent detection using Bio-Rad ChemiDoc MP Imager under the channels Cy5 for ABP **6** and Cy2 for ABP **4** 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 ABP **6** (100 nM, for 3 h) in medium. Cells were harvested with or without 10 mM ABP **4** with the above describe methods. Homogenates were denatured, resolved on SDS-PAGE and detected for ABP **6** labeling by fluorescent scanning.

4.4.11 GAA detection in fibroblast lysates of control and Pompe patients

10 μg protein from each lysate were labeled with 1 μM ABP **4** for 30 min at 37 °C at either pH 4.0 or 7.0, before subsequent SDS-PAGE and fluorescent scanning using a Typhoon FLA9500 Imager (GE Healthcare) using λ_{EX} = 635 nm laser and $\lambda_{EM} \geq$ 665 nm filter, and 100 mm as pixel size. After scanning, proteins on wet slab gels were transferred to PVDF membranes and blocked with 5 % (w/v) BSA in TBST. For GAA detection, the membranes were incubated firstly with mouse polyclonal anti GAA (kind gift from the Kornfeld lab, Washington University of St. Louis, MO, USA) and subsequently with goat anti mouse Alexa647 (Life Technologies, Carlsbad, CA, USA). Blots were scanned on the Typhoon FLA 9500 Imager using 633 nm laser and LPR filter, and 100 mm as pixel size. Rabbit anti β-actin (Cell Signaling, Danvers, MA, USA) and goat anti rabbit Alexa647 (Invitrogen) were used for loading control.

4.5 References

1 Kuo CL (2019) General introduction, this thesis.

2 Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V & Henrissat B (2008) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res* **37**(Database issue), D233–238.

3 Moreland RJ, Jin X, Zhang XK, Decker RW, Albee KL, Lee KL, Cauthron RD, Brewer K, Edmunds T & Canfield WM (2005) Lysosomal acid alpha-glucosidase consists of four different peptides processed from a single chain precursor. *J Biol Chem* **280**, 6780–6791.

4 Wisselaar HA, Kroos MA, Hermans MM, van Beeumen J & Reuser AJ (1993) Structural and functional changes of lysosomal acid alpha-glucosidase during intracellular transport and maturation. *J Biol Chem* **268**, 2223–2231.

5 Hers, HG (1963) alpha-Glucosidase deficiency in generalized glycogenstorage disease (Pompe's disease). *Biochem J* **86**, 11−16.

6 Ausems MG, Verbiest J, Hermans MP, Kroos MA, Beemer FA, Wokke JH, Sandkuijl LA, Reuser AJ & van der Ploeg AT (1999) Frequency of glycogen storage disease type II in The Netherlands: implications for diagnosis and genetic counselling. *Eur J Hum Genet* **7**, 713−716.

7 Reuser AJ, Kroos MA, Hermans MM, Bijvoet AG, Verbeet MP, Van Diggelen OP, Kleijer WJ & Van der Ploeg AT (1995) Glycogenosis type II (acid maltase deficiency). *Muscle Nerve Suppl* **3**, S61−69.

8 Kishnani PS, Hwu WL, Mandel H, Nicolino M, Yong F & Corzo D (2006) A retrospective, multinational, multicenter study on the natural history of infantile-onset Pompe disease. *J Pediatr* **148**, 671−676.

9 Winkel LP, Hagemans ML, van Doorn PA, Loonen MC, Hop WJ, Reuser AJ & van der Ploeg AT (2005) The natural course of non-classic Pompe's disease; a review of 225 published cases. *J Neurol* **252**, 875−884.

10 Umapathysivam K, Hopwood JJ & Meikle PJ (2005) Correlation of acid alpha-glucosidase and glycogen content in skin fibroblasts with age of onset in Pompe disease. Clin Chim Acta **361**, 191–198.

11 Kishnani PS, Corzo D, Nicolino M, Byrne B, Mandel H, Hwu WL, Leslie N, Levine J, Spencer C, McDonald M, Li J, Dumontier J, Halberthal M, Chien YH, Hopkin R, Vijayaraghavan S, Gruskin D, Bartholomew D, van der Ploeg A, Clancy JP, Parini R, Morin G, Beck M, De la Gastine GS, Jokic M, Thurberg B, Richards S, Bali D, Davison M, Worden M A, Chen YT & Wraith JE (2007) Recombinant human acid α-glucosidase: major clinical benefits in infantile-onset Pompe disease. *Neurology* **68**, 99−109.

12 van der Ploeg AT, Clemens PR, Corzo D, Escolar DM, Florence J, Groeneveld GJ, Herson S, Kishnani P S, Laforet P, Lake SL, Lange DJ, Leshner RT, Mayhew JE, Morgan C, Nozaki K, Park DJ, Pestronk A, Rosenbloom B, Skrinar A, van Capelle CI, van der Beek N A, Wasserstein M & Zivkovic S (2010) A randomized study of alglucosidase alfa in late-onset Pompe's disease. *N Engl J Med* **362**, 1396−1406.

13 Sly WS Quinton BA McAlister WH & Rimoin DL (1973) Beta glucuronidase deficiency: report of clinical, radiologic, and biochemical features of a new mucopolysaccharidosis. *J Pediatr* **82**, 249–257.

14 Tomatsu S, Montaño AM, Dung VC, Grubb JH & Sly WS (2009) Mutations and polymorphisms in GUSB gene in mucopolysaccharidosis VII (Sly Syndrome). *Hum Mutat* **30**, 511–519.

15 Michikawa M, Ichinose H, Momma M, Biely P, Jongkees S, Yoshida M, Kotake T, Tsumuraya Y, Withers SG, Fujimoto Z & Kaneko S (2012) Structural and biochemical characterization of glycoside hydrolase family 79 β-glucuronidase from Acidobacterium capsulatum. *J Biol Chem* **287**, 14069–14077.

16 Rivara S, Milazzo FM & Giannini G (2016) Heparanase: a rainbow pharmacological target associated to multiple pathologies including rare diseases. *Future Med Chem* **8**, 647–680.

17 Vlodavsky I, Beckhove P, Lerner I, Pisano C, Meirovitz A, Ilan N & Elkin M (2012) Significance of heparanase in cancer and inflammation. *Cancer Microenviron* **5**, 115–132.

18 Davies GJ & Williams SJ (2016) Carbohydrate-active enzymes: sequences, shapes, contortions and cells. *Biochem Soc Trans* **44**, 79−87.

19 Jiang J, Kuo CL, Wu L, Franke C, Kallemeijn WW, Florea BI, van Meel E, van der Marel GA, Codée JD, Boot RG, Davies GJ, Overkleeft HS & Aerts JM (2016) Detection of Active Mammalian GH31 α-Glucosidases in Health and Disease Using In-Class, Broad-Spectrum Activity-Based Probes. *ACS Cent Sci* **2**, 351−358.

20 Jiang J, Kallemeijn WW, Wright DW, van den Nieuwendijk AMCH, Rohde VC, Folch EC, van den Elst H, Florea BI, Scheij S, Donker-Koopman WE, Verhoek M, Li N, Schürmann M, Mink D, Boot RG, Codée JDC, van der

Marel GA, Davies GJ, Aerts JMFG & Overkleeft HS (2015) In vitro and in vivo comparative and competitive activity-based protein profiling of GH29 α-l-fucosidases. *Chem Sci* **6**, 2782−2789.

21 Witte MD, Kallemeijn WW, Aten J, Li KY, Strijland A, Donker-Koopman WE, van den Nieuwendijk AM, Bleijlevens B, Kramer G, Florea BI, Hooibrink B, Hollak CE, Ottenhoff R, Boot RG, van der Marel GA, Overkleeft HS, Aerts JM (2010) Ultrasensitive in situ visualization of active glucocerebrosidase molecules. *Nat Chem Biol* **6**, 907-913.

22 Davies GJ, Planas A & Rovira C (2012) Conformational Analyses of the Reaction Coordinate of Glycosidases. *Acc Chem Res* **45**, 308−316.

23 Speciale G, Thompson AJ, Davies GJ & Williams SJ (2014) Dissecting conformational contributions to glycosidase catalysis and inhibition. *Curr Opin Struct Biol* **28**, 1−13.

24 Petryszak R, Burdett T, Fiorelli B, Fonseca NA, Gonzalez-Porta M, Hastings E, Huber W, Jupp S, Keays M, Kryvych N, McMurry J, Marioni JC, Malone J, Megy K, Rustici G, Tang AY, Taubert J, Williams E, Mannion O, Parkinson HE & Brazma A (2014) Expression Atlas update--a database of gene and transcript expression from microarray- and sequencing-based functional genomics experiments. *Nucleic Acids Res* **42**(Database issue), D926-32.

25 Brot FE, Bell CE, Jr, & Sly WS (1978) Purification and properties of beta-glucuronidase from human placenta. *Biochemistry* **17**, 385–391.

26 Oshima A, Kyle JW, Miller RD, Hoffmann JW, Powell PP, Grubb JH, Sly WS, Tropak M, Guise KS & Gravel RA (1987) Cloning, sequencing, and expression of cDNA for human beta-glucuronidase. *Proc Natl Acad Sci USA* **84**, 685–689.

27 Poole B & Ohkuma S (1981) Effect of weak bases on the intralysosomal pH in mouse peritoneal macrophages. *J Cell Biol* **90**, 665−669.

28 Kallemeijn WW, Li KY, Witte MD, Marques AR, Aten J, Scheij S, Jiang J, Willems LI, Voorn-Brouwer TM, van Roomen CP, Ottenhoff R, Boot RG, van den Elst H, Walvoort MT, Florea BI, Codée JD, van der Marel GA, Aerts JM & Overkleeft HS (2012) Novel activity-based probes for broad-spectrum profiling of retaining β-exoglucosidases in situ and in vivo. *Angew Chem Int Ed Engl* **51**, 12529−12533.

29 Willems LI, Beenakker TJ, Murray B, Scheij S, Kallemeijn WW, Boot RG, Verhoek M, Donker-Koopman WE, Ferraz MJ, van Rijssel ER, Florea BI, Codée JD, van der Marel GA, Aerts JM & Overkleeft HS (2014) Potent and selective activity-based probes for GH27 human retaining α-galactosidases. *J Am Chem Soc* **136**, 11622−11625.

30 Jiang J, Beenakker TJ, Kallemeijn WW, van der Marel GA, van den Elst H, Codée JD, Aerts JM & Overkleeft HS (2015) Comparing Cyclophellitol N-Alkyl and N-Acyl Cyclophellitol Aziridines as Activity-Based Glycosidase Probes. *Chemistry* **21**, 10861−10869.

31 Wu L, Jiang J, Jin Y, Kallemeijn WW, Kuo CL, Artola M, Dai W, van Elk C, van Eijk M, van der Marel GA, Codée JDC, Florea BI, Aerts JMFG, Overkleeft HS & Davies GJ (2017) Activity-based probes for functional interrogation of retaining β-glucuronidases. *Nat Chem Biol* **13**, 867−873.

32 Kishnani PS, Steiner RD, Bali D, Berger K Byrne BJ, Case LE, Crowley JF, Downs S, Howell RR, Kravitz RM, Mackey J, Marsden D, Martins AM, Millington DS, Nicolino M, O'Grady G, Patterson MC, Rapoport DM, Slonim A, Spencer CT, Tifft CJ & Watson MS (2006) Pompe disease diagnosis and management guideline. *Genet Med* **8**, 267−88.

33 Kumar AB, Masi S, Ghomashchi F, Chennamaneni NK, Ito M, Scott CR, Turecek F, Gelb MH & Spacil Z (2015) Tandem Mass Spectrometry Has a Larger Analytical Range than Fluorescence Assays of Lysosomal Enzymes: Application to Newborn Screening and Diagnosis of Mucopolysaccharidoses Types II, IVA, and VI. *Clin Chem* **61**, 1363−71.

34 Zhang H, Kallwass H, Young SP, Carr C, Dai J, Kishnani PS, Millington DS, Keutzer J, Chen YT & Bali D (2006) Comparison of maltose and acarbose as inhibitors of maltase-glucoamylase activity in assaying acid alpha-glucosidase activity in dried blood spots for the diagnosis of infantile Pompe disease. *Genet Med* **8**, 302−306.

35 Lin, N, Huang, J, Violante, S, Orsini, JJ, Caggana, M, Hughes, EE, Stevens, C, DiAntonio, L, Chieh Liao, H, Hong, X, Ghomashchi, F, Babu Kumar, A, Zhou, H, Kornreich, R, Wasserstein, M, Gelb, M H & Yu, C (2017) Liquid Chromatography-Tandem Mass Spectrometry Assay of Leukocyte Acid α-Glucosidase for Post-Newborn Screening Evaluation of Pompe Disease. *Clin Chem* **63**, 842−851.

36 Lahav D, Liu B, van den Berg RJBHN, van den Nieuwendijk AMCH, Wennekes T, Ghisaidoobe AT, Breen I, Ferraz MJ, Kuo CL, Wu L, Geurink PP, Ovaa H, van der Marel GA, van der Stelt M, Boot RG, Davies GJ, Aerts JMFG & Overkleeft HS (2017) A Fluorescence Polarization Activity-Based Protein Profiling Assay in the

Discovery of Potent, Selective Inhibitors for Human Nonlysosomal Glucosylceramidase. *J Am Chem Soc* **139**, 14192−14197.

37 Larsbrink J, Izumi A, Hemsworth GR, Davies GJ & Brumer H (2012) *J Biol Chem* **287**, 43288−43299.

38 Kuo CL, van Meel E, Kytidou K, Kallemeijn WW, Witte M, Overkleeft HS, Artola ME & Aerts JM (2018) Activity-Based Probes for Glycosidases: Profiling and Other Applications. *Methods Enzymol* **598**, 217−235.

39 Baici A, Schenker P, Wächter M & Rüedi P (2009) 3-Fluoro-2,4-dioxa-3-phosphadecalins as inhibitors of acetylcholinesterase. A reappraisal of kinetic mechanisms and diagnostic methods. *Chem Biodivers* **6**, 261–282.

40 Kabsch W (2010) XDS. *Acta Crystallogr. D Biol Crystallogr* **66**, 125–132.

41 Evans PR & Murshudov GN (2013) How good are my data and what is the resolution? A*cta Crystallogr D Biol Crystallogr.* **69**, 1204–1214.

42 Winter G (2010) xia2: an expert system for macromolecular crystallography data reduction. *J Appl Crystallogr* **43**, 186–190.

43 Vagin A & Teplyakov A (2010) Molecular replacement with MOLREP. *Acta Crystallogr D Biol Crystallogr* **66**, 22–25.

44 Emsley P, Lohkamp B, Scott,WG & Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486–501.

45 Murshudov GN, Skubák P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F & Vagin AA (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* **67**, 355–367.

46 Painter J & Merritt EA (2006) Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr D Biol Crystallogr* **62**, 439–450.

47 Lebedev AA, Young P, Isupov MN, Moroz OV, Vagin AA & Murshudov GN (2012) JLigand: a graphical tool for the CCP4 template-restraint library. *Acta Crystallogr D Biol Crystallogr* **68**, 431–40.

48 McNicholas S, Potterton E, Wilson KS & Noble MEM (2011) Presenting your structures: the CCP4mg molecular-graphics software. *Acta Crystallogr D Biol Crystallogr* **67**, 386–394.

49 Li N, Kuo CL, Paniagua G, van den Elst H, Verdoes M, Willems LI, van der Linden WA, Ruben M, van Genderen E, Gubbens J, van Wezel GP, Overkleeft HS & Florea BI (2013) Relative quantification of proteasome activity by activity-based protein profiling and LC-MS/MS. *Nat Protoc* **8**, 1155–68.

APPENDIX

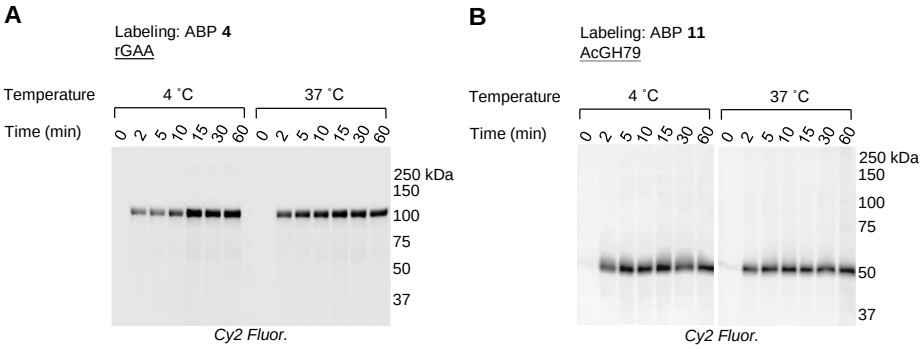


Figure 4.S1. Labeling kinetics of ABPs towards recombinant enzymes at 4 °C or at 37 °C. A) ABP **4** labeling towards rGAA. B) ABP **11** labeling towards AcGH79.

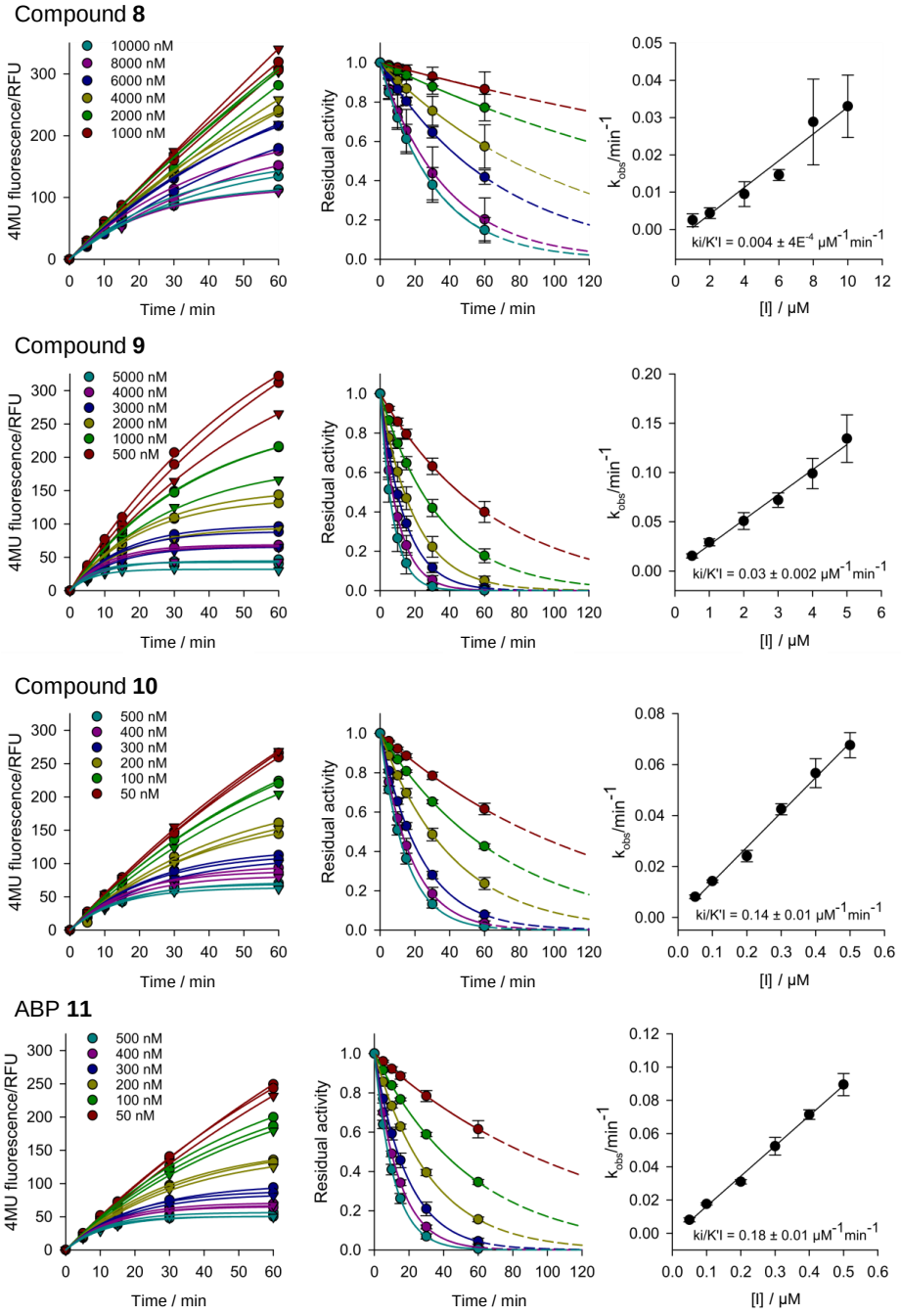


Figure 4.S2 (continued, 1 of 2).

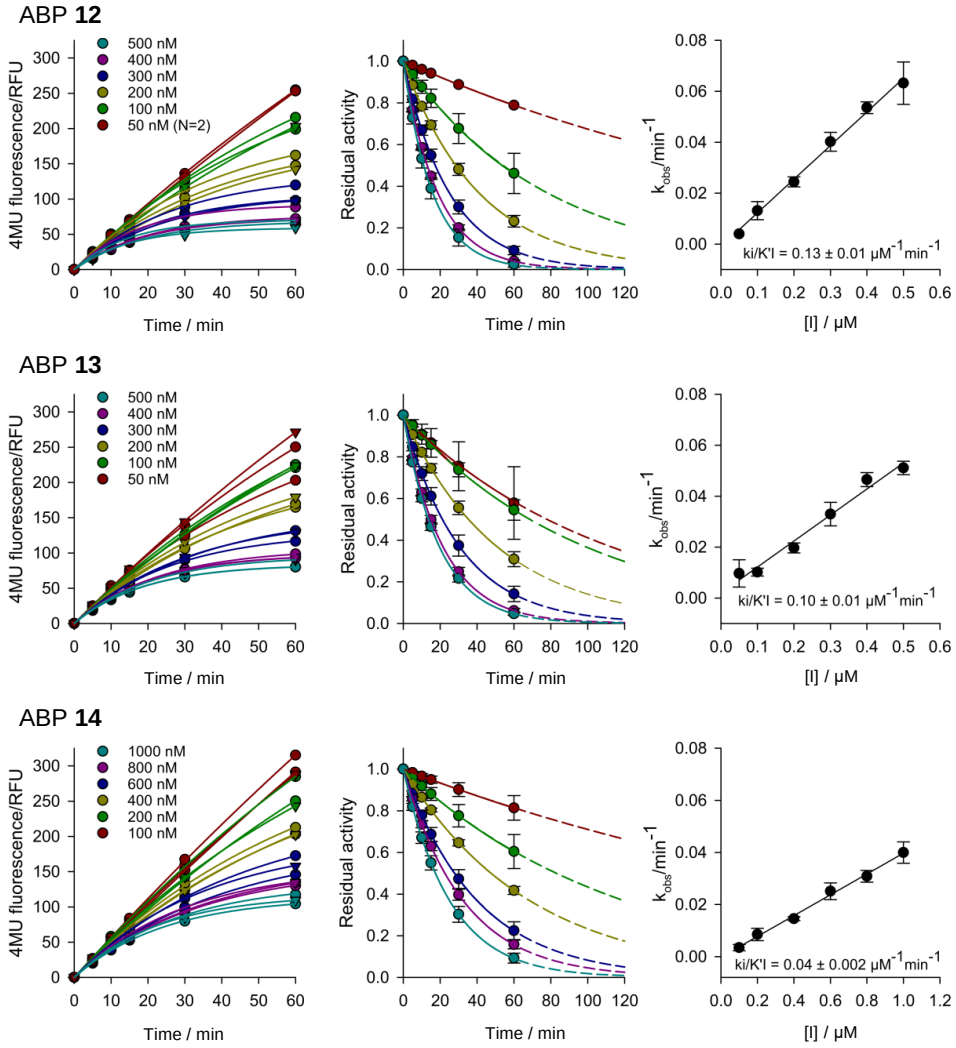


Figure 4.S2. Kinetic parameter measurements for compound 8–14 towards AcGH79. Left panels: progressing curves of relative fluorescent unit (RFU) for each sample incubated with the indicated inhibitor concentration for different incubation time. Middle panels: inhibition curves for each sample incubated with the indicated inhibitor concentration through the indicated incubation time, relative to the activity from the untreated sample at 60 min incubation time. Right panel: k_{obs} vs inhibitor concentration ([I]) plots for each compound.

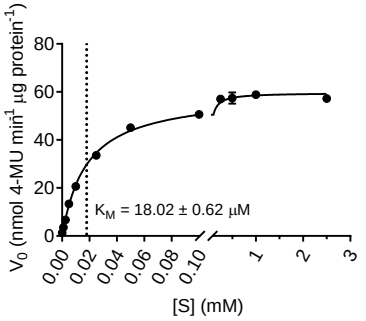
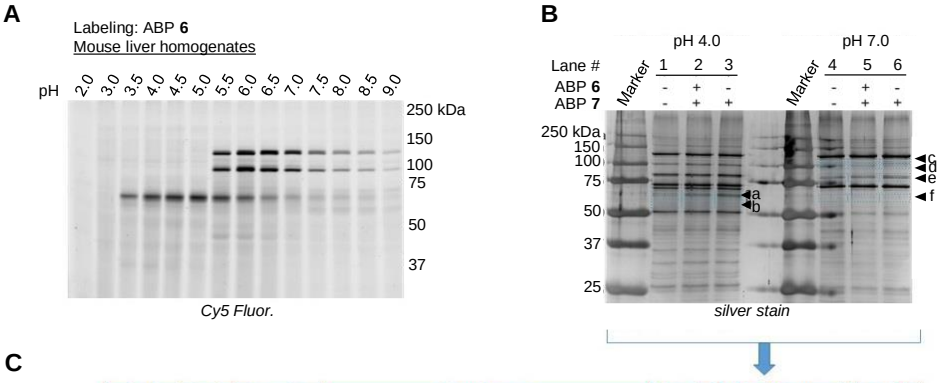


Figure 4.S3. Michaelis-Menten plot for AcGH79 using 4-MU-β-glc as substrate.

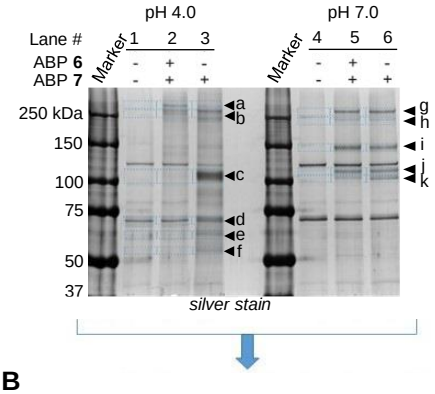


C

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 4.S4. Target identification for ABP 6 and 7 using chemical proteomics in mouse liver homogenates. A) ABP **6** labeling at various pHs. B) Silver stain of gels containing samples labeled with(out) ABPs at pH 4.0 or 7.0, and affinity-enriched for biotin. C) List of identified glycosidases by LC-MS measurement by LC-MS-based proteomics at positions shown in B.

A



C

pH 4.0	pH 7.0
a = MGAM, Sis	**g** = Sis
b = Sis, MGAM	**h** = Sis
c = Sis, MGAM	**i** = Lct
d = Sis, MGAM, GAA	**j** = Sis, MGAM
e = Sis, MGAM	**k** = GANAB, Sis
f = MGAM, Sis, GBA	

B

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						

Figure 4.S5. Target identification for ABP 6 and 7 using chemical proteomics in mouse intestine homogenates. A) Silver stain of gels containing samples labeled with(out) ABPs at pH 4.0 or 7.0, and affinity-enriched for biotin. B) List of identified glycosidases by LC-MS measurement by LC-MS-based proteomics at positions shown in A. C) Table of the most abundant glycosidase identified at each gel position shown in A.

Table 4.S1. Apparent IC_{50} values for compounds 1–7 towards recombinant human GBA (rGBA).

Compounds	IC_{50} (nM) (rGBA)
1 (JJB307)	155,000 ± 1,770
2 (CF21)	41,900 ± 490
3 (CF22)	603 ± 28.0
4 (JJB382)	593 ± 214
5 (JJB347)	1,080 ± 63.6
6 (JJB383)	816 ± 526
7 (JJB384)	2,060 ± 219

Table 4.S2. Crystallographic data collection and refinement statistics.

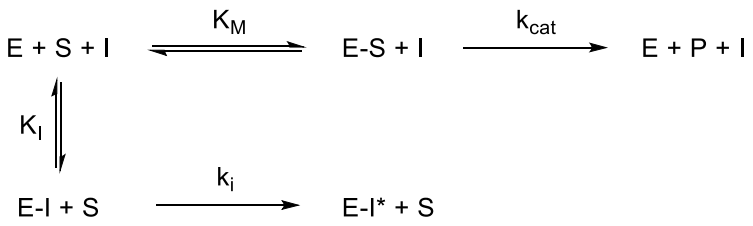
	CjAgd31B–**3**	AcGH79(wt)–**10**	AcGH79(E287Q)–**10**
PDB	5123	5G0Q	5L77
Data collection			
Space group	P622	C2	C2
Cell dimensions			
a, b, c (Å)	198.0, 198.0, 103.0	148.5, 44.9, 82.2	150.2, 45.1, 82.0
α, β, γ (°)	90, 90, 120	90, 114.8, 90	90, 115.1, 90
Resolution (Å)	85.74-1.95	42.60-1.60	28.90-1.24
R_{merge}	0.13 (1.64)	0.058 (0.664)	0.045 (0.489)
$I/\sigma I$	19.6 (2.3)	12.1 (1.8)	13.7 (1.9)
Completeness (%)	100 (100)	99.5 (99.7)	89.8 (49.7)
Redundancy	20.0 (20.1)	4.0 (3.7)	3.9 (2.7)
Refinement			
Resolution (Å)	85.74-1.95	42.60-1.60	28.90-1.24
No. reflections	82110	56355	120177
R_{work}/ R_{free}	0.17/0.19	0.12/0.17	0.13/0.16
No. atoms			
Protein	6299	3434	3477
Ligand/ion	102	36	56
Water	616	294	486
B-factors			
Protein	33.8	19.82	17.87
Ligand/ion	64.8	28.07	33.05
Water	39.7	30.29	30.95
R.m.s. deviations			
Bond lengths (Å)	0.012	0.020	0.016
Bond angles (°)	1.57	1.86	1.73

Value in parenthesis is for highest-resolution shell.

4.S2. Supporting note

Modeling irreversible enzyme inactivation by ABPs in presence of substrate.

Inactivation of enzyme by ABP is assumed to follow the following model (**Main text** Ref. 39):



K_M is the equilibrium constant for E-S formation (i.e. Michaelis constant; the concentration of S in which product formation is at 1/2 V_{max}), k_{cat} is the rate constant for P formation, K_I is the equilibrium constant for E-I formation (i.e. inhibition constant), k_i is the rate constant for E-I* formation (i.e. inactivation rate constant).

In the experiments I is the ABP, S is the substrate 4-MU-GlcA, E is AcGH79 and P is the measured product 4-MU. [I] and [S] are both >> $[E_{tot}]$ which is fixed at 260 pM ($[E_{tot}]$ = [E]+[EI]+[ES]+[EI*]). [S] is 2.5 mM.

In this model, the substrate and ABP compete reversibly for the enzyme active site, followed by irreversible reactions to give either P, or the inactivated enzyme complex E-I*.

Under these conditions, initial enzyme reaction rate V_0 is given by:

$$V_0 = {dP}/{dt} = \frac{k_{cat}[S][E_{tot}]}{K_M\left(1+{[I]}/{K_I}\right)+[S]} \qquad (1)$$

At time t:

$$V_t = \frac{k_{cat}[S][E_{tot} - EI^*]}{K_M\left(1+{[I]}/{K_I}\right)+[S]} \qquad (2)$$

Analogously, the rate of enzyme inactivation at time t is given by:

$$dEI^*/dt = \frac{k_i[I][E_{tot} - EI^*]}{K_I\left(1+{[S]}/{K_M}\right)+[I]} = k_{obs}[E_{tot} - EI^*] \quad (3)$$

Where k_{obs} is defined as the observed inactivation rate constant:

$$k_{obs} = \frac{k_i[I]}{K_I\left(1+{[S]}/{K_M}\right)+[I]} \quad (4)$$

Here, the appeared equilibrium constant for I is dependent on the type and the amount of S applied in the assay (K_M and [S]), and is defined as K'_I:

$$K_I\left(1+{[S]}/{K_M}\right) = K'_I \quad (5)$$

$$k_{obs} = \frac{k_i[I]}{K'_I+[I]} \quad (6)$$

For deriving k_{obs} experimentally, [I] and [S] can be both set at >> [E_{tot}]. Under this condition, k_{obs} is a pseudo-first order inactivation rate constant (dependent on [I]). The extent of enzyme inactivation at time t is quantified by the residual activity V_t/V_0, where V_t is the enzyme reaction rate at time t. Thus, V_t/V_0 for each value of [I] can be described by the first-order rate equation:

$$V_t/V_0 = e^{-k_{obs}t} \quad (7)$$

$$V_t = V_0 e^{-k_{obs}t} \quad (8)$$

The amount of product [P] formed at time t is the integral of V_t, with respect to t, between the interval [0, t]:

$$[P] = \int_0^t [V_t]dt = {V_0}/{k_{obs}}\,(1 - e^{-k_{obs}t}) \quad (9)$$

Fitting 4-MU fluorescence (which is $\propto$ [P]) *vs.* time using (9) allows k_{obs} to be derived for each value of [I]. Next, the derived k_{obs} values at each [I] can be plotted against each [I], and the resulting plot of k_{obs} vs [I] can be fitted using (6), which is an analogue of Michaelis-Menton equation, to derive K'_I and k_i.

When [I] << K'_I, as observed from the data in this chapter, K'_I and k_i cannot be accurately calculated. Instead, (4) is well approximated by:

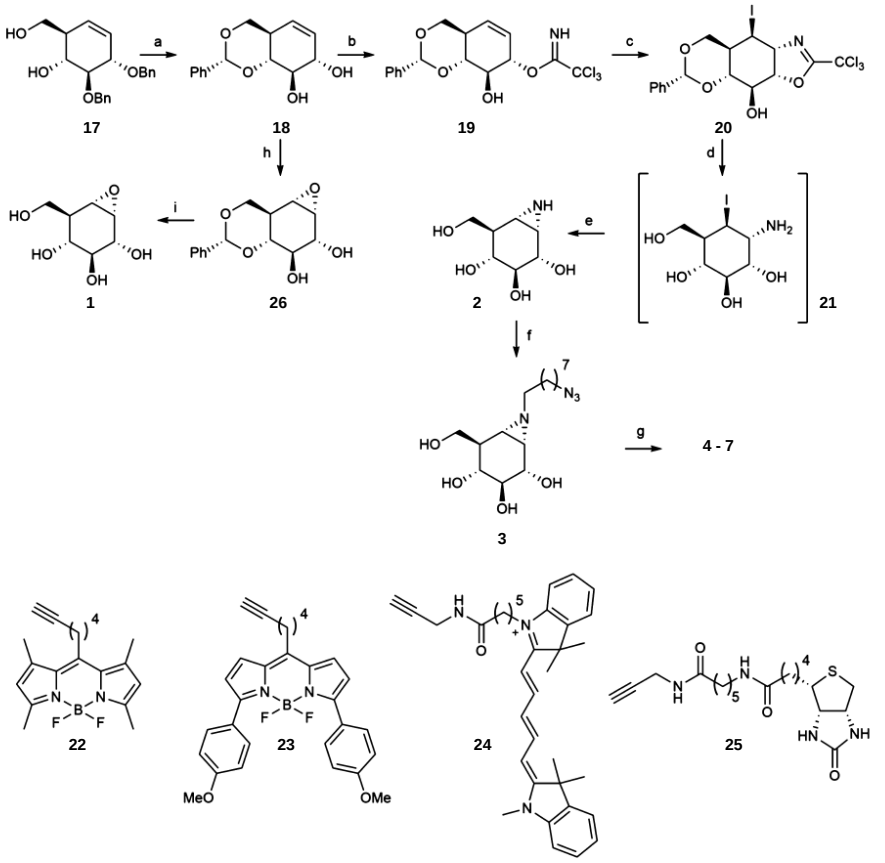
$$k_{obs} \approx {k_i}/{K'_I}\,[I] \qquad (10)$$

Thus, a subsaturated plot of k_{obs} against [I] is better fitted using linear function (10), allowing an estimate of k_i/K'_I to be derived from the gradient of the slope.

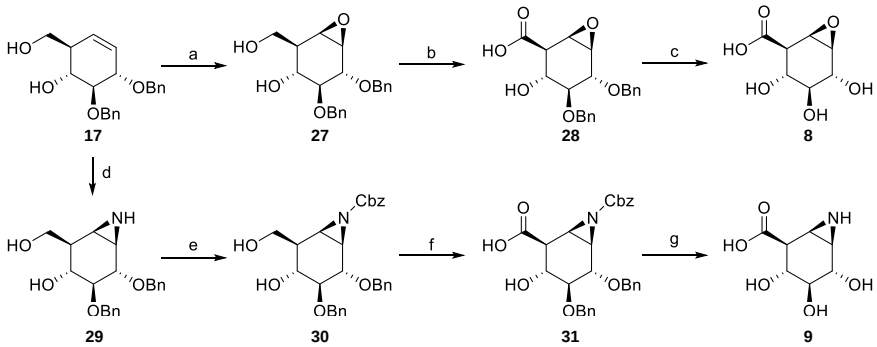
k_i/K_I is derived from k_i/K'_I by applying (5), where [S] is known and K_M can be derived by a separate Michaelis-Menton experiment for S on the given enzyme.

4.S3 Synthetic schemes for compounds 1-15 (Department of Bio-organic synthesis, Leiden University)

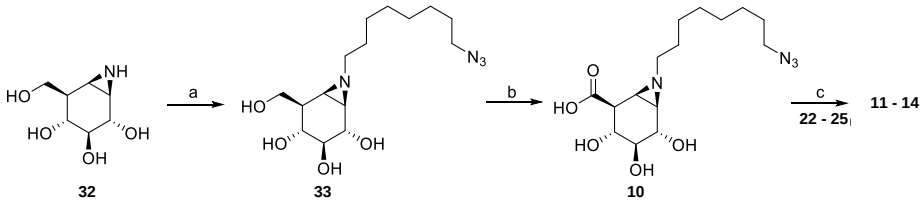
For complete synthetic method and characterization data, the reader is directed to the supporting information in the published articles.[Main text refs 19, 31]



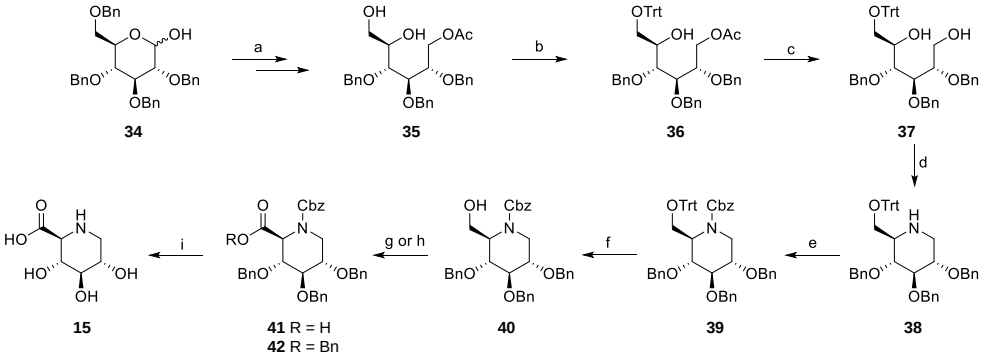
Scheme 4.S1. Synthesis of 1,6-epi-cyclophellitol 1, the cyclophellitol aziridine inhibitors 2, 3, and ABPs 4−7. Reagents and conditions: (a) (i) Li, NH_3, THF, −60 °C, 57 %; (ii) $PhCH(OMe)_2$, CSA, DMF, 61 %. (b) CCl_3CN, DBU, DCM, 0 °C. (c) $NaHCO_3$, I_2, H_2O, two step yield 41 %. (d) 37 % HCl aq, dioxane. (e) $NaHCO_3$, MeOH, two step yield 63 %. (f) 1-azido-8-iodooctane, K_2CO_3, DMF, 80 °C, 39 %. (g) **4**, **5**, **6**, or **7**, $CuSO_4$, sodium ascorbate, 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 %.



Scheme 4.S2. Synthesis of uronic cyclophellitol epoxide 8 and aziridine 9. Reagents and conditions: (a) *m*-CPBA, Na_2HPO_4 (aq., 1 M), NaH_2PO_4 (aq., 1.0 M), DCM, 50 °C, 55 %; (b) BAIB, TEMPO, DCM/^{t}BuOH/H_2O (4:4:1), 0 °C, 79 %; (c) H_2, $Pd(OH)_2/C$, MeOH, 98 %; (d) i) CCl_3CN, DBU, DCM, 0 °C; ii) I_2, $NaHCO_3$, H_2O; iii) 37 % HCl aq., dioxane, 60 °C; iv) $NaHCO_3$, MeOH, 45 %; (e) Cbz-Cl, Pyridine, THF, 75 %; (f) BAIB, TEMPO, DCM/^{t}BuOH/H_2O (4:4:1), 0 °C, 60 %; (g) NH_3, Li, THF, -60 °C, 60 %.



Scheme 4.S3. Synthesis of compounds 10-14. Reagents and conditions**:** (a) K_2CO_3, DMF, 1-azido-8-iodooctane, 80 °C, 50 %; (b) TEMPO, NaClO, NaBr, NaOH, H_2O, 15 %; (c) $CuSO_4$, sodium ascorbate, DMF, 22 % **11**, 13 % **12**, 17 % **13**, 13 % **14.**



Scheme 4.S4. Synthesis of 2S, 3R, 4R, 5S-trihydroxypipecolic acid 15. Reagents and conditions: (a) Supplementary Ref. [1] (b) DMAP, TEA, TrtCl, DCM, 3 h, 75 %; (c) Na, MeOH, 3 h, 91 %; (d) i): $(COCl)_2$, DMSO, DCM, –65 °C, 2 h; ii): TEA, -65 to 5 °C, 2 h; iii): NH_4HCO_3, $NaBH_3CN$, 3 Å mol. sieves, MeOH, 0 °C, 1 h, then 0 °C to RT, 20 h, 85 %; (e) CbzCl, THF, TEA, RT, 20 h, 72 %; (f) *p*-TsOH, DCM, 0 °C to RT, 3 h, 83%; (g) TEMPO, BAIB, DCM/H_2O (2/1), 0 °C, 5 h, **41**: 71 %; (h) i): TEMPO, BAIB, DCM/H_2O (2/1), 0 °C, 5 h; ii): BnBr, Cs_2CO_3, DMF, RT, 3 h, **42**: 75 %; (i) H_2, $Pd(OH)_2$/C, AcOH/H_2O(4/1), RT, 16 h, 57 %.

6.S4 Supplementary Reference

1 La Ferla B, Bugada P & Nicotra F (2006) Synthesis of the dimethyl ester of 1-deoxy-L-idonojirimycin-1-methylenphosphonate: A new approach to iminosugar phosphonates. *J Carb Chem* **25**, 151–162.

CHAPTER 5

Development and application of activity-based probes for α-L-iduronidase

Manuscript published as:

Artola M, **Kuo CL**, McMahon SA, Oehler V, Hansen T, van der Lienden M, He X, van den Elst H, Florea BI, Kermode AR, van der Marel GA, Gloster TM, Codée JDC, Overkleeft HS & Aerts JMFG (2018) New Irreversible α-l-Iduronidase Inhibitors and Activity-Based Probes. *Chemistry* **24**, 19081–19088.

ABSTRACT

Human α-L-iduronidase is a retaining exo-acting glycosidase responsible for catalyzing the lysosomal turnover of glycosaminoglycans, and its hereditary deficiency underlies the lysosomal storage disorder mucopolysaccharidosis type I (MPS I). New covalent inhibitors and ABPs of α-L-iduronidase has been synthesized based on the cyclophellitol scaffold, and these were shown to react covalently and irreversibly in an activity-based manner with human recombinant α-L-iduronidase (rIDUA, Aldurazyme®), albeit with a lower potency and slower inhibitory kinetics compared to earlier generated ABPs towards other retaining glycosidases. The structures of IDUA when complexed with the inhibitors in a noncovalent transition-state-mimicking form and a covalent enzyme-bound form provide insights into its conformational itinerary. Inhibitors **1–3** adopt a half-chair conformation in solution ($^{4}H_{3}$ and $^{3}H_{4}$), as predicted by DFT calculations, which is different from the conformation of the Michaelis complex observed by crystallographic studies. Consequently, **1–3** may need to overcome an energy barrier in order to switch from the $^{4}H_{3}$ conformation to the transition state ($^{2,}{}^{5}B$) binding conformation before reacting and adopting a covalent $^{5}S_{1}$ conformation, explaining their lower potency and slower inhibitory kinetics. Nevertheless, rIDUA can be efficiently labeled with fluorescent Cy5 ABP **2**, which allows monitoring of the delivery of this therapeutic recombinant enzyme to lysosomes, as is intended in enzyme replacement therapy for the treatment of MPS I patients.

5.1 Introduction

Human α-L-iduronidase (IDUA), which belongs to the GH39 family in the Carbohydrate Active EnZyme (CAZy) classification,[1, 2] hydrolyzes terminal non-reducing α-L-iduronic acid residues in glycosaminoglycans (GAGs), including dermatan sulfate (DS) and heparan sulfate (HS), through a two-step Koshland double-displacement mechanism (**Fig. 5.1A**).[3–6] The active site of the enzyme contains a carboxylic acid/carboxylate pair that acts as an acid/base (Glu188) and a nucleophilic (Glu299) catalyst. Protonation of the exocyclic oxygen by the acid/base residue and concomitant nucleophilic attack at the anomeric carbon by Glu299 leads to S_N2 displacement of the aglycon, yielding a covalent enzyme–substrate complex with inversion of stereochemistry at the anomeric carbon. In the next step, a water molecule enters the enzyme active site, where it is deprotonated by the acid/base (Glu188) and then hydrolyzes the enzyme–substrate intermediate in a reversal of steps, again with inversion of anomeric configuration. The conformational change of IDUA from Michaelis complex to transition state (TS) and enzyme–inhibitor covalent complex has recently been suggested to follow a $^{2}S_{0} \rightarrow {}^{2,5}B \rightarrow {}^{5}S_{1}$ itinerary.[5, 7, 8] This catalytic pathway was predicted on the basis of the structures of Michaelis complexes with (2*R*, 3*R*, 4*R*, 5*S*)-2-carboxy-3,4,5-trihydroxypiperidine (IdoA-DNJ) and 5-fluoro-α-L-idopyranosyluronic acid fluoride (5F-IdoAF) as reversible inhibitors and the 2-deoxy-2-fluoro-α-L-idopyranosyluronic acid (2F-IdoA)–enzyme covalent complex intermediate.

Mutations in the gene encoding IDUA may result in mucopolysaccharidosis type I (MPS I), which is a rare autosomal inherited recessive disease that leads to toxic accumulation of HS and DS. MPS I is a devastating disease that affects around 1 in 100,000 individuals and is classified as attenuated MPS I and more severe MPS I (traditionally categorized from less to more severe form as Scheie, Hurler–Scheie, or Hurler disease) to distinguish between disease severity and therapeutic options.[8] Children with severe MPS I are treated at a young age by hematopoietic cell transplantation (HCT).[9] Enzyme replacement therapy (ERT) with recombinant human α-L-iduronidase (rIDUA, Aldurazyme®) is considered as a treatment for non-neurological manifestations of MPS I.[10] There is consensus among treating clinicians that

the impact of ERT with rIDUA is greatest when the treatment is initiated early in the disease progression. An obvious prerequisite for effectiveness is the successful targeting of infused rIDUA to lysosomes in multiple cell types, and for this purpose a detailed understanding of rIDUA targeting is still urgently needed. Besides MPS I, lysosomal α-L-iduronidase is indirectly involved in two other inherited lysosomal storage disorders, mucolipidosis II (ML II) and III a/b. Here, a deficiency in the generation of mannose-6-phosphate (M6P) moieties in N-linked glycans of newly formed lysosomal enzymes impairs their correct routing to lysosomes and therefore these, including α-L-iduronidase, are largely erroneously secreted.[11]

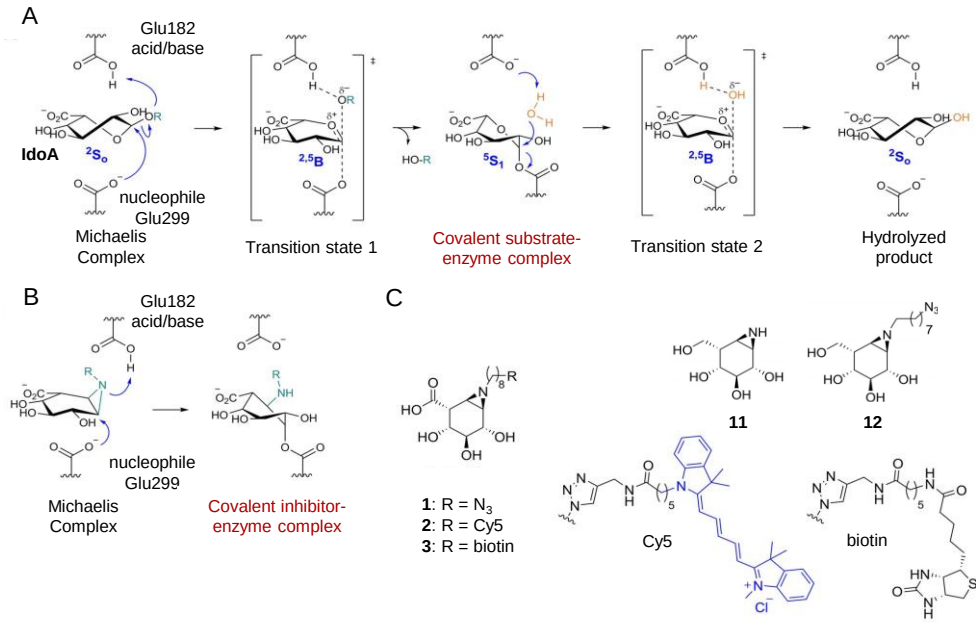


Figure 5.1. Reaction mechanisms and structures of compounds used in this chapter. A) Koshland double-displacement mechanism employed by retaining α-L-iduronidase, showing the $^{2}S_{0}$ → $^{2,5}B$ → $^{5}S_{1}$ conformational reaction itinerary from the Michaelis complex, transition state 1, covalent substrate–enzyme intermediate, and transition state 2, to the hydrolyzed product. B) Proposed inhibition mechanism of aziridine-based inhibitor **1** and ABPs **2** and **3**. C) Chemical structures of α-L-iduronic-configured mechanism-based irreversible inhibitor **1** and ABPs **2** and **3**, and the idose-configured compound **11** and **12** described in this work.

Herein, the characterization and application of new irreversible IDUA inhibitors and activity-based probes (ABPs) bearing an α-L-iduronic configured cyclophellitol aziridine as an electrophilic "warhead" is reported. Functionalization of the aziridine with a Cy5 fluorophore

(for gel and/or cell imaging) or biotin (for chemical proteomics studies) afforded valuable tools (**Fig. 5.1B, C**) for the study of α-L-iduronidase *in vitro* and *in situ*, structural analysis, and for monitoring rIDUA uptake and trafficking to lysosomes, as is revealed in this chapter.

5.2 Results

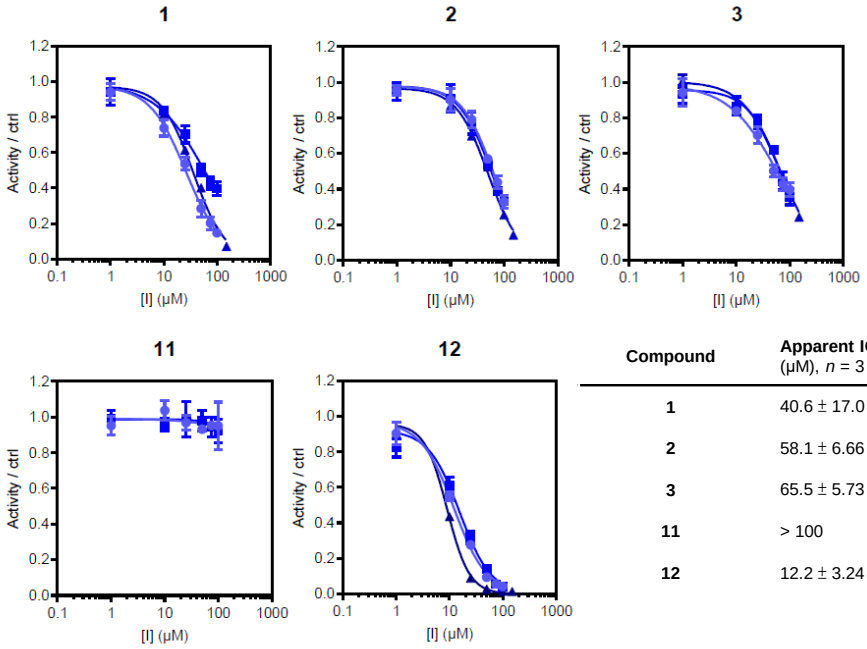
5.2.1 Synthesis of α-L-iduronic-configured inhibitors and ABPs

The synthetic strategy for compound **1**, ABP **2** and ABP **3** (F**ig. 5.1C**) (Department of Bio-organic Synthesis, Leiden University) is described in detail in section **4.S1.2** in the **Appendix**. Briefly, to obtain the desired α-L-iduronic-configured inhibitors and ABPs, the α-L-iduronic-configured cyclohexene **8** was produced in three steps, followed by benzylation of free alcohol and aziridination with 3-amino-2-(trifluoromethyl)quinazolin-4(3*H*)-one (CF_3-Q-NHOAc), following strategy described by Llebaria and coworkers[12] and the Overkleeft group[13] (**Scheme 5.S1**). After removal of benzyl and CF_3-Q groups, the obtained idose-configured aziridine **11** (**Fig. 5.1C**) was alkylated to afford the azide-containing compound **12** (**Fig. 5.1C**), which was further oxidized at C-6 to afford the α-L-iduronic-configured cyclophellitol aziridine compound **1** (**Fig. 5.1C**). Finally, click reaction with a Cy5- or biotin-substituted alkyne on compound **1** afforded the desired ABP **2** or **3** (**Fig. 5.1C**).

5.2.2 *In vitro* inhibition and labeling of recombinant human α-L-iduronidase with compounds **1–3**

The *in vitro* inhibition potencies of compounds **1–3** was examined by incubating them for 60 min at various inhibitor concentrations with human recombinant α-L-iduronidase (rIDUA, Genzyme) at pH 4.5, followed by 30 min incubation with the fluorogenic substrate 4-methylumbelliferyl-a-iduronide (4-MU-IdoA). The assay mixture was supplemented with BSA, Triton X-100, NaCl, and chloride salt of divalent cations, by adapting methods described in literatures[14, 15], which resulted in a 3.5-fold increase in rIDUA activity (**Fig. 5.S1**). Compounds **1–3** inhibited rIDUA with apparent half-maximum inhibitory concentrations in the micromolar range (**Fig. 5.2**). Intermediate **11** showed no activity (**Fig. 5.2**), in line with the role of the

carboxylate group at C5 for IDUA binding in a positively charged enzymatic pocket formed by the Arg363 and Lys264 side chains and the main-chain NH groups of Gly305 and Trp306.[8] Surprisingly, alkyl aziridine **12** inhibited rIDUA on a par with compound **1–3** (**Fig. 5.2**), suggesting that the carboxylate group may not be essential for binding if an N-alkyl group is presented. These results are also in line with previous findings that N-alkyl aziridines display improved binding potency towards glycosidases.[16]



Compound	Apparent IC_{50} (μM), n = 3 ± SD
1	40.6 ± 17.0
2	58.1 ± 6.66
3	65.5 ± 5.73
11	> 100
12	12.2 ± 3.24

Figure 5.2. *In vitro* inhibition of recombinant α-L-iduronidase (rIDUA) by compounds 1-3, 11 and 12 at 1 hour incubation time. Compounds **1-3**, **12** = triplicate sets of experiment, each with technical triplicates; **11** = duplicate sets of experiment. Error range in inhibition curves = SD from technical triplicates. Error range in IC_{50} values = SD from triplicate (or duplicate) experimental sets.

Cy5 ABP **2** labeled rIDUA in a concentration- and time-dependent manner, consistent with the irreversible inhibition mechanism of these analogues, with optimal labeling at 50 μM and about 120–180 min incubation, as visualized by SDS-PAGE (**Fig. 5.3A**). The optimum pH for labeling with ABP **2** was determined as 4.5–5.0, consistent with the reported optimum pH for enzymatic activity (**Fig. 5.3B**).[2] The stability of the covalent enzyme–inhibitor complex was

also tested, and it was observed that rIDUA remained inactivated for at least 100 h (**Fig. 5.3C**). In addition, competitive activity-based protein profiling (ABPP) showed competition in a

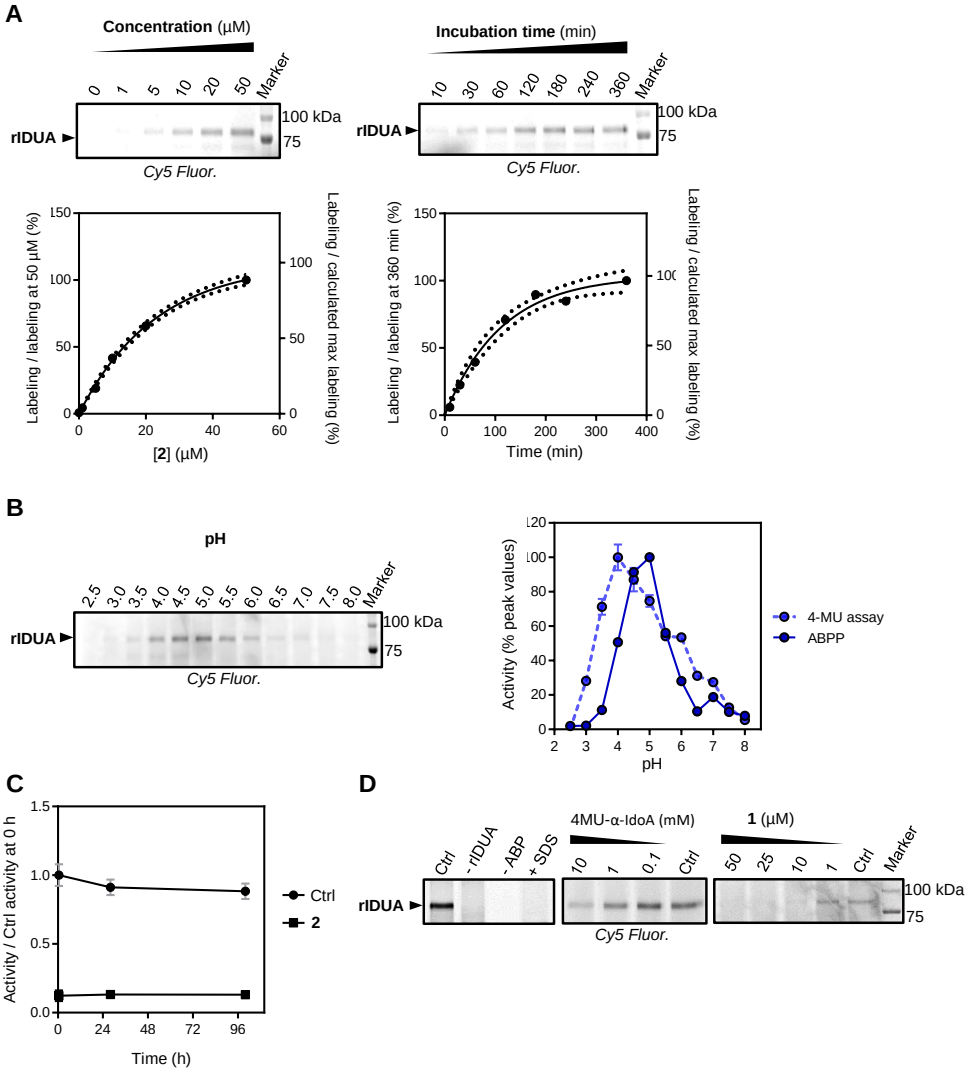


Figure 5.3. Labeling of rIDUA with ABPs 2. A) Cy5 ABP **2** labels rIDUA (10 ng) in a concentration-dependent (left, at 4 h incubation time) and time-dependent (right, at 50 µM) manner. Labeling signals were quantified (below each gel image) and fitted with a one-phase association equation. Dotted lines represent 95 % CI. B) ABP **2** labels rIDUA in a pH-dependent manner. The quantified labeling signals were compared to data obtained with 4-MU substrate assay (below; error range =SD from technical triplicates). C) Irreversible inhibition of rIDUA by ABP **2**. Error range = SD from technical triplicates. D) ABP **2** labels rIDUA and negative controls: without rIDUA, without ABP **2** or with SDS (left panel), competition with 4-MU-α-L-iduronide (middle panel) or 1 (right panel).

concentration-dependent manner with 4-MU-IdoA as well as with inhibitor **1**, illustrating the applicability of this probe for the screening of new inhibitors (**Fig. 5.3D**).

Next, kinetic parameters for rIDUA labeling/inhibition by ABP **2** was determined. Due to the fluorescent nature and the slow labeling on rIDUA by ABP **2** (Fig. 4.3A, right panel), it was envisioned that kinetic parameters could be derived by monitoring the progress of rIDUA labeling by ABP **2** at various concentrations (5–60 mm) and different incubation times (30–150 min) using an SDS-PAGE-based ABPP method. This would simplify the derivatization of kinetic parameters commonly applied for irreversible inhibitors (see Experimental procedures **5.4.9** for more detailed explanation). Indeed, when labeling of rIDUA by ABP **2** (**Fig. 5.S2**) was quantified (**Fig. 5.4A**), the initial binding constant (K_I) and an inactivation rate constant (k_{inact}) could be determined (**Fig. 5.4B**). The calculated k_{inact}/K_I value was 1.38 x 10^{-4} μM^{-1} min^{-1},

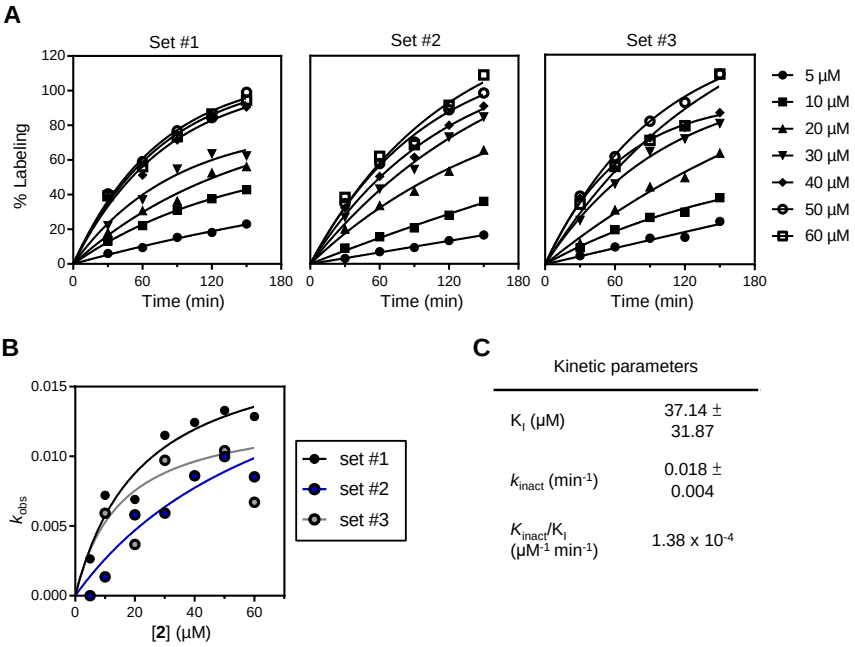


Figure 5.4. Kinetic parameters for ABP 2 towards rIDUA. A) Percentages of rIDUA labeling at different time points and at different concentrations of ABP **2**. Data were quantified from three sets of fluorescent gels containing rIDUA labeled with ABP **2** under the depicted conditions to derive a rate constant *k* for each ABP 2 concentration. B) Left, *k* vs. [inhibitor] plot. Data were curve-fitted with the Michaelis–Menten equation to obtain kinetic parameters. Right, calculated kinetic parameters for ABP **2** labeling of rIDUA. Error range = SD from the three sets.

showing this compound to be a less potent inactivator than other ABPs for related glycosidases, such as β-glucuronidases[16], β-glucosidases[17, 18], α-glucosidases[19], and α-galactosidase[20].

To further demonstrate the covalent binding to the catalytically active amino acid, biotinylated ABP **3** was incubated with rIDUA, and peptide mass fingerprinting-based proteomic was utilized to detect for the presence of ABP **3** on the active site peptide of rIDUA. After ABP incubation, rIDUA was denatured and digested by chymotrypsin, and the resulting ABP **3**-labeled peptides were affinity-enriched and analyzed by nanoscale liquid chromatography coupled with tandem mass spectrometry (nano-LC-MS/MS). A fragment of the IDUA nucleophile (Glu299) covalently attached to ABP **3** was detected by MS/MS fragmentation of the 7-amino-acid peptide containing the nucleophilic residue (**Fig. 5.5**, **5.S3** and **5.S4**).

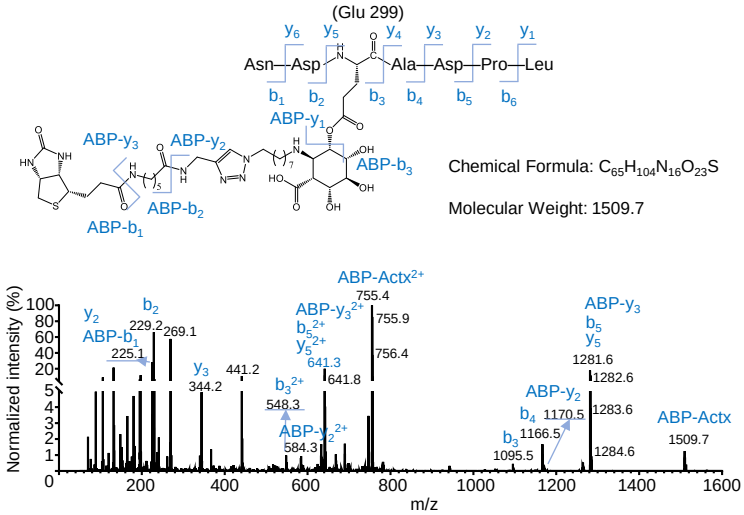


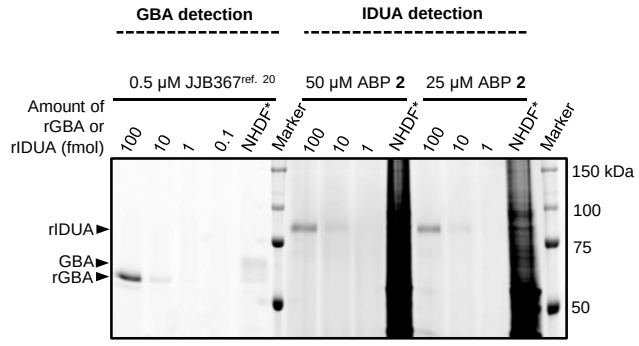
Figure 5.5. Proteomic detection of ABP 3-labeled rIDUA peptides. MS/MS pattern of a sample containing rIDUA Asn297–Leu303 active site peptide labeled with biotin ABP 3 at Glu299 is shown, and peaks corresponding to the detected fragments are annotated. *Actx* = active site peptide.

5.2.3 Activity-based protein profiling of IDUA in homogenates of fibroblasts

It was attempted to use ABP **2** to visualize endogenous IDUA in lysates of cultured normal human dermal fibroblasts (NHDF) and concentrated human urine, with and without

pre-purification by concanavalin A beads. Unfortunately, no clear labeling on rIDUA was observed with ABP **2** at 1–10 μM concentration, and many non-specific bands were detected at higher concentrations, that could not be competed by inhibitor **1** nor ABP **3** (data not shown). To investigate whether the lack of specific IDUA labeling in NHDF lysates is due to low enzyme amount, the specific glycosidase activity and ABP labeling of IDUA vs. β-glucocerebrosidase (GBA) with those of the corresponding Cy5 ABPs (**2** vs. ABP JJB367[21]) was compared in recombinant enzymes and NHDF lysates. It was observed that while it is theoretically possible to detect IDUA in fibroblast lysates (calculated amount of IDUA in lysates was only 1.5 lower than that of GBA in the same lysates), its detection in NHDF lysates is still challenging due to nonspecific labeling of other proteins at the required concentration of ABP **2** (25–50 mm) (**Fig. 5.6**). On the other hand, measurement of ABP **2**'s inhibition on IDUA in NHDF lysates showed

A



B

	Specific activity (nmol h^{-1})		**Calculated enzyme amount** (fmol)
	Recombinant enzyme (per 100 fmol)	NHDF lysate (per 10 μg)	NHDF lysate (per 10 μg)
IDUA	2.44 ± 0.11	0.40 ± 0.01	16.39
GBA	8.62 ± 1.26	2.12 ± 0.10	24.59
GBA/ IDUA	***3.53***	***5.3***	***1.50***

Figure 5.6. Detection of endogenous IDUA in human fibroblasts (NHDF) lysates. A) Comparing endogenous enzyme amount in human fibroblasts (NHDF) lysates for IDUA vs glucocerebrosidase (GBA), by comparing measured specific activity presented in lysates vs from known amount of rIDUA or rGBA with 4-methylumbelliferyl-glycoside substrates. B) Comparing ABP labeling of known amounts of rGBA and rIDUA vs ABP labeling in NHDF lysates. *10 μg protein was loaded.

an apparent IC_{50} value of 23.5 μM (**Fig. 5.S5**), comparable to the value towards rIDUA. Therefore, it was concluded that direct detection of IDUA in biological samples by ABP **2** via SDS-PAGE is not feasible due to a combination of this compound's moderate activity and a lack of relative substrate specificity. Future optimization of enrichment method for endogenous IDUA is warranted for its eventual application in diagnostic purposes.

5.2.4 Conformations of α-L-ido-aziridine **11** and α-L-idoA-aziridine by DFT calculations

The conformational preferences of α-L-ido-aziridine **11** and α-L-idoA methylated aziridine was studied (Department of Bio-organic Synthesis, Leiden University), as simplified representations of inhibitors **1–3**, in order to elucidate the involvement of conformation on inhibitor potency and binding kinetics. A conformer distribution search in Spartan 14[22] and further optimization with Gaussian 09[23] by utilizing B3LYP/6–311G(d, p)/PCM(H_2O) (for details, see section **5.S3** in the **Appendix**) showed that the $^4H_3$ conformation of α-L-ido-aziridine **11** is greatly favored, with variations in the geometry about the C5–C7 bond (**Table 5.S2**). Conversely, α-L-idoA-aziridine showed both $^4H_3$ and $^3H_4$ (the latter 1.4 kcal mol^{-1} higher in energy) as relevant conformations. Interestingly, the $^{2,5}B$ boat conformation was also found as a relevant geometry for α-L-idoA-aziridine, albeit with an energetic cost of 8.0 kcal mol^{-1}. In addition, coupling constants (*J*) were calculated for the lowest-energy conformations, and these were in excellent agreement with experimental NMR data (**Table 5.S3**, **5.S4**).

5.2.5 Structural analysis of IDUA interactions with inhibitors **1–3**

In order to study the mechanism of action of inhibitors **1–3**, their conformations upon binding to IDUA were analyzed by crystallographic studies by collaboration with the Gloster group (University of St Andrews). Compounds were applied in solution for various durations to raIDUA crystals (IDUA recombinantly expressed in the seeds of a cgl (complex glycan deficient) mutant of *Arabidopsis thaliana*). Data were collected from a crystal soaked with ABP **1** for 24 h to 2.02 Å resolution (**Table 5.S1**), which revealed the structure of raIDUA in a covalent complex with **1** (**Fig. 5.7A**). The aziridine nitrogen is displaced by nucleophilic attack of the active site carboxylate to form a trans-2-amino ester (with the rest of the R group not visible in the electron

density, presumably disordered due to its inherent flexibility; this region of the structure is exposed to the solvent). Interestingly, the pseudo-glycoside was observed in a $^{5}S_{1}$ skew-boat conformation, which differs slightly from the distorted $^{2,5}B$ boat conformation reported for the previously described irreversible inhibitor 2-deoxy-2-fluoro-α-L-ido-pyranosyluronic acid

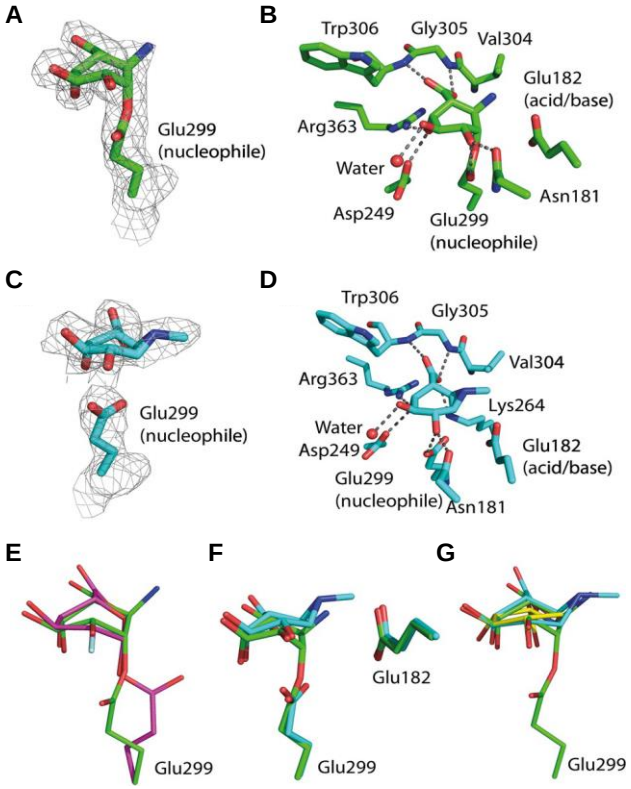


Figure 5.7. Structural insights into raIDUA complexed with ABPs. A) Structure of raIDUA complexed with a fragment of compound **1**, which is covalently linked to the nucleophile Glu299. The maximum likelihood/s_A weighted $2F_{obs}–F_{calc}$ electron density map (gray) is contoured at 1.2 sigma. B) Structure of raIDUA covalently complexed with a fragment of compound **1**, illustrating the active site residues that interact with the pseudo-glycoside. C) Structure of raIDUA complexed with a fragment of ABP **3**. The nucleophile Glu299 is shown. The maximum likelihood/s_A weighted $2F_{obs}–F_{calc}$ electron density map (gray) is contoured at 1.0 sigma. D) Structure of raIDUA complexed with a fragment of ABP **3**, illustrating the active site residues that interact with the pseudo-glycoside. E) Superposition of raIDUA covalently complexed with fragments of compound **1** (green) and 2F-IdoA (pink; PDB code 4KH2[8]). F) Superposition (based on alignment of protein main-chain atoms) of raIDUA complexed with a fragment of compound **1** (covalent, green) and a fragment of ABP **3** (transition state, cyan). G) Superposition (based on alignment of C3 and C4 atoms of each molecule) of raIDUA complexed with a fragment of compound **1** (covalent, green), a fragment of ABP **3** (transition state, cyan), and IdoA-DNJ (Michaelis complex, yellow; PDB code 4KGL).[8]

(2F-IdoA) covalently bound to IDUA.[8] The observed $^{5}S_{1}$ conformation of the covalent inhibitor **1**–enzyme complex supports predictions for the conformational itinerary followed by α-L-iduronidase GH39 (**Fig. 5.1A**).[7] The carboxylate group of the pseudo-iduronic acid forms bidentate hydrogen bonds with the main-chain nitrogen atoms of Gly305 and Trp306, the C4 hydroxyl group forms hydrogen bonds with Arg363 and Asp349, the C3 hydroxyl group interacts with Asp349 and a water molecule, and the C2 hydroxyl group forms hydrogen bonds with Asn181 and the nucleophile Glu299 (**Fig. 5.7B**). In the covalent complex between IDUA and 2F-IdoA, the nucleophile Glu299 is rotated by around 90 °C compared to the position observed in the complex here with the fragment of **1**,[8] and the fluoro group at C2 may preclude an interaction with Oε2 of Glu299, causing it to rotate. However, the inhibitor **1**–IDUA complex presented here, bearing a hydroxyl group at C2 and showing an interaction with Glu299, is more likely to represent what occurs during catalysis (**Fig. 5.7E**).

In an attempt to fully define the conformational inhibition of compounds **1–3**, raIDUA crystals were soaked with the ABPs for shorter durations. Data collected to 2.39 Å resolution (**Table 5.S1**) on a crystal soaked with ABP **3** for 45 min revealed electron density in the active site of raIDUA consistent with the unreacted cyclophellitol aziridine **3** (**Fig. 5.7C**). A methyl group on the cyclophellitol aziridine was visible, but the rest of the R group was not evident and presumably disordered. Interestingly, the pseudo-glycoside was observed in a $^{2,5}B$ conformation, which is the predicted transition state for GH39 α-L-iduronisase.[7] The majority of the interactions with active-site residues were the same as those described for the covalent complex with raIDUA (**Fig. 5.7D**), although a shift in position of the glycoside indicated that the carboxylate group additionally interacted with Lys264. The hydroxyl group at C2 forms a hydrogen bond with the nucleophile Oε2 of Glu299, but at a surprisingly short distance of 2.4 Å, suggesting a tight interaction. This close proximity results in a distance between the pseudo-anomeric carbon and Oε1 of only 2.9 Å. These tight interactions, together with the $^{2,5}B$ conformation of the pseudo-glycoside, suggest that the pseudo-glycoside at the transition state was being observed; such structural observations are rare using wild-type enzymes, but here it was possible due to the slow inactivation kinetics of **3**. The importance of the interaction

between the glycoside and the C2 hydroxyl group supports work by others;[24, 25] indeed, interactions at the 2-position were estimated to contribute 18 kJ mol^{-1} binding energy to stabilization of the transition state for a β-glucosidase during the glycosylation step of the catalysis.[24] Based on this work, it was postulated that a hydrogen bond formed between the C2 hydroxyl group and the nucleophile would be optimal at the transition state, as the two groups come into close proximity during formation of the covalent glycosyl–enzyme bond.[24]

Superimposition of the main chain atoms for the two complexes revealed a shift in the position of the cyclophellitol aziridine to accommodate formation of the covalent bond (**Fig. 5.7F**). This engendered movement of between 0.2 and 0.4 Å at C5, C4 and C3, 0.5 Å at C2, 0.6 Å at the carbon at the position of the endocyclic oxygen, and 1.1 Å at the pseudo-anomeric carbon. These structures, together with the previously reported structure of IDUA in complex with the inhibitor IdoA-DNJ, in which the pseudo-glycoside was observed in a $^{2}S_{0}$ conformation (predicted Michaelis complex conformation), allows the full conformational itinerary for IDUA to be structurally defined. The structures of IDUA with IdoADNJ (Michaelis complex) and **3** (transition state complex) and the fragment of compound **1** (covalent complex) were overlapped at the C3 and C4 atoms (**Fig. 5.7G**). This clearly shows the electrophilic migration from the Michaelis complex in a $^{2}S_{0}$ conformation, through the transition state in a $^{2,5}B$ conformation, to the covalent intermediate in a $^{5}S_{1}$ conformation. At the pseudo-anomeric carbon, there is a displacement of 0.74 Å from the Michaelis complex to the covalent intermediate. This is accompanied by a small (0.23 Å) movement at C2, but larger movement at C5 (0.72 Å) and the atom at the position of the endocyclic oxygen (0.70 Å), presumably to bring about the required migration at the anomeric position.

5.2.6 Visualizing rIDUA uptake in normal and patient cells

Finally, the question whether ABP **2** could be used to study rIDUA cellular uptake and lysosomal internalization was investigated. The majority of therapeutic glycosidases are amended with mannose 6-phosphate (M6P) residues for their recognition by M6P receptors (MPRs) on the plasma membrane and consequent transport to the lysosomes. In order to track the rIDUA

within cells, NHDF and fibroblasts of MPS I and ML II patients were fed with pre-labeled rIDUA-ABP **2**, and after fixation, cells were analyzed by confocal fluorescence microscopy. After 16 h incubation, punctate fluorescence structures were observed, which colocalized with the signals from the antibody-labeled lysosomal-associated membrane protein 1 (LAMP1), indicating lysosomal uptake of labeled rIDUA by all analyzed fibroblast models (**Fig. 5.8A**). The fluorescence from labeled rIDUA was absent in cells without the addition of rIDUA, as well as cells pre-incubated with M6P, which blocks M6P receptors (**Fig. 5.8A**). This finding was recapitulated in a parallel experiment, in which identically treated cells were lysed and detected for Cy5 fluorescence by SDS-PAGE and fluorescent scanning, revealing that only cells treated with ABP **2**-labeled rIDUA and without M6P pre-incubation contained bands corresponding to rIDUA (**Fig. 5.8B**). Interestingly, two extra bands that were smaller than the original rIDUA (85 kDa) were observed in the NHDF lysates, whereas only one band of 74 kDa and one band of 85 kDa was observed in lysates of MPS I or ML II fibroblasts, respectively. The 74 kDa and lower molecular bands were possibly rIDUA processed in the lysosomes of NHDF,[26] and this difference in labeling patterns could therefore indicate difference in the lysosomal function or endosomal uptake between normal and patient cells. Altogether, these results clearly demonstrated that rIDUA lysosomal uptake is mediated by MPRs, in agreement with earlier literature,[27] and that ABP **2** can be used to study the trafficking, localization, and lysosomal processing of rIDUA within cells derived from MPSI and MLII patients.

5.3 Discussion

In this chapter, the characterization and application of newly synthesized α-L-iduronidase inhibitors and ABPs were discussed. α-L-ido-configured cyclophellitol aziridine was synthesized as a key intermediate for generation of the first α-L-iduronidase ABPs. The α-L-ido-configured cyclohexene **8** could be an interesting starting point for the development of new IDUA inhibitors or chaperones. With the inhibitors described herein, ABPP studies on rIDUA has been conducted, showing that ABP **2** irreversibly labels rIDUA in a concentration- and time-dependent manner, with optimum labeling at pH 4.5–5.0. The covalent rIDUA inhibition has been further demonstrated by nano-LC-MS/MS, detecting a 7-amino-acid peptide fragment of

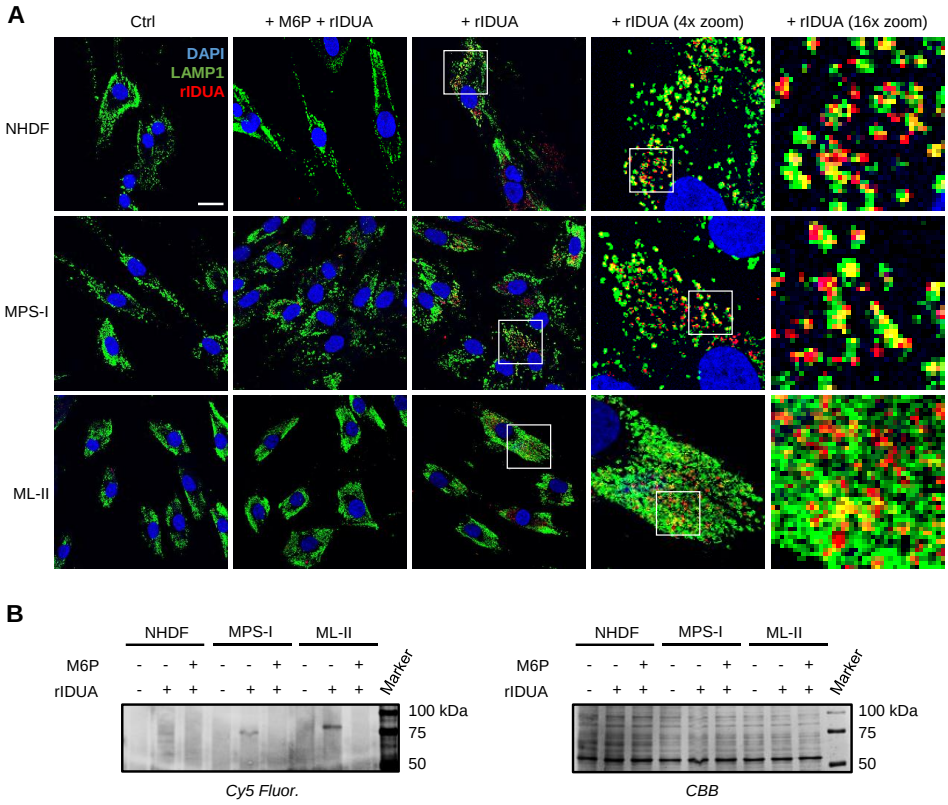


Figure 5.8. rIDUA visualization in human fibroblasts by confocal fluorescence microscopy and SDS-PAGE. A) From top to bottom: NHDF, human normal dermal fibroblasts; MPS I, patient fibroblasts with mucopolysaccharidosis type I, and ML II, patient fibroblasts with mucolipidosis type II. From left to right: cells were incubated without (Ctrl) or with ABP **2**-prelabeled rIDUA (+rIDUA, with successive zoomed-in images from areas within the indicated white squares), or pre-treated with man-nose-6-phosphate prior to rIDUA incubation (+M6P +rIDUA). Color legend: nuclei were stained with DAPI (blue), lysosomes with immunostaining of lysosomal-associated membrane protein 1 (LAMP1) (green), and rIDUA was labeled with ABP **2** (red). Scale bar =25 mm. B) Lysates of cells with identical treatment, detected for Cy5 fluorescence by fluorescent scanning on the wet slab gel, after samples were resolved by SDS-PAGE (left panel) The same gel was stained for Coomassie Brilliant Blue (*CBB*) for analysis of protein loading amount (right panel).

rIDUA containing the nucleophilic residue bound to ABP **3**.

It was found that the α-L-iduronidase inhibitors and ABPs exhibited much lower inhibitory potency and slower binding kinetics compared to earlier generated cyclophellitol aziridines conformers, such as the β-gluco-aziridine ABP JJB367 for GBA[20]. This was explained by DFT

calculations, which showed that inhibitors **1–3** adopt mainly a half-chair conformation ($^{4}H_{3}$ and $^{3}H_{4}$) in solution, with the $^{2,5}B$ conformation also as relevant geometry at an extra cost of 8.0 kcal mol^{-1}. Furthermore, crystallographic studies showed that compounds **1–3** bind IDUA in a $^{2,5}B$ boat conformation in the Michaelis complex. As a consequence, the half-chair conformations ($^{4}H_{3}$ and $^{3}H_{4}$) of inhibitors **1–3**, as predicted by DFT calculations, differ from any of the reaction itinerary conformations of α-L-iduronidase depicted in **Fig. 5.1A**, and thus inhibitors **1–3** may need to overcome this 8.0 kcal mol^{-1} energetic barrier to adopt the Michaelis complex transition state ($^{2,5}B$) conformation. Therefore, it is likely that the lower potency and slower inhibition kinetics of α-L-idoA-aziridine analogues on rIDUA is a manifestation of this energetic barrier.

Structural evidence for covalent addition of the ABPs to the nucleophilic residue of raIDUA was also provided by crystallographic studies. In the process, the conformation of the cyclophellitol aziridine at the transition state and in the covalent intermediate was defined, which supports earlier predictions concerning the conformational itinerary followed by α-L-iduronidase.[7] The insights gained through these studies should help in the design of closer conformational $^{2,5}B$ analogues by the use of different electrophilic traps or reactive species for the generation of improved inhibitors or molecular chaperones, with the end goal being the future provision of improved therapies for MPS I patients.

IDUA labeling with ABP **2** in complex biological samples—such as cell lysates, was not successful due to lower potency and lack of relative selectivity of this analogue at the applied concentration (low μM) compared with previously reported glycosidase probes. Future optimization on ABP labeling conditions and IDUA purification procedures from the desired sample are warranted, in order to utilize the ABPs for studying endogenous IDUA from clinical or laboratory samples. On the other hand, it was demonstrated that ABP **2** can be used to study the localization and trafficking of therapeutic rIDUA within cultured cells. It was shown that the rIDUA–ABP **2** complex is recognized by MPRs and internalized in lysosomes, and possibly processed differently between normal and patient fibroblasts.

5.4 Experimental procedures

5.4.1 Materials

Recombinant human iduronidase (rIDUA) was obtained from Genzyme (Aldurazyme®). Normal human dermal fibroblasts (NHDF) were obtained from Lonza. Human patient fibroblasts (MPS-I and ML-II) were obtained with consent from donors from the Academic Medical Center in Amsterdam, the Netherlands. 4-methylumbelliferyl α-L-iduronide (4-MU-α-L-IdoA) was purchased from Glycosynth. Pierce™ Bicinchoninic acid (BCA) protein assay kit and Pierce™ Polyacrylamide Spin Desalting Columns 7K MWCO was acquired from Thermo Fisher Scientific. All other chemicals were obtained from commercial sources.

5.4.2 Cell culture and lysate preparation

Fibroblasts were cultured in Dulbecco's modified Eagle's medium: Nutrient Mixture F-12 (DMEM/F-12, Invitrogen), containing 10 % (v/v) heat-inactivated fetal calf serum, 200 μg/mL penicillin, 200 μg/mL streptomycin, at 37 °C and 5 % CO_2. Confluent Fibroblasts were subcultured at a ratio of 1:4 each week. To prepare lysates, cells were washed three times with phosphate-buffered saline (PBS), detached by scraping in the presence of ice-cold lysis 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 tablet (Roche ,version 12)), and collected in Eppendorf tubes. The collected suspension was vortexed vigorously, incubated on ice for 1 h, freeze-thawed once using liquid nitrogen, and stored at –80 °C. Concentration of lysates was determined using BCA kit.

5.4.3 Iduronidase activity assay using 4-MU-α-L-IdoA as substrate

Samples containing enzyme were diluted in assay buffer (150 mM citrate/Na_2HPO_4, pH 4.5, supplemented with 0.1 % (v/v) Triton X-100, 0.15 M NaCl, 33 mM $CaCl_2$, 33 mM $MgCl_2$), and load onto a black flat-bottom 96-well plate (Greiner) in triplicates of 25 μL per well. Negative controls (blank) were prepared by substituting enzyme samples with buffer. For reaction, samples were added with 25 μL of 360 μM 4-MU-α-L-IdoA[28] (prepared in assay buffer supplemented with 0.1 % (w/v) bovine serum albumin (BSA)), and incubated at 37 °C for 30

min to 4h. Reaction was stopped by adding 200 μL 1 M Glycine-NaOH (pH 10.3), and fluorescence was measured using a LS-55 fluorometer (PerkinElmer) at λ_{ex} = 366 nm and λ_{em} = 445 nm.

5.4.4 Apparent IC_{50} values of inhibitors and ABPs on recombinant human iduronidase and cell lysates

Inhibitors and ABPs were diluted at various concentrations in 12.5 μL assay buffer and incubated with 12.5 μL rIDUA (5.8 ng, or 70 fmol) or fibroblast lysates (50 μg protein), both diluted in assay buffer, at 37°C for 1 h (rIDUA) or 4 h (fibroblasts lysates), at 1 % (v/v) DMSO concentration. This was followed by iduronidase activity assay (30 min incubation time) described in the previous section. The detected 4-MU fluorescence at each concentration for each compound were normalized to the fluorescence from the control sample without inhibitor, and data were fitted with [inhibitor] *vs* response – variable slope (four parameters) function in GraphPad Prism 7.0 software to obtain apparent IC_{50} values.

5.4.5 Labeling and SDS-PAGE of recombinant human iduronidase

To prepare for labeling, rIDUA stock (0.58 μg/μL in PBS) was diluted with assay buffer to final concentration of 10 ng (120 fmol) in 14 μL assay buffer, if not otherwise specified. ABP **2** stock (15 mM in DMSO) was diluted subsequently with DMSO and then assay buffer, to 8x of intended assay concentration obtaining 4 % (v/v) DMSO. For labeling, 2 μL of ABP **2** dilution was incubated with 14 μL enzyme dilution at 37 °C for the intended time periods. Labeling was terminated by denaturing samples with 4 μL sample buffer (5x Laemmli buffer, containing 0.3 M Tris-HCl pH 6.8, 50 % (v/v) 100 % glycerol, 8 % (w/v) dithiothreitol (DTT), 10 % (w/v) sodium dodecyl sulfate (SDS), 0.01 % (w/v) bromophenol blue) and heated at 98 °C for 5 minutes. SDS-PAGE and fluorescence detection procedures followed the previously described methods.[21]

5.4.6 Labeling of recombinant human iduronidase at different pH

Assay buffers with pH 2.5–8.0 were used to dilute rIDUA and ABP **2**. The experiment was

performed with 10 ng rIDUA at 10 μM ABP **2** (assay concentration) in 16 μL volume and 1 h incubation time, followed by denaturation, SDS-PAGE, and fluorescence scan.

5.4.7 Reversibility of labeling of ABP **2** on recombinant human iduronidase

5 μL rIDUA stock (0.58 μg/μL in PBS) was diluted with 20 μL assay buffer without Triton X-100, and incubated with either 25 μL ABP **2** (150 μM and 1 % (v/v) DMSO) or 25 μL assay buffer containing 0.1 % (w/v) DMSO for 1 h at 37 °C. Thereafter, 45 μL of sample was passed through a desalting column, and diluted with 216 μL assay buffer (with Triton X-100) to a final enzyme concentration of 1 ng/μL. For assessing the reversibility of labeling/inactivation of ABP **2** on rIDUA, both samples were kept at 4 °C until subjecting to iduronidase activity assay using 4-MU-α-IdoA at the following time points after desalting: 0.5, 28, and 100 h. Iduronidase activity assay was performed with 1 μL sample in triplicates, and assay buffer was used as blank.

5.4.8 Competitive ABPP experiments

For the competitive ABPP (cABPP) experiment using 4-MU-α-IdoA as an active-site directed inhibitor, 10 ng rIDUA, or no-enzyme blank (PBS), was prepared in 7 μL assay buffer (pH 4.5). This was pre-incubated with 7 μL assay buffer containing 4-MU-α-IdoA (0.2 -20 mM) on ice for 5 min, then incubated with 1 μL ABP **2** (750 μM, 8 % (v/v) in assay buffer) or DMSO control (8 % (v/v) in assay buffer) at 37 °C for 2 h. In the same experiment, 7 μL SDS (4 % (w/v)) was added to the 7 μL enzyme, pre-incubated at 98 °C for 2 min, and incubated with ABP **2** at 37 °C for 2 h. For the cABPP experiment with compound **1**, 10 ng rIDUA was prepared in 12 μL assay buffer, pre-incubated with 2 μL compound **1** (2.8 mM – 70 μM, 18.6% DMSO in assay buffer) at 37 °C for 2 h, then incubated with 2 μL ABP **2** (100 μM, 4 % DMSO in assay buffer) at 37 °C for 2 h. Samples were then denatured, and subjected to SDS-PAGE and fluorescence scan.

5.4.9 Kinetic parameters of ABP **2**

The kinetic parameters of ABP **2** for rIDUA was determined by an SDS-PAGE-based assay, in which the intensity of fluorescent signal from ABP **2**-labeled rIDUA on the wet slab gel is proportional to the extent of rIDUA inactivation, which was normalized to the signal from a

control group with saturate labeled rIDUA. rIDUA was labeled with ABP **2** for different time periods and at different ABP **2** concentrations (which were >> enzyme concentration), and the kinetic parameters k_{inact} (pseudo first-order inactivation rate constant) and K_I (inhibition constant) were obtained using non-linear curve fitting. To perform the experiment, three sets of rIDUA were diluted in series of 1.5 mL Eppendorf tubes to 10 ng (120 fmol) per 14 μL, and labeled with 2 μL ABP **2** (diluted in DMSO and assay buffer to various concentrations and [DMSO] fixed at 8 % (v/v)) for 30, 60, 90, 120, or 150 min. Concentrations in assay were 5 μM, 10 μM, 20 μM, 30 μM, 40 μM, 50 μM, or 60 μM for ABP **2**, 7.5 nM for rIDUA, and 1 % (v/v) for DMSO. Reaction was terminated by incubating samples with sample buffer at 98°C for 5 min. Samples were subjected to SDS-PAGE and fluorescence scan, and Cy5 fluorescence from the ABP **2**-labeled rIDUA was quantified using ImageQuant software (GE Healthcare). The incubation condition for maximum labeling (complete inactivation) on rIDUA for ABP **2** was determined to be 60 μM and 150 min, and a control sample with this condition was loaded on every gel to allow normalization of signals form samples in the same gel. After normalization, the value at each ABP **2** concentration were plotted in a time *vs* % labeling (inactivation) graph, and the plotted data were fitted with one-phase exponential association function to obtain the rate constant k at each ABP **2** concentration. Finally, in a second plot the obtained k values were plotted against ABP **2** concentrations, and the data were fitted with a Michaelis-Menten equation to obtain k_{inact} and K_I values for ABP **2** on rIDUA for each set (n = 3). All non-linear curve-fitting was performed using GraphPad Prism 7.0 software.

5.4.10 LC-MS/MS identification of rIDUA active site peptide

A total of 10 μg rIDUA was diluted in assay buffer and incubated with either 75 μM ABP **3** or DMSO (negative control) for 1 h at 37 °C in 100 μL volume ([DMSO] = 0.5 % (v/v)). The samples were then added with 100 μg BSA (100 μL), and followed by chloroform/methanol precipitation and reduction/alkylation procedures described previously.[29] Consequently, the samples were dissolved in 2 % (w/v) SDS and diluted with 50 mM Tris-HCl (pH 7.8) to a final SDS concentration of 0.05 % (w/v). The samples were then concentrated with size exclusion columns (Amicon 10 k) to a volume of 74 μL, and digested O/N at 25 ºC with 1.1 μg

Chymotrypsin (Promega) in the presence of 10 mM $CaCl_2$. Digested peptides were pulled-down using 50 µL of Streptavidin paramagnetic beads (MyOne T1, ThermoFisher) in 1 mL of pull-down buffer (50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 0.5 % (w/v) SDS) for 1 h at RT under vigorous shaking. The beads were washed stringently following previously described procedures[29] and eluted with 100 µL of elution buffer (25 % (v/v) acetonitrile, 5 % (v/v) formic acid, 70 H_2O, 10 µM biotin) for 30 min at 37°C. Afterwards, acetonitrile in the supernatant were evaporated using a Speedvac at 45°C, and this was followed by desalting using StageTips. The eluate was evaporated and reconstituted in 20 µL of LC-MS sample solution (95: 3: 0.1, H_2O: acetonitrile: formic acid) for LC-MS/MS analysis. Peptide samples were analyzed with a two-hour gradient of 5 % to 25 % acetonitrile on nano-LC, hyphenated to an LTQ Orbitrap and identified by manual search for the theoretical m/z values of the active site peptide and its MS/MS fragments labeled with ABP **3**.

5.4.11 Comparing specific activity and ABP labeling between GBA and IDUA

For measurement of GBA activity, 100 fmol of rGBA (Cerezyme/Imiglucerase, Genzyme) or NHDF lysates (10 µg protein) were diluted in 25 µL of GBA buffer (150 mM citric acid/Na_2HPO_4, 0.1 % (v/v) Triton X-100, 0.2 % (w/v) sodium taurocholate) and incubated with 100 µL GBA substrate mixture (3.75 mM 4-MU-β-D-glucopyranoside (Glycosynth), in GBA buffer, 0.1 % (w/v) BSA) for 30 min at 37 °C. IDUA activity was measured following the methods described in previous section. For ABP labeling of GBA, rGBA (100 – 0.1 fmol) or NHDF lysates (10 µg protein) were diluted in 14 µL GBA buffer, and incubated with 0.5 µM JJB367[21] for 30 min at 37 °C ([DMSO] = 0.5 (v/v)). For ABP labeling of IDUA, rIDUA (1–100 fmol) or NHDF lysates (10 µg protein) were diluted in 14 µL assay buffer and incubated with either 50 µM or 25 µM ABP **2** for 4 h at 37 °C ([DMSO] = 0.5 % (v/v)). Samples were denatured, and proceeded to SDS-PAGE and fluorescent detection.

5.4.12 Recombinant expression and purification of IDUA in seeds of *Arabidopsis thaliana* (University of St Andrews)

Recombinant human α-L-iduronidase (abbreviated to raIDUA to distinguish from rIDUA

obtained from Genzyme) was produced in seeds of *Arabidopsis thaliana cgl* (***c***omplex ***gl***ycan deficient) line A4.7[30] in which the seeds (T3 generation) accumulated raIDUA to 7.2 ± 0.6 % total soluble protein (9.8 μg/mg dry seeds). raIDUA was purified to homogeneity from the T3 seeds using concanavalin A-sepharose and anti-IDUA affinity chromatography as described previously[31, 32]. In human IDUA, there are six N-linked glycosylation sites. The oligosaccharide structures at each site of rIDUA secreted from a Chinese hamster ovary (CHO) cell line have been determined by mass spectrometry[33]. The raIDUA expressed in the seeds of the *cgl* mutant of *Arabidopsis* has much reduced complexity in these N-linked glycans, the majority of which are non-matured, high mannose N-glycans[30, 34].

5.4.13 Crystallization of raIDUA (University of St Andrews)

raIDUA was further purified by size exclusion chromatography using an S200 10/300 column (GE Healthcare) equilibrated in 20 mM Tris, pH 7.0, 500 mM NaCl, and 0.02 % sodium azide. The fractions containing raIDUA were buffer exchanged using a PD10 desalt column (GE Healthcare) into 20 mM dimethylglutaric acid, pH 6.0, 0.2 M NaCl, 5 % (v/v) glycerol, and 5 % (v/v) ethanol and concentrated to 10 mg/ml for crystallization. Crystallization was performed in a 24 well plate using hanging-drop vapor diffusion. The rhomboid-plate shaped crystals grew at room temperature from 0.1 M HEPES, pH 7.5, 0.26 M sodium potassium tartrate, 20 % (w/v) polyethylene glycol 3350 (optimized from the crystallization condition reported by Bie *et al.*[8]). Crystals were soaked in mother liquor containing a minute amount of solid **1** for 24 hours or **3** for 45 minutes, before being harvested. Crystals were cryo-protected in a solution containing the mother liquor plus 30 % glycerol prior to vitrification in liquid nitrogen.

5.4.14 Data collection and processing for raIDUA crystals (University of St Andrews)

X-ray diffraction data were collected at Diamond Light Source (DLS) on beamlines I03 and I04; the data processing and refinement statistics can be found in **Table 4.S1**. Diffraction data were processed either using the FastDP pipeline[35] (which utilizes XDS[36] with Aimless[37]) or Xia2[38] (also with XDS[36] with Aimless[37]). Molecular replacement was performed using Phaser[39] with Protein Data Bank (PDB) entry 4JXO as the search model. Refinement was performed using

REFMAC5[40] and manual model building was done using Coot[41]. Structures were validated using PDB_REDO[42]. Models for the fragments of **1** and **3** were built in JSME[43] and the libraries generated with PRODRG[44].

5.4.15 Fibroblasts uptake of ABP **2**-labeled rIDUA

rIDUA was diluted to 58 ng/μL and labeled with ABP **2** at 75 μM for 1 h at 37 °C in a volume of 40 μL. After incubation, unbound ABP **2** in solution was removed by passing through desalting columns (Pierce™ Polyacrylamide Spin Desalting Columns 7K MWCO), and the eluate were diluted in assay buffer without Triton X-100, to a final rIDUA concentration of 10 ng/μL. For the uptake experiment, human normal and patient fibroblasts were sub-cultured 1 day before treatment in 12-well plates (1 mL culture medium per well) with or without glass coverslips. Cells were then pre-treated with 4 mM mannose-6-phosphate (M6P, diluted in water) or same volume of water for 1 h, followed by treating with ABP **2**-labeled rIDUA (100 ng/mL culture medium) for 20 h, without removing M6P. A control group was included for each cell type, treated only with water and assay buffer. Confocal microscopy analysis was carried out following a previously described procedure[45], where samples were fixed with 4 % formaldehyde, permeabilized with 0.1 % (w/v) saponin and 2 % (w/v) BSA, and immuno-stained for the lysosomal membrane protein LAMP1 using mouse anti human LAMP1 (Southern Biotech) as primary antibody, and donkey anti mouse Alexa488 (Molecular Probes) as secondary antibody. The coverslips were mounted to microscopy slides using ProLong™ Diamond Antifade Mountant with DAPI (Thermo Fisher), and scanned using a Leica SP8 confocal microscope for DAPI, Alexa488, and Cy5 fluorescence with a 40x oil-immersed objective. Pictures at each fluorescence channel were captured at 1024 x 1024 resolution, with $n = 3$ frame averages.

5.5 References

1 Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM & Henrissat B (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res* **42**, D490–495.

2 Nieman CE, Wong AW, He S, Clarke L, Hopwood JJ & Withers SG (2003) Family 39 alpha-l-iduronidases and beta-D-xylosidases react through similar glycosyl-enzyme intermediates: identification of the human iduronidase nucleophile. *Biochemistry* **42**, 8054–8065.

3 Archer LD, Langford-Smith KJ, Bigger BW & Fildes JE (2014) Mucopolysaccharide diseases: a complex interplay between neuroinflammation, microglial activation and adaptive immunity. *J Inherit Metab Dis* **37**, 1–12.

4 Clarke LA (2008) The mucopolysaccharidoses: a success of molecular medicine. *Expert Rev Mol Med* **10**, e1.

5 Maita N, Tsukimura T, Taniguchi T, Saito S, Ohno K, Taniguchi H & Sakuraba H (2013) Human α-L-iduronidase uses its own N-glycan as a substrate-binding and catalytic module. *Proc Natl Acad Sci U S A* **110**, 14628–14633.

6 Koshland DE (1953) Stereochemistry and the mechanism of enzymatic reactions. *Biol Rev* **28**, 416–436.

7 Speciale G, Thompson AJ, Davies GJ & Williams SJ (2014) Dissecting conformational contributions to glycosidase catalysis and inhibition. *Curr Opin Struct Biol* **28**, 1–13.

8 Bie H, Yin J, He X, Kermode AR, Goddard-Borger ED, Withers SG & James MN (2013) Insights into mucopolysaccharidosis I from the structure and action of α-L-iduronidase. *Nat Chem Biol* **9**, 739–745.

9 Grewal SS, Wynn R, Abdenur JE, Burton BK, Gharib M, Haase C, Hayashi RJ, Shenoy S, Sillence D, Tiller GE, Dudek ME, van Royen-Kerkhof A, Wraith JE, Woodard P, Young GA, Wulffraat N, Whitley CB & Peters C (2005) Safety and efficacy of enzyme replacement therapy in combination with hematopoietic stem cell transplantation in Hurler syndrome. *Genet Med* **7**, 143–146.

10 Clarke LA (1993) Mucopolysaccharidosis Type I. *GeneReviews*. Seatle, WA: University of Washington.

11 Coutinho MF, Prata MJ & Alves S (2012) Mannose-6-phosphate pathway: a review on its role in lysosomal function and dysfunction. *Mol Genet Metab* **105**, 542–550.

12 Harrak Y, Barra CM, Delgado A, Castaño AR & Llebaria A (2011) Galacto-configured aminocyclitol phytoceramides are potent in vivo invariant natural killer T cell stimulators. *J Am Chem Soc* **133**, 12079–12084.

13 Jiang J, Kallemeijn WW, Wright DW, van den Nieuwendijk AMCH, Rohde VC, Folch EC, van den Elst H, Florea BI, Scheij S, Donker-Koopman WE, Verhoek M, Li N, Schürmann M, Mink D, Boot RG, Codée JDC, van der Marel GA, Davies GJ, Aerts JMFG & Overkleeft HS (2015) In vitro and in vivo comparative and competitive activity-based protein profiling of GH29 α-l-fucosidases. *Chem Sci* **6**, 2782–2789

14 Clements PR, Brooks DA, Saccone GT & Hopwood JJ (1985) Human alpha-L-iduronidase. 1. Purification, monoclonal antibody production, native and subunit molecular mass. *Eur J Biochem* **152**, 21–28.

15 Schuchman EH, Guzman NA, Takada G & Desnick RJ (1984) Human alpha-L-iduronidase. II. Comparative biochemical and immunologic properties of the purified low and high uptake forms. *Enzyme* **31**, 166–175.

16 Wu L, Jiang J, Jin Y, Kallemeijn WW, Kuo C-L, Artola M, Dai W, van Elk C, van Eijk M, van der Marel GA, Codee JDC, Florea BI, Aerts JMFG, Overkleeft HS & Davies GJ (2017) Activity-based probes for functional interrogation of retaining [beta]-glucuronidases. *Nat Chem Biol* **13**, 867–873.

17 Witte MD, Kallemeijn WW, Aten J, Li KY, Strijland A, Donker-Koopman WE, Van Den Nieuwendijk AMCH, Bleijlevens B, Kramer G, Florea BI, Hooibrink B, Hollak CEM, Ottenhoff R, Boot RG, Van Der Marel GA, Overkleeft HS & Aerts JMFG (2010) Ultrasensitive in situ visualization of active glucocerebrosidase molecules. *Nat Chem Biol* **6**, 907–913.

18 Kallemeijn WW, Li KY, Witte MD, Marques ARA, Aten J, Scheij S, Jiang J, Willems LI, Voorn-Brouwer TM, Van Roomen CPAA, Ottenhoff R, Boot RG, Van Den Elst H, Walvoort MTC, Florea BI, Codée JDC, Van Der Marel GA, Aerts JMFG & Overkleeft HS (2012) Novel activity-based probes for broad-spectrum profiling of retaining β-exoglucosidases in situ and in vivo. *Angew Chemie - Int Ed* **51**, 12529–12533.

19 Jiang J, Kuo CL, Wu L, Franke C, Kallemeijn WW, Florea BI, Van Meel E, Van Der Marel GA, Codée JDC, Boot RG, Davies GJ, Overkleeft HS & Aerts JMFG (2016) Detection of active mammalian GH31 α-glucosidases in health and disease using in-class, broad-spectrum activity-based probes. *ACS Cent Sci* **2**, 351–358.

20 Willems LI, Beenakker TJ, Murray B, Scheij S, Kallemeijn WW, Boot RG, Verhoek M, Donker-Koopman WE, Ferraz MJ, van Rijssel ER, Florea BI, Codée JD, van der Marel GA, Aerts JM & Overkleeft HS (2014) Potent

and selective activity-based probes for GH27 human retaining α-galactosidases. *J Am Chem Soc* 2014 **136**, 11622–11625.

21 Schröder SP, van de Sande JW, Kallemeijn WW, Kuo CL, Artola M, van Rooden EJ, Jiang J, Beenakker TJM, Florea BI, Offen WA, Davies GJ, Minnaard AJ, Aerts JMFG, Codée JDC, van der Marel GA & Overkleeft HS (2017) Towards Broad Spectrum Activity-Based Glycosidase Probes: Synthesis and Evaluation of Deoxygenated Cyclophellitol Aziridines. *Chem Commun* **53**, 12528–12531.

22 Shao Y, Molnar LF, Jung Y, Kussmann J, Ochsenfeld C, Brown ST, Gilbert AT, Slipchenko LV, Levchenko SV, O'Neill DP, DiStasio RA Jr, Lochan RC, Wang T, Beran GJ, Besley NA, Herbert JM, Lin CY, Van Voorhis T, Chien SH, Sodt A, Steele RP, Rassolov VA, Maslen PE, Korambath PP, Adamson RD, Austin B, Baker J, Byrd EF, Dachsel H, Doerksen RJ, Dreuw A, Dunietz BD, Dutoi AD, Furlani TR, Gwaltney SR, Heyden A, Hirata S, Hsu CP, Kedziora G, Khalliulin RZ, Klunzinger P, Lee AM, Lee MS, Liang W, Lotan I, Nair N, Peters B, Proynov EI, Pieniazek PA, Rhee YM, Ritchie J, Rosta E, Sherrill CD, Simmonett AC, Subotnik JE, Woodcock HL 3rd, Zhang W, Bell AT, Chakraborty AK, Chipman DM, Keil FJ, Warshel A, Hehre WJ, Schaefer HF 3rd, Kong J, Krylov AI, Gill PM & Head-Gordon M (2006) Advances in methods and algorithms in a modern quantum chemistry program package. *Phys Chem Chem Phys* **8**, 3172–3191.

23 Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Men-nucci B, Petersson GA, Nakatsuji H, Caricato M, Li X, Hratchian HP, Izmaylov AF, Bloino J, Zheng G, Sonnenberg JL, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven H, Montgomery Jr JA, Peralta JE, Ogliaro F, Bearpark M, Heyd JJ, Brothers E, Kudin KN, Staroverov VN, Ko-bayashi R, Normand J, Raghavachari K, Rendell A, Burant JC, Iyen-gar SS, Tomasi J, Cossi M, Rega N, Millam JM, Klene M, Knox JE, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Martin RL, Morokuma K, Zakrzewski VG, Voth GA, Salvador P, Dannenberg JJ, Dapprich S, Daniels AD, Farkas O, Foresman JB, Ortiz JV, Cio-slowski J & Fox DJ (2009) Gaussian 09, Revision D.01 [Computer software]. Wallingford CT: Gaussian, Inc.

24 Namchuk MN & Withers SG (1995) Mechanism of Agrobacterium beta-glucosidase: kinetic analysis of the role of noncovalent enzyme/substrate interactions. *Biochemistry* **34**, 16194–16202.

25 Tanaka KSE, Zhu J, Huang XC, Lipari F & Bennet AJ (2000) Glycosidase-catalyzed hydrolysis of 2-deoxyglucopyranosyl pyridinium salts: effect of the 2-OH group on binding and catalysis. *Can J Chem* **78**, 577–582.

26 Unger EG, Durrant J, Anson DS & Hopwood JJ (1994) Recombinant alpha-L-iduronidase: characterization of the purified enzyme and correction of mucopolysaccharidosis type I fibroblasts. *Biochem J* **304**, 43–49.

27 Hasilik A & Neufeld EF (1980) Biosynthesis of lysosomal enzymes in fibroblasts. Synthesis as precursors of higher molecular weight. *J Biol Chem* **255**, 4937–4945.

28 Ou L, Herzog TL, Wilmot CM & Whitley CB (2014) Standardization of alpha-L-iduronidase enzyme assay with Michaelis-Menten kinetics. *Mol. Genet. Metab* **111**, 113–115.

29 Li N, Kuo CL, Paniagua G, van den Elst H, Verdoes M, Willems LI, van der Linden WA, Ruben M, van Genderen E, Gubbens J, van Wezel GP, Overkleeft HS, Florea BI. (2013) Relative quantification of proteasome activity by activity-based protein profiling and LC-MS/MS. *Nat Protoc*, **8**, 1155–1168.

30 He X, Pierce O, Haselhorst T, von Itzstein M, Kolarich D, Packer NH, Gloster TM, Vocadlo DJ, Qian Y, Brooks D & Kermode AR (2013) Characterization and downstream mannose phosphorylation of human recombinant alpha-L-iduronidase produced in Arabidopsis complex glycan-deficient (cgl) seeds. *Plant Biotechnol J* **11**, 1034–1043.

31 He X, Haselhorst T, von Itzstein M, Kolarich D, Packer NH, Gloster TM, Vocadlo DJ, Clarke LA, Qian Y & Kermode AR (2013) Production of alpha-L-iduronidase in maize for the potential treatment of a human lysosomal storage disease. *Nat Commun* **3**, 1062.

32 He X, Haselhorst T, von Itzstein M, Kolarich D, Packer NH & Kermode AR (2012) Influence of an ER-retention signal on the N-glycosylation of recombinant human alpha-L-iduronidase generated in seeds of Arabidopsis. *Plant Mol Biol* **79**, 157–169.

33 Zhao KW, Faull KF, Kakkis ED & Neufeld EF (1997) Carbohydrate structures of recombinant human alpha-L-iduronidase secreted by Chinese hamster ovary cells. *J Biol Chem* **272**, 22758–22765.

34 Pierce OM, McNair GR, He X, Kajiura H, Fujiyama K & Kermode AR (2017) N-glycan structures and downstream mannose-phosphorylation of plant recombinant human alpha-L-iduronidase: toward development of enzyme replacement therapy for mucopolysaccharidosis. *Plant Mol Biol* **95**, 593–606.

35 Winter G & McAuley KE (2011) Automated data collection for macromolecular crystallography. *Methods* **55**, 81–93.

36 Kabsch W (2010) XDS. *Acta Crystallogr D Biol Crystallogr* **66** , 125–132.

37 Evans P R & Murshudov GN (2013) How good are my data and what is the resolution? *Acta Crystallogr D Biol Crystallogr* **69**, 1204–1214.

38 Winter G. (2010) Xia2: An Expert System for Macromolecular Crystallography Data Reduction. *J. Appl. Crystallogr.* **43**, 186–190.

39 McCoy AJ (2007) Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr D Biol Crystallogr* **63**, 32–41.

40 Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F & Vagin AA (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* **67**, 355–367.

41 Emsley P, & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126–2132.

42 Joosten RP, Long F, Murshudov GN & Perrakis A (2014) The PDB_REDO server for macromolecular structure model optimization. *IUCrJ* **1**, 213–220.

43 Bienfait B & Ertl P (2013) JSME: a free molecule editor in JavaScript. *J Cheminform* **5**, 24.

44 Schuttelkopf AW & van Aalten DM (2004) PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr D Biol Crystallogr* **60**, 1355–1363.

45 Kuo CL, van Meel E, Kytidou K, Kallemeijn WW, Witte M, Overkleeft HS, Artola ME & Aerts JM (2018) Activity-Based Probes for Glycosidases: Profiling and Other Applications. *Methods Enzymol* **598**, 217– 235.

APPENDIX

5.S1. Supporting Figures and Tables

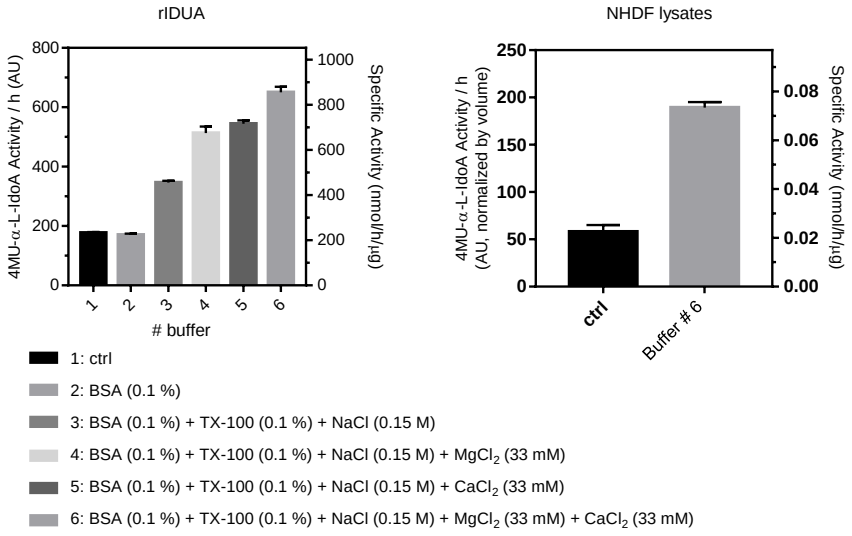


Figure 5.S1. Optimization of assay buffers for IDUA activity. Error range = SD from technical triplicates.

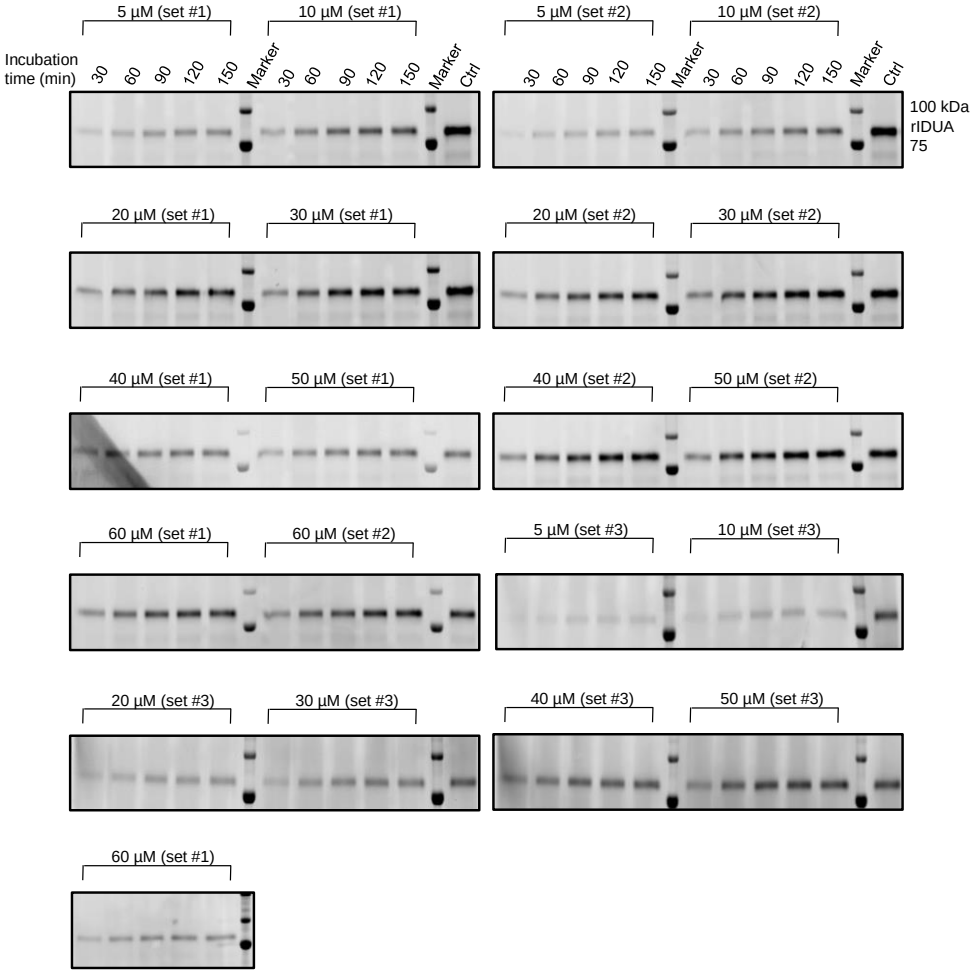


Figure 5.S2. SDS-PAGE gels used for labeling kinetic studies of ABP 2 with rIDUA. rIDUA was labeled at different incubation time periods (30–150 min) and at different ABP **2** concentrations (5–60 µM), before SDS-PAGE and fluorescent detection and quantification. The labeling experiment was performed in triplicate sets.

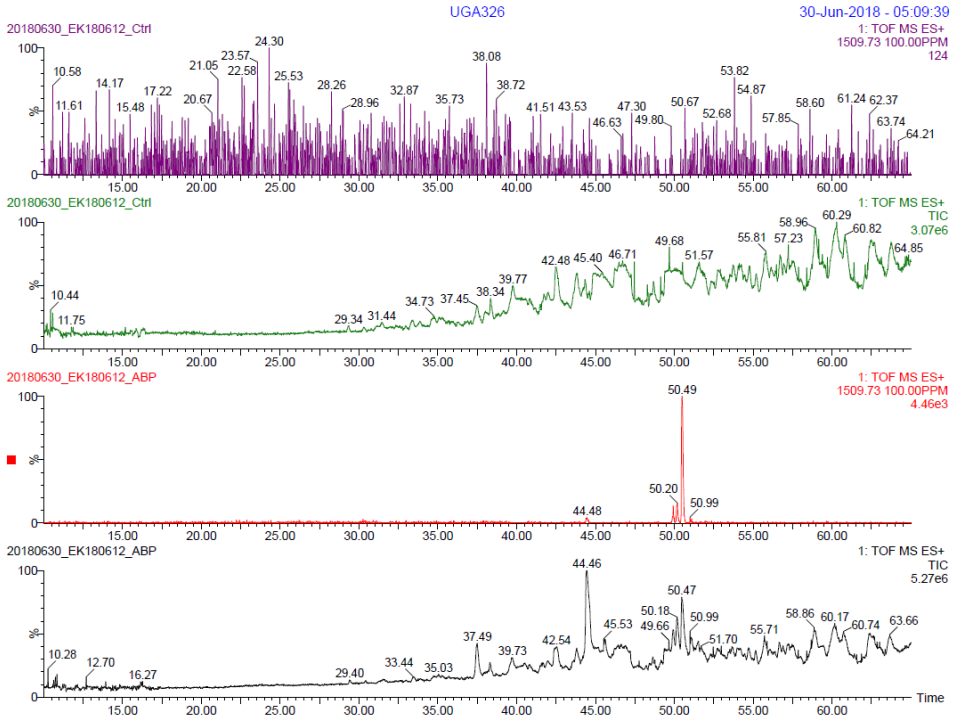
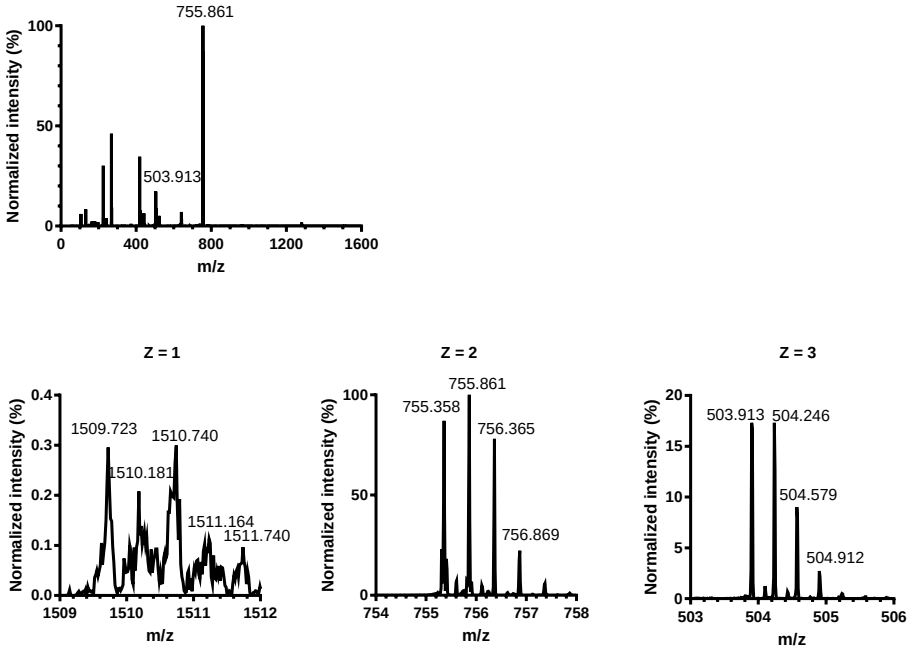


Figure 5.S3. Chromatogram of sample containing rIDUA active site peptide labeled without (Ctrl) or with ABP 3.

A



B

z	Theoretical		Experimental		Delta	ppm	
1	1509.725	1510.729	1509.723	1510.740	0.9999	1.52	7.28
2	755.3663	755.868	755.3578	755.8612	0.0534	11.25	9.00
3	503.9133	504.2478	503.9127	504.2455	0.3328	1.19	4.56

Figure 5.S4. LC-MS/MS identification of IDUA peptides labeled with ABP 3. A) Mass spectrum of sample containing rIDUA active site peptide labeled with ABP **3**. B) Theoretical and experimental m/z values for rIDUA active site peptide labeled with ABP **3**.

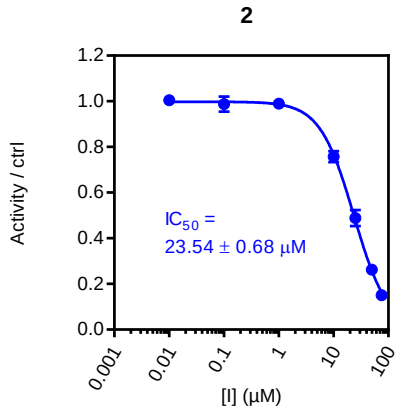


Figure 5.S5. Inhibition curve of ABP 2 on IDUA activity from NHDF lysates. Error range = SD from technical triplicates.

Table 5.S1. Data collection and refinement statistics for ralDUA in complex with fragments of 1 and 3.

	ralDUA in covalent complex with fragment of **1** (PDB: 6I6R)	ralDUA in non-covalent complex with fragment of **3** (PDB: 6I6X)
Data collection		
Space group	$P2_12_12_1$	$H3$
Cell dimensions		
a, *b*, *c* (Å)	206.9, 69.9, 93.7	259.2, 259.2, 71.0
α, β, γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 120.0
Resolution (Å)	66.25-2.02 (2.05-2.02)	29.73-2.39 (2.44-2.39)*
R_{merge}	0.164 (1.047)	0.124 (0.866)
R_{pim}	0.080 (0.50)	0.061 (0.433)
$I / \sigma I$	5.4 (1.3)	10.2 (1.3)
Completeness (%)	94.4 (97.8)	99.6 (94.0)
Redundancy	4.9 (5.0)	5.2 (4.8)
$CC_{1/2}$	0.99 (0.54)	0.995 (0.669)
Refinement		
Resolution (Å)	66.25-2.02	29.73-2.39
No. reflections	80558	66351
R_{work} / R_{free}	0.21 / 0.24	0.21 / 0.23
No. atoms		
Protein	9658	9818
Ligand	26	14
Water	250	294
B-factors		
Protein	38.6	43.2
Ligand	45.3	59.8
Water	35.2	39.2
Overall		
R.M.S. deviations		
Bond lengths (Å)	0.012	0.010
Bond angles (°)	1.56	1.45

*Values in parentheses are for highest-resolution shell.

5.S2 Synthesis of α-L-iduronic-configured inhibitors and ABPs (Department of Bio-organic Synthesis, Leiden University)

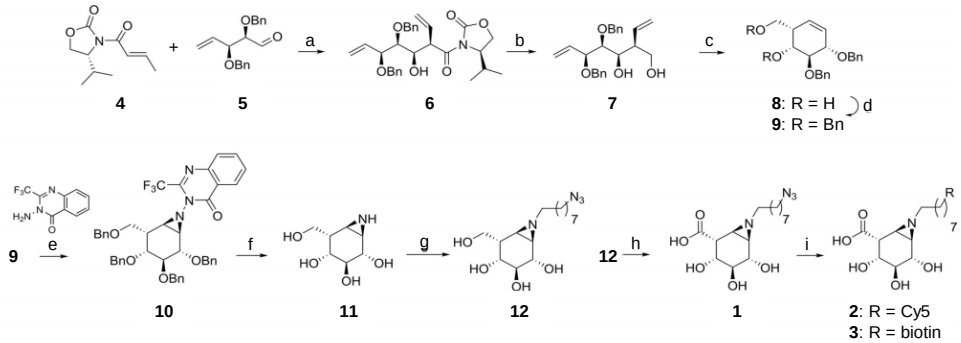
5.S2.1 General experimental details

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. Triethylamine was dried

over KOH and distilled before using. 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 analytical thin-layer chromatography (TLC) using Merck aluminum sheets pre-coated with silica gel 60 with detection by UV absorption (254 nm) and by spraying with a solution of $(NH_4)_6Mo_7O_{24}\cdot H_2O$ (25 g/L) and $(NH_4)_4Ce(SO_4)_4\cdot H_2O$ (10 g/L) in 10 % sulfuric acid followed by charring at ~150 °C or by spraying with an aqueous solution of $KMnO_4$ (7 %) and K_2CO_3 (2 %) followed by charring at ~150 °C. Column chromatography was performed manually using either Baker or Screening Device silica gel 60 (0.04 - 0.063 mm) or a Biotage Isolera™ flash purification system using silica gel cartridges (Screening devices SiliaSep HP, particle size 15-40 μm, 60A) in the indicated solvents. ^1H NMR and ^{13}C NMR spectra were recorded on 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 chloroform residual solvent peak or tetramethylsilane (TMS) as internal standard. Coupling constants are given in Hz. All given ^{13}C spectra are proton decoupled. The following abbreviations are used to describe peak patterns when appropriate: s (singlet), d (doublet), t (triplet), qt (quintet), m (multiplet), br (broad), ar (aromatic), app (apparent). 2D NMR experiments (HSQC, COSY and NOESY) were carried out to assign protons and carbons of the new structures and assignation follows the general numbering shown in cyclohexene **9**. High-resolution mass spectra (HRMS) of intermediates were recorded with a LTQ Orbitrap (Thermo Finnigan) and final compounds were recorded with an apex-QE instrument (Bruker). Optical rotations were measured on a Anton Paar MCP automatic polarimeter (Sodium D-line, λ = 589 nm). LC/MS analysis was performed on an LCQ Advantage Max (Thermo Finnigan) ion-trap spectrometer (ESI+) coupled to a Surveyor HPLC system (Thermo Finnigan) equipped with a C18 column (Gemini, 4.6 mm x 50 mm, 3 μm particle size, Phenomenex) equipped with buffers A: H_2O, B: acetonitrile (MeCN) and C: 1% aqueous TFA, or an Agilent Technologies 1260 Infinity LCMS with a 6120 Quadrupole MS system equipped with buffers A: H2O, B: acetonitrile (MeCN) and C: 100 mM NH4OAc. For reversed-phase HPLC-MS purifications an Agilent Technologies 1200 series prep LCMS with a 6130 Quadrupole MS system was used equipped with buffers A: 50 mM NH_4HCO_3 in H2O and B: MeCN.

5.S2.2 Synthetic strategies

Synthetic strategies to assemble cyclophellitol derivatives often involve different configurations of functionalized cyclohexenes as starting materials. Based on the synthesis of D-galacto- and L-fuco-configured cyclohexenes described by Llebaria and coworkers[1] and the Overkleeft group,[2] it was reasoned that reaction of debenzylated aldehyde **5**[3] with chiral Evans' oxazolidinone **4** should provide the L-ido-configured cyclohexene by syn-aldol addition (**Scheme 4.S1**). Indeed, asymmetric aldol condensation catalyzed by dibutylboryl triflate at low temperatures (–78 °C to –20 °C) proceeded stereo-selectively to provide the desired aldol product **6** in 60 % yield. During the reaction, the non-reactive terminal alkene derived from isomerization of acrylamide **4** was observed as a major side product. Reduction of oxazolidinone **6** with $LiBH_4$ followed by Grubbs II-catalyzed metathesis afforded the desired L-ido-configured cyclohexene **8** in excellent yield.



Scheme 5.S1. Synthesis of α-L-iduronic-configured inhibitors and ABPs 1–3. Reagents and conditions: a) DBBT, Et_3N, CH_2Cl_2, –78 °C to –20 °C, 5 h, 60 %; b) $LiBH_4$, THF, RT, 2 h, 99 %; c) Grubbs II catalyst, CH_2Cl_2, 40 °C, 18 h, 98 %; d) BnBr, TBAI, NaH, DMF, RT, 18 h, 79 %; e) $PhI(OAc)_2$, CH_2Cl_2, RT 48 h, 43 %; f) Li, NH_3, THF, –60 °C, 1 h, 93 %; g) 8-azido-1-iodooctane, K_2CO_3, DMF, 55 °C, 24 h, **12**: 22 %; h) TEMPO, NaBr, NaOCl, H_2O, 0 °C, 3 h, 14 %; i) $CuSO_4$, NaAsc, RT, 18–48 h, **2**: 22 %, **3**: 34 %.

Thereafter, olefin aziridination of L-ido-configured cyclohexene **8** was attempted, with the specific aim of obtaining the α-stereoisomer. Llebaria and co-workers have recently reported the first *N*-aminoaziridine covalent glycosidase inhibitors, which were prepared by stereoselective hydrogen-bonding-guided aziridination using 3-amino-2-ethylquinazolin-4(3*H*)-one (Et-Q-NH2).[2] Because such hydrogen-bond-mediated aziridination of cyclohexene **8** would generate

the undesired β-diastereoisomer, the free alcohol groups in **8** were benzylated with benzyl bromide and sodium hydride to generate cyclohexene **9**. When the direct azidirination of **9** was performed with in situ generated CF_3-Q-NHOAc complex, the desired a-aziridine **10** was obtained in 43% yield together with 32% recovered starting material (**Scheme 4.S1**). This result implied that hydrogen bonding is not required for a productive aziridination, provided that the double bond is freely accessible. Removal of the CF_3-Q and benzyl groups was achieved in one step by Birch reduction using lithium and liquid ammonia at –78 °C. After quenching the reaction with H_2O, CF_3-Q-derived impurities precipitated and were filtered off. Aziridine **11** was purged of lithium hydroxide by cation-exchange chromatography with Amberlite H^+ resin, and the fully deprotected cyclitol aziridine was obtained in 93 % yield (**Scheme 4.S1**). The α-L-configuration of aziridine **11** was established by ^{1}H NMR analysis (**Section 4.S2**), and the experimental coupling constants were compared with the corresponding calculated values obtained from DFT calculations (**Section 4.S3**). Aziridine **11** was then alkylated with 8-azido-1-iodooctane and K_2CO_3 or acylated with 8-azidooctanoic acid and EEDQ to afford intermediates **12** or **13**, respectively, which were purified by reversed-phase column chromatography. Oxidation of C-6 proved to be challenging due to instability of the aziridine under acidic or basic conditions. Aziridine **1** was obtained in 14 % yield by oxoammonium-catalyzed oxidation, maintaining the reaction and HPLC-MS purification at basic pH. Final click reaction with Cy5- and biotin-substituted alkynes afforded the desired ABPs **2** and **3**.

5.S2.3 Synthesis and characterization data of compounds **1–3**

For detail the reader is referred to the supporting information published in Artola *et al.*[4]

5.S3. DFT Calculations (Department of Bio-organic Synthesis, Leiden University)

5.S3.1 Geometry optimization

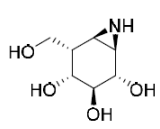
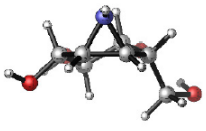
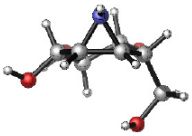
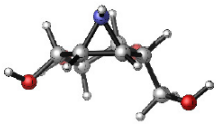
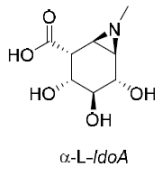
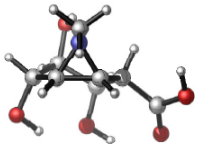
By using the conformer distribution search option included in the Spartan 14 program[5], exclusively the $^4H_3$ conformation of the structure was found. Only notable variations of the geometry were found at the C5-C7 bond, including multiple rotamers which were significant higher in terms of energy. For α-L-idoA aziridine the $^4H_3$ conformation was also found as lowest energy conformer, but in this case the $^3H_4$ was only 1.4 kcal/mol higher in terms of energy.

All calculations were performed with DFT as level of theory in combination with the B3LYP hybrid functional. A conformer distribution search option included in the Spartan 14 program[5], in gas-phase with the use of 6-31G(d) as basis set, was used as starting point for the geometry optimization. All generated structures were further optimized with Gaussian 09[6] at 6-311G(d, p). Optimization was done in gas-phase and subsequently corrections for solvent effects were done by the use of a polarizable continuum model using water as solvent parameter. The free Gibbs energy of the computed conformations was calculated using Equation (1) in which ΔE_{gas} is the gas-phase energy (electronic energy), ΔG^T_{RRHO} (T= 298.15 K and pressure= 1 atm.) is the sum of corrections from the electronic energy to free Gibbs energy in the rigid-rotor-harmonic-oscillator approximation (RRHO) also including zero-point-vibrational energy, and ΔG^T_{solv} is their corresponding free solvation Gibbs energy.

$$\Delta G^T_{aq} = \Delta E_{gas} + \Delta G^T_{gas,RRHO} + \Delta G_{solv} = \Delta G^T_{gas} + \Delta G_{solv} \quad (1)$$

The used free energies include unscaled zero-point vibrational energies. Visualization of the conformations of interest was done with CYLview.[7]

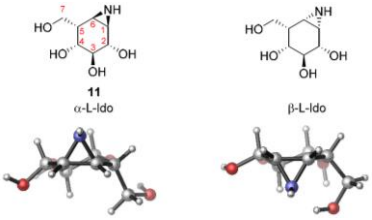
Table 5.S2. Geometry optimization of compound 11 and the *N*-methylated α-L-iduronic configured cyclophellitol aziridine performed by of DFT calculation.

 α-L-*Ido* **11**			
	$^4H_3$ ***tg***	$^4H_3$ ***gg***	$^4H_3$ ***gt***
ΔG^T_{aq} **(kcal/mol)**	0.0	1.5	1.6
Geometry	D1: -45.4 ° D3: -49.6 ° D5: -2.3 °	D1: -43.4 ° D3: -49.8 ° D5: -3.1 °	D1: -45.6 ° D3: -49.5 ° D5: -3.0 °
 α-L-*IdoA*			
	$^4H_3$	$^3H_4$	$^{2,5}B$
ΔG^T_{aq} **(kcal/mol)**	0.0	1.4	8.0
Geometry	D1: -44.1 ° D3: -53.5 ° D5: -2.3 °	D1: 51.7 ° D3: 43.8 ° D5: 0.9 °	D1: -52.1 ° D3: 48.0 ° D5: 2.0 °

5.S3.2 NMR calculations

Based on the optimized lowest energy structure the spin-spin coupling constants were calculated according to the work of Rablen and Bally[8] with the use of 6-311g(d, p) u+1s as basis set and PCM(H_2O) as solvent model and a scaling factor of 0.92. The calculated total nuclear spin-spin coupling terms were used as calculated spin-spin coupling constants.

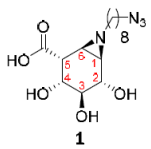
Table 5.S3. Experimental coupling constants of α-idose configured cyclophellitol aziridine 12 compared to DFT calculated coupling constants.



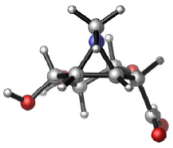
H-H Coupling	Exp. $^3J_{(H,H)}$ (Hz)	α-L-*ido* DFT calc. $^3J_{(H,H)}$ (Hz)	β-L-*ido* DFT calc. $^3J_{(H,H)}$ (Hz)
H1-H6	6.0	5.8	6.8
H1-H2	n.d.	3.1	2.9
H2-H3	7.4	7.5	8.2
H3-H4	10.5	10.2	9.3
H4-H5	5.6	6.3	9.8
H5-H6	1.6	1.3	5.9
H5-H7a	4.4	2.8	0.9
H5-H7b	11.1	10.3	4.3

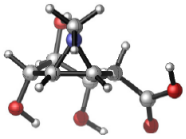
Coupling constants were determined by ^{1}H NMR experiments (exp.). n.d.: values not determine due to very small coupling constant (J < 1 Hz).

Table 5.S4. Experimental coupling constants of 1 compared to DFT calculated coupling constants of the methylated α-iduronic configured cyclophellitol aziridine.



1





H-H Coupling	Exp. $^3J_{(H,H)}$ (Hz)	α-L-*idoA* DFT calc. $^3J_{(H,H)}$ (Hz) $^4H_3$ (0.0 kcal/mol)	α-L-*idoA* DFT calc. $^3J_{(H,H)}$ (Hz) $^3H_4$ (1.4 kcal/mol)	α-L-*idoA* Boltzmann-weighted average $^3J_{(H,H)}$ (Hz)
H1-H6	n.d.	6.0	6.2	6.0
H1-H2	-	0.0	1.7	0.3
H2-H3	6.8	7.2	2.9	6.4
H3-H4	9.2-9.4	10.6	4.8	9.5
H4-H5	4.3-5.5	6.5	3.7	6.0
H5-H6	n.d.	1.7	0.8	1.5

Coupling constants were determined by ^{1}H NMR experiments (exp.). n.d.: values not determine due to peak overlap.

5.S4 Supplementary References

1 Hansen FG, Bundgaard E & Madsen R (2005) A short synthesis of (+)-cyclophellitol. *J Org Chem* **70**, 10139–10142.

2 Harrak Y, Barra CM, Delgado A, Castaño AR & Llebaria A (2011) Galacto-configured aminocyclitol phytoceramides are potent in vivo invariant natural killer T cell stimulators. *J Am Chem Soc* **133**, 12079–12084.

3 Jiang J, Kallemeijn WW, Wright DW, van den Nieuwendijk AMCH, Rohde VC, Folch EC, van den Elst H, Florea BI, Scheij S, Donker-Koopman WE, Verhoek M, Li N, Schürmann M, Mink D, Boot RG, Codée JDC, van der Marel GA, Davies GJ, Aerts JMFG & Overkleeft HS (2015) In vitro and in vivo comparative and competitive activity-based protein profiling of GH29 α-l-fucosidases. *Chem Sci* **6**, 2782–2789.

4 Artola M, Kuo CL, McMahon SA, Oehler V, Hansen T, van der Lienden M, He X, van den Elst H, Florea BI, Kermode AR, van der Marel GA, Gloster TM, Codée JDC, Overkleeft HS & Aerts JMFG (2018) New Irreversible α-l-Iduronidase Inhibitors and Activity-Based Probes. *Chemistry* **24**, 19081–19088.

5 Spartan '10 [Computer software]. (2010) Irvine, CA: Wavefunction Inc.

6 Gaussian 03 [Computer software]. (2004) Gaussian, Inc., Wallingford CT.

7 Legault CY (2009) CYLview 1.0b [Computer software]. Québec, Montreal, Canada: Univ Sherbrooke. Retrieved from http//www.cylview.org).

8 Bally T & Rablen PR (2011) Quantum-chemical simulation of 1H NMR spectra. 2. Comparison of DFT-based procedures for computing proton-proton coupling constants in organic molecules. *Org. Chem* **76**, 4818–4830.

CHAPTER 6

Activity-based probes for retaining exo-mannosidases

Based on:

Kuo CL, Beenakker TJM, Lahav D, Hissink C, Armstrong Z, Wu L, Johnson R, de Boer C, Artola, M, Florea B, Boot RG, Codée JDC, van der Marel GA, Davies GJ, Aerts JMFG & Overkleeft HS. *To be submitted.*

ABSTRACT

Mannosidases are key enzymes in eukaryotic N-linked glycan synthesis and degradation. Among these hydrolases, retaining exo-α-mannosidases (Glycoside Hydrolase (GH) family 38) and exo-β-mannosidase (GH family 2) are implicated in pathologies such as cancer and lysosomal storage diseases. To generate tools to profile these enzymes, mechanism-based suicide inhibitors and activity-based probes (ABPs) were synthesized and characterized. These compounds, based on α-mannose or β-mannose configured cyclophellitol scaffolds, are micromolar inactivators of their expected target enzymes. The ABPs label the target mannosidases in mechanism-based manner, as confirmed by SDS-PAGE-based fluorescent detection, kinetic studies, and protein crystallography. Proteomics revealed that the α-mannose configured ABP equipped with biotin labels all five GH38 retaining exo-α-mannosidases in mouse tissue extracts. Similarly, the α-mannose configured Cy5 ABP labels all five human GH38 α-mannosidases cloned and individually expressed in human cell lines. The unique molecular weight and pH optimum of each α-mannosidase allows their simultaneous activity-based profiling in complex biological samples such as cell lysates and tissue extracts. β-Mannose configured ABP labels the GH2 β-mannosidase (MANBA) in mouse kidney extracts. It additional labels retaining β-glucosidases, but specific visualization of MANBA is still feasible by pre-incubating the samples with β-glucosidase inhibitors. In conclusion, the novel ABPs described here enable the simultaneous visualization of all retaining exo-mannosidases in complex biological samples. They should assist future screening for small molecule inhibitors/activators of these highly relevant enzymes.

6.1 Introduction

Protein N-linked glycosylation takes place in all domains of life. In the ER of eukaryotes, it plays essential roles in folding ,quality control and subsequent transport of newly formed N-linked glycoproteins to the Golgi apparatus.[1, 2] This finely orchestrated process is carried out by lectins recognizing specific glycan structures, and by glycosidases residing at various subcellular locations that modify the glycans. A key component of the N-linked glycan is mannose. This sugar is abundant in the $Glc_3Man_9GlcNAc_2$ glycan that is transferred from the dolichol donor to nascent polypeptide in the ER (**Fig. 6.1**, step 1−2). Maturation of N-linked glycoprotein is accompanied by removal of specific mannose residues from their N-linked glycans. Several mannosidases are involved in this process. These enzymes differ in subcellular location and substrate specificity, and are classified into Glycoside Hydrolase (GH) family 2, 38, 47, and 99 (**Fig. 6.1**, bottom right) based on the Carbohydrate active enzyme (CAZy) database[3].

In man, the GH47 family comprises four inverting exo-α-1,2-mannosidases located in the ER and Golgi complex, and three additional ER-dependent α-mannosidase-like proteins (EDEMs) that have putative α-mannosidase activity. The ER-α-mannosidase I (MAN1B1) trims one mannose from the protein N-linked glycan, which signals the protein to either pass the ER quality control checkpoint (**Fig. 6.1**, step 3−4)[4, 5] or enter the ER-associated degradation pathway in a process facilitated by the EDEMs (**Fig. 6.1**, step 5−6).[6, 7] The three Golgi GH47 mannosidases, Ia (MAN1A1)[8], Ib (MAN1A2)[9], and Ic (MAN1C1)[10] further trim the α-1,2-linked mannoses from glycans of glycoproteins arriving at the *cis*-Golgi (**Fig. 6.1**, step 7−8), enabling downstream processes such as protein N-linked hybrid- or complex-type glycan synthesis (**Fig. 6.1**, step 11−12 and below) or the mannose-6-phosphate mediated endosomal/lysosomal targeting (**Fig. 6.1**, step 13 and below)[11]. The GH99 endo-α-1,2-mannosidase (MANEA) is also residing at the *cis*-Golgi. It recognizes glycoproteins that still contain terminal glucose residues and catalyzes the one-step endo-glycosidic hydrolysis of these glucoses together with the adjacent mannose. This provides an additional pathway for glycoprotein maturation (**Fig. 6.1**, step 9−10).[12, 13]

The other two mannosidase families are GH38 α-mannosidase (MAN2A1, MAN2A2, MAN2B1, MAN2B2, MAN2C1) and GH2 β-mannosidase (MANBA), both composed of

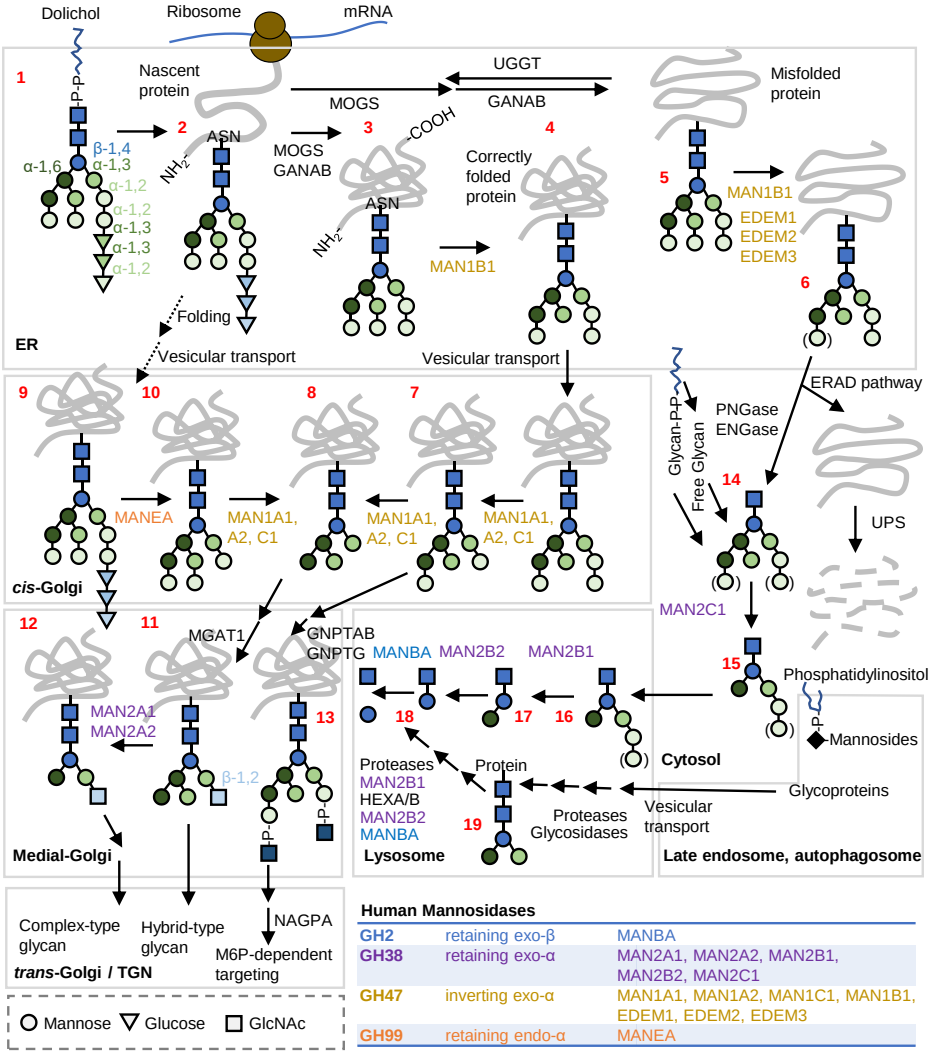


Figure 6.1. Human mannosidases in N-linked glycoprotein synthetic and degradation pathways. 1−4, trimming of glucoses and one mannose on correctly folded proteins in the ER. 5−6, mannosidases responsible for initiating the ER-associated degradation (ERAD) pathway. 7−8, mannose trimming in the *cis*-Golgi. 9−10, action of the endo-mannosidase on glycoproteins arriving at the Golgi with terminal glucoses. 11−12, mannose trimming by Golgi mannosidase II and IIx (MAN2A1, A2) for complex-type glycan formation. 13, glycan processing for the mannose-6-phophate (M6P)-dependent protein targeting pathways. 14−15, glycan degradation by cytosolic mannosidase (MAN2C1). 16−19, actions of lysosomal retaining exo-mannosidases (MAN2B1, B2, and MANBA) on glycans from the cytosol or endosomal compartments. *UPS*, ubiquitin-proteasome system; *MOGS*, ER α-glucosidase I; *GANAB*, ER α-glucosidase II; *UGGT*, UDP-glucose:glycoprotein glucosyltransferase; *GNPTAB/GNPTG*, GlcNAc-phosphotransferase α, β/γ subunits; *NAGPA*, GlcNAc1-1-phosphodiesterase; *MGAT1*, α-1,3-mannosyl-glycoprotein 2-β-GlcNAc transferase.

retaining glycosidases. The human GH38 α-mannosidase family consists of five members, all of which require metal ions (zinc or cobalt) for catalysis but differ in subcellular location and substrate specificity. Two of these, Golgi mannosidase II[14] and IIx[15] (MAN2A1 and MAN2A2, E.C. 3.2.1.114), remove the two outer α-1,3- and α-1,6-linked mannoses on the protein $GlcNAcMan_4GlcNAc_2$ glycans, thus allowing further synthesis of complex-type glycans (**Fig. 6.1**, step 11−12).[16] These two enzymes arose from a recent mammalian gene duplication event, exhibiting overlapping substrate specificity[17] but differ in expression levels depending on tissue type.[18, 19] The neutral α-mannosidase (MAN2C1) is a cytosolic enzyme degrading soluble N-linked glycans released from glycoprotein or glycolipids into $Man_5GlcNac$, thus allows the trimmed glycan to be further degraded in the lysosome (**Fig. 6.1**, step 14−15).[20, 21] It depends on Co^{2+} for catalysis, but is also activated by Fe^{2+} and Mn^{2+}.[22] The lysosomal GH38 α-mannosidase MAN2B1 acts on α-1,2- and α-1,3-linked mannoses on glycans derived from protein N-linked glycans or glycolipids (E.C.3.2.1.24),[23] as well as on glycans attached to glycoproteins (**Fig. 6.1**, step 16 and 19).[24] It does not cleave the core α-1,6-linked mannose,[24] which is specifically hydrolyzed by the lysosomal GH38 core-specific α-mannosidase (MAN2B2, E.C. 3.2.1.114) (**Fig. 6.1**, step 17 and 19),[25, 26] also known as the epididymis α-mannosidase.[27] Finally, the lysosomal GH2 β-mannosidase (MANBA) releases the last mannose residue from GlcNAc (E.C. 3.2.1.25) and completes the mannose catabolism (**Fig. 6.1**, step 18 and 19).[28, 29]

Both the GH38 and GH2 mannosidases receive continuous interest as therapeutic targets for human diseases. The Golgi α-mannosidase MAN2A1 has been linked to the progression of several cancers.[30] Clinical trials had been conducted using the inhibitor swainsonine, albeit unsuccessful due to adverse side effects attributed to its concomitant inhibition of other mannosidases [31, 32] Inherited deficiency of the lysosomal α-mannosidase MAN2B1 deficiency underlies the lysosomal storage disorder α-mannosidosis (OMIM: 248500) in man.[33] Affected individuals accumulate in lysosomes undegraded mannose-containing oligosaccharides, causing variable degree and progression of mental and respirational impairment, hearing loss, Hurler-like facial distortion, and reoccurring infections among paitents.[33] A recombinant α-mannosidase-based enzyme replacement therapy has been recently approved in the EU.[34, 35] Deficiency in the lysosomal β-mannosidase MANBA also causes a lysosomal storage disorder, β-mannosidosis (OMIM: 248510). This disease is firstly described in goat[36] and is rare in human. Human patients generally have milder symptoms (e.g. mental retardation) compared to the affected lifestock.[37] Last but not least, the cytosolic neutral α-mannosidase MAN2C1 is

involved in tumorigenesis.[38, 39] The cause for this process is believed to be independent of MAN2C1's catalytic activity, but rather through its direct association with tumor suppressive proteins and thereby causes their inactivation during tumorigenesis.[40, 41]

In the past, GH38 α-mannosidase activities are distinguished from those of GH47 α-mannosidases based on their different cation preference (Zn^{2+} or Co^{2+} for GH38 *vs* Ca^{2+} for GH47) and inhibitor sensitivity (furanose-based inhibitors such as swainsonine and mannostatin A for GH38; pyranose-based inhibitors such as 1-deoxymannojirimycin for GH47).[42] Because GH38 and GH2 mannosidases are retaining glycosidases employing the Koshland double displacement catalytic mechanism (**Fig. 6.2A, B**),[43−45] it is envisioned that their activity can be selectively measured over the GH47 enzymes (which are inverting glycosidases) by activity-based protein profiling with compounds harboring a mannose-configured scaffold that covalently becomes trapped at the catalytic nucleophile of the enzyme upon the initial nucleophilic attack (**Fig. 6.2C, D**). Similar approaches using configurational isomers of cyclophellitols and cyclophellitol aziridines have been recently demonstrated to be useful tools in labeling their targeted glycosidases (General Introduction and Chapter 6, this thesis). This chapter presents the characterization of the α- or β-mannose configured cyclophellitol aziridines inhibitors and ABPs in their inhibitory potency and labeling towards GH38 and GH2 mannosidases, and discusses the potential application of the ABPs in the study of mannosidase biology and associated diseases.

6.2 Results

6.2.1 Synthesis of compounds used in this chapter

The synthesis of compounds **1**−**7** and **8**−**13** was performed at the Department of Bio-organic Synthesis at Leiden University and followed the strategies reported for the α- and β-mannose configured cyclophellitols **1** and **8**,[46] the β-mannose configured cyclophellitol aziridine **9**,[46] and the α-mannose configured cyclophellitol aziridine **2**[47] (**Fig. 6.3** and **Scheme 6.S1**−**2**). N-alkylation of **2** or **9** with 1-azido-8-iodooctane yielded compound **3** and **10**, which were subsequently appended with either a BODIPY, Cy5, or a biotin tag (compounds **4**−**6** and **11**−**13**) using the Cu(I)-catalyzed azide-alkyne Huisgen [2 + 3]-cycloaddition ("click" reaction).[48, 49] N-acylation of **3** was also attempted,[48] with which successfully yielded the Cy5 compound **7**.

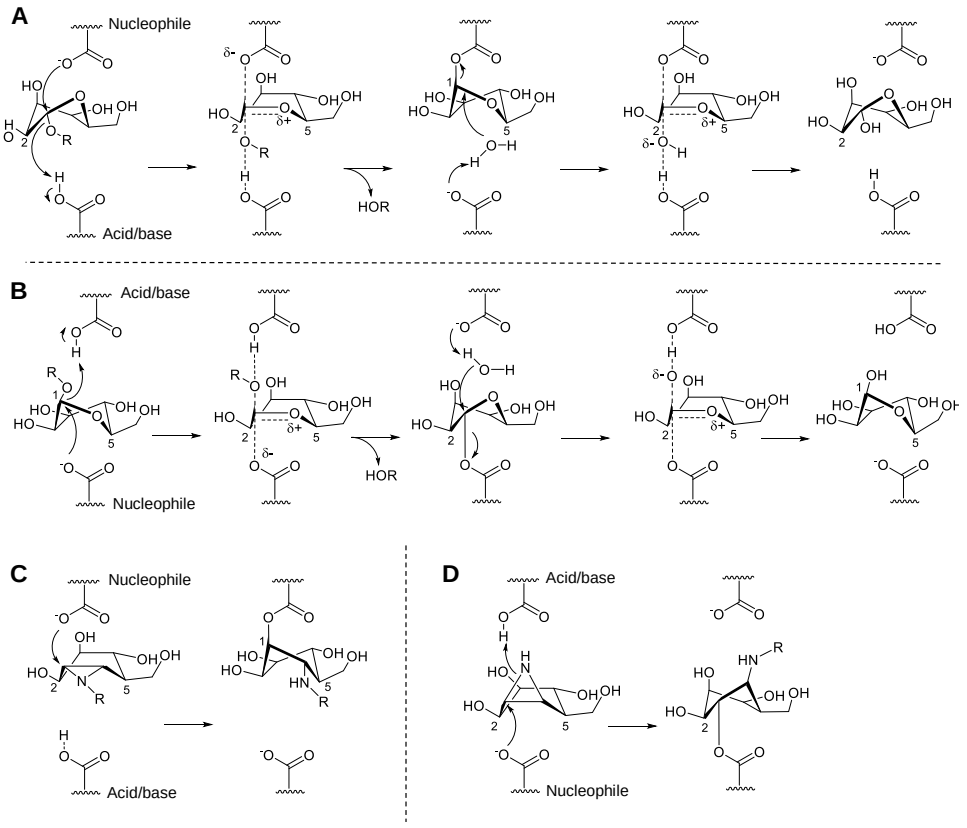


Figure 6.2. Reaction mechanisms of retaining exo-mannosidases. A) Reaction itinerary by GH38 α-mannosidases. B) Reaction itinerary by GH382 β-mannosidase. C) Proposed reaction mechanism for α-mannose configured cyclophellitol aziridine. D) Proposed reaction mechanism for β-mannose configured cyclophellitol aziridine. Numbers shown correspond to pyranose numbering of the carbons.

6.2.2 *In vitro* activity of compound **1**−**7** on GH38 α-mannosidases

The α-mannose configured compounds **1**−**7** were tested for *in vitro* inhibitory potency by enzymatic assay of commercially available Jack bean (*Canavalia ensiformis*) GH38 α-mannosidase (E.C. 3.2.1.24). This enzyme exhibits a similar catalytic profile and zinc ion dependency to the human lysosomal broad-specificity α-mannosidase (MAN2B1).[50] The assay was performed at an acidic pH of 4.5 and at 37 °C, by means of a 30 min incubation of enzyme with compounds and the substrate, 4-methylumbelliferyl-α-D-mannopyranoside (4-MU-α-man).

It turned out that all the tested compounds were α-mannosidase inhibitors, exhibiting low- to mid-micromolar apparent IC_{50} values (**Table 6.1**, **Fig. 6.S1A**). The most potent of the series were the biotin **6**, Cy5 **5**, and N-alkyl azide compound **3** (apparent IC_{50} = 2−4 μM); they were followed by the unsubstituted α-mannose configured cyclophellitol aziridine **2**, and the epoxide **1**, with the later having similar value to the one reported by Tatsuta et al.[51] BODIPY ABP **4** was the least potent of the series, having apparent IC_{50} value over 50 μM. The N-acylated Cy5 ABP **7** required more steps to synthesize but was equally potent as the N-alkylated Cy5 ABP **5**, similar

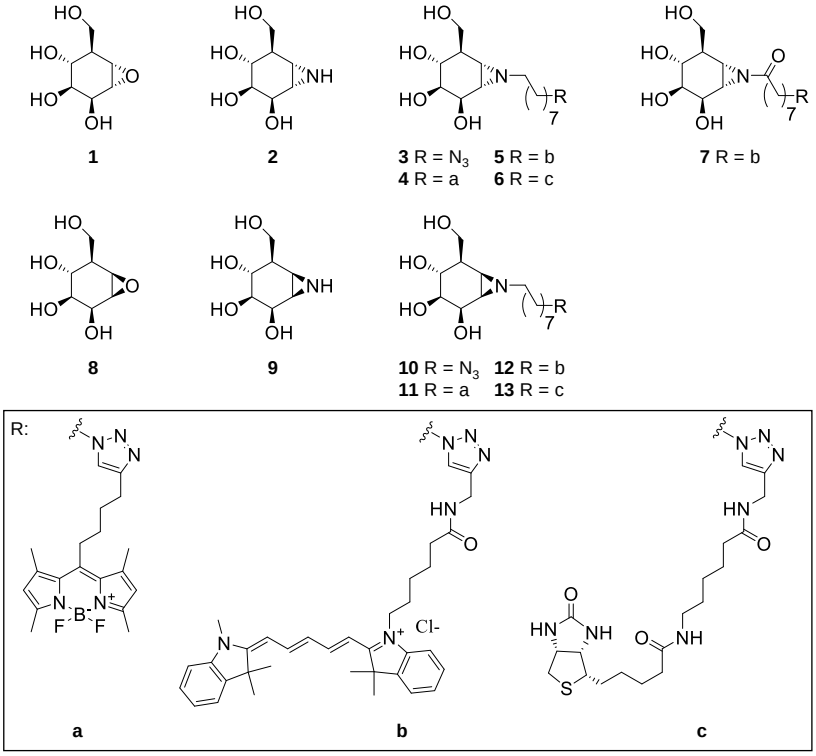


Figure 6.3. Structures of compounds used in this study.

Table 6.1 Apparent IC_{50} values of α-mannose configured cyclophellitol and aziridines towards Jack bean GH38 α-mannosidase. A) Apparent IC_{50} values at 30 min compound incubation time. B) Apparent IC_{50} values of **5** at different incubation time.

A

Compounds	IC_{50} (μM)
1 (TB440)	35.4 ± 6.70
2 (TB450)	7.19 ± 0.12
3 (TB481)	3.54 ± 0.48
4 (TB521)	> 50
5 (TB482)	3.43 ± 0.68
6 (TB484)	2.22 ± 0.24
7 (TB480)	3.86 ± 1.30

B

Incubation time (min)	IC_{50} (μM)
10	4.94 ± 0.13
30	2.84 ± 0.17
120	1.13 ± 0.06

to the general trend observed on earlier reported ABPs towards other glycosidases.[52] Next, time dependency of inhibition by ABP **5** was examined. The apparent IC_{50} values for the N-alkyl ABP **5** gradually were found to decrease from 4.94 μM to 1.13 μM with incubation times increasing from 10 min to 120 min (**Table 6.1B, Fig. 6.S1B**), hinting to irreversible inhibition.

Next, SDS-PAGE-based fluorescent readout was used to directly visualize the covalent ABP labeling of Jack bean α-mannosidase. In the first experiment, 3 μM ABP **5** was incubated with the enzyme for 30 min at 37 °C at various pHs. The results showed that the ABP covalently labeled the enzyme in a pH-dependent manner: one band between 60 and 70 kDa was found to be most prominently labeled at acidic pHs (**Fig. 6.4A**, left), corresponding to the known size of the large subunit (66 kDa) of Jack bean α-mannosidase bearing the active site.[51] Removing zinc ion from the reaction mixture did not affect the ABP labeling during the 30 min incubation time (**Fig. 6.4A**, left). Compared to enzymatic activity, the ABP labeling had a slight shift in pH optimum (0.5−1 unit)—with maximal enzymatic activity occurring at pH 4.5−5.0 and maximal ABP labeling at pH 5.0−6.0 (**Fig. 6.4A**, right). The labeling potency of ABP **5** was next compared to that of ABP **7** (N-acyl Cy5) at pH 5.5: both labeled the enzyme and were equally potent (**Fig. 6.4B**). Saturation of labeling occurred at around 3 μM for both ABPs, with the calculated concentration for 50 % labeling being 0.3 μM (**Fig. 6.4B**, right). This value was ten-fold lower than the apparent IC_{50} values (around 3 μM), which might be resulted from different assay pHs (5.5 during ABP labeling *vs* 4.5 during inhibitory IC_{50} determination). Time-dependency of labeling by ABP **5** was next examined, at 3 μM ABP **5** and pH 5.5. It was found

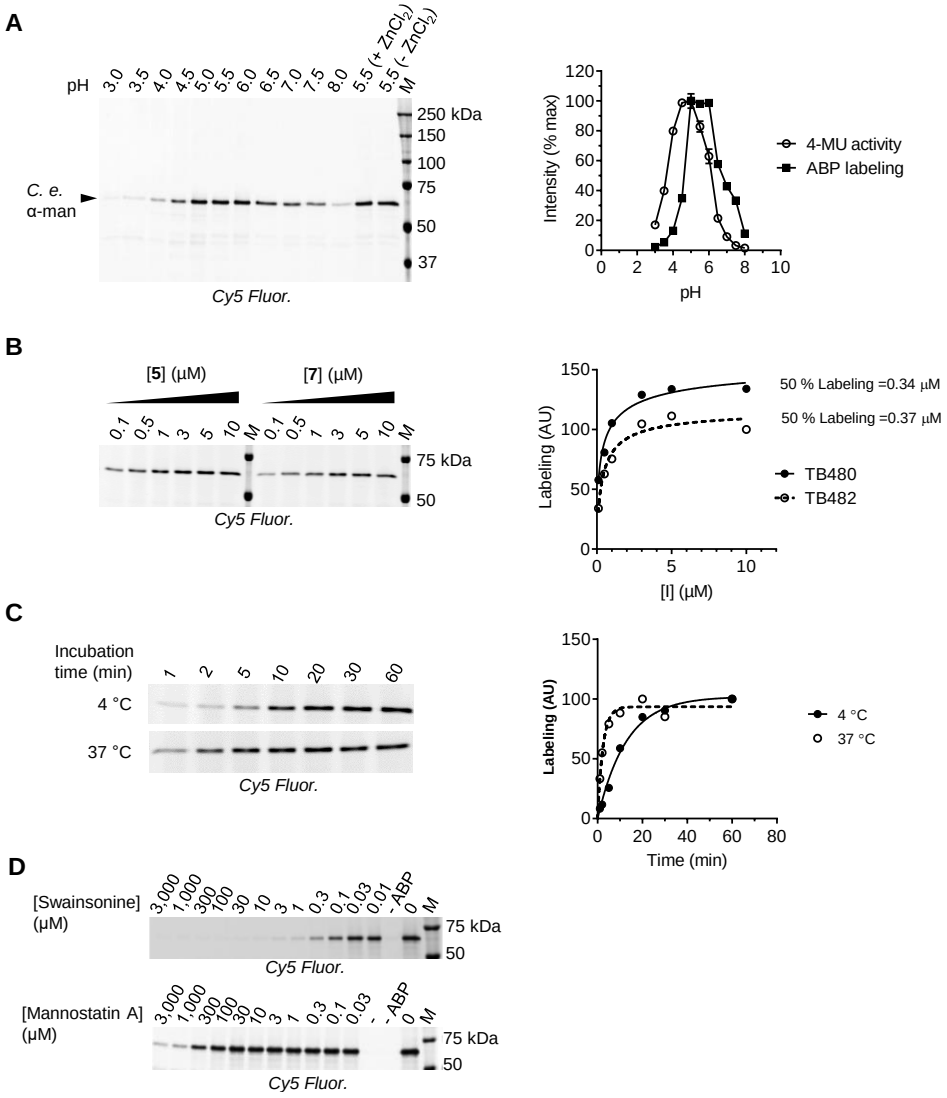


Figure 6.4. *In vitro* labeling of compounds towards Jack bean (*Canavalia ensiformis*) GH38 α-mannosidase. A) pH-dependent labeling of **5** (left) and comparison of quantified ABP labeling with measured 4-MU activity across different pHs. Error ranges = ± SD from technical triplicates. B) Labeling of **5** and **7** at different ABP concentrations (left) and band quantification (right). C) Time-dependent labeling of **5** at two different temperatures (left) and band quantification (right). D) Competitive ABP labeling of **5** against pre-incubation of swainsonine and mannostatin A. -, empty lane.

that labeling increased with incubation time, and reached saturation within 10 min at 37 °C (60 min for labeling at 4 °C) (**Fig. 6.4C**). Finally, to determine if the labeling occurred at the enzymes' active site pocket, the enzyme was pre-incubated with two known GH38 α-mannosidase inhibitors, swainsonine and mannostatin A, followed by a short ABP labeling period of 10 min. As expected, labeling of ABP **5** towards the enzyme was abolished by both inhibitors, suggesting active-site pocket occupancy of the ABPs in the enzyme (**Fig. 6.4D**).

The active site occupancy and labeling of the compound towards GH38 α-mannosidase was further examined in detail by structural analysis perform at the University of York. For this experiment, protein crystals of the *Drosophila melanogaster* GH38 MAN2A1 (Golgi α-mannosidaseII) homologue (dGMII)[30] were incubated with the bare aziridine compound **2**, and solved for structure by protein X-ray crystallography. In the resolved structure, a covalent glycosidic bond was clearly visible between the C1 of compound **2** and the catalytic nucleophile Asp204 (**Fig. 6.5**) of dGMII. Compound **2** adopted an $^{1}S_{5}$, conformation (**Fig. 6.5**, right), matching the known substrate conformation at the covalent enzyme-substrate intermediate (**Fig. 6.2A**, middle).[45]

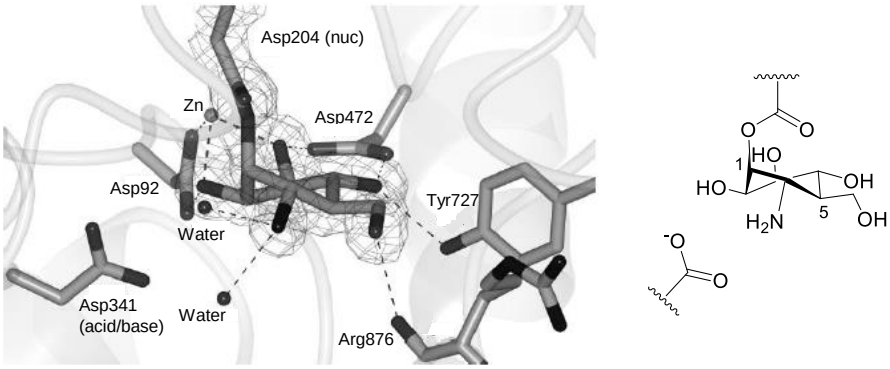


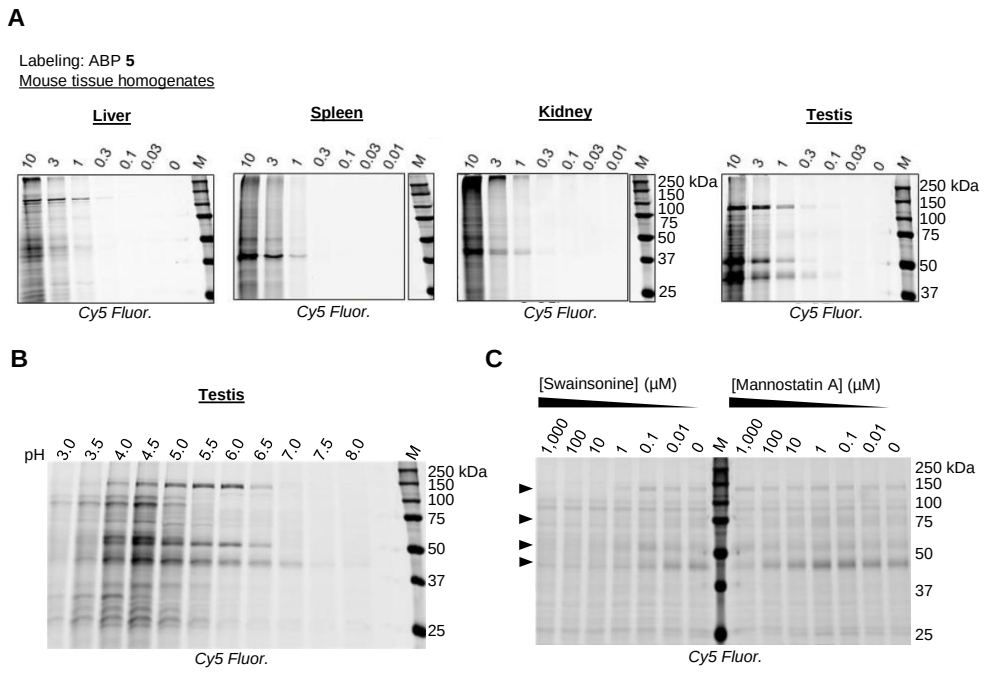
Figure 6.5. Structure of *Drosophila melanogaster* GH38 α-mannosidase (dGMII) in complex with compound 2. Right, skeletal structure of compound **2** bound to the nucleophile, showing the conformation observed from the crystal structure (left).

6.2.3 Target detection and identification of the α-mannose configured ABPs in complex biological samples

Activity-based protein profiling (ABPP) was performed with ABP **5** in mouse tissue extracts. Titration of ABP concentration from 0.03 to 10 μM at pH 5.5 in homogenates of liver,

spleen, kidney, and testis enabled the detection of distinct bands in these samples. At 1−3 μM ABP **5**, a sharp band around 140 kDa was detected in liver and testis extracts, whilst two bands around 50 kDa and 45 kDa were clearly visible in spleen, kidney, and testis extracts (**Fig. 6.6A**). As testis extracts contain multiple bands, a pH titration was performed herewith from pH 3.0 to 8.0. The experiment revealed that the labeling of the different bands had a distinct pH optimum; the 140 kDa band was optimally labeled at pH 6.0, while the 50 and 45 kDa bands were optimally labeled at pH 4.5 (**Fig. 6.6B**). Several other minor bands were also detected, most of which also showed a pH optimum of 4.5. To determine whether these bands were GH38 α-mannosidases, a competitive ABPP (cABPP) experiment was employed. Mouse testis extracts were pre-incubated with swainsonine or mannostatin A for 30 min at pH 4.5, and then incubated with 3 μM ABP **5** for 10 min. The 140 kDa, 50 kDa, 45 kDa, and an additional 76 kDa bands were abolished by swainsonine pre-incubation, and the 45 kDa band was additionally competed away

Figure 6.6. ABPP in mouse tissue extracts with ABP 5. A) Concentration titration of ABP **5** in



homogenates of liver, spleen, kidney, and testis. B) Titration of labeling pH with ABP **5** in mouse testis homogenates. C) cABPP of swainsonine or mannostatin A with ABP **5** in mouse testis homogenates. Arrows, bands that were abolished by swainsonine pre-incubation.

with 1 mM mannostatin A (**Fig. 6.6C**). The other minor bands were not abolished by inhibitor

pre-incubation, pointing to non-specific labeling (**Fig. 6.6C**). MANB1 is cleaved in the lysosome into five fragments that contain the 42 kDa peptide A (with catalytic active site), the 10 kDa peptide B, the 24 kDa peptide C, and peptide D and E; the latter two do not form disulfide bridges to the other three fragments.[23, 32] Thus, the 76, 50, and 45 kDa bands, all showing a labeling pH optimum of 4.5, could correspond to MAN2B1's peptide ABC, AB, and A, respectively. The 140 kDa with a labeling pH optimum of 6.0 and being abolished by swainsonine pre-incubation is likely MAN2A1 and/or A2.

A parallel experiment using the biotin ABP **6** was performed in mouse testis extracts to verify the labeling of GH38 enzymes. Samples incubated with ABP 6 at either pH 4.5 or 6.0 were subjected to biotin affinity enrichment and tryptic digestion (both on-bead digestion and in-gel digestion), and finally LC-MS-based protein identification. While silver stain of the gel containing the affinity-enriched samples yielded few, if any, distinct bands (data not shown), the on-bead digest protocol identified MAN2B1 and MAN2B2 in the sample labeled by ABP **6** at pH 4.5, and all five GH38 α-mannosidases in the sample labeled at pH 6.0 (**Table 6.2**). No other glycosidases were detected in these samples.

Table 6.2. List of identified glycosidases by LC-MS-based proteomics in samples of mouse testis extracts incubated with ABP 6. PLGS, ProteinLynx Global Server.

Condition	Accession	Entry	PLGS Score	Peptides	Theoretical peptides	Coverage (%)
pH 4.5	O09159	MAN2B1_MOUSE	175	9	64	9
	O54782	MAN2B2_MOUSE	306	28	57	20
pH 6.0	P27046	MAN2A1_MOUSE	391	28	84	22
	Q8BRK9	MAN2A2_MOUSE	63	6	78	6
	O09159	MAN2B1_MOUSE	175	6	64	9
	O54782	MAN2B2_MOUSE	306	16	57	16
	Q91W89	MAN2C1_MOUSE	119	12	63	14

To shed definitive light on the labeled mannosidases, in a final experiment all five human GH38 α-mannosidases were individually cloned and expressed in HEK293T cells, and the cell lysates were labeled at various pHs with ABP **5**. Lysates of cells transfected with MAN2A1 and MAN2A2 shown bands at around 140 kDa and labeled optimally at pH 5.5−6.0 (**Fig. 6.7**). Lysates of cells transfected with MAN2B1 showed multiple bands optimally labeled

at pH 4.0−5.5, with molecular weights of 130 kDa, 65 kDa, 45 kDa, and 30 kDa. This labeling pattern possibly reflects the complex processing and maturation of MAN2B1 in the lysosome (**Fig. 6.7**). Lysates of cells transfected with MAN2B2 showed a weaker 130 kDa band and a prominent 50 kDa band, both of which had pH optimum around 4.0−5.5 (**Fig. 6.7**). Lysates of cells transfected with MAN2C1 showed a band just above 100 kDa and most prominently at pH 6.5. It is also prominently labeled at pH 7.5, which is different from other GH38 enzymes expressed in HEK293T cells (**Fig. 6.7**).

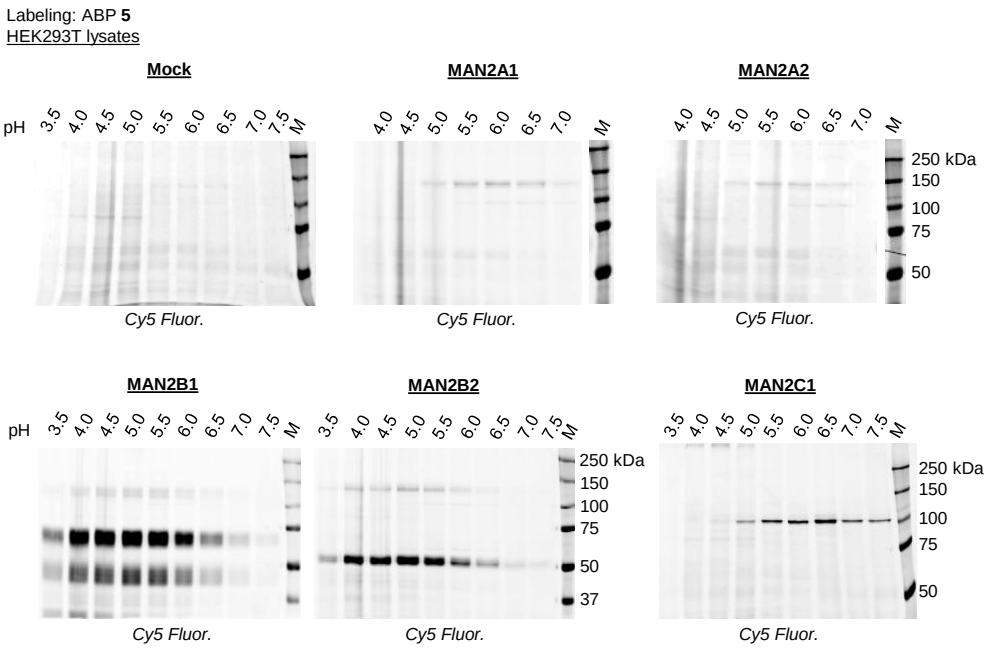


Figure 6.7. ABPP in lysates of HEK293T cells transfected with human GH38 α-mannosidases.

While it was not possible to discriminate MAN2A1 from MAN2A2 in the ABPP setup due to their similar pH range and molecular weight, the other four enzymes were readily identifiable by ABP labeling at different pH values. It was observed that MAN2A1/A2 were most prominently expressed in mouse testis and less in HEK203T cells and mouse brain and epididymis; MAN2B1 and MAN2B2 were expressed in all the samples, whilst MAN2C1 was only observed in mouse brain (**Fig. 6.8**). In mouse epididymis extracts, the 65 kDa and 45 kDa

MAN2B1 seemed to have multiple forms that differ in molecular weights, in contrast to the sharp bands observed in mouse brain extracts.

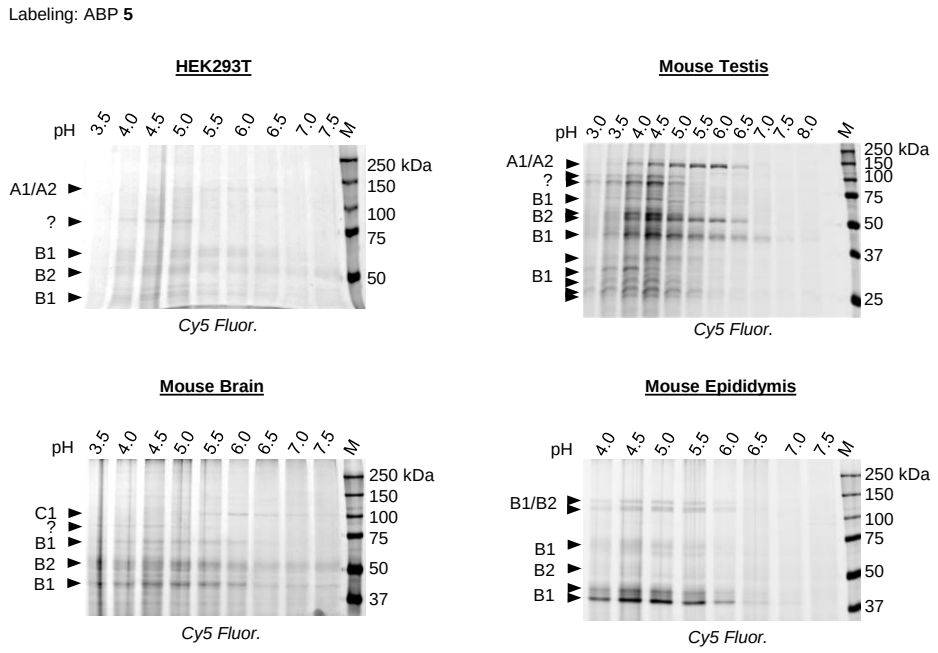
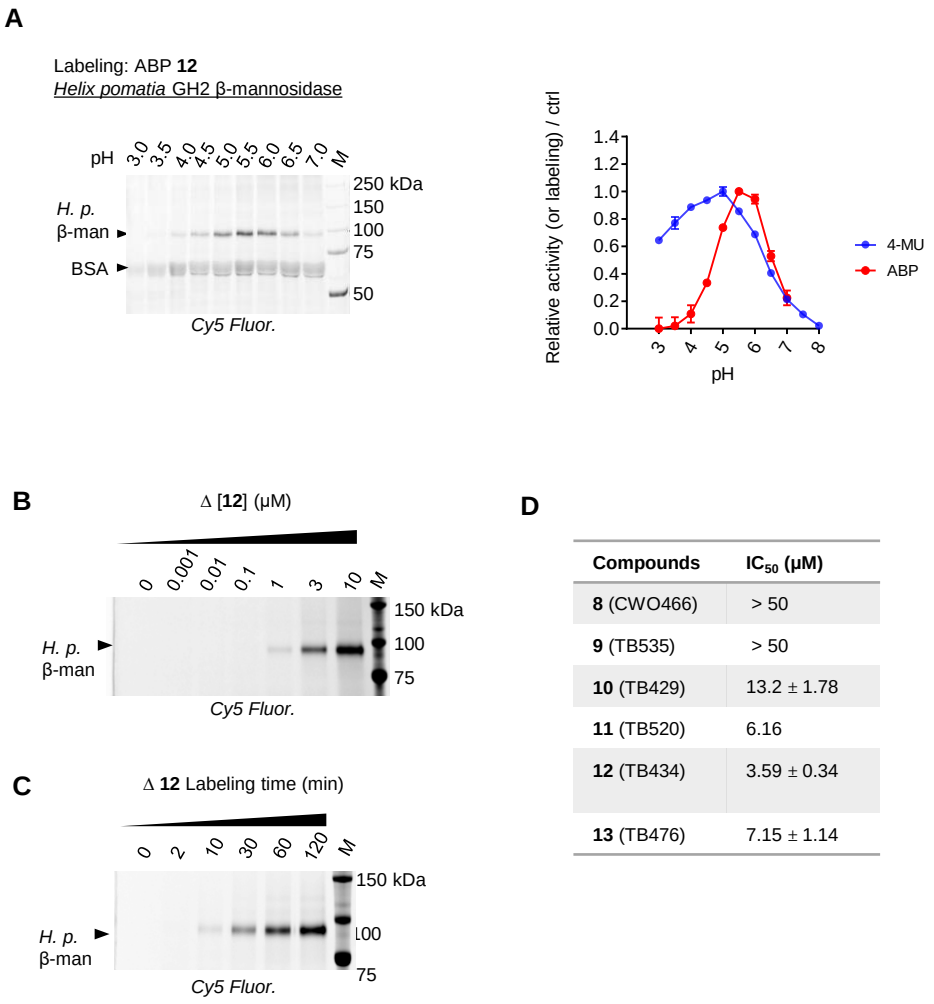


Figure 6.8. Assigning GH38 α-mannosidases in cell lysates and tissue extracts labeled with ABP 5.

6.2.4 *In vitro* activity of compound **8**−**13** towards GH2 β-mannosidase

Next, the β-mannose configured compounds were characterized regarding inhibitory potency and labeling characteristics using commercially available GH2 β-mannosidase from Roman snail (*Helix pomatia*).[54] Initial apparent IC_{50} measurements were performed by means of 30 min incubation of enzyme with compounds and the substrate 4-MU-β-D-(4-MU-β-Man) at pH 4.2, the optimum pH for the enzyme.[55] However, none of the compounds inhibited the enzyme, even at the highest concentration (50 μM). To verify this observation, labeling at various pH values was examined. In this experiment, 10 μM of the Cy5 ABP **12** was incubated with the enzyme for 1 h across a range of pH from 3.0 to 7.0 in the presence of bovine serum albumin (BSA, which stabilized the enzyme at pH > 3.5 (**Fig. 6.S3A, B**)), and samples were subjected to SDS-PAGE-based fluorescence detection. It turned out that a prominent band was detected

from pH 4.0 to 7.0 at just below 100 kDa (**Fig. 6.9A**, left, **Fig. 6.S3C**), which matched the known molecular weight (94 kDa) of the snail GH2 enzyme.[54] Band quantification showed that optimum labeling occurred at pH 5.5, and the labeling decreased rapidly at higher and lower pH (**Fig. 6.9A**, right, red). Similar to the α-mannose configured Cy5 ABP **5**, the relative ABP labeling



Compounds	IC_{50} (μM)
8 (CWO466)	> 50
9 (TB535)	> 50
10 (TB429)	13.2 ± 1.78
11 (TB520)	6.16
12 (TB434)	3.59 ± 0.34
13 (TB476)	7.15 ± 1.14

Figure 6.9. Labeling and inhibitory potency of compounds towards GH2 β-mannosidase from Roman snail (*Helix pomatia*). A) ABP **12** labeling at various pHs (left) and comparison of the quantified band intensity with relative enzymatic activity across pHs (right). B) Labeling of ABP **12** at various ABP concentration. C) Labeling of ABP **12** at various incubation time. D) Apparent IC_{50} values of compounds **8**–**13**.

intensity was lower than the enzymatic activity at acidic pH. This effect was not due to the instability of the ABP at acidic pH, as labeling with ABPs pre-incubated at different pH values over a period of 0–60 min showed identical intensity (**Fig. 6.S3D**). At pH higher than 5.5, the relative labeling by ABP **12** generally matched the enzymatic activity.

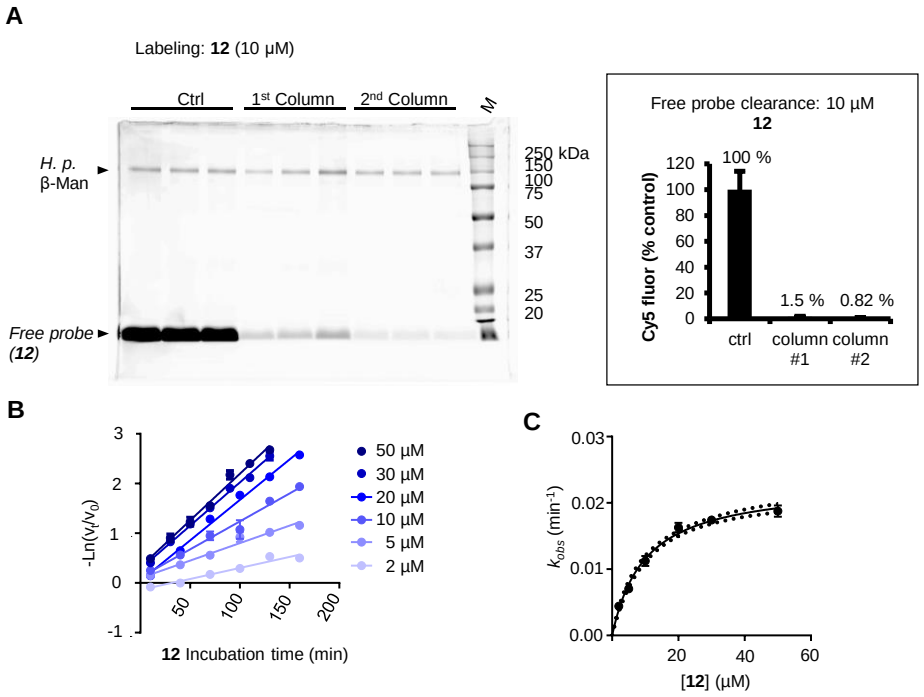
Using the determined optimal labeling pH (5.5), ABP **12** was next incubated with the β-mannosidase at various ABP concentration and labeling time. Labeling increased with both ABP concentration and labeling time, and saturate labeling in the experiments was observed at 10 μM ABP (1 h incubation) (**Fig. 6.9B**) and 120 min incubation (10 μM ABP) (**Fig. 6.9C**). Using these conditions (pH 5.5, 2 h incubation), IC_{50} measurement was performed again for all the β-mannose configured compounds. The results showed that at these reaction conditions, the compounds did inhibit the enzyme's hydrolysis of 4-MU-β-man (**Fig. 6.9D**). ABP **12** was the most potent of the series (apparent IC_{50} = 3.6 μM), followed by the BODIPY ABP **11**, the biotin ABP **13**, and the alkyl azide **10**. The bare epoxide **8** and aziridine **9** turned out not to inhibit the enzyme at the highest compound concentration tested (50 μM) (**Fig. 6.S3**).

6.2.3 Kinetic parameters determination for ABP **12** towards GH2 β-mannosidase

In previous chapters, a continuous method (simultaneous incubation of inhibitor and fluorogenic substrate) has been utilized for assessing the kinetic parameters of the β-glucuronidase ABP towards its target enzyme (Chapter 4, this thesis). However, due to its high potency and fast binding kinetics, only the pseudo-first order kinetic parameter (k_{inact}/K_I) could be obtained (Chapter 4, this thesis). A gel-based method has also been employed for the determination of kinetic parameters K_I (inhibition constant) and k_{inact} (inactivation rate constant) for the slower-reacting α-iduronidase ABP (Chapter 5, this thesis). While it completely circumvented the influence of added substrates, it was more laborious than the fluorogenic substrate assay. An alternative approach was used here, in which free ABP in the reaction mixture was removed by chromatography spin column, before the samples were added with 4-MU substrates and measured for fluorescence in 96-well plates. The extra step in removing free ABPs before adding substrates should theoretically prevent any influence from the substrates during ABP labeling, making the determination of kinetic parameters more accurate.

To verify this approach in free ABP removal, the snail GH2 enzyme was firstly labeled with 10 μM ABP **12** for 2 h at pH 5.5, and subjected to polyacrylamide desalting spin column

for two times. The extent of removal of unbound ABPs was checked by SDS-PAGE-based fluorescence detection. The gel showed that the unbound ABPs, which were migrated to the bottom of the gel, were mostly removed after consecutive spin column (**Fig. 6.10A**). Interestingly, the first spin column removed 98.5 % of the unbound ABPs in the samples, while the second column only removed less than half of the remaining unbound ABPs (**Fig. 6.10A**, right). The enzyme amount was partially affected by the spin columns, but a yield of 73 % was still obtained after the consecutive columns (**Fig. 6.S4**). With the method in place, the enzyme was next



D

Compound	K_i (µM)	k_{inact} (min^{-1})	k_{inact}/K_i ($\mu M^{-1}min^{-1}$)
12	9.92 ± 0.94	0.023 ± 0.001	0.0023

Figure 6.10. Determination of inhibitory kinetics of ABP 12 towards *Helix pomatia* GH2 β-mannosidase. A) Assessing free-probe clearance by consecutive desalting columns using SDS-PAGE-based fluorescence scanning (left); quantified relative band intensity from the free probe (right). Error ranges = SD from *n* = 2 experiments. B) Inverted logarithmic plot of relative enzymatic reaction rate of various ABP concentration at different incubation time C) Michaelis-Menten plot of the derived k_{obs} at various ABP concentration. D) Derived kinetic parameters. Error ranges = SD from *n* = 3 experiments.

incubated with ABP **12** at a range of ABP concentrations (2−50 μM), each performed in a series of incubation time (0−160 min). Calculation of the observed inhibition rate constant (k_{obs}) at each ABP concentration (**Fig. 6.10B**) and the subsequent plotting of the k_{obs} values at each ABP concentration (**Fig. 6.10C**) led to a K_I (9.92 μM) and k_{inact} (0.023 min^{-1}) (**Fig. 6.10D**). The calculated pseudo first-order inhibition rate constant (k_{inact} / K_I) is 0.0023 μM^{-1} min^{-1} (**Fig. 6.10D**), which is 10^{-4} times lower than the values for the Cy5 ABP for β- glucuronidase but 1 order higher than the Cy5 ABP for α-iduronidase. The K_I (9.92 μM) was 100-times lower than the K_M of the substrate 4-MU-β-man (0.91 mM)[54], suggesting that ABP **12** has a much higher affinity than the fluorogenic substrate.

6.2.5 Target detection and identification of the β-mannose configured ABPs in mouse kidney homogenates

Chemical proteomics was firstly employed for target identification for the β-mannose configured biotin ABP **13** in mouse kidney extracts, a tissue that is high in MANBA mRNA expression (according to BioGPS dataset GeneAtlas MOE430[56]). The experiment was performed with or without 1 h pre-incubation with 5 μM ABP **11** (BODIPY green). A negative control was also included in which DMSO replaced the ABPs. Analysis of the identified proteins showed that the β-mannosidase MANBA was the only glycosidase identified in samples incubated with ABP **13**, while no glycosidases were detected in the DMSO sample (**Table 6.3**). The sample with ABP **11** pre-incubation showed a reduced PLGS (ProteinLynx Global Server) score and a reduced number of total identified peptide assigned to MANBA, indicating that ABP **11** partially blocked the labeling of the biotin ABP **13** towards MANBA at the tested incubation condition (5 μM, 1 h).

Table 6.3. List of identified glycosidases by LC-MS-based proteomics in samples of mouse kidney extracts incubated with ABP 13. PLGS, ProteinLynx Global Server.

Sample	Accession	Entry	PLGS Score	Peptides	Theoretical peptides	Coverage (%)
# 1 (ABP **13**)	Q8K2I4	MANBA_MOUSE	2825	32	72	33
# 2 (ABP **11** → ABP **13**)	Q8K2I4	MANBA_MOUSE	733	18	72	22
# 3 (DMSO)	N/A	N/A	N/A	N/A	N/A	N/A

Next, the Cy5 ABP **12** was incubated with mouse kidney extracts and followed by SDS-PAGE-based fluorescence detection, to test if MANBA can be specifically visualized. Various ABP concentration (pH 5.5, 2 h), incubation time (pH 5.5, 3 µM), and pH (3 µM, 2 h) were tested. The results showed a distinct band at just below 100 kDa, that was visible at over 1 µM **12** (**Fig. 6.11A**), over 10 min incubation (**Fig. 6.11B**), and between pH 4.5 to 6.0 (optimally at pH 5.0 and 5.5) (**Fig. 6.11C**). The labeling did not saturate in intensity at 10 µM ABP and 120 min incubation, consistent with the labeling and inhibitory potency results with the *Helix pomatia*

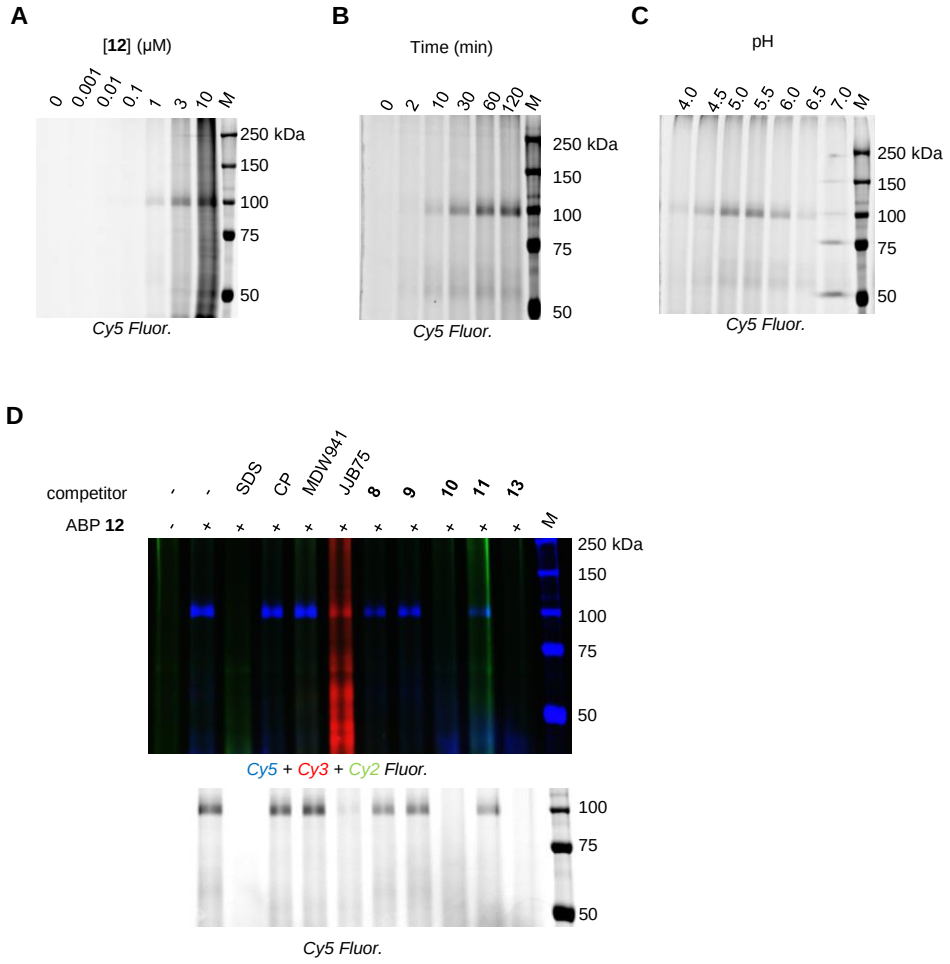


Figure 6.11. cABPP in mouse kidney homogenates. Samples were pre-incubated with competitors or SDS, followed by incubation with ABP **12**. CP, cyclophellitol. JJB75, cyclophellitol aziridine BODIPY-red ABP labeling retaining β-glucosidases.[52, 58]

GH2 β-mannosidase. In addition to the ~100 kDa band, a minor band at ~55 kDa was also noted in the gels. To further verify the identity of both bands, a cABPP experiment was conducted in which mouse kidney extracts were pre-incubated with inhibitors of glucocerebrosidase (GBA, which has a molecular weight of around 55 kDa), retaining β-glucosidases (including GBA2, that has a molecular weight at around 100 kDa), and compounds **8**−**11** and **13**, followed by ABP **12** incubation. The upper band was not abolished by pre-incubation with the GBA and GBA2 inhibitor cyclophellitol (CP, Chapter 2, this thesis) as well as the GBA-specific ABP MDW941[57], and only marginally reduced by compound **8**, **9**, **11**; it was abolished by pre-incubation with the retaining β-glucosidase ABP JJB75[52, 58], compound **10** and **13**, and SDS (**Fig. 6.11**). The labeling on the lower band was not intense enough in this experiment, but cABPP experiment and activity measurements in HEK293T lysates and recombinant enzymes revealed that GBA and GBA2 were both labeled by **12** (**Fig. 6.S5**). Together, these results suggest that the ~100 kDa band in mouse kidney homogenate was likely MANBA. Specific visualization of MANBA was obtainable in samples relatively abundant in MANBA and low in GBA and GBA2, or with pre-incubation with GBA and GBA2 inhibitor.

6.3 Discussion

This chapter describes the characterization of α- and β-mannose configured cyclophellitol aziridine ABPs' activities towards the GH38 α-mannosidases and GH2 β-mannosidase. Both enzyme families are retaining exo-glycosidases, and are involved in metabolism within the protein N-linked glycan pathways. Continuous interest is placed on a number of these enzymes in relation to diseases such as cancer and lysosomal storage disorders. As such, tools to visualize and profile the activity of each of these enzymes would be of great value in laboratory and clinical studies.

The α-mannose configured cyclophellitol and cyclophellitol aziridine compounds (except for the green BODIPY ABP **4**), are low micromolar inhibitors towards the Jack bean (*Canavalis ensiformis*) GH38 α-mannosidase. The alkyl Cy5 ABP **5**'s inhibitory potency increases with prolonged incubation time, and it labels the enzyme in a pH-, concentration-, and time-dependent manner. Its labeling of the enzyme is abolished by pre-incubation of swainsonine and mannostatin A, both known inhibitors of GH38 enzyme. These results suggest the ABP is a mechanism-based irreversible inhibitor of GH38 α-mannosidases. This conclusion is further

substantiated by structural analysis of the *Drosophila* GH38 enzyme dGMII in complex with compound **2**, which shows a glycosidic linkage between the anomeric carbon of the compound and the enzymes catalytic nucleophile, and that the compound adopts the $^{1}S_{5}$ conformation—identical to the one adopted by the enzyme's natural substrates. By gel-based fluorescent ABPP and chemical proteomics, it is shown that the ABPs offer in-class labeling of all five of the GH38 α-mannosidases in mouse tissue extracts, and in HEK293T cells expressing each of the cloned human GH38 enzymes. By tuning the labeling pH, individual mannosidases can be simultaneously profiled on gel in one experiment. The human MAN2B1 is optimally labeled at pH 4.0 to 5.5, and has multiple molecular weight forms (130 kDa, 65 kDa, 45 kDa, 30 kDa) which is likely a result of its complex lysosomal processing.[23, 32] The human MAN2B2 is processed from 130 kDa to 50 kDa. The 50 kDa form was not reported from previous literatures, which all isolated the enzyme in culture medium or tissue fluid instead of from homogenates of cells or whole tissue. Here, it is shown that this 50 kDa form has a pH profile similar to MAN2B1, and is ubiquitously expressed in all mouse tissue tested, as well as in HEK293T cell. Therefore, it is likely that this 50 kDa form of MAN2B2 is the mature lysosomal form, and conveniently, it can be readily distinguished from MAN2B1 based on an ABPP gel based on molecular weight—the first assay that offers simultaneous activity readout of the two enzymes in complex samples. The ABPP assay further identifies endogenous MAN2C1 in mouse brain extracts, which has a molecular weight of around 100 kDa and a pH range between 5.0 to 7.5 (optimally at pH 6.5). It also identifies endogenous MAN2A1 and MAN2A2 in mouse testis and possibly liver extracts, but due to their similar molecular weight it is not possible to discriminate between the two. The best ABP concentration for labeling is 3 μM, as higher concentrations results in more aspecific labeling that complicates interpretation of the results.

The ability of the α-mannose configured ABPs to label all five GH38 α-mannosidases invites two major applications. One would be gel-based fluorescent activity profiling for individual retaining exo-mannosidases across different sample types, to study the processing and activity of these enzymes in different cell/tissue types and in healthy vs. disease context. For example, it could be used to study MAN2C1's involvement in various cancer samples, or be used to monitor the bio-distribution and processing of therapeutic MAN2B1 (or MAN2B1 induced by other potential therapies) in α-mannosidosis patients. Another application is the screening for inhibitors/activators that are specific towards one of the five GH38 enzymes. For example, specific inhibitors for MAN2A1 would offer valuable leads for treating a number of cancers, a

research direction that has been actively pursued.[59, 60] Chaperone/activators for MAN2B1, on the other hand, could lead to alternative therapies for α-mannosidosis. A setup involving fluorescence polarization (FluoPol) ABPP coupled to automated high-throughput screening is currently being investigated (Daniel Lahav, ongoing investigation).

The N-alkylated β-mannose configured cyclophellitol and cyclophellitol aziridine compounds also inhibit their target enzyme—GH2 β-mannosidase—at low micromolar range. The bare cyclophellitol **9** and bare cyclophellitol aziridine **10**, surprisingly, did not inhibit the GH2 β-mannosidase. On the other hand, the Cy5 ABP **12** labels the commercial GH2 enzyme from *Helix pomatia* in a mechanism-based manner., It is noted that the ABP does not avidly label the enzyme between pH 4.0 to 5.0, in contrast to activity measurement using the fluorogenic substrates. As both the enzyme (at 0.1 % (w/v) BSA) and the ABP are stable at this pH range, this noted difference might reflect the intrinsic difference of the enzyme in its reactivity towards the two types of artificial compounds under low pH. Proteomics with the biotin ABP **13** in mouse kidney extracts identified the mouse GH2 enzyme MANBA as the only glycosidase targets. ABPP with the Cy5 ABP **12** in mouse kidney homogenates confirmed that MANBA (around 96 kDa) is labeled, while ABPP in other sample types revealed that the β-glucosidases GBA and GBA2 are also targeted. Interestingly, MANBA in mouse kidney is also labeled by the β-glucose configured cyclophellitol aziridine ABP at the experimental condition (200 nM, 2 h incubation). Despite this, by pre-incubation of specific β-glucosidase inhibitors such as cyclophellitol, MANBA can still be selectively visualized over the β-glucosidases. Next, a novel experimental protocol for determining inhibition kinetic parameters for irreversible inhibitors has also been setup, which is relatively free of substrate during inhibitor incubation and thus should offer a better estimate of inhibition kinetic parameters when compare to the commonly used continuous methods. Application of the β-mannose configured ABPs might not be only restricted to the study in the rare lysosomal storage disorder β-mannosidosis. For instance, it could also be applied in studying other human diseases with abnormal MANBA activity, such as kidney disease.[61]

In conclusion, the presented study in this chapter established that the mannose configured cyclophellitol aziridine ABPs are valuable tools to profile individual retaining exo-mannosidases in complex biological samples. They invite future application in laboratory and clinical investigation of these enzymes in both health and disease.

6.4 Experimental procedures

6.4.1 General materials and methods

All chemicals and reagents were purchased from Sigma Aldrich, if not otherwise stated. HEK293T cells were purchased from ATCC and handled according to the published methods.[62] GH38 α-mannosidase from Jack bean (*Canavalia ensiformis*) and GH2 β-mannosidase from Roman snail (*Helix pomatia*) were purchased from Sigma Aldrich; recombinant GH38 α-mannosidase from *Drosophila melanogaster* (dGMII) was generated according to the methods described in the Appendix section **6.S2.1** and **6.S2.2**. Protein concentration in samples was determined using the Pierce™ BCA kit from Thermo Fisher; for the commercial enzymes, the enzyme stocks were firstly desalted using Pierce™ 7k polyacrylamide desalting spin column (Thermo Fisher) before subjecting to BCA assay; protein concentrations was 1.45 μg (13 pmol) μL^{-1} for *C. e.* α-mannosidase and 2.41 μg (25.6 pmol) μL^{-1} for *H. p.* β-mannosidase. DMSO concentration in samples was kept at 0.5 to 1 % (v/v) during inhibitor/ABP incubation. Coomassie stain were carried out as loading control for SDS-PAGE experiments. HEK293T cells (ATCC CRL-3216) were cultured in DMEM high glucose (Sigma-Aldrich) supplemented with 10 % FCS, 100 units/mL penicillin/streptomycin, and 1 mM Glutamax at 37 °C and at 7 % CO_2.

6.4.2 IC_{50} determination for compound **1**−**7** towards *Canavalia ensiformis* GH38 α-mannosidase

The enzyme (1.45 ng, or 13 fmol) was equilibrated in 12.5 μL McIlvaine buffer (150 mM citric acid/Na_2HPO_4, pH 4.5) for 5 min on ice, and incubated with 12.5 μL inhibitor dilutions (prepared in McIlvaine buffer) for 30 min at 37 °C in black flat-bottom medium-binding 96-well plates (Greiner) in triplicates. Samples were next incubated with 100 μL of substrate mixture (10 mM 4-methylumbelliferyl (4-MU)-α-D-mannopyranoside (4-MU-α-man) in McIlvaine buffer) for 30 min at 37 °C, or for 0−120 min for compound **5**. Reaction was quenched by adding 200 μL 1M Glycine-NaOH (pH 10.3) to the samples. 4-MU fluorescence was measured in the plates using a LS55 fluorometer (PerkinElmer) at λ_{Ex} = 366 nm and λ_{Em} = 445 nm. Measured values were subtracted with background (no enzyme) values, normalized against control values (no inhibitor, with enzyme), and plotted against inhibitor concentrations. IC_{50} values were calculated with Prism 7.0 (GraphPad), using one-phase-exponential decay function. Standard deviations were obtained from 2 sets of calculated IC_{50} values.

6.4.3 Enzymatic activity measurement at various pH

Canavalia ensiformis GH38 α-mannosidase (13 fmol) or *Helix pomatia* GH2 β-mannosidase (340 fmol) were equilibrated in 25 μL McIlvaine buffer for 10 min at 37 °C at various pH (+ 1 mM $ZnCl_2$ for *Canavalia ensiformis* GH38 α-mannosidase), followed by 30 min incubation with 100 μL substrate mixtures (10 mM 4-MU-α-D-mannosylpyranoside (Glycosynth) or 2 mM 4-methylumbelliferyl (4-MU)-β-D-mannosylpyranoside (Glycosynth), 0.1 % (w/v) BSA) at 37 °C. After quenching the reaction by adding 200 μL stop buffer, fluorescence from samples were detected and quantified following methods in a previous section (**6.4.2**).

6.4.4 ABPP with *Canavalia ensiformis* GH38 α-mannosidase

13 fmol of enzyme was equilibrated in 10 μL McIlvaine buffer (+ 1 mM $ZnCl_2$; pH 5.5 if not otherwise indicated) for 5 min on ice, and incubated with ABP **5** (3 μM during incubation, if not otherwise indicated) or ABP **7** (0.1−10 μM during incubation) for 30 min (if not otherwise indicated) at 37 °C. For cABPP experiment, same enzyme dilution was pre-incubated with swainsonine (Cayman Chemical) or mannostatin A (Santa Cruz) at 0.01−3,000 μM for 30 min at 37 °C, followed by ABP **5** incubation (3 μM) for 10 min at 37 °C. After ABP incubation, proteins were denatured and separated by SDS-PAGE, detected and analysed according to the previously described methods.[62, 63]

6.4.5 ABPP with ABP **5** in mouse tissue extracts

Mouse tissue extracts were generated by homogenizing mouse testis sample in four to five volumes of KPi buffer (25 mM K_2HPO_4 / KH_2PO_4, pH 6.5, 0.1 % (v/v) Triton X-100, Protease inhibitor cocktail (EDTA-free, Roche)) in a 2.0 mL screw-cap Eppendorf tube, using 1mm sterile glass beads and FastPrep 24 homogenizer (MP Biomedicals) at 6 m s^{-1} rpm for 20 sec for 3 times. The homogenates were centrifuged at 16,000 rpm for 10 min at 4 °C, and the supernatant was collected and measured for protein concentration using BCA method. Samples were stored in aliquots at −20 °C. For ABPP, samples (40 μg protein) were diluted in 10 μL McIlvaine buffer (150 mM, pH 5.5, if not otherwise stated) and incubated with 5 μL ABP **5** (diluted in McIlvaine buffer) to a final ABP concentration of 3 μM (if not otherwise stated) for 30 min at 37 °C. For competitive ABPP, mouse testis extracts (40 μg protein) were diluted in 10 μL McIlvaine buffer (pH 5.5) and pre-incubated with 0−1,000 μM swainsonine or mannostatin

A in 12.5 µL volume for 30 min at 37 °C, followed by ABP **5** incubation (3 µM) in 15 µL volume for 10 min at 37 °C. Samples were denatured and analyzed by gel-based fluorescent ABPP as described previously.

6.4.6 Proteomics

For pull-down with ABP **6**, 2.5 mg protein from mouse testis extracts were diluted with McIlvaine buffer (750 mM, pH 4.5 or 6.0) to a total volume of 500 µL, and incubated with 10 µM ABP **6** at 37 °C for 1 h. For control, DMSO was used in place of ABP. For pull-down with ABP **13**, 4.0 mg total protein from the kidney extracts were diluted with McIlvaine buffer (750 mM, pH 5.5) in a total volume of 500 µL, and incubated with 10 µM ABP **13** at 37 °C for 1 h, with or without prior pre-incubation with 5 µM ABP **11**. After ABP incubation, samples were denatured with SDS, subjected to chloroform/methanol precipitation (C/M), reduction/alkylation, C/M again, and pull-down with 200 µL streptavidin beads in a volume of 10.2 mL pull-down buffer for O/N at 4 °C following the previously described procedures.[64] Afterwards, half of the samples were subjected to on-bead digestion, and half to in-gel digestion, and desalted using stage-tips according to the described procedures.[64] For LC/MS analysis, 1 µL of sample was injected with phase A (0.1 % (v/v) formic acid in MilliQ H_2O) on a C18 column (Acquity UPLC M-Class 300 µm x 50 mm, packed with BEH C18 material of 1.7 µm diameter and 300 Å pore size particles), eluted with a 50 min gradient of 10 % to 60 % phase B (0.1 % (v/v) formic acid in ACN), followed by 10 min equilibration to 1% phase B at a flow of 0.4 µL/min, hyphenated with Electro-spray ionization (ESI) via Nano-spray source with ESI emitters (New Objectives) fused silica tubing 360 µm OD x 25 µm ID tapered to 5 ± 0.5 µm (5 nL/cm void volume) to a Synapt G2Si mass spectrometer (Waters) operating with Masslynx for acquisition and Ent3 software for polymer envelope signal deconvolution. The following settings in positive resolution mode were used: source temperature of 80 °C, capillary voltage 4.5 kV, nano flow gas of 0.25 Bar, purge gas 250 L/h, trap gas flow 2.0 ml/min, cone gas 100 L/h, sampling cone 25V, source offset 25, trap CE 32V, scan time 3.0 sec, mass range 400-2400 m/z. Lock mass acquiring was done with a mixture of Leu Enk (556.2771) and Glu Fib (785.84265), lockspray voltage 3.5 kV, Glufib fragmentation was used as calibrant. The PLGS (Waters) program was used for data analysis, protein ID or extraction of mgf files for further Mascot (Matrix Science) search analysis. The identification results were exported as Excel file including protein accession numbers, mass of the protein, pI, PLGS score, and % coverage of the protein

by amino acids identified by LC/MS.

6.4.7 Cloning and transient expression of human GH38 α-mannosidases in HEK203T cells

The coding sequences from human GH38 α-mannosidases were PCR-amplified from total cDNA's from human Gaucher spleen using primers listed in **Table 6.S1**. Primers were designed based on NCBI reference sequence NM_002372.3 (MAN2A1), NM_001320977.1 (MAN2A2), NM_000528.4 (MAN2B1), NM_001292038.1 (MAN2B2), and NM_006715.4 (MAN2C1), and cloned into pDNOR-221 and subcloned into pcDNA3.1/Zeo using the Gateway system (Invitrogen). Correctness of the construct was verified by sequencing. Sub-confluent HEK293T cells were transfected with the generated plasmids (or vector plasmids) by the PEI method with a plasmid:PEI ratio 1:3. Media was refreshed 24 h later, and 72 h after transfection, cells were washed three times with phosphate buffered saline (PBS) buffer and collected in KPi buffer. The cell suspension was incubated on ice for 30 min, and stored at −80 °C.

6.4.8 ABPP in lysates of cells expressing the cloned GH38 α-mannosidases

Lysates (20 μg protein) were diluted in 10 μL McIlvaine buffer (150 mM, pH 3.5−7.5, 1 mM $ZnCl_2$), and incubated with 5 μL ABP **5** (diluted in DMSO and 150 mM McIlvaine buffer pH 3.5−7.5, 1 mM $ZnCl_2$) at a final ABP concentration of 3 μM for 30 min at 37 °C. Samples were denatured and analyzed by gel-based fluorescent ABPP as described previously.

6.4.9 Stability test for *Helix pomatia* GH2 β-mannosidase

For the effect of supplements, 340 pmol enzyme was diluted in 25 μL McIlvaine buffer (150 mM) with or without BSA (0.1 % (w/v)), Triton X-100 (0.1 % (v/v)), sodium taurocholate (0.2 % (w/v)), or the combinations of these for 0−120 min at pH 5.5, before subjecting to enzymatic assay (with 100 μL 2 mM 4-MU-β-D-mannopyranoside, 30 min incubation at 37 ºC, pH 5.5, 0.1 % (w/v) BSA). For the effect of pH, same enzyme dilutions were prepared in McIlvaine buffer (+ 0.1 % (w/v) BSA) at various pH values, and incubated for 0−60 min at 37 ºC. Samples were next subjected to the substrate assay (prepared with McIlvaine buffer at matching pH) before activity readout.

6.4.10 ABPP with *Helix pomatia* GH2 β-mannosidase

7.3−51.2 pmol of enzyme was equilibrated in 10 μL McIlvaine buffer (+ 0.1 % w/v BSA, pH

5.5, if not otherwise indicated) for 5 min on ice, and incubated with ABP **12** (3 μM during incubation, if not otherwise indicated) for 30 min (if not otherwise indicated) at 37 °C. After ABP incubation, samples were subjected to SDS-PAGE-based fluorescence detection followed the previously described methods.[62, 63]

6.4.11 IC_{50} determination for compound **8**−**13** towards *Helix pomatia* GH2 β-mannosidase

The enzyme (32.2 ng, or 340 fmol) was equilibrated in 12.5 μL McIlvaine buffer (150 mM, pH 5.5, + 0.1 % v/v Triton X-100) for 5 min on ice, and incubated firstly with the compounds (12.5 μL) for 2 h at 37 °C and secondly with substrates (2 mM 4-MU-β-man) for 1 h at 37 °C. Fluorescence detection and data analysis followed the procedures described in a previous section (**6.4.2**).

6.4.12 Determination of kinetic parameters of ABP **12** towards *helix pomatia* GH2 β-mannosidase

0.9 μg (9.6 pmole) of the enzyme was equilibrated in 120 μL McIlvaine buffer pH 5.5 (150mM, 0.1% (v/v) Triton X-100) and pre-warmed to 37°C for 10 min. For inhibitor incubation, enzymes were combined with equal volume of pre-warmed ABP (prepared in McIlvaine buffer) at 2−50 μM ABP concentrations during incubation, and incubated on a thermoshaker at 37°C. Thereafter, 30 μL aliquots were taken at different time points (0−160 min) from each sample, and subjected to free-ABP removal using polyacrylamide desalting spin column (Pierce). The eluent were diluted 8x with McIlvaine buffer pH 4.5 (150 mM, 0.1% (v/v) Triton-X100), and loaded onto a 96-well plate in 25 μL triplicates. For substrate incubation, 100 μL of pre-warmed substrate mixture (2mM 4-MU-β-D-mannosylpyranoside , dissolved in 150 mM McIlvaine buffer pH 4.5 + 0.1% (v/v) Triton X-100) was added to the samples, and allowed incubation at 37°C for 1 h before subsequent activity measurement. The measured activity was converted to $-\mathrm{Ln}(v_t/v_0)$, with v_t being the rate of substrate hydrolysis at the given time point for a given ABP concentration, and v_0 being the rate at the given time point without ABP incubation. k_{obs} was next determined from the $-\log(V_t/V_0)$ *vs* time plot, using linear regression (GraphPad Prism 7.0). The subsequent k_{obs} *vs* [ABP **12**] plot was used to determine k_{inact} and K_I for ABP **12** towards the enzyme, using the Michaelis-Menten curve-fitting (GraphPad Prism 7.0). For fluorescent gel-based assessment in enzyme yield and free ABP removal, 153.6 pmol enzyme was diluted in 90 μL McIlvaine buffer pH 5.5, and this was incubated with 40 μL ABP 12 for 2 h at 37 ºC, at a

final ABP concentration of 10 μM or 1 μM. Thereafter, 100 μL from each sample was desalted using the spin column, and 90 μL from the eluate was desalted again. Next, 10 μL from the eluates (desalted once or twice) and the un-desalted sample were diluted with 140 μL McIlvaine buffer pH 5.5, and 15 μL from these were loaded onto 10 % polyacrylamide gels for SDS-PAGE and fluorescence detection. The electrophoresis was stopped when the dye front has migrated to about 1 cm from the bottom of the gel.

6.4.13 ABPP using ABP **12** in mouse kidney extracts

25 μg total protein from mouse kidney extracts was diluted in McIlvaine buffer (150 mM, pH 5.5, if not otherwise stated) in a total volume of 10 μL, and incubated with 5 μL ABP **12** (diluted in DMSO and McIlvaine buffer) at final ABP concentration of 3 μM (if not otherwise stated) for 2 h (if not otherwise stated) at 37 °C. cABPP was performed with pre-incubating the extracts with SDS (2 % (w/v)), cyclophellitol (3 μM), ABP MDW941[57] (3 μM), ABP JJB75[52, 58] (3 μM), and compound **8**−**9** (50 μM), **10** (50 μM), **11** (3 μM), and **13** (50 μM) in a volume of 12.5 μL for 2 h at 37 °C, followed by ABP **12** incubation (3 μM) in a volume of 15 μL for 2 h at 37 °C. Samples were denatured and analyzed by gel-based fluorescent ABPP as described previously.

6.4.13 ABPP using ABP **12** in lysates of GBA2-overexpressing HEK293T cells

18.9 μg of lysates from GBA2-overexpressing HEK293T cells (Chapter 2)[62] were pre-incubated with 1 μM of cyclophellitol, the cyclophellitol aziridine BODIPY-red ABP JJB75, compound **8**, **10**, **11**, **13**, or 2 % (w/v) SDS (with 5 min boiling at 98 ºC when pre-incubation completed) at pH 5.5 for 1 h at 37 ºC, and next incubated with 1 μM ABP **12** at pH 5.5 for 2 h at 37 ºC. Samples were denatured and subjected to SDS-PAGE and fluorescence detection. The gel was stained with Coomassie Brilliant Blue G250 for assessing total protein loading amount.

6.4.14 Apparent IC_{50} values of ABP **12** towards recombinant GBA or GBA2

Assays were performed with recombinant GBA, or GBA2 from lysates of overexpressing HEK293T cells using ABP **12** according to the described methods in Chapter 2.[62]

6.5 References

1 Stanley P, Taniguchi N & Aebi M (2017) N-Glycans. In Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, Darvill AG, Kinoshita T, Packer NH, Prestegard JH, Schnaar RL & Seeberger PH (Eds) *Essentials of Glycobiology [Internet]. 3rd edition*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

2 Schwarz F & Aebi M (2011) Mechanisms and principles of N-linked protein glycosylation. *Curr Opin Struct Biol* **21**, 576–582.

3 Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM & Henrissat B (2014) The Carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res* **42**, D490–D495.

4 Avezov E, Frenkel Z, Ehrlich M, Herscovics A & Lederkremer GZ (2008) Endoplasmic reticulum (ER) mannosidase I is compartmentalized and required for N-glycan trimming to Man5-6GlcNAc2 in glycoprotein ER-associated degradation. *Mol Biol Cell* **19**, 216–225.

5 Wu Y, Swulius MT, Moremen KW & Sifers RN (2003) Elucidation of the molecular logic by which misfolded alpha 1-antitrypsin is preferentially selected for degradation. *Proc Natl Acad Sci USA* **100**, 8229–8234.

6 Helenius A & Aebi M (2004) Roles of N-linked glycans in the endoplasmic reticulum. *Annu Rev Biochem* **73**, 1019–1049.

7 Słomińska-Wojewódzka M & Sandvig K (2015) The Role of Lectin-Carbohydrate Interactions in the Regulation of ER-Associated Protein Degradation. *Molecules* **20**, 9816—9846.

8 Lal A, Schutzbach JS, Forsee WT, Neame PJ & Moremen KW (1994) Isolation and expression of murine and rabbit cDNAs encoding an alpha 1,2-mannosidase involved in the processing of asparagine-linked oligosaccharides. *J Biol Chem* **269**, 9872–9881.

9 Herscovics A, Schneikert J, Athanassiadis A & Moremen KW (1994) Isolation of a mouse Golgi mannosidase cDNA, a member of a gene family conserved from yeast to mammals. *J Biol Chem* **269**, 9864–9871.

10 Tremblay LO & Herscovics A (2000) Characterization of a cDNA encoding a novel human Golgi alpha 1, 2-mannosidase (IC) involved in N-glycan biosynthesis. *J Biol Chem* **275**, 31655–31660.

11 Braulke T & Bonifacino JS (2009) Sorting of lysosomal proteins. *Biochim Biophys Acta* **1793**, 605–614.

12 Lubas WA & Spiro RG (1987) Golgi endo-alpha-D-mannosidase from rat liver, a novel N-linked carbohydrate unit processing enzyme. *J Biol Chem* **262**, 3775–3781.

13 Lubas WA & Spiro RG (1988) Evaluation of the role of rat liver Golgi endo-alpha-D-mannosidase in processing N-linked oligosaccharides. *J Biol Chem* **263**, 3990–3998.

14 Moremen KW & Robbins PW (1991) Isolation, characterization, and expression of cDNAs encoding murine alpha-mannosidase II, a Golgi enzyme that controls conversion of high mannose to complex N-glycans. *J Cell Biol* **115**, 1521–1534.

15 Misago M, Liao YF, Kudo S, Eto S, Mattei MG, Moremen KW & Fukuda MN (1995) Molecular cloning and expression of cDNAs encoding human alpha-mannosidase II and a previously unrecognized alpha-mannosidase IIx isozyme. *Proc Natl Acad Sci USA* **92**, 11766–11770.

16 Shah N, Kuntz DA & Rose DR (2008) Golgi alpha-mannosidase II cleaves two sugars sequentially in the same catalytic site. *Proc Natl Acad Sci USA* **105**, 9570–9575.

17 Akama TO, Nakagawa H, Wong NK, Sutton-Smith M, Dell A, Morris HR, Nakayama J, Nishimura S, Pai A, Moremen KW, Marth JD & Fukuda MN (2006) Essential and mutually compensatory roles of alpha-mannosidase II and alpha-mannosidase IIx in N-glycan processing in vivo in mice. *Proc Natl Acad Sci USA* **103**, 8983–8988.

18 Fukuda MN, Masri KA, Dell A, Luzzatto L & Moremen KW (1990) Incomplete synthesis of N-glycans in congenital dyserythropoietic anemia type II caused by a defect in the gene encoding alpha-mannosidase II. *Proc Natl Acad Sci USA* **87**, 7443–7447.

19 Fukuda MN & Akama TO (2002) *In vivo* role of alpha-mannosidase IIx: ineffective spermatogenesis resulting from targeted disruption of the Man2a2 in the mouse. *Biochim Biophys Acta* **1573**, 382–387.

20 Suzuki T, Hara I, Nakano M, Shigeta M, Nakagawa T, Kondo A, Funakoshi Y & Taniguchi N (2006) Man2C1, an alpha-mannosidase, is involved in the trimming of free oligosaccharides in the cytosol. *Biochem J* **400**, 33–41.

21 Suzuki T & Harada Y (2014) Non-lysosomal degradation pathway for N-linked glycans and dolichol-linked

oligosaccharides. *Biochem Biophys Res Commun* **453**, 213–219.

22 Kuokkanen E, Smith W, Mäkinen M, Tuominen H, Puhka M, Jokitalo E, Duvet S, Berg T & Heikinheimo P (2007) Characterization and subcellular localization of human neutral class II alpha-mannosidase [corrected]. *Glycobiology* **17**, 1084–1093.

23 Liao YF, Lal A & Moremen KW (1996) Cloning, expression, purification, and characterization of the human broad specificity lysosomal acid alpha-mannosidase. *J Biol Chem* **271**, 28348–28358.

24 Damme M, Morelle W, Schmidt B, Andersson C, Fogh J, Michalski JC & Lübke T (2010) Impaired lysosomal trimming of N-linked oligosaccharides leads to hyperglycosylation of native lysosomal proteins in mice with alpha-mannosidosis. *Mol Cell Biol* **30**, 273–83.

25 De Gasperi R, Daniel PF & Warren CD (1992) A human lysosomal alpha-mannosidase specific for the core of complex glycans. *J Biol Chem* **267**, 9706–9712.

26 Park C, Meng L, Stanton LH, Collins RE, Mast SW, Yi X, Strachan H & Moremen KW (2005) Characterization of a human core-specific lysosomal alpha 1,6-mannosidase involved in N-glycan catabolism. *J Biol Chem* **280**, 37204–37216.

27 Okamura N, Tamba M, Liao HJ, Onoe S, Sugita Y, Dacheux F & Dacheux JL (1995) Cloning of complementary DNA encoding a 135-kilodalton protein secreted from porcine corpus epididymis and its identification as an epididymis-specific alpha-mannosidase. *Mol Reprod Dev* **42**, 141–148.

28 Alkhayat AH, Kraemer SA, Leipprandt JR, Macek M, Kleijer WJ & Friderici KH (1998) Human beta-mannosidase cDNA characterization and first identification of a mutation associated with human beta-mannosidosis. *Hum Mol Genet* **7**, 75–83.

29 Samra ZQ & Athar MA (2008) Cloning, sequence, expression and characterization of human beta-mannosidase. *Acta Biochim Pol* **55**, 479–490.

30 van den Elsen JM, Kuntz DA & Rose DR (2001) Structure of Golgi alpha-mannosidase II: a target for inhibition of growth and metastasis of cancer cells. *EMBO J* **20**, 3008–3017.

31 Goss PE, Reid CL, Bailey D & Dennis JW (1997) Phase IB clinical trial of the oligosaccharide processing inhibitor swainsonine in patients with advanced malignancies. *Clin Cancer Res* **3**, 1077–1086.

32 Heikinheimo P, Helland R, Leiros HK, Leiros I, Karlsen S, Evjen G, Ravelli R, Schoehn G, Ruigrok R, Tollersrud OK, McSweeney S & Hough E (2003) The structure of bovine lysosomal alpha-mannosidase suggests a novel mechanism for low-pH activation. *J Mol Biol* **327**, 631–44.

33 Malm D & Nilssen Ø (2008) Alpha-mannosidosis. *Orphanet J Rare Dis* **3**, 21.

34 Ceccarini MR, Codini M, Conte C, Patria F, Cataldi S, Bertelli M, Albi E & Beccari T (2018) Alpha-Mannosidosis: Therapeutic Strategies. *Int J Mol Sci* **19**, E1500.

35 Harmatz P, Cattaneo F, Ardigò D, Geraci S, Hennermann JB, Guffon N, Lund A, Hendriksz CJ & Borgwardt L (2018) Enzyme replacement therapy with velmanase alfa (human recombinant alpha-mannosidase): Novel global treatment response model and outcomes in patients with alpha-mannosidosis. *Mol Genet Metab* **124**, 152–160.

36 Bedilu R, Nummy KA, Cooper A, Wevers R, Smeitink J, Kleijer WJ & Friderici KH (2002) Variable clinical presentation of lysosomal beta-mannosidosis in patients with null mutations. *Mol Genet Metab* **77**, 282–290.

37 Jones MZ & Dawson G (1981) Caprine beta-mannosidosis. Inherited deficiency of beta-D-mannosidase. *J Biol Chem* **256**, 5185–5188.

38 Tian Y, Ju JY, Zhou YQ, Liu Y & Zhu LP (2008) Inhibition of alpha-mannosidase Man2c1 gene expression suppresses growth of esophageal carcinoma cells through mitotic arrest and apoptosis. *Cancer Sci* **99**, 2428–2434.

39 Xiang ZG, Jiang DD, Liu Y, Zhang LF & Zhu LP (2010) hMan2c1 transgene promotes tumor progress in mice. *Transgenic Res* **19**, 67–75.

40 He L, Fan C, Kapoor A, Ingram AJ, Rybak AP, Austin RC, Dickhout J, Cutz JC, Scholey J & Tang D (2011) α-Mannosidase 2C1 attenuates PTEN function in prostate cancer cells. *Nat Commun* **2**, 307.

41 Wang L & Suzuki T (2013) Dual functions for cytosolic α-mannosidase (Man2C1): its down-regulation causes mitochondria-dependent apoptosis independently of its α-mannosidase activity. *J Biol Chem* **288**, 11887–11896.

42 Daniel PF, Winchester B & Warren CD (1994) Mammalian alpha-mannosidases—multiple forms but a common purpose? *Glycobiology* **4**, 551–566.

43 Numao S, Kuntz DA, Withers SG & Rose DR (2003) Insights into the mechanism of Drosophila melanogaster

Golgi alpha-mannosidase II through the structural analysis of covalent reaction intermediates. *J Biol Chem* **278**, 48074–48083.

44 Tailford LE, Offen WA, Smith NL, Dumon C, Morland C, Gratien J, Heck MP, Stick RV, Blériot Y, Vasella A, Gilbert HJ & Davies GJ (2008) Structural and biochemical evidence for a boat-like transition state in beta-mannosidases. *Nat Chem Biol* **4**, 306-312.

45 Davies GJ, Planas A, Rovira C (2012) Conformational analyses of the reaction coordinate of glycosidases. *Acc Chem Res* **45**, 308–316.

46 Wong C-S (2015) Chapter 5. In *The synthesis of mannose-derived bioconjugates and enzyme inhibitors* (Doctoral dissertation) (pp 141−116). Retrieved from Leiden University Repository.

47 Ritzen B, van Oers MC, van Delft FL & Rutjes FP (2009) Enantioselective chemoenzymatic synthesis of trans-aziridines. *J Org Chem.* 74, 7548-7551.

48 Beenakker TJM (2018) Chapter 4. In *Design and development of conformational inhibitors and activity-based probes for retaining glycosidases* (Doctoral dissertation) (pp 47−60). Retrieved from Leiden University Repository.

49 Beenakker TJM (2018) Chapter 5. In *Design and development of conformational inhibitors and activity-based probes for retaining glycosidases* (Doctoral dissertation) (pp 61−74). Retrieved from Leiden University Repository.

50 Howard S, Braun C, McCarter J, Moremen KW, Liao YF & Withers SG (1997) Human lysosomal and jack bean alpha-mannosidases are retaining glycosidases. *Biochem Biophys Res Commun* **238**, 896−898.

51 Tatsuta K, Niwata Y, Umezawa K, Toshima K & Nakata M (1991) Syntheses and enzyme inhibiting activities of cyclophellitol analogs. *J Antibiot (Tokyo)* **44**, 912−914.

52 Jiang J, Beenakker TJ, Kallemeijn WW, van der Marel GA, van den Elst H, Codée JD, Aerts JM & Overkleeft HS (2015) Comparing Cyclophellitol N-Alkyl and N-Acyl Cyclophellitol Aziridines as Activity-Based Glycosidase Probes. *Chemistry* **21**, 10861−10869.

53 Howard S, He S & Withers SG (1998) Identification of the active site nucleophile in jack bean alpha-mannosidase using 5-fluoro-beta-L-gulosyl fluoride. *J Biol Chem* **273**, 2067−2072.

54 McCleary BV (1983) beta-D-Mannosidase from Helix pomatia. Carbohydr Res 111, 297−310.

55 Remen L & Vasella A (2002) *Helv Chim Acta* **85**, 1118−1127.

56 Wu C, Jin X, Tsueng G, Afrasiabi C & Su AI (2016) BioGPS: building your own mash-u of gene annotations and expression profiles. *Nucleic Acids Res* **44**(D1), D313−316.

57 Witte MD, Kallemeijn WW, Aten J, Li KY, Strijland A, Donker-Koopman WE, van den Nieuwendijk AM, Bleijlevens B, Kramer G, Florea BI, Hooibrink B, Hollak CE, Ottenhoff R, Boot RG, van der Marel GA, Overkleeft HS & Aerts JM (2010) Ultrasensitive in situ visualization of active glucocerebrosidase molecules. *Nat Chem Biol* **6**, 907−913.

58 Kallemeijn WW, Li KY, Witte MD, Marques AR, Aten J, Scheij S, Jiang J, Willems LI, Voorn-Brouwer TM, van Roomen CP, Ottenhoff R, Boot RG, van den Elst H, Walvoort MT, Florea BI, Codée JD, van der Marel GA, Aerts JM & Overkleeft HS (2012) Novel activity-based probes for broad-spectrum profiling of retaining β-exoglucosidases *in situ* and *in vivo*. *Angew Chem Int Ed Engl* **51**, 12529−12533.

59 Howard E, Cousido-Siah A, Lepage ML, Schneider JP, Bodlenner A, Mitschler A, Meli A, Izzo I, Alvarez HA, Podjarny A & Compain P (2018) Structural basis of outstanding multivalent effects in jack bean α-mannosidase inhibition. *Angew Chem Int Ed Engl* **57**, 8002−8006.

60 Gerber-Lemaire S & Juillerat-Jeanneret L (2010) Studies toward new anti-cancer strategies based on alpha-mannosidase inhibition. *Chimia (Aarau)* **64**, 634−639.

61 Ko YA, Yi H, Qiu C, Huang S, Park J, Ledo N, Köttgen A, Li H, Rader DJ, Pack MA, Brown CD & Susztak K (2017) Genetic-Variation-Driven Gene-expression changes highlight genes with important functions for kidney disease. *Am J Hum Genet* **100**, 940−953.

62 Kuo CL, Kallemeijn WW, Lelieveld LT, Mirzaian M, Zoutendijk I, Vardi A, Futerman AH, Meijer AH, Spaink HP, Overkleeft HS, Aerts JMFG & Artola M (2019) *In vivo* inactivation of glycosidases by conduritol B epoxide and cyclophellitol as revealed by activity-based protein profiling. *FEBS J* **286**, 584−600.

63 Kuo CL, van Meel E, Kytidou K, Kallemeijn WW, Witte M, Overkleeft HS, Artola ME & Aerts JM (2018) Activity-based probes for glycosidases: profiling and other applications. *Methods Enzymol* **598**, 217−235.

64 Li N, Kuo CL, Paniagua G, van den Elst H, Verdoes M, Willems LI, van der Linden WA, Ruben M, van Genderen E, Gubbens J, van Wezel GP, Overkleeft HS & Florea BI (2013) Relative quantification of proteasome activity

by activity-based protein profiling and LC-MS/MS. *Nat Protoc* **8**, 1155−1168.

APPENDIX

6.S1. Supporting Figures and Tables

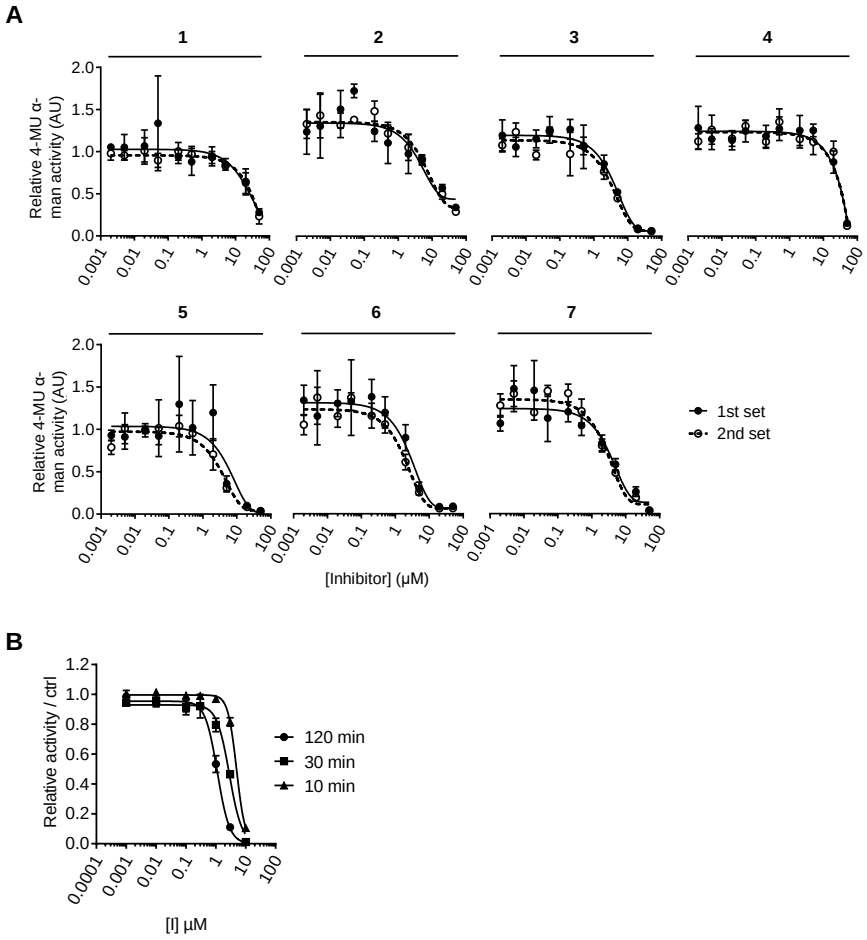


Figure 6.S1. Inhibition curves of compound 1−7 towards Jack bean (*Canavalis ensiformis*) GH38 α-mannosidase for apparent IC_{50} determination. A) 30 min incubation. B) 30−120 min incubation. Error range = ± SD (*n* = 3 technical replicates).

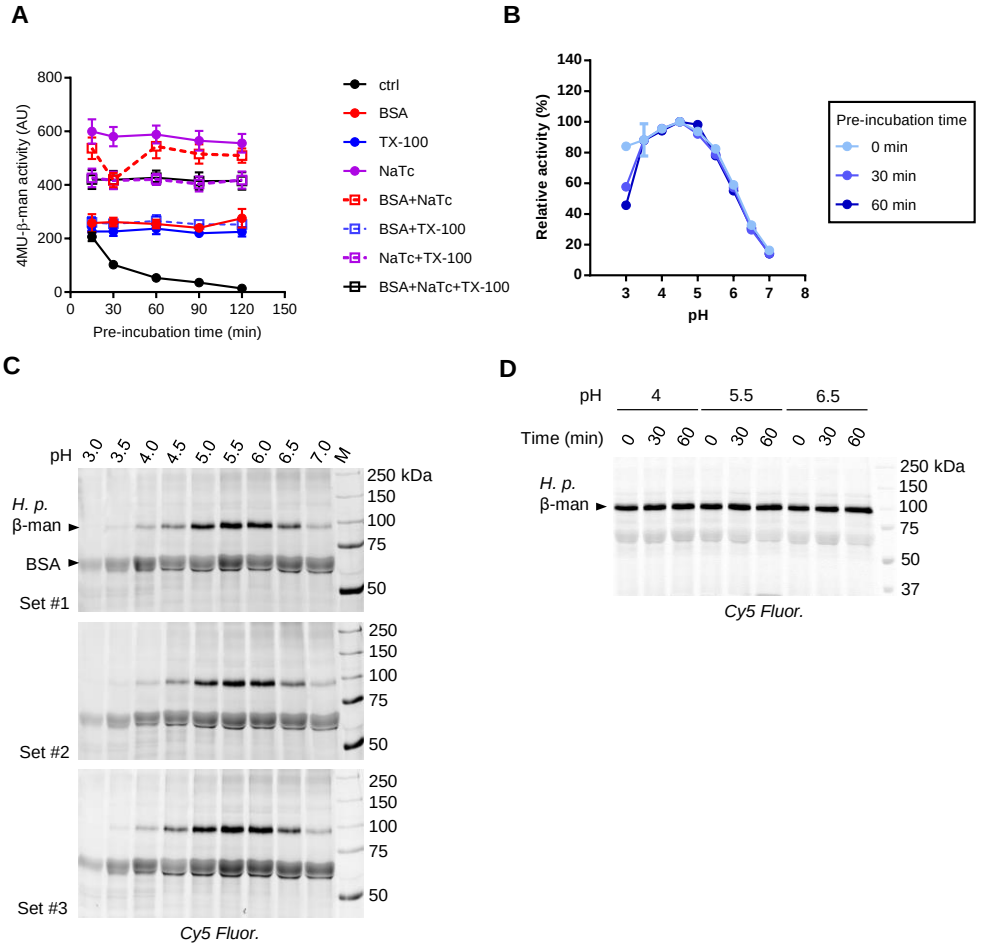


Figure 6.S2. Characterization of ABP labeling conditions for *Helix pomatia* GH2 β-mannosidase. A) Measured enzyme activity (30 min with 4-MU β-man substrate assay at pH 5.5) with or without supplements, over the indicated pre-incubation periods at pH 5.5. B) Effect of pH on enzymatic activity at various pH and over different pre-incubation periods, in the presence of 0.1 % (w/v) BSA. C) ABP **12** labeling at various pH value, in the presence of 0.1 % (w/v) BSA. D) Effect of pre-incubating ABP **12** at pH 4.0, 5.5, and 6.5 for 0, 30, or 60 min at 37 ºC on its labeling towards the enzyme (at 10 µM [ABP], pH 5.5, 1 h 37 ºC). Error range = ± SD (*n* = 3, technical replicates).

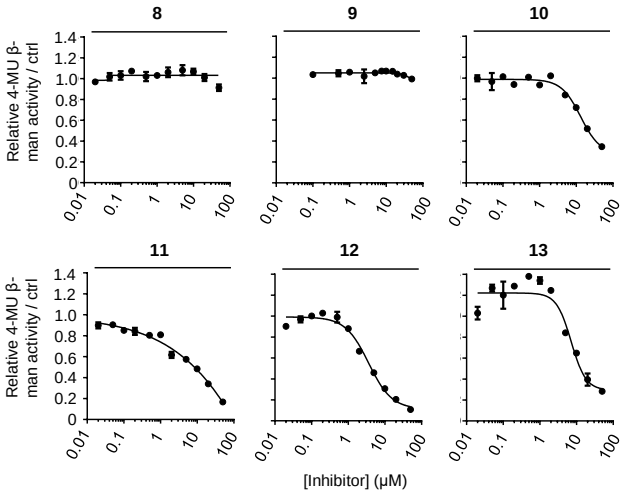


Figure 6.S3. Inhibition curves of compound 8−13 for apparent IC_{50} determination in snail (*Helix pomatia*) GH2 β-mannosidase. Error range = ± SD, *n* = 3 technical replicates.

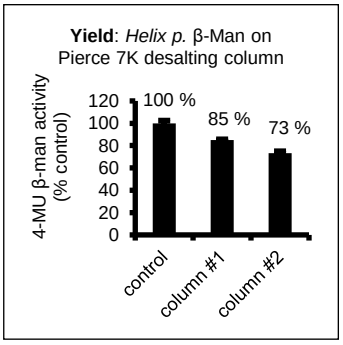
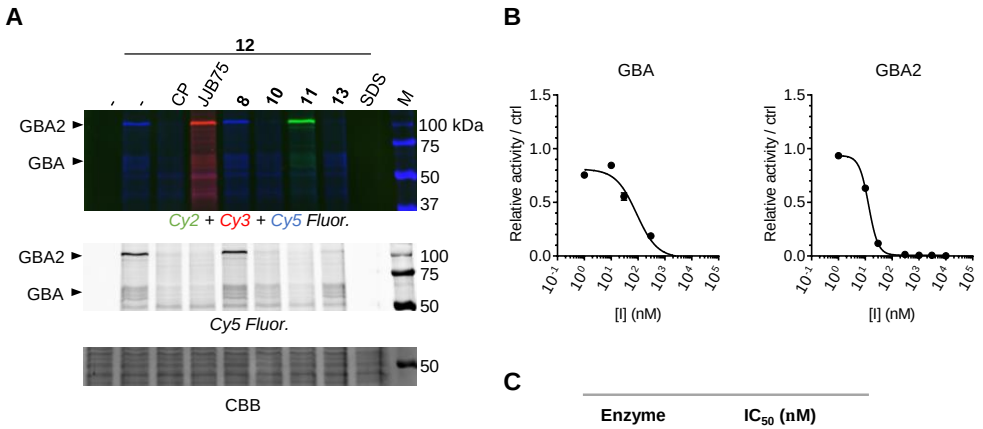


Figure 6.S4. Activity of snail (*Helix pomatia*) GH2 β-mannosidase after desalting. Error range = ± SD, *n* = 3 technical replicates.



Enzyme	IC_{50} (nM)
GBA	87.8 ± 19.4
GBA2	13.4 ± 0.232

Figure 6.S5. Reactivity of ABP 12 towards retaining β-glucosidases. A) cABPP in lysates of GBA2-overexpressed HEK293T cells, using pre-incubation of cyclophellitol (CP), ABP JJB75, compounds **8**, **10**, **11**, **13**, and 2 % (w/v) SDS. *CBB*, Coomassie Brilliant Blue staining. B) Inhibition curves on GBA, GBA2, and GBA3. C) Apparent IC_{50} values. Error range = ± SD (*n* = 3 technical replicates).

Table 6.S1. Primers used for coning human GH38 α-mannosidases

Protein	Primer	Sequence
MAN2A1	Forward	GGGGACAAGTTTGTACAAAAAAGCAGGCTACCACCATGAAGTTAAGCCGCCAGTTCAC
	Reverse	GGGGACCACTTTGTACAAGAAAGCTGGGTCTCACCTCAACTGGATTCGGAATG
MAN2A2	Forward	GGGGACAAGTTTGTACAAAAAAGCAGGCTACCACCATGAAGCTGAAAAAGCAGGTGAC
	Reverse	GGGGACCACTTTGTACAAGAAAGCTGGGTCCTAACCCAAGCGGAGGCGAAAGG
MAN2B1	Forward	GGGGACAAGTTTGTACAAAAAAGCAGGCTACCACCATGGGCGCCTACGCGCGGGCTTC
	Reverse	GGGGACCACTTTGTACAAGAAAGCTGGGTCCTAACCATCCACCTCCTTCCATTG
MAN2B2	Forward	GGGGACAAGTTTGTACAAAAAAGCAGGCTACCACCATGGGGCAGCTGTGCTGGCTGC
	Reverse	GGGGACCACTTTGTACAAGAAAGCTGGGTCTCACTGCTGTTGAAAGTGAATAA
MAN2C1	Forward	GGGGACAAGTTTGTACAAAAAAGCAGGCTACCACCATGGCGGCTGCGCCGGCCTTG
	Reverse	GGGGACCACTTTGTACAAGAAAGCTGGGTCTCAGTGTGGCGGAGGCTGAAG

6.S2. Supporting experimental procedures for protein X-ray crystallography (University of York)

6.S2.1 Cloning and expression of dGMII

A plasmid containing cDNA encoding for the dGMII gene was obtained with kind permission from Dr. Sean Sweeny (University of York). From this plasmid, cDNA encoding for amino acids 76-1108 of dGMII (to remove the N-terminal cytosolic, transmembrane, and stalk domains) was subcloned into the pOMNIBac vector (Geneva Biotech), behind a honeybee mellitin secretion peptide, 6xHis tag, and TEV cleavage site. Recombinant bacmid was produced using the Tn7 transposition method in DH10EMBacY (Geneva Biotech)[1], and purified using the PureLink miniprep kit (Invitrogen) following standard protocols. V1 baculovirus was produced by transfection of bacmid into low passage adherent Sf21 cells (Invitrogen) using FuGENE HD transfection reagent (ProMega), at a ratio of 2 µg DNA to 4.5 µL FuGENE. V1→V2 virus amplification was carried out using suspension Sf21 cells, using the YFP marker present in EMBacY baculovirus to determine optimum amplification prior to harvesting (typically ~60 % cells fluorescent). For expression, *T. Ni* cells (Invitrogen) were infected with V2 baculovirus at a multiplicity of infection (MOI) > 1, and infection followed using the EMBacY YFP marker to determine optimum timepoint for harvesting (typically 72 h, with > 80 % cells fluorescent). All insect cells used tested negative for mycoplasma contamination.

6.S2.2 Purification of dGMII

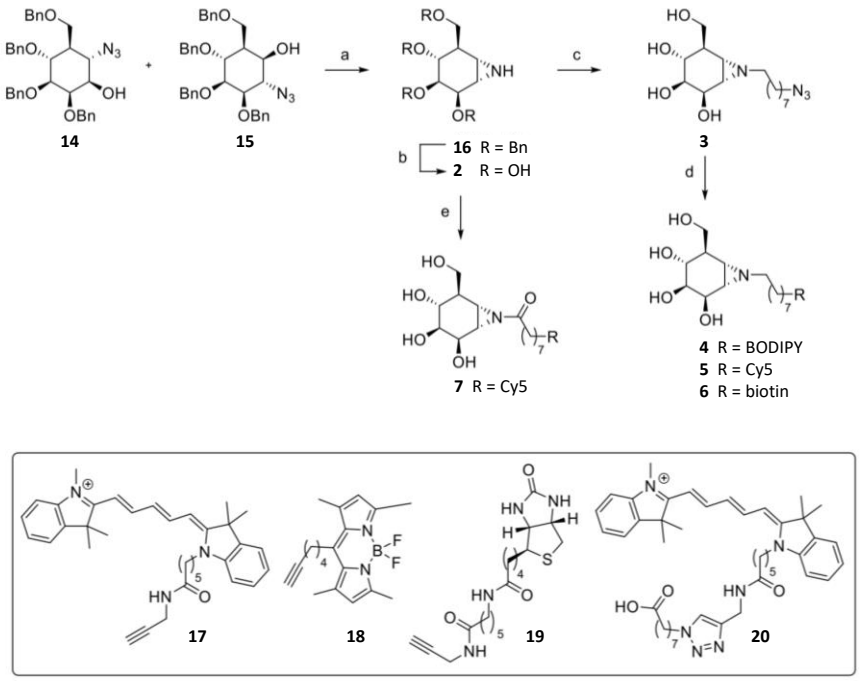
Harvested High Five cultures were spun for 15 minutes at 200 g to remove cells and spun again at 4000 g for 1 hour to remove insoluble cellular components. AEBSF (to a final concentration of 0.1 mM) and DDT (to a final concentration of 1 mM) were added to the supernatant. Clarified supernatant was loaded on to 2 x 5mL HiTrap Blue HP columns (GE Healthcare) pre-equilibrated in blue agarose buffer A (20 mM HEPES pH 7.4, 100 mM NaCl, 1 mM DTT). Loaded columns were washed with 5 column volumes of blue agarose buffer A and eluted using a linear gradient of blue agarose buffer B (20 mM HEPES pH 7.4, 2 M NaCl, 1 mM DTT) over 20 column volumes. HiTrap Blue fractions containing dGMII (as determined by SDS-PAGE) were pooled and diluted approximately 5-fold in HisTrap buffer A (50 mM HEPES pH 8.0, 0.5 M NaCl, 30 mM Imidazole, 1 mM DTT) and loaded on to a 1 mL HisTrap FF Crude column (GE Healthcare) The loaded HisTrap column was washed with 10 column volumes of HisTrap

buffer A and dGMII eluted using a linear gradient of HisTrap buffer B (50 mM HEPES pH 8.0, 0.5 M NaCl, 0.5 M imidazole, 1 mM DTT) over 20 column volumes, followed by 10 column volumes of 100 % Buffer B. HisTrap fractions containing dGMII were pooled and concentrated to less than 2 ml by centrifugation using VivaSpin 30,000 MW concentrator (GE Healthcare). Pooled dGMII was rediluted to ~2 mL in 1x AcTEV (Invitrogen) reaction buffer before addition of 5 µL AcTEV protease to remove the N-HisTEV tag. Digests were typically carried out overnight at ambient temperature, and reaction progress assessed by comparison to positive and negative controls using SDS-PAGE. Upon completion of the AcTEV digest, as indicated by SDS-PAGE, dGMII was purified using a S200 16/600 size exclusion chromatography (SEC) column equilibrated in SEC buffer (50 mM HEPES pH 7.4, 200 mM NaCl, 1 mM DTT), at 1 ml min-1. SEC fractions containing dGMII were pooled, concentrated to 10 mg ml-1 and stored in 20 µl aliquots at −80 °C.

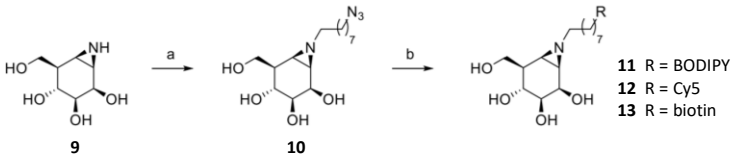
6.S2.3 3D Crystallography of dGMII

Initial crystallization conditions were screened using JCSG+ HT-96, PACT premier HT-96 (both Molecular Dimensions), Index HT and PEG/Ion HT (both Hampton Research) commercial screens. Hits were optimized, scaled up to maxi 48-well plates and a dGMII seed stock generated from crystals grown in 0.1 M imidazole pH 7.0 and 10 % (w/v) PEG 3350 using the Seed Bead protocol (Hampton Research). The above screens were repeated with seeding using an Oryx8 (Douglas Instruments) and additional hits optimized further. Diffraction quality dGMII crystals were grown in maxi 48-well plates using sitting drop vapor diffusion in 0.1 M sodium succinate pH 7.0 and 10 % (w/v) PEG 3350. Crystals were cryoprotected using cryoprotectant solution (mother liquor supplemented with 25% v/v ethylene glycol) prior to flash freezing in liquid N_2 for data collection. Dataset collection and processing followed the previous methods.[2] Apo dGMII was solved by molecular replacement with PDB model 1HWW[main text ref 30] using MolRep[3], followed by alternating rounds of manual model building and refinement using Coot and REFMAC5 respectively[4, 5]. For ligand complexes, dGMII crystals were soaked in solutions of **2** (1 mM) in dGMII cryoprotectant solution for ~3 hours before flash freezing in liquid N_2 for data collection. Complexes were solved by molecular replacement with the apo dGMII structure, followed by rounds of manual model building and refinement using Coot and REFMAC5. Generation of crystal structure figures and ligand coordinates followed the described methods.[2]

6.S3. Synthetic strategies for compounds used in this chapter (Department of Bio-organic Synthesis, Leiden University)



Scheme 6.S1. Synthetic strategy of compound 2−7. Reagents and conditions: a) PPh_3 polymer-bound on styrene-divinylbenzene copolymer, CH_3CN, reflux, overnight, 54 %; b) Li, THF, NH3 (l), −60 °C, 75 min; c) 1-azido-8-iodooctane (see Chapter 6), K_2CO_3, DMF, 80 °C, overnight, 38 % over 2 steps; d) **17**[6], **18**[7] or **19**[8], $Cu.SO_4 \cdot 5H_2O$, sodium ascorbate, DMF, RT, **4** 35 %, **5** 8.5 %, **6** 16 %; e) **20**[6], EEDQ, DMF, 0 °C, 2.5 h, 8 % over 2 steps.



Scheme 6.S2. Synthetic strategy of compound 9−13. Reagents and conditions: a) 1-azido-8-iodooctane (see Chapter 4), K_2CO_3, DMF, 80 °C, overnight, 64 %; b) **17**[6], **18**[7] or **19**[8], $Cu.SO_4 \cdot 5H_2O$, sodium ascorbate, **11** 17 %, 12 36 % and **13** 32 %.

Supporting References

1 Bieniossek C, Imasaki T, Takagi Y & Berger I (2012) MultiBac: expanding the research toolbox for multiprotein complexes. Trends Biochem Sci **37**, 49−57.

2 Wu L, Jiang J, Jin Y, Kallemeijn WW, Kuo CL, Artola M, Dai W, van Elk C, van Eijk M, van der Marel GA, Codée JDC, Florea BI, Aerts JMFG, Overkleeft HS & Davies GJ (2017) Activity-based probes for functional interrogation of retaining β-glucuronidases. *Nat Chem Biol* **13**, 867−873.

3 Vagin A & Teplyakov A (2010) Molecular replacement with MOLREP. *Acta Crystallogr D Biol Crystallogr* **66**(Pt 1), 22−25.

4 Emsley P, Lohkamp B, Scott WG & Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**(Pt 4), 486−501.

5 Murshudov GN, Skubák P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F & Vagin AA (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* **67**(Pt 4), 355−367.

6 Beenakker TJM (2018) Chapter 2. In *Design and development of conformational inhibitors and activity-based probes for retaining glycosidases* (Doctoral dissertation) (pp 13−27). Retrieved from Leiden University Repository.

7 Verdoes M, Hillaert U, Florea BI, Sae-Heng M, Risseeuw MD, Filippov DV, van der Marel GA & Overkleeft HS (2007) Acetylene functionalized BODIPY dyes and their application in the synthesis of activity based proteasome probes. *Bioorg Med Chem Lett* **17**, 6169−6171.

8 Inoue M, Tong W, Esko JD & Tor Y (2013) Aggregation-mediated macromolecular uptake by a molecular transporter. *ACS Chem Biol* **8**, 1383−1388.

CHAPTER 7

β-Galactose Configured Cyclophellitol Aziridine as Activity-Based Probes for Retaining exo-β-Galactosidases

Based on:

Kuo CL, Beenakker TJM, Profijt R, Marques ARA, Groenewegen N, Offen WA, Sarris AJC, Boot RG, Codée JDC, van der Marel GA, Davies GJ, Overkleeft HS & Aerts JMFG. *To be submitted.*

ABSTRACT

The enzymes β-galactosidase (GLB1) and galactocerebrosidase (GALC) are both retaining exo-glycosidases important in glycoconjugate metabolism, and their hereditary deficiency in man can lead to lysosomal storage disorders with still unmet medical needs. Chemical tools to study these enzymes have been developed in the past, including the cyclophellitol-based activity-based probes (ABPs) for GALC, and broad spectrum (4-deoxy) cyclophellitol aziridine ABPs for both β-glucosidases and β-galactosidases. However, β-galactose-configured N-tagged cyclophellitol aziridines—which might offer specific ABPs for β-galactosidases—have not been generated due to earlier synthetic challenges. Here, ABPs containing such scaffold are evaluated for their reactivities. The analysis shows that the ABPs exhibit expected mechanism-based inhibition and labeling of recombinant bacterial GLB1 homologue, and both GLB1 and GALC in cell lysates, culture medium, and tissue extracts. The Cy5-tagged ABP additionally labels in mouse intestine extracts the dietary enzyme lactase-phlorizin hydrolase (LPH). Pull-down experiments with biotin conjugated ABP and mouse kidney extracts, followed by LC-MS analysis identified GLB1 and GALC as major glycosidase targets, as well as two other putative β-galactosidases GLB1-lile protein 1 and 2 (GLB1L and GLB1L2) that have yet unknown biological functions. The β-galactose-configured N-tagged cyclophellitol aziridine ABPs still label β-glucosidases with considerable affinity, illustrating the limitation of this scaffold for specific β-galactosidase labeling. Nevertheless, by pre-incubating samples with β-glucosidase inhibitors, different β-galactosidases can be simultaneously visualized by gel-based ABP detection. The novel ABPs probes can be employed in future fundamental and clinical research on reactive β-galactosidases

7.1 Introduction

The monosaccharide galactose is incorporated by higher eukaryotes into a variety of glycoconjugates such as glycoproteins, glycolipids, and glycosaminoglycans (GAGs). Among the two possible galactose anomers in glycoconjugates, the predominant one is β-galactose. In humans and other mammals, β-galactose is transferred from UDP-α-galactose to specific acceptors in the Golgi apparatus by the action of dedicated β-1,3- and β-1,4-galactosyltransferases.[1] It is removed from glycoconjugates predominantly by two distinct lysosomal β-galactosidase: acid β-galactosidase (GLB1, E.C. 3.2.1.24) degrading oligosaccharides (glycosaminoglycans and glycoproteins) and gangliosides (such as GM1a and GA1),[2] and galactocerebrosidase (GALC, E.C. 3.2.1.46) degrading mainly galactosylceramide.[3]

GLB1 belongs to Glycoside Hydrolase (GH) family 35.[4] Synthesized as glycosylated 85 kDa precursor, it is targeted to the lysosome through the mannose-6-phosphate (M6P) dependent pathway; proteolytic cleavage at its C terminus results in the mature enzyme containing a large 64 kDa subunit and a small 20 kDa subunit that remain associated with each other without disulfide linkage.[5, 6] The 64 kDa subunit contains a TIM barrel domain containing the catalytic glutamates Glu188 and Glu268, while the 20 kDa subunit is required for the stabilization and functioning of the enzyme.[6, 7, 8] Similar to other lysosomal glycosidases, GLB1's catalysis of glycolipids is assisted by activator proteins, in this case both saposin B and the GM2 activator protein.[9] GLB1 forms in lysosomes a complex together with cathepsin A (PPCA, *CSTA* gene) and neuraminidase (NEU1), resulting in a stable and efficient catalytic machinery for degrading the glycosphingolipid GM1 ganglioside.[10] Alternative splicing of the *GLB1* gene produces a protein named elastin-binding protein (EBP), which does not possess catalytic activity and is instead transported to the extracellular matrix to act as a lectin that modulates elastin fiber formation.[11, 12] Another lysosomal β-galactosidase, GALC, belongs to the GH59 family.[4] It is synthesized as an 80 kDa precursor containing four N-linked glycans, and targeted either directly to the lysosomes via the M6P-mediated pathways or by re-uptake of the secreted proteins through M6P-dependent or -independent pathways.[13, 14] In the lysosome, it is processed to mature enzyme consisting of an N-terminal 50 kDa subunit (containing the TIM barrel domain harboring catalytic residues glu198 and Glu274) and a C-terminal 30 kDa subunit comprising the lectin domain bound with a calcium ion.[13, 15, 16] With the assistance by saposin A,

[17, 18] GALC specifically degrades galactosylceramide that is particularly abundant in myelin, kidney, and epithelial cells of intestine and colon.[19]

The importance of GLB1 and GALC activity is illustrated by diseases caused by their deficiency. Deficiency in GLB1 may result in three distinct lysosomal storage disorders, namely GM1 gangliosidosis, Morquio B syndrome, and galactosialidosis. GM1 gangliosidosis is caused by mutations in the *GLB1* gene and is characterized by an elevated cellular level of GM1 ganglioside, causing lysosomal vacuolization in lymphocytes and neuronal cells, which in turn leads to demyelination and neuronal cell death in the periphery as well as in the central nervous system.[20] Patients are classified into infantile (type I, OMIM # 230500), late infantile/juvenile (type two, OMIM # 230600), and adult onset forms (type III, OMIM # 230650) according to disease feature and progression.[20] Disease severity is inversely related to residual GLB1 activity.[21] Although the molecular mechanisms leading to disease pathogenesis is not completely understood, it has been suggested that abnormal accumulation of GM1 ganglioside in the membranes linking the ER and mitochondria can drive Ca^{2+} efflux from the ER to mitochondria, leading to unfolded protein response (UPR) and mitochondrial stress, followed by apoptosis and ultimately neuronal death.[22-24] Different mutations in *GLB1* gene may lead to another disease entity named Morquio B syndrome (mucopolysaccharidosis type IVB, MPSIVB, OMIM # 253010), where the primary accumulated substrates are oligosaccharides deriving from keratin sulfate and glycoproteins.[2] No central nervous system involvement is observed in these patients, and the major affected tissues is the skeletal system.[2] The third disease related to GLB1 deficiency is galactosialidosis (OMIM # 256540), and is in fact caused by primary deficiency of cathepsin A (PPCA, gene = *CTSA*) that normally forms a protein complex with GLB1 and neuraminidase (NEU1) in the lysosome. Cathepsin A deficiency causes premature degradation of both GLB1 and NEU1, leading to their secondary deficiencies in patients.[10] Deficiency in GALC or its activator protein saposin A, forms the basis of Krabbe disease (globoid cell leukodystrophy, GLD, OMIM # 245200).[3, 25] About 85 % of Krabbe disease patients develop the infantile-onset form that presents with developmental delay and severe neurological damages, with death usually within a few years.[26] Later-onset forms have higher residual GALC activity and milder symptoms, and may have life spans up to the seventh decade of life. The primarily accumulated substrate is galactosylsphingosine (psychosine) in macrophages (globoid cells) and neural cells—particularly oligodendrocytes and Schwann cells, and this cytotoxic compound is believed to cause demyelination and cell death in both the central and peripheral nervous

system.[27, 28] No therapy is yet available for these β-galactose-related lysosomal storage disorders besides bone marrow transplantation (BMT) and hematopoietic stem cell transplantation (HCT) of Krabbe disease, which is only effective before the onset of symptoms (which is usually not the case). Several therapeutic approaches are presently studied including substrate reduction therapies (SRT),[29] enzyme replacement therapies (ERT),[30, 31] pharmacological chaperon therapies (PCT),[32, 33] gene therapies,[34-36] and lysosomal re-acidification therapy[37].

Mechanistically, both GLB1 and GALC are retaining exo-glycosidases using the Koshland double-displacement catalytic mechanism. A covalent glycosidic bond is formed during the reaction itinerary, which makes both enzymes amenable to activity-based glycosidase profiling (general introduction, this thesis). In the past, activity-based probes (ABPs) based on the β-galactose configured cyclophellitol scaffold have been generated (**Fig. 7.1A**, LWA487) and enabled specific labeling of GALC in mouse tissues extracts at high micromolar ABP concentration.[38]

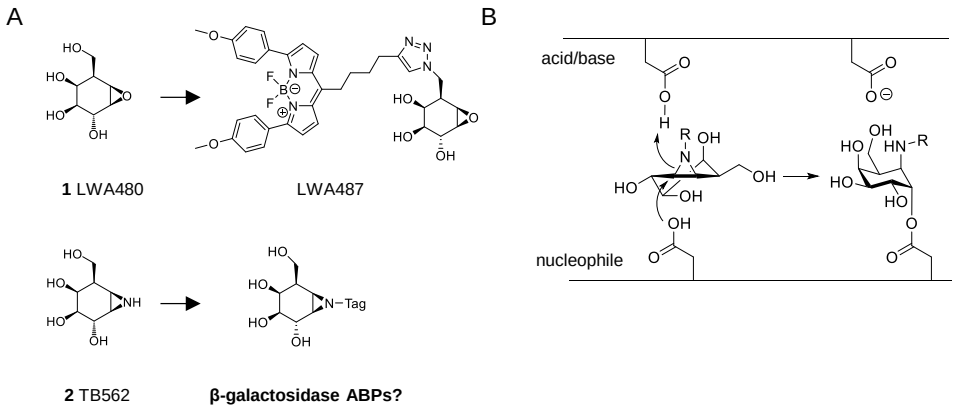


Figure 7.1 Strategies for activity-based labeling towards exo-galactosidases. A) Inhibitor and ABP for GALC (LWA480 and LWA487) and potential inhibitor and ABPs for β-galactosidases B) Proposed reaction mechanism of the β-galactose configured cyclophellitol aziridines towards β-galactosidases.–

On the other hand, ABPs based on the 4-deoxy cyclophellitol aziridine scaffold offers broad spectrum ABPs (**Fig. 7.2**, SYD215) that label both β-glucosidases and β-galactosidases.[39] To develop ABPs specifically towards β-galactosidases, attempts have been made to synthesize the N-alkylated or N-acylated β-galactose configured cyclophellitol aziridines (**Fig. 7.1A**). This was initially hampered by instability of compound intermediates during synthesis.[40] This chapter aims to study the recently available N-alkylated β-galactose configured cyclophellitol

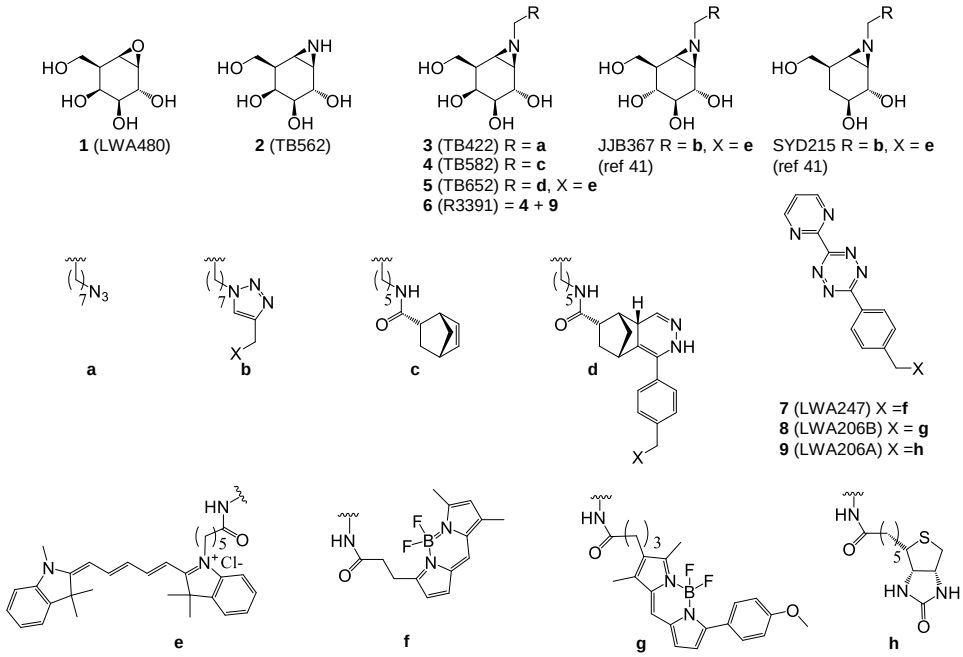


Figure 7.2. Structures of compounds used in this chapter.

aziridine compounds. These ABPs were ultimately generated through an alternative click chemistry involving the reaction between the norbornene-bearing cyclophellitol aziridine and the tetrazine-bearing reporter groups. This chapter reports on a detailed investigation of these compounds regarding inhibitory potency, glycosidase labeling, and target specificity.

7.2 Results

7.2.1 Synthetic strategies for β-galactose configured cyclophellitol aziridines

Compounds were synthesized at the Department of Bio-organic Synthesis, Leiden University. The synthesis of β-galactose configured cyclophellitol **1** and cyclophellitol aziridine **2** has been previously described.[40, 41] Acylation of **2** was unsuccessful due to the rapid decomposition of the products, but alkylation with 8-iodooctylazide was successful, leading to the N-alkylated compound **3**. However, normal Cu(I) catalyzed click reaction with an alkyne-fluorophore did not yield the expected products.[41] Therefore, an alternative strategy was tested, based on the inverse electron demand Diels-Alder (IEEDA) reaction between a norbornene and

a tetrazine.[42] Following this alternative approach, the cyclophellitol aziridine N-alkyl norbornene **4** was firstly synthesized by alkylation of **2** with a norbornene handle, and subsequent click reaction with tetrazine conjugates afforded the desired Cy5 ABP **5**[41] and biotin ABP **6**. In addition, due to the initial unavailability of **5** and **6**, the BODIPY-FL, BODIPY-TMR, and the biotin probe were generated *in situ* by incubation of **4** with compound **7**, **8**, or **9** in DMSO at RT for some period (30 min to overnight), and the mixture was used for several labeling experiments.

7.2.2 *In vitro* inhibition and labeling of compounds **2**-**7** towards β-galactosidases

Compounds were tested for inhibitory potency towards GLB1 (present in human fibroblasts lysates) and mouse GALC (collected culture medium of HEK293T cells overexpressing this protein).[38] The presence of GLB1 and GALC in these materials was first tested with 4-methylumbelliferyl β-D-galactopyranoside (4-MU-β-gal) substrate assay at various pH, and by using 11 μM $AgNO_3$ to selectively inhibit GLB1 over GALC.[43] The β-galactosidase activity in both materials had a pH optimum of 4.5, but at pH 6.0 the culture medium retained 60 % activity, in contrast to the 20 % retained by β-galactosidase in the fibroblast lysates (**Fig. 7.S1A**). In the presence of $AgNO_3$, no activity remained in the fibroblast lysates, while the culture medium retained full β-galactosidase activity (**Fig. 7.S1B**), indicating that the fibroblast lysates contained only GLB1 and that the culture medium contained only GALC. Compounds were next assessed for their inhibitory potency on both enzymes using 4-MU β-gal assay. The GLB1 and GALC enzyme preparations were incubated with substrate and compounds. Compounds **2** and **3** exhibited low nanomolar IC_{50} values towards both GLB1 and GALC; the epoxide **1**[44] and bulkier compound **4**, **5**, and **6** were about ten-fold less potent, but were still nanomolar inhibitors (**Table 7.1, Fig. 7.S2**). For comparison, the 4-deoxy Cy5 ABP SYD215 was a high nanomolar inhibitor for both enzymes, and was eight- to ten-fold less potent than the Cy5 ABP **5**; the β-glucose configured Cy5 ABP JJB367 was a further 50- to 100-fold less reactive, showing apparent IC_{50} values in the micromolar range (**Table 7.1, Fig. 7.S2**).

The compounds also inhibited the bacterial GH35 β-galactosidase from *Cellvibrio japonicus* (CjGH35A, an homologue to human GLB1[45]) with high potency (**Table 7.S1, Fig. 7.S2C**)), and kinetic studies of the compounds towards this enzyme revealed that the N-alkyl azide compound **2** has the highest affinity (lowest K_I) towards the enzyme and the highest inactivation rate constant (k_{inact} / K_I) of 21.71 min^{-1} μM^{-1} (**Table 7.2**, **Fig. 7.S3**, **S4**), with the later value comparable to that of cyclophellitol ABPs towards glucocerebrosidase (GBA).[46] It was also

Table 7.1. Apparent IC_{50} values (nM) of compounds towards retaining exo β-galactosidases. Error range = ± SD, *n* = 2 biological replicates.

	GLB1	GALC
JJB367	7,270 ± 658	53,500 ± 2,770
SYD215	138 ± 24.5	472 ± 37.6
1 (LWA480)	21.7 (ref 44)	39.1 (ref 44)
2 (TB562)	2.55 ± 0.59	5.57 ± 0.36
3 (TB422)	2.55 ± 0.47	12.0 ± 2.27
4 (TB582)	57.8 ± 3.05	98.6 ± 20.8
5 (TB652)	14.6 ± 0.98	61.0 ± 6.89
6 (R3391)	4.79 ± 4.09	9.95 ± 5.73

Table 7.2. Kinetic data for compounds 2-5 towards CjGH35A. Error range = ± SD, *n* = 3 biological replicates.

	K_{inact} (min^{-1})	K_I (μM)	k_{inact}/K_I (min^{-1} μM^{-1})
2 (TB562)	0.99 ± 0.11	1.90 ± 0.39	0.52 ± 0.06
3 (TB422)	0.56 ± 0.11	0.03 ± 0.01	21.71 ± 4.08
4 (TB582)	2.19 ± 0.55	3.37 ± 1.20	0.65 ± 0.16
5 (TB652)	2.03 ± 0.63	0.11 ± 0.04	18.98 ± 5.88

observed that while the maximum potential reaction rate (k_{inact}) of compound **2**-**5** was similar, their affinity (K_I) towards CjGH35A was quite different—with the bare aziridine **2** and the N-alkyl norbornene **4** less favorable than the N-alkyl azide **3** and the Cy5 ABP **5** (**Table 7.2**).

A crystal structure of CjGH35A soaked with compound **2** was resolved and revealed that the compound bound to the enzyme at its catalytic nucleophile and adopted a $^4C_1$ conformation (**Fig. 7.3**).

Next, ABP labeling was attempted in human fibroblast lysates and the culture medium of GALC-overexpressing HEK293T cells. Due to the initial unavailability of the Cy5 ABP **5**, the experiment was performed by firstly pre-incubating the norbornene compound **4** with the tetrazine-BODIPY compounds **8**, and incubating the mixture with the culture medium of

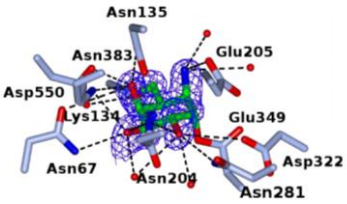


Figure 7.3. Crystal structure of CjGH35A in complex with 2. The map shown is a Fo-Fc map, with phases calculated prior to the inclusion of ligand in the refinement, contoured at 3σ. Carbon atoms are colored green for the ligand and ice blue for the side chains. The interacting residues are annotated, including catalytic residues Glu349 (nucleophile) and Glu205 (acid/base).

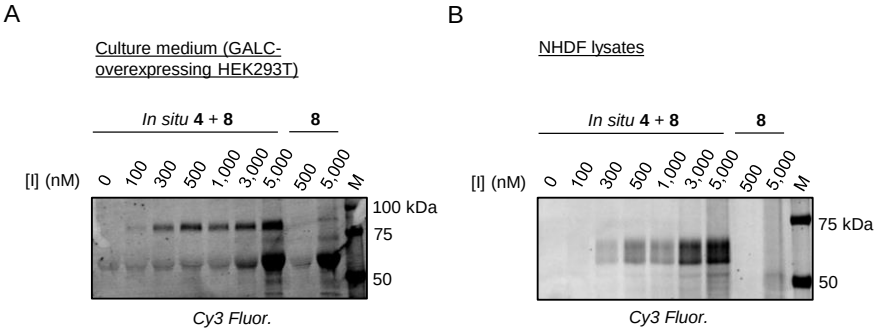


Figure 7.4. *In vitro* ABP labeling using *in situ* generated probe (4 + 8) at different probe concentration. A) Labeling in culture medium of HEK293T cells overexpressing mouse GALC. B) Labeling in human fibroblast (NHDF) lysates.

GALC-overexpressing cells. A band of about 80 kDa in size was observed, (**Fig. 7.4A**), which could correspond to the secreted form of mouse GALC. In human fibroblast lysates, a smear of bands between 50 and 75 kDa was detected (**Fig. 7.4B**), which indicated the labeling of glycoprotein(s).

Because this labeling pattern resembled that of cyclophellitol ABPs towards GBA (Chapter 1 and 2, this thesis), the reactivity of compound **2**-**6** was also examined towards the recombinant human GBA (Imiglucerase) by enzymatic assay. The results showed that all compounds inhibited GBA with nanomolar potency, except for the aziridine compound **2** and the norbornene compound **4**, both being low micromolar inhibitors of GBA **(Table. 7.S1**, **Fig. 7.S5)**. The compounds were also tested for their reactivity towards another retaining β-glucosidase, GBA2. It turned out that they were about two- to ten-fold less potent towards GBA2 compared to GBA **(Table. 7.S1**, **Fig. 7.S5)**. The labeling by the *in situ* generated

BODIPY-TMR probe (**4** + **8**) towards GBA was confirmed in NHDF lysates using a competitive ABPP (cABPP) setup, where its labeling was partially abrogated by pre-incubating the lysates with the GBA-specific ABP ME569 (chapter 3, this thesis) (**Fig. 7.S6**). To further verify if GLB1 and GALC were labeled by the probe, an additional cABPP was performed in mouse tissue extracts (expected to express both GLB1 and GALC) pre-incubated with cyclophellitol (known β-glucosidase inhibitor), compound **3** (inhibitor for both β-glucosidase and β-galactosidase) or **4** (β-galactosidase inhibitor at lower concentrations), subsequently labeled with the GBA-specific ABP ME569, and finally with the *in situ* generated probe. The probe labeled a ~50 kDa band and a ~60 kDa band that overlapped with GBA labeling by ME569 (lane 2 of both gels, **Fig. 7.5A**). Pre-incubation with **3** or **4** (the later at 500 nM), but not cyclophellitol, abolished the labeling by the *in situ* generated BODIPY-FL probe (**Fig. 7.5A**). Contrastingly, cyclophellitol and **3**, but not **4**, abolished the labeling by ME569 (**Fig. 7.5A**). In

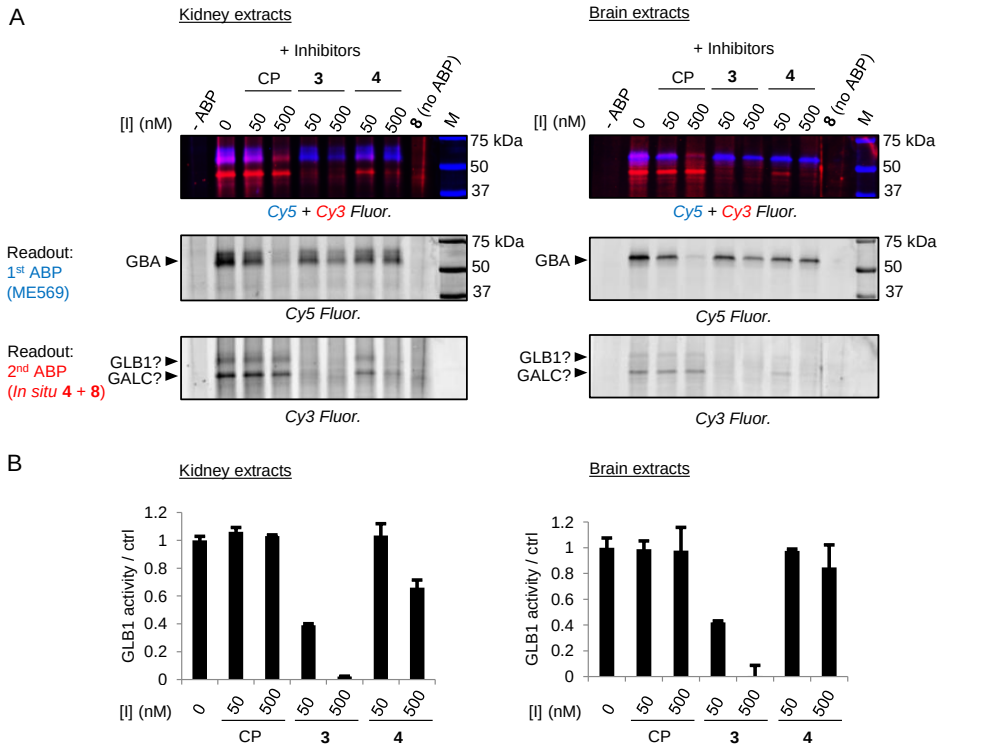


Figure 7.5. Labeling on GLB1 and GBA by the *in situ* generated probe (4 + 8) in extracts of mouse kidney (left) or brain (right). A) Competitive ABPP pre-incubated firstly with cyclophellitol (CP), **3**, or **4**, and secondly with ABP ME569. B) GLB1 activity in the identically treated samples measured by fluorogenic substrate assay. Error range = ± SD, *n* = 3 technical replicates.

parallel, enzymatic assay performed in the identically-treated samples showed that GLB1 activity was abolished by **3** and (at 500 nM) **4** (**Fig. 7.5B**), but not by cyclophellitol. Together, these results strongly suggested that the β-galactose-configured N-tagged cyclophellitol aziridine ABP allows gel-based fluorescence detection for both GLB1 (known size = 64 kDa) and GALC (known size = 50 kDa).

Next, the β-galactose-configured Cy5 ABP **5** was employed for evaluating with mouse kidney extracts (known to contain both GALC and GLB1) its labeling properties. The labeling was performed in samples pre-incubated with the broad-spectrum β-glucosidase BODIPY ABP JJB70, in order to specifically visualize the β-galactosidases (over GBA and GBA2). Labeling occurred in an irreversible and mechanism-based manner: optimal labeling for both GLB1 and GALC occurred at 1 μM ABP 5 (**Fig. 7.6A**), 30 min incubation time (**Fig. 7.6 B**), and at pH from 4.0 to 5.0 (**Fig. 7.6 C**). The labeling was partially abrogated by pre-incubating the samples with 4-MU β-gal, consistent with labeling occurring in the active site of the enzymes (**Fig. 7.6D**).

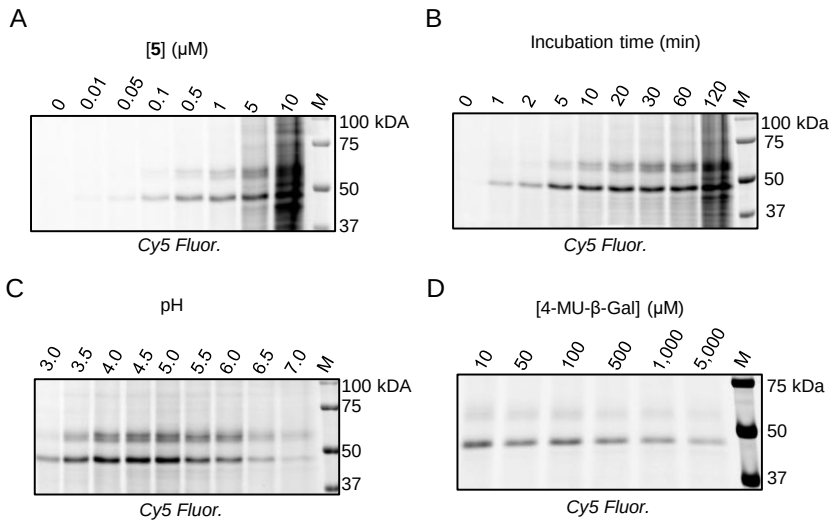


Figure 7.6. Labeling of ABP 5 in mouse kidney extracts. (A) Concentration-dependent labeling. (B) Time-dependent labeling. (C) pH-dependent labeling. (D) Competitive ABPP with 4-MU-β-gal.

7.2.3 Glycosidase target analysis by chemical proteomics and gel-based ABPP in mouse tissue extracts

Glycosidase targets of the biotin ABP **6** in mouse kidney extracts were next examined. For

this purpose, biotin ABP **6** was generated from a mixture of 5 μM **4** and 10 μM **9** reacted O/N. The ABP was with mouse kidney extracts at pH 4.5, and samples were subjected to LC/MS-based protein identification form the streptavidin-enriched and on-bead trypsin-digested peptides. Five glycosidases were identified in samples incubated with the *in situ* generated biotin probe (**4** and **9**), and none in the control (no probe) or the competitive sample (pre-incubated with the Cy5 ABP **5**). The identified glycosidases included the expected targets GALC, GLB1, the β-glucosidase GBA, as well as two other targets: the GLB1-like protein 1 and 2 (GLB1L and GLB1L2, respectively) (**Table 7.3**). The latter two proteins are putative enzymes that are also classified onto GH35 family but with yet unknown biological functions. Both proteins contain catalytic amino acids identical to those of GLB1 (**Fig. 7.S7**). Despite having similar molecular weight to the GLB1 precursor protein (around 70 kDa without N-glycosylation), their post-translational processing, cellular localizations, and tissue-dependent expression are predicted to be different to those of GLB1 (**Table 7.S3)**. To examine whether these features would allow their visualization by the β-galactose configured Cy5 ABP **5**, mouse tissue extracts from brain, epididymis, testis, duodenum, and colon were pre-incubated with the broad-spectrum β-glucosidase ABP JJB70 (to prevent labeling on β-glucosidases by ABP **5**), and next incubated with ABP **5** for gel-based fluorescence detection. The gel showed that each tissue had a distinct labeling profile, and that additional bands were identified at molecular weight or pH range different from those of GLB1 and GALC. In extracts of mouse brain, epididymis,

Table 7.3. List of identified glycosidases in mouse kidney extracts labeled and pulled-down with ABP 6 by LC-MS.

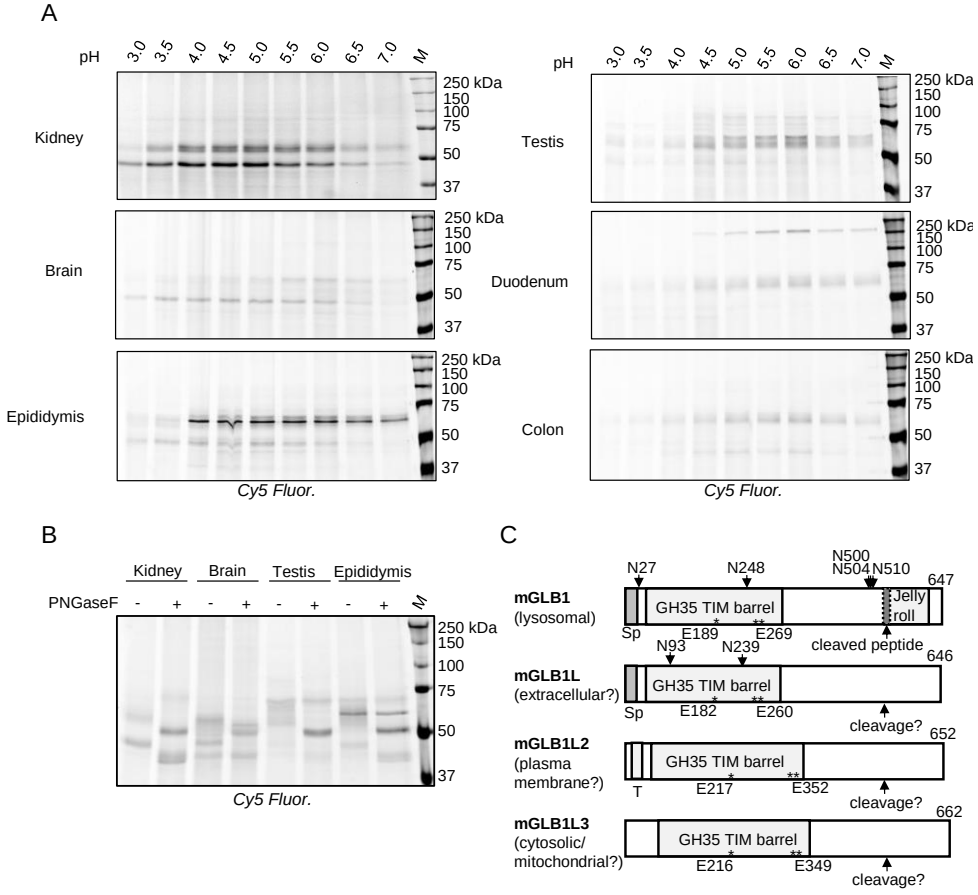
Rank	Accession	Entry	Description	PLGS Score	Coverage (%)
11	P54818	GALC_MOUSE	Galactocerebrosidase OS=Mus musculus GN=GALC PE=1 SV=2	4177	18
13	P23780	BGAL_MOUSE	Beta-galactosidase OS=Mus musculus GN=Glb1 PE=1 SV=1	2851	31
22	Q8VC60	GLB1L_MOUSE	Beta-galactosidase-1-lile protein OS=Mus musculus GN=Glb1l PE =1 SV=1	1783	22
47	P17439	GLCM_MOUSE	Glucosylceramidase OS=Mus musculus GN=Gba PE=1 SV=1	634	27
62	Q3UPY5	GLBL2_MOUSE	Bet-galactosidase-1-lile protein 2 OS=Mus musculus GN=Glb1l2 PE=1 SV=1	367	23

testis, duodenum, and colon, a number of bands at around 60 to 70 kDa displayed pH optimum at 6.0, instead of 4.5 from the one in kidney extracts (**Fig. 7.7A**). In duodenum extracts, an additional band between 150 and 250 kDa was detected optimally at pH 6.0, but it was not presented in extracts of colon. It was reasoned that the later higher band could be the intestinal dietary enzyme lactase phlorizin-hydrolase (LPH), as it contains a lactase pocket that hydrolyzes a β-galactose from the disaccharide lactose.[47] Labeling with ABP **5** in cells expressing either the wild-type human LPH, or LPH mutated at the catalytic nucleophile of phlorizin hydrolase (pocket III) or lactase (IV) revealed that indeed LPH was labeled, and that the labeling largely occurred at the lactase pocket (**Fig. 7.S8**).

The bands detected in mouse kidney, brain, epididymis, and testis were further examined by treating the samples with or without PNGase F, an enzyme that removes the protein N-linked glycans. The treatment in kidney extracts shifted the 64 kDa band (putatively GLB1) to just above 50 kDa, and the 50 kDa band (GALC) to duplet bands at around 40 kDa (**Fig. 7.7B**). In brain extracts, deglycosylation of the 64 kDa band resulted in two bands just above kDa (**Fig. 7.7B**). In testis and epididymis extracts, the ~70 kDa band was not affected by deglycosylation, as well as the prominent ~64 kDa band in the epididymis extracts (**Fig. 7.7B**). Together with the pH experiment, these results suggested that the Cy5 ABP **5** also likely labeled the GLB1-like proteins, as some of the bands between 50 kDa and 75 kDa either has a pH optimum of 6.0 (suggesting non-lysosomal localization), or were not N-glycosylated (GLB1L2 and possibly GLB1L3, also points to non-lysosomal localization). In agreement with these observation, bioinformatics analysis (**Fig. 7.7C and Table 7.S3**) predicts that in contrast to GLB1L (which is likely secreted to the extracellular milieu), the other two GLB1-like proteins are not N-glycosylated. GLB1L2 is predicted to contain no signal peptide, but instead a short transmembrane domain close to its N-terminus. Accordingly, it could be an one-pass plasma membrane protein having its catalytic pocket (C-terminal) facing the extracellular side. GLB1L3 is also predicted to have no signal peptide, and could be a cytosolic/peroxisomal protein; it does have possible N-glycan sites, but if it does not contain an ER-targeting signal, N-glycosylation is not likely to occur. Furthermore, since the GLB1-like proteins are possibly not targeted to the lysosomes, they are unlikely to undergo a similar lysosomal proteolytic cleavage as that of GLB1. This would make their molecular weight larger than the mature, deglycosylated GLB1. The latter corresponds to the band observed just above 50 kDa seen in the PNGase F + lanes in **Fig. 7.7B**. In all the tested tissue homogenates at least one band with higher molecular weight was observed

that was insensitive to PNGase F treatment. These bands are possibly GLB1-like proteins.

Figure 7.7. *In vitro* labeling with ABP 5 in mouse tissue extracts. A) pH-dependent labeling in mouse



tissue extracts. B) Labeling in mouse tissue extracts (pH 6.0) with or without deglycosylation by PNGaseF. C) Comparing domain structure, catalytic sites, and N-glycosylation sites for mouse GLB1 and GLB1-like proteins. *Sp*, signal peptide; *T*, transmembrane domain; *, acid/base; **, nucleophile; downward arrows, N-glycan sites; GH35 TIM barrel, Glycohydrolase family 35 domain (pfam01301); *Jelly roll*, β-galactosidase 4_5 domain (pfam13364).

7.3 Discussion

In this chapter, the newly synthesized β-galactose configured cyclophellitol aziridine ABPs were examined for their mechanism-based labeling of GLB1 and GALC—two retaining exo-β-galactosidases whose lysosomal deficiency underlies the lysosomal storage diseases such as GM1-gangliosidosis, Morquio B syndrome, galactosialidosis, and Krabbe disease.

It is shown that the ABPs successfully label both GLB1 and GALC in mouse kidney extracts in a mechanism-based manner. They exhibit nanomolar inhibitory potency against both enzymes, as well as the bacterial GH35 enzyme CjGH35A. Towards the later enzyme, they also exhibit fast inhibition kinetics, resembling the values of previously reported cyclophellitol-based inhibitors and ABPs towards GBA.[46] When incubated with crystals of CjGH35A, Compound **2** forms a glycosidic bond with the enzyme, as seem in the resolved structure. This further verified its mechanism-based labeling towards β-galactosidases. While all the novel compounds (including the ABPs) also concomitantly inhibited GBA, compound **4** (β-galactose configured cyclophellitol aziridine N-alkyl norbornene) is relatively inactive towards this enzyme (IC_{50} ratio of GBA/GLB1 = 200, highest among the tested compounds). This allows it to be used as a tool to selectively block β-galactosidase activity over β-glucosidases (at a carefully chosen concentration, such as 0.5 μM), a feature exploited in this chapter to verify targets of the newly-generated ABPs. With its concomitant inhibition of β-glucosidases in mind, GLB1 and GALC can be visualized by the Cy5 ABP after pre-incubation of GBA inhibitor, as demonstrate in this chapter. Thus, the Cy5 ABP can be used to rapidly and simultaneously profile GLB1 and GALC activity, as well as their post-translational modification and proteolytic processing. This would make it a very useful tool in the study and diagnosis of lysosomal storage disorders caused by GLB1 or GALC deficiency.

Chemical proteomics with the biotin ABP **6** in mouse kidney extracts revealed that, besides the expected target GLB1, GALC, and GBA, two other putative GH35 β-galactosidases were identified. These are the GLB1-like protein 1 and 2 (GLB1L and GLB1L2), which share homology to GLB1 and contain highly-conserved amino acid sequences near and at the catalytic residues (**Fig. 7.S7**). Gel-based ABPP in a variety of mouse tissue extracts also identifies several proteins labeled by the Cy5 ABP **5** that exhibits properties distinct from the lysosomal protein GLB1 and GALC, such as a more neutral pH optimum (around 6.0), no N-linked glycosylation, and different tissue-dependent expression profile. These observations are corroborated with bioinformatics analysis, which shows that both the mouse and human GLB1L and GLB1L2 are likely not lysosomal proteins, and that GLB1L2 lacks a signal peptide nor possesses any N-glycosylation sites (**Table 7.S3**). It is noteworthy that there exists another GH35 protein, the GLB1-like protein 3 (GLB1L3) in both mouse and human, predicted to be located in the cytosol, mitochondria, or peroxisomes (**Table 7.S3**). This protein was not identified by chemical proteomics in mouse kidney extracts, where it is probably poorly expressed (according to

BioGPS expression database[48], MOE430 dataset[49]). Nevertheless, due to the presence of conserved amino acid sequence at regions homologous to the GLB1's catalytic site (**Table 7.S3**), it is envisioned that future pull-down experiment in tissues highly expressing this protein (such as testis[49]) might also reveal that GLB1L3 also is a *bona fide* β-galactosidase.

Previously, a so-called senescence-associated β-galactosidase (SA-β-gal) is reported in a wide range of senescent cells, from which the β-galactosidase activity can be selectively detected over the non-senescent cells at pH 6.0.[50] The detected pH optimum at 6.0 for GLB1L and GLB1L2 (and possibly also GLB1L3) makes one to speculate the possibility that they might be the SA-β-gal. However, it has been shown with compelling evidence that the identity of SA-β-gal is in fact the lysosomal GLB1, and that upon its upregulation (both mRNA and protein) during senescence its suboptimal activity at pH 6.0 can be readily detected.[51] In fact, such senescence-dependent upregulation is not unique to GLB1, as other lysosomal glycosidases, such as α-L-fucosidase and β-glucuronidase, are also similarly upregulated.[52] Recently, a cell-permeable tellurophene-containing ABP labeling proximal nucleophile has been developed for measuring SA-β-gal activity using an MS-based detection method.[53] It can be envisioned that the here-described cyclophellitol aziridine fluorescent ABPs could also be used to simultaneous profile activity from multiple β-galactosidase across a pH range during senescence, using the simpler gel-based method.

Because the molecular processing, enzymatic activity, and biology of GLB1-like proteins are still largely unknown, the ability of ABP **5** and **6** to profile their activity could offer exciting research directions. These include understanding their cell biology and biochemistry, and potential involvement senescence and in the lysosomal storage diseases caused by GLB1 or GALC deficiency. Additionally, the Cy5 ABP labels the dietary enzyme lactose-phlorizin hydrolase (LPH) in mouse duodenum extracts and in lysates of cells overexpressing this protein. This probe might be also useful to monitor the activity of this enzyme in other research context, such as its age-dependent expression[54] and in diseases such as lactose-intolerance.[55]

To conclude, the novel β-galactose configured cyclophellitol aziridine compounds are true mechanism-based irreversible inhibitors for both GLB1 and GALC. They concomitantly inactivate GBA and GBA2 at higher concentrations, but by pre-incubating the samples with β-glucosidase inhibitors their visualization with the β-galactose ABPs can be prevented. The fluorescent and biotinylated ABPs are useful to investigate activities of GLB1 and GALC across

cell lysates and tissue homogenates. They also labeled the dietary enzyme LPH, and additionally revealed that the GLB1-like proteins are functional β-galactosidases that have different post-translational processing, pH optimum, and tissue-dependent expression to that of GLB1. They should be valuable tools in the future study of β-galactosidases in health and disease.

7.4 Experimental procedures

7.4.1 Chemicals for biological assays

Cyclophellitol (CP) and the β-glucosidase ABPs were synthesized as described earlier.[39, 46] Chemicals were obtained from Sigma-Aldrich, if not otherwise indicated. Protein concentration was measured using Pierce™ BCA assay kit (Thermo Fisher).

7.4.2 Cell culture

HEK293T cells (ATCC, CRL-3216) were cultured in DMEM high glucose (Sigma-Aldrich) supplemented with 10 % FCS, 0.1 (w/v) penicillin/streptomycin, and 1 % (v/v) Glutamax at 37 °C at 7 % CO_2, and sub-cultured at 1:10 ratio twice a week. Normal human dermal fibroblasts (NHDF, Cambrex-Lona, CC-2511) were cultured in HAMF12-DMEM medium (Sigma-Aldrich) supplemented with 10 % FCS and 0.1 (w/v) penicillin/streptomycin at 5 % CO_2, and sub-cultured at 1:3 ratio once per 1-2 weeks. Culture medium was refreshed every 2-3 days for NHDF.

7.4.3 Enzyme activity assays

Recombinant murine GALC was cloned and expressed in HEK293 cells following previously described procedures.[38] The produced protein which was secreted to the culture medium (DMEM high glucose, Gibco) was directly used (5 µL volume) in enzyme activity studies. Human fibroblast lysates were prepared in KPi buffer (25 mM K_2HPO_4/KH_2PO_4, pH 6.5, 0.1 % (v/v) Triton X-100, protease inhibitor cocktail (Roche, version 12, EDTA-free)), and 5.5 µg total protein was used. β-galactosidase from *Cellvibrio japonicus* (CjGH35A)[45] was prepared in KPi buffer, and 8.9 ng protein was used. Enzyme activity assays were performed in 96-well plates (Greiner, black, flat-bottomed, medium-binding). All samples were diluted in McIlvaine buffer (150 mM citrate/Na_2HPO_4, pH 4.5) in a total volume of 25 µL, before incubated with 100 µL substrate mixture (1 mM 4-methylumbelliferyl β-D-galactopyranoside (4-MU-β-gal) in 150 mM McIlvaine buffer (pH 4.5) with 0.2 M NaCl) for 30 min at 37 °C. Reactions were stopped and analyzed as described earlier.[56] To discriminate GLB1 and GALC activity, samples were incubated with McIlvaine buffer with or without 11 µM $AgNO_3$[43], before incubated with substrates. To determine the apparent IC_{50} values, enzymes were equilibrated in 12.5 µL McIlvaine buffer (pH 4.5) and incubated with a range of inhibitor dilutions (12.5 µL) for 30 min at 37 °C, before subsequent incubation with substrates. All assays were performed in triplicate wells, and IC_{50} assays were additionally performed in experimental duplicates per compounds.

IC_{50} values were calculated as described earlier.[56] KPi buffer and culture medium form control HEK293T cells were used as negative control for GLB1 and GALC, respectively.

7.4.4 Inhibition kinetics

An enzyme stock was prepared by diluting CjGH35A to 0.71 ng/μL in McIlvaine buffer (pH 4.5, 0.1 % (w/v) BSA), and pre-warmed at 37℃ for 10 min in a thermoshaker. A series of 2 mL Eppendorf tubes containing compounds **2**-**5** (162.5 μL, various compound concentrations) was added with 4-MU-β-gal (1300 μL, reaction [4-MU-β-gal] = 2345 μM), and identically pre-warmed. The t=0 samples were prepared by taking 112.5 μL from the [compound + 4-MU-β-gal] to the first two columns of a 96-well plate (duplo for each compound concentration), after which stop buffer (200 μL, glycine-NaOH 1M, pH 10.3) and lastly the enzyme (12.5 μL) was added. Then, enzyme (137.5 μL) was added (t=0) to each of the 2 mL Eppendorf tubes (containing mixtures of compounds and substrates). At t=2, 4, 6, 8 and 10 min, an aliquot of 125 μL from each of the reaction mixture was taken in duplo and added to the plate containing stop buffer. After 10 minutes, the plate was measured with a LS-55 Fluorescence Spectrometer (PerkinElmer) using BL Studio with excitation at 366 nm and emission at 445 nm. Kinetic experiments were performed in duplo sets, each set with technical duplicates. Obtained values were plotted against incubation time, and fitted with one phase exponential association function to derive k_{obs} values. The obtained k_{obs} values were plotted against compound concentration in a second graph, and curve-fitted with Michaelis-Menten function to derive K_I' and k_{inact}. K_I was derived from K_I' using the function $K_I (1 + [4\text{-MU-}\beta\text{-gal}]/K_M) = K_I'$, where [4-MU-β-gal] = 2345 μM and K_M = 487.4 μM. The K_M was determined in a separate experiment by incubating the enzyme (same concentration as the kinetic experiments) with a series of substrate concentrations for 10 minutes at 37 °C. The reaction was stopped by adding stop buffer and samples were measured for 4-MU fluorescence. The results were plotted against substrate concentration, and K_M was determined with Michaelis-Menten equation. Results were processed and analyzed using GraphPad Prism 7.0.

7.4.5 Labeling using *in situ* generated probes

The *in situ* generated probe was prepared by incubating 1 mM **4** (in DMSO) with 1 mM **7** or **8** (both in DMSO) for at least 15 min at RT, unless otherwise stated. The produced probe had a concentration of 0.5 mM, assuming the click reaction proceeded to completion. Human

fibroblast lysates (5 μg protein) and culture medium of HEK293T cells (5 μL volume) were diluted in 150 mM McIlvaine buffer (pH 4.5) in a total volume of 10 μL, followed by incubation with 5 μL of the *in situ* generated probe mixture (diluted in McIlvaine buffer) for 30 min at 37 °C, at a final probe concentration of 1 μM. For competitive activity-based protein profiling (cABPP), human fibroblasts lysates were diluted as described above, and pre-incubated with 2.5 μL ME569[57] or the *in situ* generated probe mixture for 30 min at 37 °C at a final probe concentration of 100 nM for ME569 and 1 μM for the *in situ* generated probe. Samples were subsequently incubated with the second probe (ME569 for samples pre-incubated with the in situ generated probe, and vice versa) for another 30 min at 37 °C with identical probe concentration during pre-incubation. Samples were denatured, resolved by SDS-PAGE, and scanned for fluorescence as described previously.[56]

7.4.6 Competitive ABPP in extracts of mouse kidney or brain

25 μg protein from extracts of mouse kidney or brain (prepared as described in chapter 8) was diluted in 10 μL McIlvaine buffer (150 mM, pH 4.5), and pre-incubated with 2 μL cyclophellitol, **3**, or **4** (diluted in DMSO and McIlvaine buffer) for 30 min at 37 °C at 50 or 500 nM concentration for compound during incubation. Samples were subsequently incubated with 2 μL ME569 (100 nM during incubation) for 30 min at 37 °C, then with 2 μL *in situ* generated probe (**4** + **8,** 1 μM of the clicked product during incubation, assuming complete reaction) or 8 (1 μM during incubation, as negative control) for 30 min at 37 °C. Samples were denatured, resolved by SDS-PAGE, and scanned for fluorescence as described previously.[56]

7.4.7 Characterization of labeling by ABP **5** in mouse kidney extracts

25 μg total protein from mouse kidney extracts was diluted in McIlvaine buffer (150 mM, various pH) in a total volume of 10 μL, and incubated with 5 μL ABP **5** prepared at different concentrations and pHs. For labeling at varying pH, kidney extracts were incubated with McIlvaine buffer (pH 3.0 – 7.0) for 15 min at 37 °C, followed by incubating with MDW933[46] (prepared in McIlvaine buffer pH 3.0 - 7.0, 5 μM during incubation) for 30 min at 37 °C to label all GBA, then with 1 μM ABP **5** (prepared in McIlvaine buffer pH 3.0 -7.0, 1 μM during incubation) for 30 min at 37 °C. For labeling at varying ABP concentration, mouse kidney extracts were equilibrated in McIlvaine buffer pH 4.5 for 5 min on ice, incubated with 5 μM MDW933 (end concentration; pH 4.5) for 30 min at 37 °C, then with varying concentration of

ABP **5** (end concentration = 10nM -10μM; pH 4.5) for 30 min at 37 °C. For labeling at varying incubation time, mouse kidney extracts were prepared as above, and incubated with 1 μM ABP **5** (end concentration; pH 4.5) at 37 °C for 1 min to 2 h without MDW933 pre-incubation.

7.4.8 Chemical proteomics

3.5 mg protein from mouse kidney extracts was diluted with McIlvaine buffer (750 mM, pH 4.5) in a total volume of 500 μL. Three samples were prepared, the first incubated with DMSO only, the second firstly with 5 μM ABP **5** and secondly with 10 μM *in situ* generated ABP **6** (1 mM **4** + 2 mM **9**, O/N incubation at RT in DMSO), and the third with firstly DMSO and secondly the 10 μM *in situ* generated ABP **6**. All incubation step was performed at 37 °C for 1 h. Samples were denatured and subjected to pull-down, on-bead tryptic digest, and LC-MS protein identification as described in Chapter 6.

7.4.9 ABPP in mouse tissue extracts with ABP **5**

20 μg total protein from extracts of mouse kidney, brain, epididymis, testis, duodenum, and intestine were diluted in 10 μL McIlvaine buffer (150 mM, various pH from 3.0 to 7.0). Samples were then pre-incubated with 2.5 μL ABP JJB70[58] (diluted in McIlvaine buffer pH 3.0 – 7.0, 200 nM during incubation) for 30 min at 37 °C, and then 2.5 μL ABP **5** (diluted in McIlvaine buffer pH 3.0 – 7.0, 1 μM during incubation) for 30 min at 37 °C, before subjected to SDS-PAGE-based fluorescent detection. For deglycosylation analysis, ABP labeled samples (60 μg protein diluted consecutively in 20 μL McIlvaine buffer pH 6.0 (5 min on ice), 5 μL ABP JJB70 (30 min, 37 °C) and 5 μL ABP **5** (30 min, 37 °C), at identical ABP concentrations as described) were firstly desalted using Pierce 7K polyacrylamide spin column (Thermo Fisher), and a 10 μL aliquot was treated with PNGase F according to the manufacturer's protocol (New England BioLabs). Non-treated samples (20 μg protein diluted in 10 μL McIlvaine buffer pH 6.0) were similarly labeled with ABPs. Both the non-treated and PNGase F-treated samples were subjected to SDS-PAGE and fluorescence detection. For labeling on recombinant LPH, wild-type of catalytic mutant LPH constructs[59] were expressed in HEK293T cells using the PEI method (Chapter 6). Cells were lysed in lysis buffer, and 20 μg total protein from the resulting lysates were labeled with 10 μM ABP **5** at pH 6.0 for 30 min at 37°C, before subjecting to SDS-PAGE and fluorescence detection.

7.4.10 Bioinformatics

Amino acid sequences (human GLB1: NP_000395.3, GLB1L: NP_001273352.1, GLB1L2: NP_612351.2, GLB1L3: NP_001073876.2; mouse (*Mus musculus*) GLB1: NP_033882.1, GLB1L: NP_083286.1, GLB1L2: NP_722498.1, GLB1L3: NP_001106794.1) were retrieved from NCBI. Multiple sequence alignment was performed with CLUSTALW.[60] Signal peptide prediction was based on SignalP 5.0[61] (DTU Bioinformatics). Transmembrane domain prediction was performed with XtalPred-RF.[62] N-glycosylation sites were predicted using NetNGlyc 1.0[63] (DTU Bioinformatics). Cellular localization was predicted using WoLF PSORT.[64]

7.5 References

1 Hennet T (2002) The galactosyltransferase family. *Cell Mol Life Sci* **59**, 1081–1095.

2 Suzuki Y & Nanba E (2010) β-Galactosidase deficiency (β-galactosidosis) GM1-gangliosidosis and Morquio B disease. In *The Metabolic and Molecular Basis of Inherited Disease*. New York, NY: McGraw-Hill.

3 Suzuki K & Suzuki Y (1970) Globoid cell leucodystrophy (Krabbe's disease): deficiency of galactocerebroside beta-galactosidase. *Proc Natl Acad Sci USA* **66**, 302–309.

4 Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM & Henrissat B (2014) The Carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res* **42**, D490–D495.

5 D'Azzo A, Hoogeveen A, Reuser AJ, Robinson D & Galjaard H (1982) Molecular defect in combined beta-galactosidase and neuraminidase deficiency in man. *Proc Natl Acad Sci USA* **79**, 4535–4539.

6 van der Spoel A, Bonten E & d'Azzo A (2000) Processing of lysosomal beta-galactosidase. The C-terminal precursor fragment is an essential domain of the mature enzyme. *J Biol Chem* **275**, 10035–10040.

7 Kreutzer R, Kreutzer M, Pröpsting MJ, Sewell AC, Leeb T, Naim HY & Baumgärtner W (2008) Insights into post-translational processing of beta-galactosidase in an animal model resembling late infantile human G-gangliosidosis. *J Cell Mol Med* **12**, 1661–1671.

8 Ohto U, Usui K, Ochi T, Yuki K, Satow Y & Shimizu T (2012) Crystal structure of human β-galactosidase: structural basis of Gm1 gangliosidosis and morquio B diseases. *J Biol Chem* **287**, 1801–1812.

9 Wilkening G, Linke T, Uhlhorn-Dierks G & Sandhoff K (2000) Degradation of membrane-bound ganglioside GM1. Stimulation by bis(monoacylglycero)phosphate and the activator proteins SAP-B and GM2-AP. *J Biol Chem* **275**, 35814–35819.

10 Galjart NJ, Gillemans N, Harris A, van der Horst GT, Verheijen FW, Galjaard H & d'Azzo A (1988) Expression of cDNA encoding the human "protective protein" associated with lysosomal beta-galactosidase and neuraminidase: homology to yeast proteases. *Cell* **54**, 755–764.

11 Morreau H, Galjart NJ, Gillemans N, Willemsen R, van der Horst GT & d'Azzo A (1989) Alternative splicing of beta-galactosidase mRNA generates the classic lysosomal enzyme and a beta-galactosidase-related protein. *J Biol Chem* **264**, 20655–20663.

12 Caciotti A, Donati MA, Boneh A, d'Azzo A, Federico A, Parini R, Antuzzi D, Bardelli T, Nosi D, Kimonis V, Zammarchi E & Morrone A (2005) Role of beta-galactosidase and elastin binding protein in lysosomal and nonlysosomal complexes of patients with GM1-gangliosidosis. *Hum Mutat* **25**, 285–292.

13 Nagano S, Yamada T, Shinnoh N, Furuya H, Taniwaki T & Kira J (1998) Expression and processing of recombinant human galactosylceramidase. *Clin Chim Acta* **276**, 53–61.

14 Wenger DA, Rafi MA & Luzi P (1997) Molecular genetics of Krabbe disease (globoid cell leukodystrophy): diagnostic and clinical implications. *Hum Mutat* **10**, 268–279.

15 Deane JE, Graham SC, Kim NN, Stein PE, McNair R, Cachón-González MB, Cox TM & Read RJ (2011) Insights into Krabbe disease from structures of galactocerebrosidase. *Proc Natl Acad Sci USA* **108**, 15169–15173.

16 Hill CH, Graham SC, Read RJ & Deane JE (2013) Structural snapshots illustrate the catalytic cycle of β-galactocerebrosidase, the defective enzyme in Krabbe disease. *Proc Natl Acad Sci USA* **110**, 20479–20484.

17 Matsuda J, Vanier MT, Saito Y, Tohyama J, Suzuki K & Suzuki K (2010) A mutation in the saposin

A domain of the sphingolipid activator protein (prosaposin) gene results in a late-onset, chronic form of globoid cell leukodystrophy in the mouse. *Hum Mol Genet* **10**, 1191–1199.

18 Hill CH, Cook GM, Spratley SJ, Fawke S, Graham SC & Deane JE (2018) The mechanism of glycosphingolipid degradation revealed by a GALC-SapA complex structure. *Nat Commun* **9**, 151.

19 Chen YQ, Rafi MA, de Gala G & Wenger DA (1993) Cloning and expression of cDNA encoding human galactocerebrosidase, the enzyme deficient in globoid cell leukodystrophy. *Hum Mol Genet* **2**, 1841–1845.

20 Brunetti-Pierri N & Scaglia F (2008) GM1 gangliosidosis: review of clinical, molecular, and therapeutic aspects. *Mol Genet Metab* **94**, 391–396.

21 Suzuki Y, Nakamura N & Fukuoka K (1978) GM1-gangliosidosis: accumulation of ganglioside GM1 in cultured skin fibroblasts and correlation with clinical types. *Hum Genet* **43**, 127–131.

22 Tessitore A, dPM M, Sano R, Ma Y, Mann L, Ingrassia A, Laywell ED, Steindler DA, Hendershot LM & d'Azzo A (2004) GM1-ganglioside-mediated activation of the unfolded protein response causes neuronal death in a neurodegenerative gangliosidosis. *Mol Cell* **15**, 753–766.

23 Takamura A, Higaki K, Kajimaki K, Otsuka S, Ninomiya H, Matsuda J, Ohno K, Suzuki Y & Nanba E (2008) Enhanced autophagy and mitochondrial aberrations in murine G(M1)-gangliosidosis. *Biochem Biophys Res Commun* **367**, 616–622.

24 Sano R, Annunziata I, Patterson A, Moshiach S, Gomero E, Opferman J, Forte M & d'Azzo A (2009) GM1-ganglioside accumulation at the mitochondria-associated ER membranes links ER stress to Ca(2+)-dependent mitochondrial apoptosis. *Mol Cell* **36**, 500–511.

25 Spiegel R, Bach G, Sury V, Mengistu G, Meidan B, Shalev S, Shneor Y, Mandel H & Zeigler M (2005) A mutation in the saposin A coding region of the prosaposin gene in an infant presenting as Krabbe disease: first report of saposin A deficiency in humans. *Mol Genet Metab* **84**, 160–166.

26 Won JS, Singh AK & Singh I (2016) Biochemical, cell biological, pathological, and therapeutic aspects of Krabbe's disease. *J Neurosci Res* **94**, 990–1006.

27 Won JS, Kim J, Paintlia MK, Singh I & Singh AK (2013) Role of endogenous psychosine accumulation in oligodendrocyte differentiation and survival: implication for Krabbe disease. *Brain Res* **1508**, 44–52.

28 Graziano AC, Parenti R, Avola R & Cardile V (2016) Krabbe disease: involvement of connexin43 in the apoptotic effects of sphingolipid psychosine on mouse oligodendrocyte precursors. *Apoptosis* **21**, 25–35.

29 Deodato F, Procopio E, Rampazzo A, Taurisano R, Donati MA, Dionisi-Vici C, Caciotti A, Morrone A & Scarpa M (2017) The treatment of juvenile/adult GM1-gangliosidosis with Miglustat may reverse disease progression. *Metab Brain Dis* **32**, 1529–1536.

30 Condori J, Acosta W, Ayala J, Katta V, Flory A, Martin R, Radin J, Cramer CL & Radin DN (2016) Enzyme replacement for GM1-gangliosidosis: Uptake, lysosomal activation, and cellular disease correction using a novel β-galactosidase:RTB lectin fusion. *Mol Genet Metab* **117**, 199–209.

31 Gupta M, Pandey H & Sivakumar S (2017) Intracellular Delivery of β-Galactosidase Enzyme Using Arginase-Responsive Dextran Sulfate/Poly-l-arginine Capsule for Lysosomal Storage Disorder. *ACS Omega* **2**, 9002–9012.

32 Schalli M, Weber P, Tysoe C, Pabst BM, Thonhofer M, Paschke E, Stütz AE, Tschernutter M, Windischhofer W & Withers SG (2017) A new type of pharmacological chaperone for G(M1)-gangliosidosis related human lysosomal β-galactosidase: N-Substituted 5-amino-1-hydroxymethyl-cyclopentanetriols. *Bioorg Med Chem Lett* **27**, 3431–3435.

33 Hossain MA, Higaki K, Saito S, Ohno K, Sakuraba H, Nanba E, Suzuki Y, Ozono K & Sakai N (2015) Chaperone therapy for Krabbe disease: potential for late-onset GALC mutations. *J Hum Genet* **60**, 539–545.

34 Baek RC, Broekman ML, Leroy SG, Tierney LA, Sandberg MA, d'Azzo A, Seyfried TN & Sena-Esteves M (2010) AAV-mediated gene delivery in adult GM1-gangliosidosis mice corrects lysosomal storage in CNS and improves survival. *PLoS One* **5**, e13468.

35 Przybilla MJ, Ou L, Tăbăran AF, Jiang X, Sidhu R, Kell PJ, Ory DS, O'Sullivan MG & Whitley CB (2019) Comprehensive behavioral and biochemical outcomes of novel murine models of GM1-gangliosidosis and Morquio syndrome type B. *Mol Genet Metab* **126**, 139–150.

36 Rafi MA, Rao HZ, Luzi P, Luddi A, Curtis MT & Wenger DA (2015) Intravenous injection of AAVrh10-GALC after the neonatal period in twitcher mice results in significant expression in the central and peripheral nervous systems and improvement of clinical features. *Mol Genet Metab* **114**, 459–466.

37 Folts CJ, Scott-Hewitt N, Pröschel C, Mayer-Pröschel M & Noble M (2016) Lysosomal Re-acidification Prevents Lysosphingolipid-Induced Lysosomal Impairment and Cellular Toxicity. *PLoS Biol* **14**, e1002583.

38 Marques AR, Willems LI, Herrera Moro D, Florea BI, Scheij S, Ottenhoff R, van Roomen CP, Verhoek M, Nelson JK, Kallemeijn WW, Biela-Banas A, Martin OR, Cachón-González MB, Kim NN, Cox TM, Boot RG, Overkleeft HS & Aerts JM (2017) A Specific Activity-Based Probe to Monitor Family GH59 Galactosylceramidase, the Enzyme Deficient in Krabbe Disease. *Chembiochem* **18**, 402–412.

39 Schröder SP, van de Sande JW, Kallemeijn WW, Kuo CL, Artola M, van Rooden EJ, Jiang J, Beenakker TJM, Florea BI, Offen WA, Davies GJ, Minnaard AJ, Aerts JMFG, Codée JDC, van der Marel GA & Overkleeft HS (2017) Towards broad spectrum activity-based glycosidase probes: synthesis and evaluation of deoxygenated cyclophellitol aziridines. *Chem Commun (Camb)* **53**, 12528–12531.

40 Willems LI, Beenakker TJM., Murray B, Gagestein B, van den Elst H, van Rijssel ER, Codée JDC, Kallemeijn WW, Aerts JMFG, van der Marel GA & Overkleeft HS (2014) Synthesis of alpha- and beta-Galactopyranose-Configured Isomers of Cyclophellitol and Cyclophellitol Aziridine, *Eur J Org Chem* **2014**, 6044–6056.

41 Beenakker TJM (2017) *Design and development of conformational inhibitors and activity-based probes for retaining glycosidases*. (Doctoral dissertation). Retrieved from Leiden University Repository.

42 Devaraj NK, Weissleder R & Hilderbrand SA (2008) Tetrazine-based cycloadditions: application to pretargeted live cell imaging. *Bioconjug Chem* **19**, 2297–2299.

43 Martino S, Tiribuzi R, Tortori A, Conti D, Visigalli I, Lattanzi A, Biffi A, Gritti A & Orlacchio A (2009) Specific determination of beta-galactocerebrosidase activity via competitive inhibition of beta-galactosidase. *Clin Chem* **55**, 541–548.

44 Marques ARA (2016) *Lysosomal glycosidases and glycosphingolipids: New avenues for research*. (Doctoral dissertation). Retrieved from UvA-DARE.

45 Larsbrink J, Thompson AJ, Lundqvist M, Gardner JG, Davies GJ & Brumer H (2014) A complex gene locus enables xyloglucan utilization in the model saprophyte Cellvibrio japonicus. *Mol Microbiol* **94**, 418–433.

46 Witte MD, Kallemeijn WW, Aten J, Li KY, Strijland A, Donker-Koopman WE, Van Den Nieuwendijk AMCH, Bleijlevens B, Kramer G, Florea BI, Hooibrink B, Hollak CEM, Ottenhoff R, Boot RG, Van Der Marel GA, Overkleeft HS & Aerts JMFG (2010) Ultrasensitive in situ visualization of active glucocerebrosidase molecules. *Nat. Chem. Biol.* **6**, 907–913.

47 Wacker H, Keller P, Falchetto R, Legler G & Semenza G (1992) Location of the two catalytic sites in intestinal lactase-phlorizin hydrolase. Comparison with sucrase-isomaltase and with other glycosidases, the membrane anchor of lactase-phlorizin hydrolase. *J Biol Chem* **267**, 18744–18752.

48 Wu C, Orozco C, Boyer J, Leglise M, Goodale J, Batalov S, Hodge CL, Haase J, Janes J, Huss JW 3rd & Su AI (2009) BioGPS: an extensible and customizable portal for querying and organizing gene

annotation resources. *Genome Biol* **10**, R130.

49 Lattin JE, Schroder K, Su AI, Walker JR, Zhang J, Wiltshire T, Saijo K, Glass CK, Hume DA, Kellie S & Sweet MJ (2008) Expression analysis of G Protein-Coupled Receptors in mouse macrophages. *Immunome Res* **4**, 5.

50 Dimri GP, Lee X, Basile G, Acosta M, Scott G, Roskelley C, Medrano EE, Linskens M, Rubelj I & Pereira-Smith O (1995) A biomarker that identifies senescent human cells in culture and in aging skin in vivo. *Proc Natl Acad Sci USA* **92**, 9363–9367.

51 Lee BY, Han JA, Im JS, Morrone A, Johung K, Goodwin EC, Kleijer WJ, DiMaio D & Hwang ES (2006). Senescence-associated beta-galactosidase is lysosomal beta-galactosidase. *Aging Cell* **5**, 187–195.

52 Hildebrand DG, Lehle S, Borst A, Haferkamp S, Essmann F & Schulze-Osthoff K (2013) α-Fucosidase as a novel convenient biomarker for cellular senescence. *Cell Cycle* **12**, 1922–1927.

53 Lumba MA, Willis LM, Santra S, Rana R, Schito L, Rey S, Wouters BG & Nitz M (2017) A β-galactosidase probe for the detection of cellular senescence by mass cytometry. *Org Biomol Chem* **15**, 6388–6392.

54 Rings EHHM, van Beers EH, Krasinski SD, Verhave M, Montgomery RK, Grand RJ, Dekker J & Büller HA (1994) Lactase; Origin, gene expression, localization, and function. *Nutri Res* **14**, 775–797.

55 Lloyd M, Mevissen G, Fischer M, Olsen W, Goodspeed D, Genini M, Boll W, Semenza G & Mantei N (1992) Regulation of intestinal lactase in adult hypolactasia. *J Clin Invest* **89**, 524–529.

56 Kuo CL, Kallemeijn WW, Lelieveld LT, Mirzaian M, Zoutendijk I, Vardi A, Futerman AH, Meijer AH, Spaink HP, Overkleeft HS, Aerts JMFG & Artola M (2019) In vivo inactivation of glycosidases by conduritol B epoxide and cyclophellitol as revealed by activity-based protein profiling. *FEBS J* **286**, 584–600.

57 Artola M, Kuo CL, Lelieveld LT, Rowland RJ, van der Marel GA, Codée JDC, Boot RG, Davies GJ, Aerts JMFG & Overkleeft HS (2019) Functionalized Cyclophellitols Are Selective Glucocerebrosidase Inhibitors and Induce a Bona Fide Neuropathic Gaucher Model in Zebrafish. *J Am Chem Soc* **141**, 4214–4218.

58 Jiang J, Beenakker TJ, Kallemeijn WW, van der Marel GA, van den Elst H, Codée JD, Aerts JM & Overkleeft HS (2015) Comparing Cyclophellitol N-Alkyl and N-Acyl Cyclophellitol Aziridines as Activity-Based Glycosidase Probes. *Chemistry* **21**, 10861–10869.

59 Kallemeijn WW, Witte MD, Voorn-Brouwer TM, Walvoort MT, Li KY, Codée JD, van der Marel GA, Boot RG, Overkleeft HS & Aerts JM (2014) A sensitive gel-based method combining distinct cyclophellitol-based probes for the identification of acid/base residues in human retaining β-glucosidases. *J Biol Chem* **289**, 35351–35362.

60 Thompson JD, Higgins DG & Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**, 4673–4680.

61 Almagro Armenteros JJ, Tsirigos KD, Sønderby CK, Petersen TN, Winther O, Brunak S, von Heijne G & Nielsen H (2019) SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat Biotechnol* **37**, 420–423

62 Slabinski L, Jaroszewski L, Rychlewski L, Wilson IA, Lesley SA & Godzik A (2007) XtalPred: a web server for prediction of protein crystallizability. *Bioinformatics* **23**, 3403–3405.

63 Gupta R & Brunak S (2002) Prediction of glycosylation across the human proteome and the correlation to protein function. *Pac Symp Biocomput*, 310–322.

64 Horton P, Park KJ, Obayashi T, Fujita N, Harada H, Adams-Collier CJ & Nakai K (2007) WoLF PSORT: protein localization predictor. *Nucleic Acids Res* **35**, W585–587.

APPENDIX

7.S1. Supplementary Tables and Figures

Table 7.S1. Apparent IC_{50} values for compounds towards CjGH35A, GBA, and GBA2. Error range = ± SD, *n* = 2 biological replicates.

Compound	CjGH35A
2	21.55 ± 2.14
3	1.22 ± 0.66
4	27.16 ± 1.57
5	33.9 ± 2.15

Compound	GBA	GBA2
2	1,560	3,290
3	14.6	164
4	5,370	> 10,000
5	85.8	752
6	119	1,830
SYD215	430	34.9

Table 7.S2. Structural determination and refinement statistics for CjGH35A in complex with 2 (University of York).

	CjGH35A in complex with **5** PDB code: 5JAW
Data collection	
Space group	P1
Cell dimensions	
a, *b*, *c* (Å)	98.9, 115.8, 116.0
α, β, γ (°)	90.2, 90.2, 90.4
Resolution (Å)	81.99-1.6 (1.63-1.60)
R_{merge}	0.080 (0.588)
R_{pim}	0.080 (0.588)
CC(1/2)	0.979 (0.482)
$I/\sigma I$	6.0 (1.2)
Completeness (%)	95.7 (94.1)
Redundancy	1.8 (1.8)
Refinement	
Resolution (Å)	1.60
No. reflections	650665
R_{work} / R_{free}	0.20/0.21
No. atoms	
Protein	33482
Ligand/ion	125
Water	1637
B-factors (Å^2)	
Protein	25
Ligand/ion	20
Water	28
R.m.s. deviations	
Bond lengths (Å)	0.018
Bond angles (°)	1.927

*Values in parentheses are for highest-resolution shell.

Table 7.S3. Comparison between GLB1 and GLB1-like proteins in A) mouse and B) human.

A

Protein	**GLB1**	**GLB1-like 1**	**GLB1-like 2**	**GLB1-like 3**
Gene	*Glb1*	*Glb1l*	*Glb1l2*	*Glb1l3*
Animo acid	647 (isoform 1)	646	652 (isoform 1)	662
Molecular weight	73.12	73.28	73.96	75.59
N-terminal signal sequence	1-24	1-23	No	No
Transmembrane domain	No	No	13-35	No
N-glycan sites	Asn27, Asn248, Asn500, Asn504, Asn510	Asn93, Asn239	No	No
Cellular localization	Lysosomal	Extracellular	Plasma membrane (outward-facing)	Mitochondrial / peroxisomal

B

Protein	**GLB1**	**GLB1-like 1**	**GLB1-like 2**	**GLB1-like 3**
Gene	*GLB1*	*GLB1L*	*GLB1L2*	*GLB1L3*
Animo acid	677 (isoform a)	654 (isoform 1)	636	653
Molecular weight	76.075	74.158	72.079	74.823
N-terminal signal peptide	1-23	1-27	No	No
Transmembrane domain	No	No	13-32	No
N-glycan sites	Asn26, Asn247, Asn464, Asn498, Asn542, Asn545, Asn555	Asn97, Asn243, Asn625	No	No
Cellular localization	Lysosomal	Mitochondrial / extracellular	Plasma membrane (outward-facing)	Mitochondrial / cytosolic

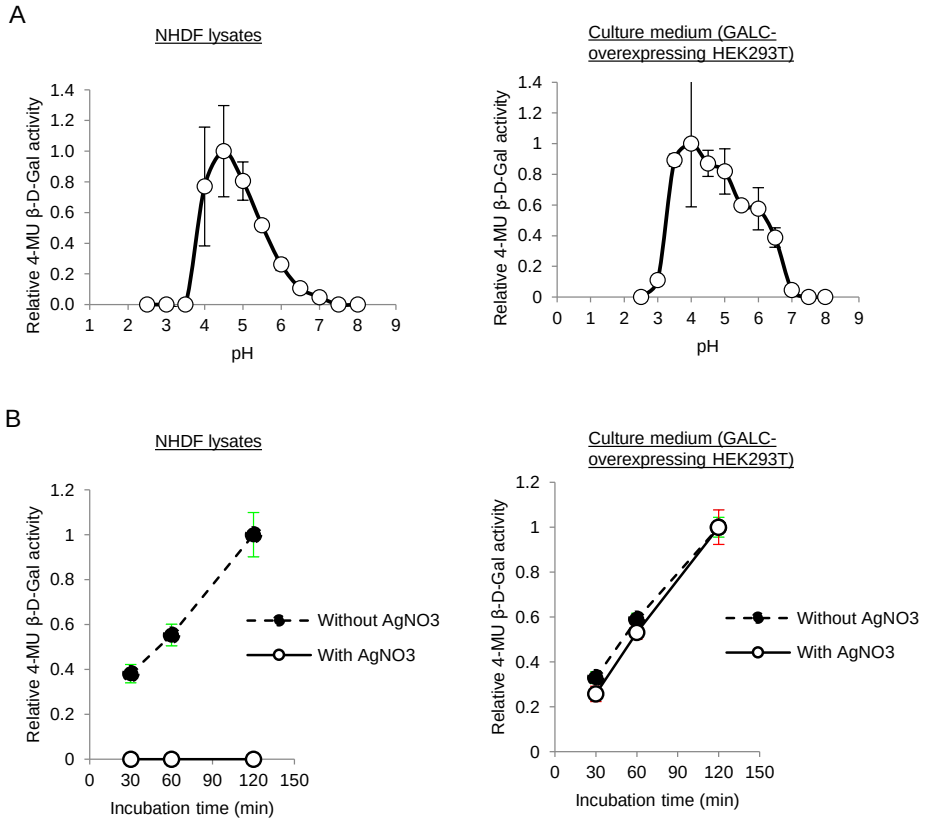


Figure 7.S1. Discriminating activities from GLB1 and GALC in different samples by enzymatic assay. A) pH-dependent β-galactosidase activity (relative to the highest measured activity). B) Effect of $AgNO_3$ (11 μM). Error range = ± SD, *n* = 3 technical replicates.

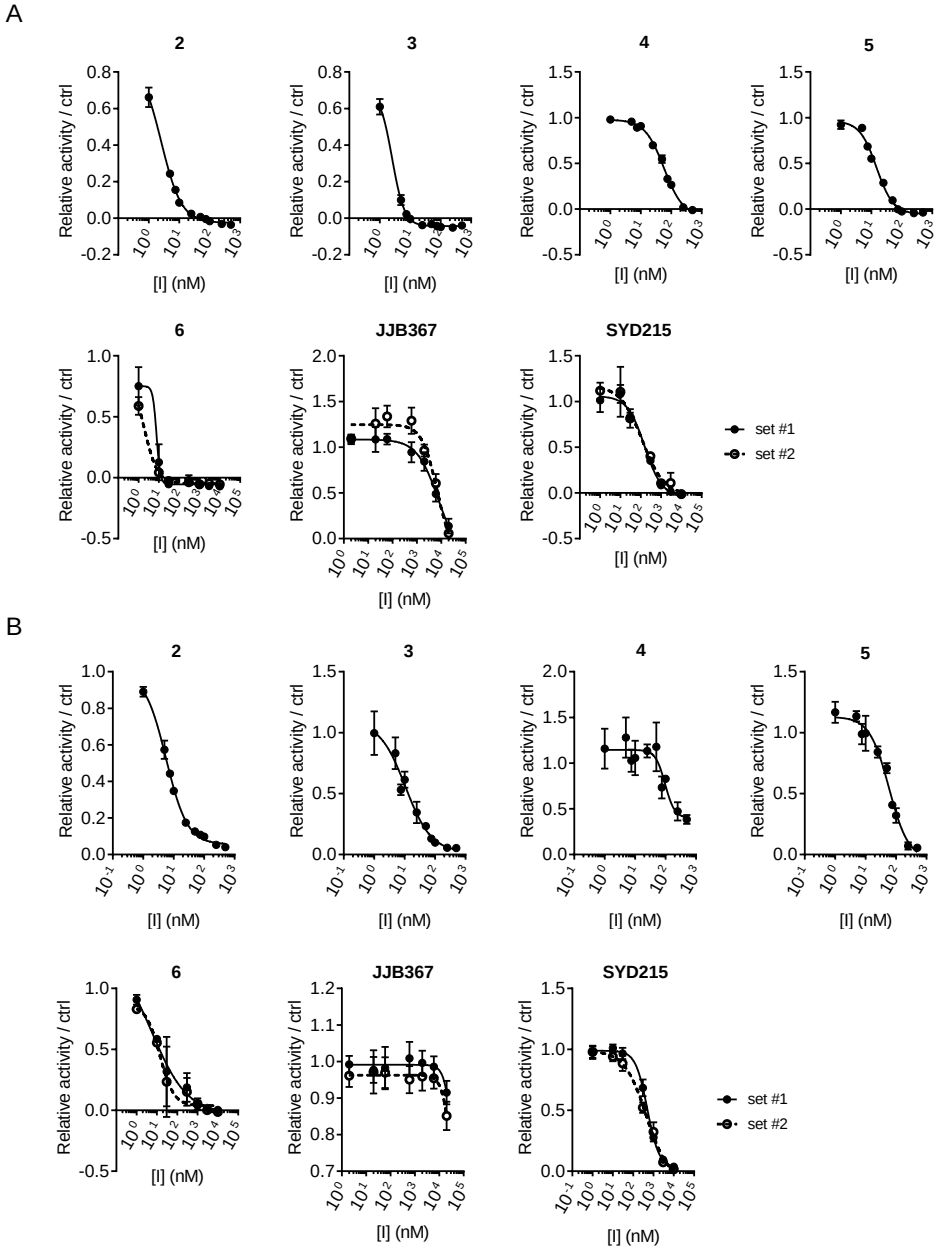


Figure 7.S2. (Continued, 1 of 2).

C

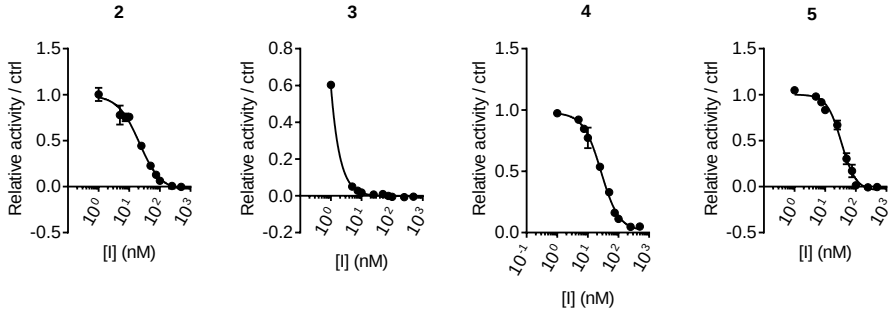


Figure 7.S2. Inhibition curves by compounds towards A) human GLB1 (hGLB1, NHDF lysates), B) mouse GALC (mGALC, from culture medium of HEK293T cells overexpressing the protein), and C) CjGH79. Error range = ± SD (*n* = 3 technical replicates).

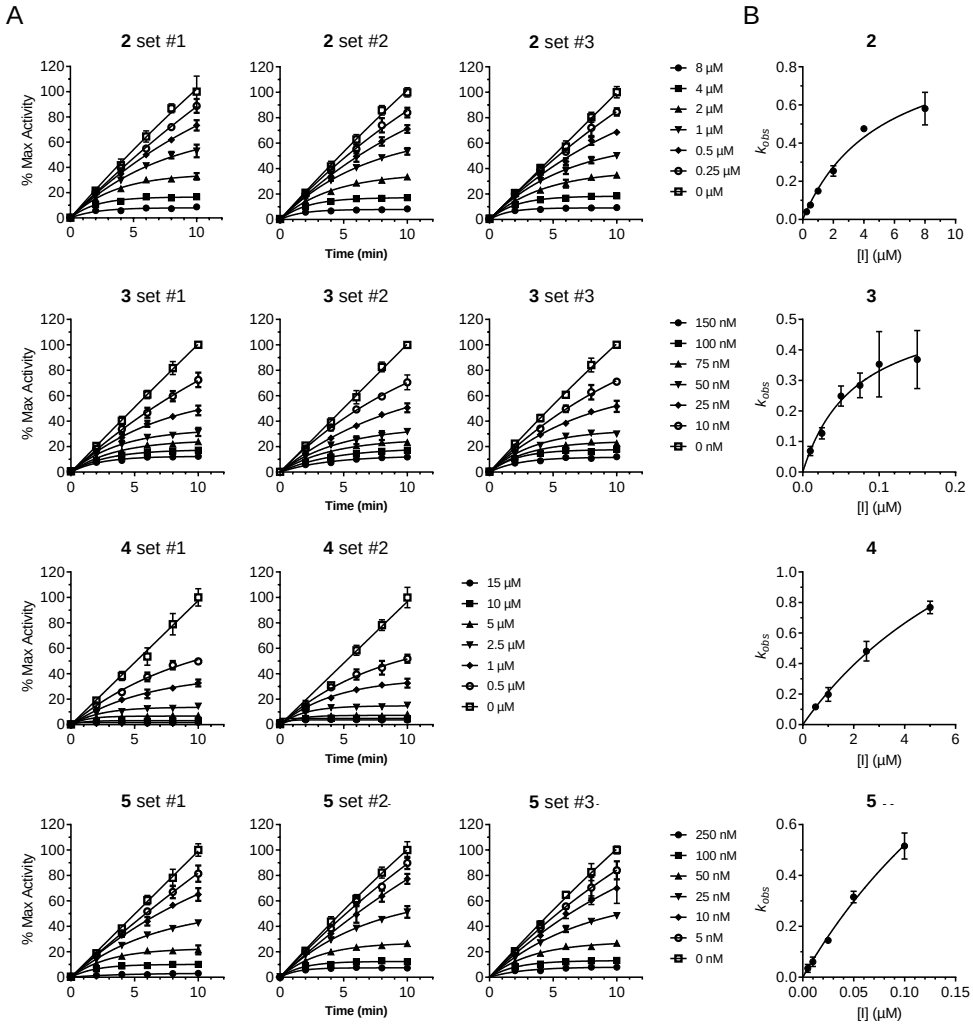


Figure 7.S3. Measurement of inhibition kinetic parameters by compounds 4-5 towards CjGH79. A) Processing curves. Error range = ± SD, *n* = 3 technical replicates. B) k_{obs} vs [I] plots, fitted with Michaelis-Menten equation. Error range = ± SD, *n* = 3 biological replicates.

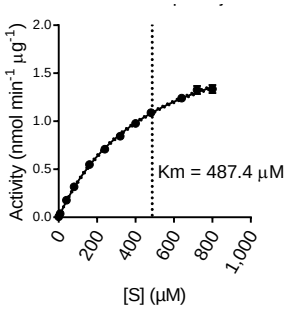


Figure 7.S4. Michaelis-Menten plot of 4-MU-β-gal towards CjGH35. Error range = ± SD, *n* = 3 technical replicates.

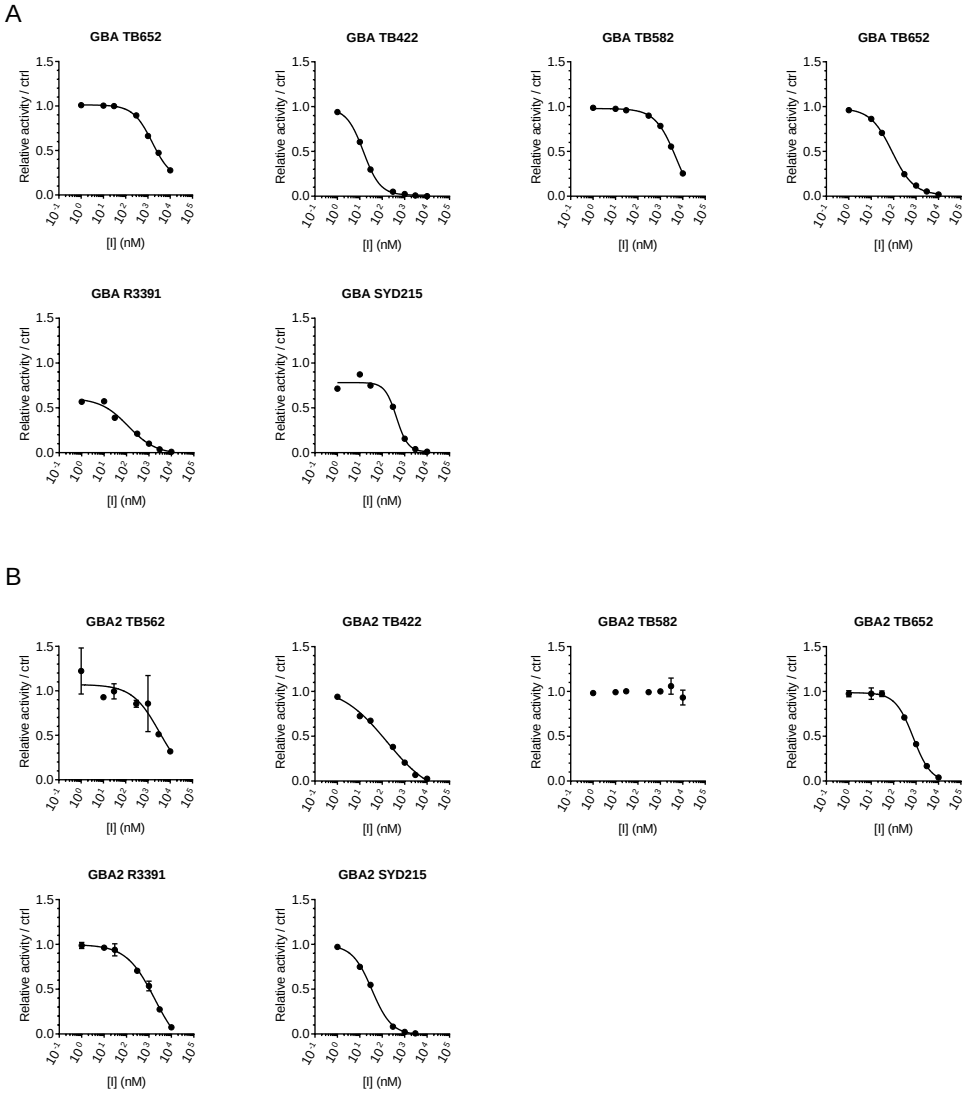


Figure 7.S5. Inhibition curves by compounds towards A) GBA (imiglucerase) and B) GBA2 (lysates of HEK293T cells overexpressing GBA2). Error range = ± SD, *n* = 3 technical replicates.

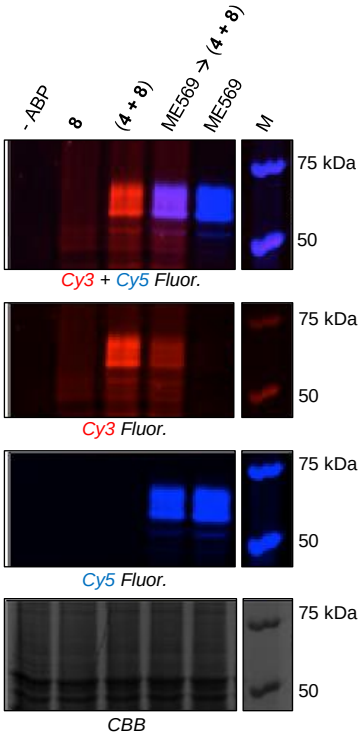


Figure 7.S6. Labeling by the in situ generated probe (4 + 8) in human fibroblast (NHDF) lysates with or without pre-incubation with the GBA-specific ABP ME569. *CBB*, Coomassie Brilliant Blue staining for assessing total protein loading.

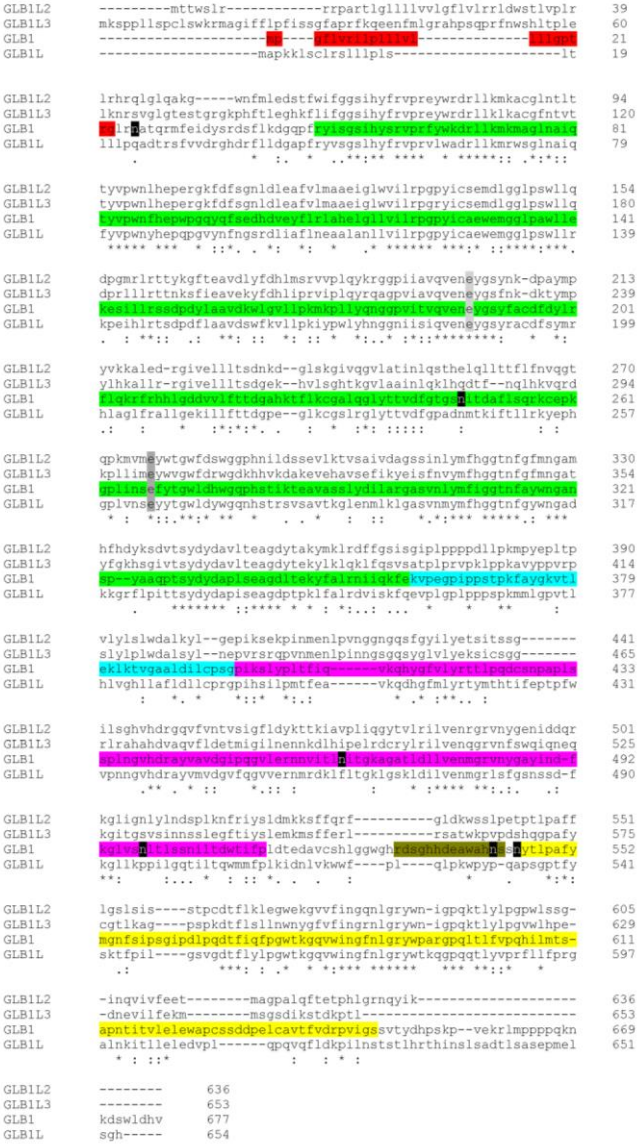


Figure 7.S7. Comparing amino acid sequences between GLB1 and GLB1-like proteins. Red, signal peptide of GLB1. Green, TIM barrel domain of GLB1. Cyan, TIM- β1 loop. Magenta, β-sheet domain 1 of GLB1. Dark yellow, cleaved peptide in mature GLB1. Yellow, β-sheet domain 2 of GLB1. Black, N-linked glycan sites of GLB1. Light gray, catalytic acid/base. Dark gray, catalytic nucleophile. Domain structures assigned based on Ohto *et al.*[main text ref 8]

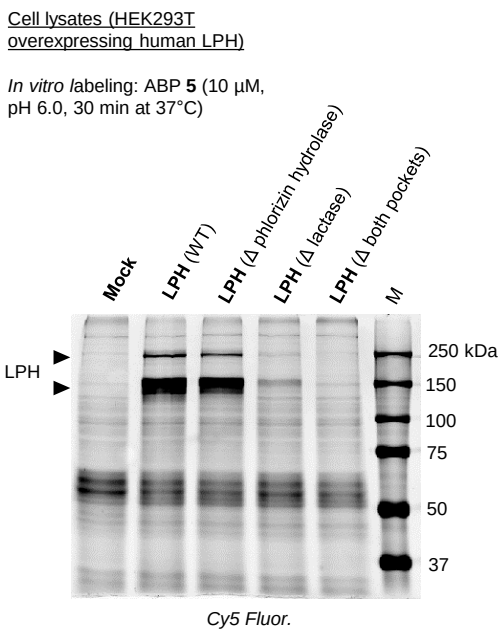
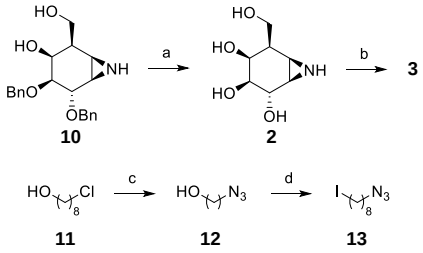
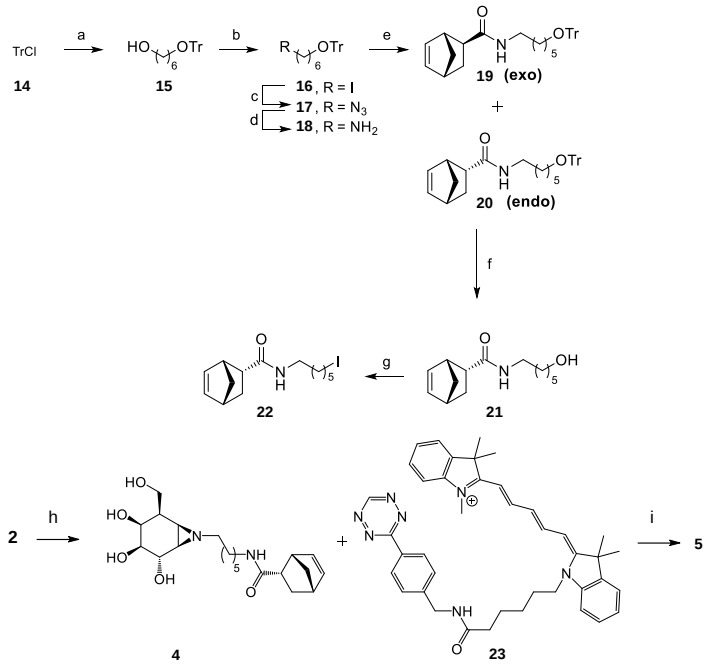


Figure 7.S8. ABP 5 labeling in HEK293T lysates containing expressed human wild-type or mutant LPH.

7.S2. Synthetic strategies for compounds used in this chapter (Department of Bio-Organic Synthesis, Leiden University)



Scheme 7.S1. Synthesis of compounds 2 and 3. Reagents and conditions: a) Li, NH_3, -60 °C, b) **10**, K_2CO_3, DMF, 80 °C, overnight, 25 %; c) NaN_3, DMF, 80 °C, quant; d) PPh_3, imidazole, I_2, THF, -20 °C to RT, 15 min, 69 %.



Scheme 7.S2. Synthesis of compounds 4 and 5. Reagents and conditions: a) 1,6-hexanediol, pyridine, CH_2Cl_2, RT, 90 min, 95 %; b) imidazole, PPh_3, I_2, Et2O, CH_3CN, RT, overnight, 80 %; c) NaN_3, DMF, 80 °C, overnight, quant; d) PPh_3 on beads, H_2O, THF, 48 h, quant; e) norbornene-OSu[1], DIPEA, DCE, RT, overnight, **19** 28 %, **20** 68 %; f) *p*-toluenesulfonic acid, CH_2Cl_2, MeOH, RT, overnight, 86 %; g) PPh_3, I_2, imidazole, THF, reflux, 1.5 h, 73 %. h) **22**, K_2CO_3, DMF, 75 °C, overnight, 12 %. i) Cy5 tetrazine **23**, MeOH, overnight, 87 %.

7.S3 Protein crystallography (University of York)

The gene for CjGH35A, cloned into a pET28a vector modified for Ligation Independent Cloning, was expressed, and the protein purified and crystallized as described previously[main text ref 45], in 2.7 M sodium acetate pH 7.2 (protein at 30 mg/ml, drop 1.2:1 μL over well). A crystal was soaked in the presence of a speck of powder of compound **2** for 70 h. The crystal was fished directly into liquid nitrogen without the need for additional cryoprotectant. Data were collected on beamline IO2 at the Diamond Light Source at wavelength 0.97950 Å, and were processed using DIALS[2] and scaled with AIMLESS[3] to 1.6 Å (Table 7.S2). The space group was P1 and the unit cell dimensions, 98.9, 115.8, 116.0 Å, and angles, 90.2, 90.2, 90.4°. The structure was solved using programs from the CCP4I2 suite. Molecular replacement was performed using Phaser[4], with the native structure, PDB entry 4d1i, as the model. The model was built manually

in Coot[5], followed by repeated cycles of REFMAC[6] employing twin refinement using observed intensities. The structure was annotated in PDB (5JAW).

7.S4 Supplementary References

1 Willems LI, Li N, Florea BI, Ruben M, van der Marel GA & Overkleeft HS (2012) Triple bioorthogonal ligation strategy for simultaneous labeling of multiple enzymatic activities. *Angew Chem Int Ed Engl* **51**, 4431–4.

2 Gildea RJ, Waterman DG, Parkhurst JM, Axford D, Sutton G, Stuart DI, Sauter NK, Evans G & Winter G (2014) New methods for indexing multi-lattice diffraction data. *Acta Crystallogr D Biol Crystallogr* **70**, 2652–2666.

3 Evans PR & Murshudov GN (2013) How good are my data and what is the resolution? *Acta Crystallogr D Biol Crystallogr* **69**, 1204–1214.

4 McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC & Read RJ (2007) Phaser crystallographic software. *J Appl Crystallogr* **40**, 658–674.

5 Emsley P, Lohkamp B, Scott WG & Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486–501.

6 Murshudov GN, Skubák P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F & Vagin AA (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* **67**, 355-–367.

General Discussion and Future Prospects

Activity-based probes for retaining exo-glycosidases: novel research tools for lysosomal storage disorders

The central theme of this thesis are molecular probes labeling retaining exo-glycosidases implicated in health and diseases such as lysosomal storage disorders (LSDs). The structure of these so-called activity-based probes (ABPs) is based on configurational isomers of a cyclophellitol or cyclophellitol aziridine scaffold, to which a reporter tag (e.g. a fluorophore or a biotin) is attached. When incubated with sample, their substrate mimicry and the latent electrophilic trap (epoxide or aziridine) allow them to react upon proximity-driven protonation with their target glycosidase's catalytic nucleophile, resulting in a stable ester bond that irreversibly inactivates the enzyme. Through the grafted reporter tag on the probe, the covalent ABP-enzyme complex can be readily visualized or affinity-enriched, allowing quantitative and qualitative analysis on the spatiotemporal distribution of the labeled enzyme *in vitro*, *in situ*, and *in vivo* by the activity-based protein profiling (ABPP) platform (**General Introduction**, this thesis).

This thesis aims to explore novel applications of glycosidase ABPs to study these enzymes' role in health and disease, in particular in the inherited lysosomal storage diseases (LSDs). It further aims to expand the ABPP platform to other retaining exo-glycosidases such as α-glucosidase, β-glucuronidase, α-L-iduronidase, α/β-mannosidases, and β-galactosidases. Different goals for the thesis work were formulated and reached: (1) setting up ABP profiling protocols for glycosidases (**Chapter 1**), (2) developing two novel *in vivo* applications for ABPs in LSD research (**Chapter 2-3**), and (3) characterizing in detail the mechanism of action, potency, target specificity, and potential applications of six novel classes of ABPs (**Chapter 4-7**). The described work builds on past research utilizing cyclophellitol ABPs labeling the lysosomal glucocerebrosidase (GBA)[1] and galactocerebrosidase (GALC)[2], and cyclophellitol aziridine N-acyl or N-alkyl variants targeting either retaining exo-β-glucosidases (GBA, GBA2, GBA3, and LPH), retaining exo-α-galactosidases, or retaining exo-α-L-fucosidases[3–6]. These ABPs, via the afore-mentioned reaction mechanism, allow sensitive and specific visualization of target enzymes across species and sample types (recombinant enzyme, cell lysates, tissue homogenates, whole cells, and tissue sections) using detection methods such as SDS-PAGE-based fluorescence scanning, fluorescence-activated cell sorting (FACS), fluorescence microscopy, and mass spectrometry (LC-MS/MS). Demonstrated applications include the diagnostic confirmation of LSDs (Gaucher disease, Fabry disease, and Krabbe disease),[1–3] the assessment of bodily distribution of labeled enzymes,[7, 8] the investigation of active-site pocket occupancy of GBA by potential small molecule chaperons,[1] and the identification of the active-site nucleophile of the

labeled enzymes.[9]

The first part of the thesis concerns Gaucher disease. **Chapter 1** systematically describes ABP protocols for gel-based and microscopy-based detection for GBA; these can be adapted to detect other lysosomal glycosidases with ABPs. Next are presented two novel *in vivo* applications of ABPs, being: (1) assessing target specificity of existing pharmacological agents for inducing cell/animal models of Gaucher disease (**Chapter 2**), and (2) generating such cell/animal models using compounds modified from the cyclophellitol-based ABPs (**Chapter 3**). Both applications exploit the cell and tissue permeability of ABPs and their highly selective inactivation of GBA by binding to the catalytic nucleophile E340. The first study established the relevance and limitations of conduritol B epoxide (CBE) and cyclophellitol (CP) in generating Gaucher cell/animal models, while the second investigation led to the identification of superior GBA inhibitors (CPs with installed biphenyl or adamantly moiety at C8).

The second part of the thesis reports on the characterization of novel cyclophellitol aziridine ABPs with a differentially configured cyclophellitol core. For this use was made of enzymatic assays to establish potency of inhibition and kinetic parameters; gel-based ABPP to assess target specificity and labeling at various conditions (concentration of ABP, incubation time, pH, temperature, sample types; chemical proteomics to identify the labeled targets, and protein X-ray crystallography (performed at the University of York) to acquire structural insights regarding the ABP-enzyme complex. It is demonstrated that the examined ABPs allow mechanism-based labeling and activity-based protein profiling (ABPP) for the corresponding LSD-relevant glycosidases, including GH31 lysosomal α-glucosidase (GAA) (**Chapter 4**), GH2 β-glucuronidase (GUSB) (**Chapter 4**), GH39 α-L-iduronidase (IDUA) (**Chapter 5**), GH38 lysosomal α-mannosidase (MAN2B1) (**Chapter 6**), GH2 β-mannosidase (MANBA) (**Chapter 6**), and GH35/GH59 lysosomal β-galactosidase/galactocerebrosidase (GLB1/GALC) (**Chapter 7**). For all ABPs, Cy5- and biotin-tagged variants are available and both variants successfully label their target enzymes. Besides the α-L-iduronidase ABPs that are less potent than other ABP classes, all ABPs are able to visualize glycosidases in the context of complex biological samples. Additional in-class (or similar class) glycosidase targets were identified for several ABPs, consistent with previous findings that cyclophellitol aziridine ABPs are broad-spectrum probes: the in-class targets for the α-glucosidase ABPs were the other GH31 α-glucosidases ER α-glucosidase II (GANAB), maltase-glucoamylase (MGAM), and sucrose-

isomaltase (SI); GH79 heparanase (HPSE) for the β-glucuronidase ABPs; four other GH38 α-mannosidases (MAN2A1, MAN2A2, MAN2B2, and MAN2C1) for the α-mannosidase ABPs; two other GH35 proteins GLB1-like protein 1 and 2 (GLB1L and GLB1L2) for the β-galactosidase ABPs. Of note, all cyclophellitol aziridine ABPs at high concentrations tend to also label GBA, with the exception of the α-L-iduronidase and α-mannosidase ABPs. Specific visualization of the target enzyme class may be achieved by pre-blocking of GBA with selective inhibitors/ABPs for this enzyme.

I. Unique research possibilities and applications offered by the glycosidase ABPs

a. Superior spatial detection

One important feature of the studied ABPs is that they allow one to directly visualize active glycosidase molecules. As such ABP labeling fundamentally differs from detecting an enzyme through measurement of its activity towards fluorogenic substrates or natural substrates. When applied *in situ* to detect the spatial distribution of active glycosidase molecules, the ABP method provides better resolution than zymography based on enzyme activity assays. A recent study reported on detection of active GBA molecules by fluorescent ABPs in human skin sections with higher resolution compared to fluorogenic substrate-based zymography[10] This study established that active GBA molecules largely reside in the extracellular lipid matrix between the viable epidermis and the lower part of stratum corneum. The *in situ* ABP detection technique, conceivably combined with the gel-based detection procedure, may be used to elucidate the (patho)physiological roles of GBA in diseases such as atopic dermatitis,[11] multiple myeloma,[12] and neurological diseases such as amyotrophic lateral sclerosis,[13] Lewy-body Dementia,[14] and Parkinson disease.[15]

b. *In situ* and *in vivo* labeling

The amphiphilic ABPs readily cross cellular membranes, which allows *in situ* and *in vivo* labeling of lysosomal glycosidases that can be coupled to either gel-based, FACS, or fluorescence microscopy-based detection.[1, 16] This approach is superior to detection of enzymes using antibodies, which do not crosses cellular membranes and report glycosidase activity. Recently, cell permeable fluorescence-quenched substrates were developed by the group of Vocadlo and shown to reveal the *in situ* activities by GBA[17, 18] Similar probes are designed by the same researchers for other glycosidases (unpublished data). By virtue of their covalent mode of

inactivation/labeling, ABPs offer several unique applications. One example is to be found in studying the life cycle of glycosidases in cells by pulse-chase experiment,[19, 20] in which the formation, post-translational modification, and degradation can be monitored. Proof of principle in this connection has been obtained for GBA.[1] It can be envisioned that pulse-chase labeling with ABPs can be exploited to systematically examine the effect of different mutations, modifier genes, or other cellular factors on the life cycle of GBA. Such studies may shed further light on Gaucher disease and Parkinson disease. The approach could be extended to other lysosomal glycosidases for which ABPs are available.

c. Distinguishing between isozymes

When coupled to gel-based detection, distinction among different enzymes or isoforms labeled by the same ABP is feasible, obviously provided that these differ in molecular weights (**Chapter 1**). Examples are the β-glucosidases GBA and GBA2, where the former is about 50-65 kDa, and the latter is about 100 kDa.[3, 9, 21] Isoforms of GBA differ in glycan composition, i.e. there are distinct GBA glycoforms between 50-65 kDa),[8] Isoforms may reflect the glycosidase's cellular location (e.g. ER, Golgi, or lysosome). Isoforms may also stem from proteolytic processing, for example the 85 kDa vs. 64 kDa mature forms of GLB1[22, 23]. In general, combining the three approaches to detect enzymes (activity assay, ABP labeling, and antibody labeling) offers the most comprehensive assessment at both protein and activity level. The combination of these three methods also assists best the biochemical confirmation of diagnosis of lysosomal storage diseases caused by defects in glycosidases.

d. Sophisticated analysis of therapeutic enzymes

Another unique advantage offered by ABPs is their potential use to track the cellular uptake and bodily distribution of therapeutic enzymes, as in enzyme replacement therapy (ERT) for LSDs. The labeling of a therapeutic enzyme with ABP is elegant and targeted, taking place at a well-defined catalytic nucleophile in the catalytic pocket. One study has compared head-to-head two registered therapeutic enzymes for ERT of Gaucher disease, imiglucerase and velaglucerase for their receptor-binding and uptake in cultured dendritic cells and macrophage, and for their bodily distribution in treated mice.[8] The enzymes were labeled with similar ABPs with distinct fluorophores, next mixed equimolar after which the uptake and bodily distribution of each was assessed in parallel.[8] In this thesis, the α-L-iduronidase ABPs was successfully

employed to monitor the cellular uptake and lysosomal targeting of labeled therapeutic IDUA molecules in cultured fibroblasts from patients of Hurler disease, mucolipidosis type II, and type III disease, using both fluorescence microscopy and gel-based methods (**Chapter 5**).[24] Similarly, a recently discovered non-glycosylated α-galactosidase from plant was pre-labeled by α-galactose configured ABPs, and the uptake and delivery to lysosomes of the fluorescent enzyme by cultured fibroblasts from a Fabry patient—via a lectin-independent pathway—was visualized.[25] More recently, correlative light and electron microscopy (CLEM) has been utilized to visualize endocytosed ABP-labeled imiglucerase in treated human fibroblasts expressing mannose receptors, revealing efficient delivery in individual lysosomes that contained endogenous GBA molecules as visualized with ABPs with other fluorophores.[26] The newly available ABPs for different LSD-related glycosidases, described in this thesis, may facilitate the assessment of distribution, processing, and efficacy of therapeutic enzymes following their administration to cells or animal models. Future installation of detection tags compatible with infrared-red or radioisotopic detection may even allow real-time monitoring of bodily distribution of the labeled enzymes in animal models or in patients. Existing examples are cysteine protease ABPs containing a ^{64}Co tag (with or without an additional Cy5 tag),[27–29] a serine protease ABP coupled to an ^{111}In tag,[30] and a GBA ABP having a 2-deoxy-2-^{18}F group.[31]

e. Multiplexing

An advantage of ABPs is the possibility offered to design multiplex readouts. Careful selection of ABPs with distinct fluorophores allows simultaneous profiling of multiple glycosidases, even in living cells or animals. Elegant simultaneous ABP profiling of endocannabinoid hydrolases in the mouse brain has been demonstrated.[32] **Chapter 2** of this thesis provides examples of multiplex ABP detection for active retaining exo-β-glucosidases, α-glucosidases, and β-glucuronidase in cultured cells. Cells were incubated with ABPs having distinct fluorophores (BODIPY green, BODIPY red, or Cy5) that label the above mentioned enzymes. Upon cell lysis and subsequent gel-based fluorescence detection, different glycosidases labeled *in vivo* were visualized in parallel. In this manner, decrease in active enzyme molecules can be simultaneously assessed. By adding a subcellular fractionation step prior to gel-based detection, it is conceivable that one can also visualize the comparable/distinct subcellular localization of different glycosidases. Furthermore, the multiplex setup coupled to gel-based detection would also allow convenient assessment of post-translational modification status (e.g.

N-linked glycosylation, proteolytic processing) for multiple enzymes. One potential application for the multiplex ABPP setup is to elucidate the molecular mechanisms underlying cellular processes affecting the transport and function of lysosomal glycosidases. For example, it has been noted that the presence of HEPES, a common pH-buffering additive to cell culture medium, could cause in cultured cells lysosomal stress and abnormality in GBA's glycosylation status.[33] The multiplex ABP incubation, coupled to subcellular fractionation technique and gel-based readout, may thereby be applied to investigate whether upon HEPES treatment GBA is specifically affected in its glycosylation status, compared to other lysosomal glycosidases (for which ABPs are available). Furthermore, it can also be used to screen for the impact of other types of stressors on the lysosomal routing and lysosomal processing/stability of GBA, compared to other lysosomal glycosidases. Impact of different types of enzyme mutations may be similarly examined in the context of delineating genotype-phenotype correlation in LSDs. Furthermore, the multiplex readout may also be used in applied research. It is in principle compatible with different types of sample, such as cells, tissues, urine and plasma, thus assisting screening and diagnostic confirmation of LSDs.

f. Investigating *in vivo* target engagement of glycosidase inhibitors

One of the most powerful applications of the cell permeable glycosidase ABPs is their use in assessing *in vivo* target engagement of (covalent) inhibitors. The *in vivo* target engagement of a drug is crucial to elicit proper therapeutic responses. The same holds for enzyme inhibitors used to induce authentic pharmacological cell/animal models of human disease. Lack of awareness of off-targets of a given small molecule can be fatal,[34] and ABPs are an excellent tool to profile potential off-targets of covalent enzyme inhibitors.[35] As mentioned, the irreversible GBA inhibitor conduritol B epoxide (CBE) with conceived specificity for GBA is commonly used for generating Gaucher disease models. It was closely examined in this thesis in cell and animal models to which extent there is interaction *in vivo* of CBE with off-target glycosidases (other β-glucosidases, α-glucosidases and β-glucuronidase; enzymes for which ABPs were available) (**Chapter 2**). Observed competition by CBE of the ABP labeling of off-target glycosidases provides evidence for undesirable co-inhibition of these enzymes. The investigation identified GBA2 and GAA as major off-target glycosidases at higher CBE concentrations. A narrow window for selective GBA inactivation exists, but when using CBE to induce a Gaucher disease model the experimenter should better check for potential concomitant inhibition of

GBA2 and GAA. The same investigation revealed that cyclophellitol (CP), a closer mimic of glucose than CBE, inhibits GBA and GBA2 with equal affinity in all the studied models.[21]

The assessment of *in vivo* target engagement is not necessarily restricted to covalent, mechanism-based inhibitors. For example, highly potent (nanomolar) reversible iminosugar inhibitors of GBA2 based could be examined for their *in vivo* selectivity for GBA2, GBA and GBA3.[36] Apparently, active-site targeted inhibitors with high affinity stay associated with the enzyme upon lysis of cells and freeze-thawing of lysates, and their ongoing active-site pocket occupancy can be readily visualized by *in vitro* applied ABPs. Obviously, such assays require considerable fine-tuning based on the features of the inhibitor tested and ABP employed in the read-out.

g. Use as mechanism-based inhibitors to inactivate enzymes on demand

The cyclophellitol and cyclophellitol aziridine ABPs can be used to selectively inhibit certain glycosidase *in vitro* or *in vivo*. The installation of the flexible hydrophobic linker (with or without a fluorophore) in fact often enhances the inhibitory potency of the compound towards their targeted glycosidases compared to just the cyclophellitol or cyclophellitol aziridine core alone;[1, 3] sometimes the specificity profile also improves as in the case of ABPs for GBA (where a bulky, hydrophobic moiety at the C8 position of cyclophellitol prevents the labeling towards GBA2).[37] For example, the cyclophellitol ABPs can be used as a more specific GBA inhibitor compared to CBE; the cyclophellitol aziridine ABPs can inhibit besides GBA also GBA2 and GBA3, while not inhibiting glucosylceramide synthase (GCS), a common off-targets by iminosugar inhibitors aiming at GBA2.[38] When combined, these potent, specific, and cell-permeable ABPs would offer convenient tools for studying the role of GBA and GBA2 in sphingolipid metabolism in healthy and disease state.

In addition, the GBA-specific ABPs have been explored as a pharmacological agent in inducing Gaucher-like models (**Chapter 3**).[37] Past research has found that the BODIPY-containing cyclophellitol ABP does not penetrate the brain well of treated mice,[7] and this is consistent with the finding in adult zebrafish (**Chapter 3**). Thus, this ABP can only generates type 1 (non-neuropathic) Gaucher model at best. Due to GBA's association with other neurological disorders such as Parkinson's disease and Lewy-Body Dementia, and that current ERT and SRT therapies do not correct neurological symptoms in Gaucher patients,[39, 40] GBA

inactivation in the animal brain is desired. Although this has been achieved with genetic GBA knockout mice[41] and fish models[42, 43], pharmacological perturbation still offers several advantages, including the experimental simplicity, tunability of the extent of GBA inactivation, and that it allows one to simply start and stop enzyme inactivation and examine the response in the starting and recovering phase.[44] Currently, only the cyclophellitol ABPs/inhibitors labeling GBA are suitable for generating an LSD model, as all others label multiple enzyme targets. The galactose configured cyclophellitol aziridine ABPs[2] might be a potential lead in creating Krabbe disease animal model. Future studies might consider replace its BODIPY fluorophore with other hydrophobic moieties (such as biphenyl or adamantyl) to achieve better brain penetration. However, due to the much lower *in vitro* inhibitory potency, the *in vivo* efficacy and concomitant inhibition towards other enzymes in animals should be thoroughly examined.

h. Discovery of novel glycosidases, including cross species

The ability of gel-based ABPP to simultaneously visualize multiple labeled targets having similar reactivity but distinct molecular properties (molecular weight, pH profile, and glycosylation status) provides a convenient platform for discovering novel glycosidases in given samples. Labeled bands (with or without affinity-enrichment by biotin ABP) can be further excised, tryptic digested, and identified by proteomics using LC-MS/MS instruments. The same approach has been widely applied in the case of ABPs targeting proteases.[45, 46] The α-galactose configured cyclophellitol aziridine ABPs were also used to investigate reactive α-galactosidases in plants, which in *Nicotiana benthamiana* led to the identification of A1.1, a non-glycosylated enzyme stable across a broad pH range, and found to degrade the non-plant sphingolipids Gb3 and lysoGb3 elevated in Fabry patients.[25] Similarly, the broad substrate specificity of some glycosidase ABPs have been be used to "fish out" novel glycosidases in other organisms such as plants.[47, 48] In this thesis, the β-galactose configured cyclophellitol aziridine ABPs identified in the mouse kidney extracts two putative β-galactosidases in the GH35 family: the GLB1-like protein 1 and 2 (GLB1L and GLB1L2) (**Chapter 7**). This is substantiated by the occurrence of ABP-labeled bands in other mouse tissue extracts showing distinct molecular weights, glycosylation profile, and pH range than GLB1 and GALC. In the future, the biology and potential contribution of GLB1-like proteins to LSDs stemming from β-galactosidase deficiency (e.g. GM1 gangliosidosis, Morquio B syndrome, Krabbe disease) should be carefully examined.

i. Detecting interaction partners of glycosidases

For proteomics identification of glycosidases, biotin-tagged ABPs are commonly used to allow enrichment of labeled enzyme with streptavidin carriers. In most of the present biotinylated glycosidase ABPs, the biotin moiety is embedded in the catalytic pocket of the enzyme and binding to streptavidin is only feasible upon denaturation of the enzyme. An exception in this connection forms one GBA ABP that is a cyclophellitol-epoxide with an extended linker preceding the biotin moiety: likely when this ABP is bound to native enzyme the biotin is accessible to streptavidin, allowing pull-down.[49] This ABP additionally contains an internal BODIPY fluorophore within the linker, making detection of the affinity-enriched ABP-enzyme complex compatible with fluorescence scanning on wet slab-gels, and thus facilitates their subsequent in-gel tryptic digest and proteomic analysis. It is conceived that such extended, dual functional extensions may be installed on other configurational isomers of cyclophellitol and cyclophellitol aziridine cores. Such novel probes could be employed to pull down native enzymes and co-purify interacting protein partners. Combined with prior sub-cellular fractionation, this could provide valuable information on the protein interaction partners for each of the glycosidases at distinct cellular locations. Identification of interacting proteins is particularly warranted for the non-lysosomal glucosylceramidase (GBA2). The exact mode of its membrane association, its precise localization, and biological functions is still enigmatic. Poorly understood is the reported association between GBA2 mutation and with locomotor dysfunction whilst pharmacological inhibition of GBA2 activity in Miglustat-treated Gaucher disease and Niemann-Pick type C patients is well tolerated.[50, 51] In the case of the lysosomal GBA, protein interaction partners that should be identified with native ABP pull-downs as positive controls include LIMP2 and saposin C. Discovery of novel GBA interacting proteins might be of interest to generate new insights and clues in the pathophysiology of Gaucher and Parkinson disease.

j. Discovery of small molecule interactors of glycosidases

ABPs can in principle be also used to identify small molecules (including peptides) interacting with the catalytic pocket. Such interactors might be developed into pharmacological chaperones to stabilize (mutant) enzyme.[52] Recently, a high-throughput *in vitro* screening platform has been developed using the glycosidase ABPs coupled to fluorescence polarization (FluoPol) assay, in which the ratio of free ABP and bound ABP can be measured by reading out the extent of polarized fluorescence (where the bound ABP emits more polarized fluorescence).[36] Identified in this way have been several potent and specific inhibitors of GBA2,

and the same approach is presently extended to GBA, α-glucosidase, and α-mannosidase (unpublished work by Daniel Lahav, Dep. Bio-organic Synthesis, Leiden University). The screening can in principle also be performed in living cells, given the cell-permeability of some ABPs. Alternatively, the fluorescent quench glycosidase substrates developed by Vocadlo and co-workers provides another useful tools that is compatible with high-throughput microscopy-based screening in living cells.[17, 18] Nevertheless, no matter the type of initial screen, ABP can be used as a confirmation tool to assess the amount of active glycosidase molecules in different tissues of treated animals, where *in vivo* imaging techniques for glycosidase activity are not yet available. This can be performed with gel-based detection for tissue homogenates and microscopy-based detection in tissue sections, as earlier discussed. The existing FluoPol and gel-based ABPP platforms are presently used to identify selective inhibitors for various glycosidases (ongoing work Daniel Lahav, Dep. Bio-organic Synthesis, Leiden University). Finally, crystallography studies on the formed Michaelis complex and covalent complex of the ABP towards the target glycosidase may provide a structural-guided design of specific inhibitors towards a given glycosidase, for example to create a better inhibitor/ABP towards GH39 α-L-iduronidase (**Chapter 5**).[24]

k. Confirming LSD diagnosis

Diagnosis of LSDs is primary based on clinical symptoms. It is often combined with genotyping and demonstration of reduced enzyme activity in the case of enzymopathies.[53] Additionally used for confirmation of diagnosis is demonstration of accumulating metabolites or elevated biomarkers.[54] For several LSDs there are now (newborn) screening programs in some parts of the world,[55] largely based on genotyping followed by metabolite analysis and/or enzyme activity assays. The cyclophellitol and cyclophellitol aziridine ABPs can assist in some cases in the biochemical confirmation of the diagnosis of an LSD caused by deficiency of a corresponding glycosidase. Available ABPs include those labeling GBA (Gaucher disease),[1] GLA (Fabry disease),[4] FUCA1 (fucosidosis),[5] GALC (Krabbe disease),[2] GAA (Pompe disease) (**Chapter 4**),[56] GUSB (Sly syndrome) (**Chapter 4**),[57] MAN2B1 (α-mannosidosis) (**Chapter 6**), MANBA (β-mannosidosis)(**Chapter 6**), and GLB1 (GM1 gangliosidosis and Morquio B syndrome) (**Chapter 7**). Of note, ABP labeling can not only visualize lack of active enzyme molecules, but also abnormalities in post-translational processing (such as glycosylation and proteolytic cleavage). The latter abnormalities can also be detected with antibodies but these are

not always available. Proof of concept for the diagnostic potential of ABPs has been provided for Gaucher disease,[1] Fabry disease,[4] Krabbe disease,[2] and Pompe disease (**Chapter 4**)[56] with patient fibroblasts. Future application might exploit the use of ABPs in other patient materials, such as dried blood spots and urine, to assist in the confirmation of LSD diagnosis.

Finally, an attractive potential application for α-galactose configured ABPs to discuss warrants discussion, that is their use to identify female heterozygotes of Fabry disease that may develop an attenuated form of the disease. Fabry disease (α-galactosidase (GLA) deficiency) is an X-linked disorder, and as the result of random X-chromosome inactivation some cells of Fabry disease heterozygotes lack the normal enzyme whereas other cells are entirely normal depending on which copy of the X-chromosome was inactivated.[58] Reliable diagnosis of Fabry disease carriers by simple enzyme activity assays is virtually impossible[59]: only the measurement of isolated hair root cells, showing random X-chromosome inactivation) allows confirmation of the status of carriers.[60] The labeling of cellular GLA with appropriate ABP might offer a convenient solution for carrier detection. Labeled cells can be sorted by FACS: concomitant demonstration of cells with normal labeling intensity and cells lacking label would point to a heterozygote. A complication is that the presently available GLA ABP also labels the enzyme N-acetylgalactosaminidase (NAGA) that is not affected in Fabry disease. A more specific ABP for GLA should be ideally designed.

II. Current limitations and future prospects

The electrophilic trap (warhead)

The broad-specificity cyclophellitol aziridine ABPs offer mechanism-based labeling of in class glycosidases. This makes them excellent tools for simultaneous gel-based profiling and for discovery of novel glycosidases. However, the high reactivity of aziridine with nucleophiles, even beyond in-class targets, also makes them unattractive for microscopic imaging, as multiple enzymes may be labeled. Cyclophellitol epoxides allow highly specific labeling of GBA and GALC. Unfortunately, its α-galactose configurational isomers did not react with retaining exo-α-galacosidases.[61] The same was observed for exo-α-glucosidase (unpublished data, Dep. Bio-organic Synthesis, Leiden University). Future studies could investigate whether the C8-tagged cyclophellitol epoxide with other extensions would offer better target specificity. Also worth considering is to generate cyclophellitol ABPs containing other types of warhead that are less

reactive than aziridine but are still compatible with installing a linker/tag at the aglycon position.

The recognition element/linker

The β-glucosidases (GBA and GBA2) are prone to become labeled by higher concentrations of cyclophellitol aziridine ABPs with various configurations of the cyclophellitol core. This was observed for ABPs with α-glucose (**Chapter 4**),[56] β-glucuronic acid (**Chapter 4**),[57] β-mannose (**Chapter 6**), and β-galactose (**Chapter 7**) configurations. For the α-glucose configured ABPs, their labeling of β-glucosidases has been attributed to the half chair ($^4H_3$) conformational mimicry of cyclophellitol aziridine to β-glucosidases' substrates at the transition-state.[62] Another explanation for the aziridine ABPs' cross-reactivity with β-glucosidases might be the presence of the N-alkyl extension (with or without a fluorophore) at the aglycon side. The major substrate of both GBA and GBA2, glucosylceramide, contains such flexible alkyl chains at the aglycon side, and it has been noted that these are important for the hydrophobic interaction of the substrate glucosylceramide with GBA during catalysis.[63] Indeed, it was noted that the bare β-galactose configured cyclophellitol aziridine core is a much weaker β-glucosidase inhibitor when compared to its N-alkylated analogues. In the future, structural modification on the alkyl chain might prevent these ABP's concomitant labeling on β-glucosidases, such as increasing the hydrophilicity, or decreasing the structural flexibility.

Interestingly, a conceptually new compound, cyclophellitol with a cyclic sulfate electrophilic warhead, has recently been developed as a selective inhibitor of α-glucosidases over β-glucosidases by mimicking the $^4C_1$ chair conformation of the Michaelis complex observed in α-glucosidases.[62] In the future, other types of nucleophilic trap favoring the $^4C_1$ chair conformation and allowing the attachment of a reporter tag at the reducing-end aglycon side should be explored for the synthesis of selective ABPs.

The detection tag

Currently used fluorophores have their limitations. The BODIPY fluorophores, while providing excellent quantum yield,[64] are usually inferior to the Cy5 fluorophore when used in samples such as cell lysates or animal tissue homogenates, as these often contain background fluorescence at overlapping wavelength. They are also not always synthetically attainable, and sometimes they are unstable during purification and storage.[65] In addition, ABPs with the BODIPY fluorophores are less brain-penetrant, which may be due to the presence and action

of p-glycoprotein or multidrug resistance protein pumps that actively excrete BODIPY-containing molecules across the blood-brain barrier.[66, 67] The Cy5 fluorophore offers better signal-to-noise ratio, but on the other hand appears less soluble in aqueous solution, and is prone to label non-specifically, particularly in cells and zebrafish (observations Dep. Medical Biochemistry, Leiden University). Similarly, oligonucleotides modified with the positively-charged Cy3 or Cy5 dye, but not the non-charged Alexa fluorophores, have been found to aspecifically accumulate at the mitochondrial membrane.[68] All of these issues might underlie the poor labeling and inactivation of enzyme by Cy5 cyclophellitol ABP in zebrafish (**Chapter 3**). Taken together, a non-charged fluorophore with good solubility in aqueous solution would be ideal for visualizing purposes. Another way to improve sensitivity of ABP detection under fluorescence microscopy is to use a two-step labeling strategy. For example, cells can be firstly treated with ABPs containing a norbornene tag (**Chapter 7**), and after fixation these can be covalently attached with tetrazene-Cy5 fluorophores *via* inverse electron-demand Diels-Alder reaction (IEDDA, **Chapter 7**). In this manner, issues concerning Cy5's solubility and aspecificity in cells might be overcome. The tetrazine can in principle also be coupled to structures containing multiple fluorophores, thus allowing signal amplification. The ABPs can also be installed with a isotopic tag, allowing real-time imaging of labeled therapeutic glycosidases using techniques such as positron emission tomography (PET)[27–29, 31] or single-photon emission computed tomography (SPECT)[30] in living animals.

For affinity-enrichment using biotinylated ABPs, the described methods in this thesis only allow enrichment of denatured proteins. As mentioned above, this is due to the inaccessibility of the biotin tag to streptavidin resin, when the glycosidase is in its native conformation. Replacing the biotin moiety with the dual functional tag with an extended linker (described in an earlier section) might allow native pull-down of glycosidases in cell lysates or tissue homogenates. The conceived advantages of this setup include: (1) "cleaner" enrichment, as endogenous biotinylated proteins would not be enriched by the streptavidin resin when they are in the native conformation (unpublished data); (2) easier confirmation of the enrichment results. After pull-down, samples can be resolved by SDS-PAGE, and the labeled glycosidases can be easily detected by fluorescence scanning. Sample resolved by SDS-PAGE can be directly used for in-gel tryptic digestion for subsequent LC-MS/MS detection; (3) co-affinity enrichment of other associated proteins with the labeled glycosidase (discussed in an earlier section). The proposed ABPs should be obviously carefully examined regarding target specificity and cell-permeability.

III. Conclusion

Over the past years, carefully designed and characterized small molecules tailored for specific proteins, such as the herein described activity-based probes labeling retaining exo-glycosidases, have opened up new research possibilities in the traditional medical biochemistry field. In the future, further development of these chemical biology tools will no doubt continue to offer novel insights on the physiological and pathological roles of relevant enzymes, as well as facilitate the development of therapies and improvement of disease diagnosis, as exemplified in this thesis for the lysosomal storage disorders.

References

1 Witte MD, Kallemeijn WW, Aten J, Li KY, Strijland A, Donker-Koopman WE, van den Nieuwendijk AM, Bleijlevens B, Kramer G, Florea BI, Hooibrink B, Hollak CE, Ottenhoff R, Boot RG, van der Marel GA, Overkleeft HS & Aerts JM (2010) Ultrasensitive in situ visualization of active glucocerebrosidase molecules. *Nat Chem Biol* **6**, 907–913.

2 Marques AR, Willems LI, Herrera Moro D, Florea BI, Scheij S, Ottenhoff R, van Roomen CP, Verhoek M, Nelson JK, Kallemeijn WW, Biela-Banas A, Martin OR, Cachón-González MB, Kim NN, Cox TM, Boot RG, Overkleeft HS & Aerts JM (2017) A Specific Activity-Based Probe to Monitor Family GH59 Galactosylceramidase, the Enzyme Deficient in Krabbe Disease. *Chembiochem* **18**, 402–412.

3 Kallemeijn WW, Witte, MD, Voorn-Brouwer TM, Walvoort MT, Li KY, Codée JD, van der Marel GA, Boot RG, Overkleeft HS & Aerts, JM (2014) A sensitive gel-based method combining distinct cyclophellitol-based probes for the identification of acid/base residues in human retaining β-glucosidases. *J Biol Chem* **289**, 35351–35362.

4 Willems LI, Beenakker TJ, Murray B, Scheij S, Kallemeijn WW, Boot RG, Verhoek M, Donker-Koopman WE, Ferraz MJ, van Rijssel ER, Florea BI, Codée JD, van der Marel GA, Aerts JM & Overkleeft HS (2014) Potent and selective activity-based probes for GH27 human retaining α-galactosidases. *J Am Chem Soc* **136**, 11622–1625.

5 Jiang J, Kallemeijn WW, Wright DW, van den Nieuwendijk AMCH, Rohde VC, Folch EC, van den Elst H, Florea BI, Scheij S, Donker-Koopman WE, Verhoek M, Li N, Schürmann M, Mink D, Boot RG, Codée JDC, van der Marel GA, Davies GJ, Aerts JMFG & Overkleeft HS (2015) In vitro and in vivo comparative and competitive activity-based protein profiling of GH29 α-l-fucosidases. *Chem Sci* **6**, 2782–2789.

6 Jiang J, Beenakker TJ, Kallemeijn WW, van der Marel GA, van den Elst H, Codée JD, Aerts JM & Overkleeft HS (2015) Comparing Cyclophellitol N-Alkyl and N-Acyl Cyclophellitol Aziridines as Activity-Based Glycosidase Probes. *Chemistry* **21**, 10861–10869.

7 Herrera Moro Chao D, Kallemeijn WW, Marques AR, Orre M, Ottenhoff R, van Roomen C, Foppen E, Renner MC, Moeton M, van Eijk M, Boot RG, Kamphuis W, Hol EM, Aten J, Overkleeft HS, Kalsbeek A & Aerts JM (2015) Visualization of Active Glucocerebrosidase in Rodent Brain with High Spatial Resolution following In Situ Labeling with Fluorescent Activity Based Probes. *PLoS One* **10**, e0138107.

8 Kallemeijn WW, Scheij S, Hoogendoorn S, Witte MD, Herrera Moro Chao D, van Roomen CP, Ottenhoff R, Overkleeft HS, Boot RG & Aerts JM (2017) Investigations on therapeutic glucocerebrosidases through paired detection with fluorescent activity-based probes. *PLoS One* **12**, e0170268.

9 Kallemeijn WW, Witte MD, Voorn-Brouwer TM, Walvoort MT, Li KY, Codée JD, van der Marel GA, Boot RG, Overkleeft HS, & Aerts JM (2014) A sensitive gel-based method combining distinct cyclophellitol-based probes for the identification of acid/base residues in human retaining β-glucosidases. *J Biol Chem* **289**, 35351–35362

10 van Smeden J, Dijkhoff IM, Helder RWJ, Al-Khakany H, Boer DEC, Schreuder A, Kallemeijn WW, Absalah S, Overkleeft HS, Aerts JMFG & Bouwstra JA (2017) In situ visualization of glucocerebrosidase in human skin tissue: zymography versus activity-based probe labeling. *J Lipid Res* **58**, 2299–2309.

11 Danso M, Boiten W, van Drongelen V, Gmelig Meijling K, Gooris G, El Ghalbzouri A, Absalah S, Vreeken R, Kezic S, van Smeden J, Lavrijsen S & Bouwstra J (2017) Altered expression of epidermal lipid bio-synthesis enzymes in atopic dermatitis skin is accompanied by changes in stratum corneum lipid composition. *J Dermatol Sci* **88**, 57–66.

12 Nair S, Branagan AR, Liu J, Boddupalli CS, Mistry PK & Dhodapkar, MV (2016) Clonal Immunoglobulin against Lysolipids in the Origin of Myeloma. *N Engl J Med* **374**, 555–561.

13 Henriques A, Huebecker M, Blasco H, Keime C, Andres CR, Corcia P, Priestman DA, Platt FM, Spedding M & Loeffler JP (2017) Inhibition of β-Glucocerebrosidase Activity Preserves Motor Unit Integrity in a Mouse Model of Amyotrophic Lateral Sclerosis. *Sci Rep* **7**, 5235.

14 Tsuang D, Leverenz JB, Lopez OL, Hamilton RL, Bennett DA, Schneider JA, Buchman AS, Larson EB, Crane PK, Kaye JA Kramer P, Woltjer R, Kukull W, Nelson PT, Jicha GA, Neltner JH, Galasko D, Masliah E, Trojanowski JQ, Schellenberg GD, Yearout D, Huston H, Fritts-Penniman A, Mata IF, Wan JY, Edwards KL, Montine TJ & Zabetian CP (2012) GBA Mutations Increase Risk for Lewy Body Disease with and without Alzheimer Disease Pathology. *Neurology* **79**, 1944–1950.

15 Schapira AHV (2015) Glucocerebrosidase and Parkinson Disease: Recent Advances. *Mol Cell Neurosci* **66**, 37–42.

16 Kuo CL, van Meel E, Kytidou K, Kallemeijn WW, Witte M, Overkleeft HS, Artola ME & Aerts JM (2018) Activity-Based Probes for Glycosidases: Profiling and Other Applications. *Methods Enzymol* **598**, 217– 235.

17 Yadav AK, Shen DL, Shan X, He X, Kermode AR & Vocadlo DJ (2015) Fluorescence-quenched substrates for live cell imaging of human glucocerebrosidase activity. *J Am Chem Soc* **137**, 1181–1189.

18 Ashmus RA, Shen DL & Vocadlo DJ (2018) Fluorescence-Quenched Substrates for Quantitative Live Cell Imaging of Glucocerebrosidase Activity. *Methods Enzymol* **598**, 199–215.

19 Erickson AH, Ginns EI & Barranger JA (1985) Biosynthesis of the lysosomal enzyme glucocerebrosidase. *J Biol Chem* **260**, 14319–14324.

20 Jonsson LM, Murray GJ, Sorrell SH, Strijland A, Aerts JF, Ginns EI, Barranger JA, Tager JM & Schram AW (1987) Biosynthesis and maturation of glucocerebrosidase in Gaucher fibroblasts. *Eur J Biochem* **164**, 171–179.

21 Kuo CL, Kallemeijn WW, Lelieveld LT, Mirzaian M, Zoutendijk I, Vardi A, Futerman AH, Meijer AH, Spaink HP, Overkleeft HS, Aerts JMFG & Artola M. (2019) In Vivo Inactivation of Glycosidases by Conduritol B Epoxide and Cyclophellitol as Revealed by Activity-Based Protein Profiling. *FEBS J* **286**, 584–600.

22 D'Azzo A, Hoogeveen A, Reuser AJ, Robinson D & Galjaard H (1982) Molecular defect in combined beta-galactosidase and neuraminidase deficiency in man. *Proc Natl Acad Sci USA* **79**, 4535–4539.

23 van der Spoel A, Bonten E & d'Azzo A (2000) Processing of lysosomal beta-galactosidase. The C-terminal precursor fragment is an essential domain of the mature enzyme. *J Biol Chem* **275**, 10035–10040.

24 Artola M, Kuo CL, McMahon SA, Oehler V, Hansen T, van der Lienden M, He X, van den Elst H, Florea BI, Kermode AR, van der Marel GA, Gloster TM, Codée JDC, Overkleeft HS & Aerts JMFG (2018) New Irreversible α-l-Iduronidase Inhibitors and Activity-Based Probes. *Chemistry* **24**, 19081–19088.

25 Kytidou K, Beekwilder J, Artola M, van Meel E, Wilbers RHP, Moolenaar GF, Goosen N, Ferraz MJ, Katzy R, Voskamp P, Florea BI, Hokke CH, Overkleeft HS, Schots A, Bosch D, Pannu N & Aerts JMFG (2018) Nicotiana benthamiana α-galactosidase A1.1 can functionally complement human α-galactosidase A deficiency associated with Fabry disease. *J Biol Chem* **293**, 10042–10058.

26 van Meel E, Bos E, van der Lienden MJC, Overkleeft HS, van Kasteren SI, Koster AJ & Aerts JMFG (2019) Localization of Active Endogenous and Exogenous GBA by Correlative Light-Electron Microscopy in Human Fibroblasts. *Traffic* **20**, 346–356.

27 Ren G, Blum G, Verdoes M, Liu H, Syed S, Edgington LE, Gheysens O, Miao Z, Jiang H, Gambhir SS, Bogyo M & Cheng Z (2011) Non-invasive imaging of cysteine cathepsin activity in solid tumors using a 64Cu-labeled activity-based probe. *PLoS One* **6**, e28029.

28 Withana NP, Saito T, Ma X, Garland M, Liu C, Kosuge H, Amsallem M, Verdoes M, Ofori LO, Fischbein M, Arakawa M, Cheng Z, McConnell MV & Bogyo M (2016) Dual-Modality Activity-Based Probes as Molecular Imaging Agents for Vascular Inflammation. *J Nucl Med* **57**, 1583–1590.

29 Withana NP, Ma X, McGuire HM, Verdoes M, van der Linden WA, Ofori LO, Zhang R, Li H, Sanman LE, Wei K, Yao S, Wu P, Li F, Huang H, Xu Z, Wolters PJ, Rosen GD, Collard HR, Zhu Z, Cheng Z & Bogyo M (2016) Non-invasive Imaging of Idiopathic Pulmonary Fibrosis Using Cathepsin Protease Probes. *Sci Rep* **6**, 19755.

30 Vangestel C, Thomae D, Van Soom J, Ides J, Wyffels L, Pauwels P, Stroobants S, Van der Veken P, Magdolen V, Joossens J, Augustyns K & Staelens S (2016) Preclinical evaluation of [(111) In]MICA-401, an activity-based probe for SPECT imaging of in vivo uPA activity. *Contrast Media Mol Imaging* 11, 448–458.

31 Phenix CP, Rempel BP, Colobong K, Doudet DJ, Adam MJ, Clarke LA & Withers SG (2010) Imaging of enzyme replacement therapy using PET. *Proc Natl Acad Sci USA* **107**, 10842–10847.

32 Janssen APA, van der Vliet D, Bakker AT, Jiang M, Grimm SH, Campiani G, Butini S & van der Stelt M (2018) Development of a Multiplexed Activity-Based Protein Profiling Assay to Evaluate Activity of Endocannabinoid Hydrolase Inhibitors. *ACS Chem Biol* **13**, 2406–2413.

33 Tol MJ, van der Lienden MJC, Gabriel TL, Hagen JJ, Scheij S, Veenendaal T, Klumperman J, Donker-Koopman WE, Verhoeven AJ, Overkleeft H, Aerts JM, Argmann CA & van Eijk M (2018) HEPES activates a MiT/TFE-dependent lysosomal-autophagic gene network in cultured cells: A call for caution. *Autophagy* **14**, 437–449.

34 Van Esbroeck ACM, Janssen APA, Cognetta AB, Ogasawara D, Shpak G, Van Der Kroeg M, Kantae V, Baggelaar MP, De Vrij FMS, Deng H, Allarà M, Fezza F, Lin Z, Van Der Wel T, Soethoudt M, Mock ED, Den Dulk H, Baak IL, Florea BI, Hendriks G, De Petrocellis L, Overkleeft HS, Hankemeier T, De Zeeuw CI, Di Marzo V, Maccarrone M, Cravatt BF, Kushner SA & Van Der Stelt M (2017) Activity-based protein profiling reveals off-target proteins of the FAAH inhibitor BIA 10-2474. *Science* **356**, 1084–1087.

35 van Rooden EJ, Florea BI, Deng H, Baggelaar MP, van Esbroeck ACM, Zhou J, Overkleeft HS & van der Stelt M (2018) Mapping in vivo target interaction profiles of covalent inhibitors using chemical proteomics with label-free quantification. *Nat Protoc* **13**, 752–767.

36 Lahav D, Liu B, Van Den Berg RJBHN, Van Den Nieuwendijk AMCH, Wennekes T, Ghisaidoobe AT, Breen I, Ferraz MJ, Kuo CL, Wu L, Geurink PP, Ovaa H, Van Der Marel GA, Van Der Stelt M, Boot RG, Davies GJ, Aerts JMFG & Overkleeft HS (2017) A fluorescence polarization activity-based protein profiling assay in the discovery of potent, selective inhibitors for human nonlysosomal glucosylceramidase. *J Am Chem Soc* **139**, 14192–14197.

37 Artola M, Kuo CL, Lelieveld LT, Rowland RJ, van der Marel GA, Codée JDC, Boot RG, Davies GJ, Aerts JMFG & Overkleeft HS (2019) Functionalized Cyclophellitols Are Selective Glucocerebrosidase Inhibitors and Induce a Bona Fide Neuropathic Gaucher Model in Zebrafish. *J Am Chem Soc* **141**, 4214–4218.

38 Wennekes T, van den Berg RJ, Donker W, van der Marel GA, Strijland A, Aerts JM, Overkleeft HS (2007) Development of adamantan-1-yl-methoxy-functionalized 1-deoxynojirimycin derivatives as selective inhibitors of glucosylceramide metabolism in man. *J Org Chem* **72**, 1088–1097.

39 Barton NW, Furbish FS, Murray GJ, Garfield M & Brady RO (1990) Therapeutic response to intravenous infusions of glucocerebrosidase in a patient with Gaucher disease. *Proc Natl Acad Sci USA* **87**, 1913–1916.

40 Weiss K, Gonzalez A, Lopez G, Pedoeim L, Groden C & Sidransky E (2015) The clinical management of Type 2 Gaucher disease. *Mol Genet Metab* **114**, 110–122.

41 Farfel-Becker T, Vitner EB & Futerman AH (2011) Animal models for Gaucher disease research. *Dis Model Mech* **4**, 746–752.

42 Uemura N, Koike M, Ansai S, Kinoshita M, Ishikawa-Fujiwara T, Matsui H, Naruse K, Sakamoto N, Uchiyama Y, Todo T, Takeda S, Yamakado H & Takahashi R (2015) Viable neuronopathic Gaucher disease model in Medaka (Oryzias latipes) displays axonal accumulation of alpha-synuclein. *PLoS Genet* **11**, e1005065.

43 Keatinge M, Bui H, Menke A, Chen YC, Sokol AM, Bai Q, Ellett F, Da Costa M, Burke D, Gegg M, Trollope L, Payne T, McTighe A, Mortiboys H, de Jager S, Nuthall H, Kuo MS, Fleming A, Schapira AH, Renshaw SA, Highley JR, Chacinska A, Panula P, Burton EA, O'Neill MJ & Bandmann O (2015) Glucocerebrosidase 1 deficient Danio rerio mirror key pathological aspects of human Gaucher disease and provide evidence of early microglial activation preceding alpha-synuclein-independent neuronal cell death. *Hum Mol Genet* **24**, 6640–6652.

44 Vardi A, Zigdon H, Meshcheriakova A, Klein AD, Yaacobi C, Eilam R, Kenwood BM, Rahim AA, Massaro G, Merrill AH, Vitner EB & Futerman AH (2016) Delineating pathological pathways in a chemically induced mouse model of Gaucher disease. *J Pathol* **239**, 496–509.

45 Fonović M & Bogyo M (2008) Activity-based probes as a tool for functional proteomic analysis of proteases. *Expert Rev Proteomics* **5**, 721–30.

46 Cravatt BF, Wright AT & Kozarich JW (2008) Activity-based protein profiling: from enzyme chemistry to proteomic chemistry. *Annu Rev Biochem* **77**, 383–414.

47 Chandrasekar B, Colby T, Emran Khan Emon A, Jiang J, Hong TN, Villamor JG, Harzen A, Overkleeft HS & van der Hoorn RA (2104) Broad-range glycosidase activity profiling. *Mol Cell Proteomics* **13**, 2787–800.

48 Husaini AM, Morimoto K, Chandrasekar B, Kelly S, Kaschani F, Palmero D, Jiang J, Kaiser M, Ahrazem O, Overkleeft HS & van der Hoorn RAL (2018) Multiplex Fluorescent, Activity-Based Protein Profiling Identifies Active α-Glycosidases and Other Hydrolases in Plants. *Plant Physiol* **177**, 24–37.

49 Kallemeijn WW, Scheij S, Voorn-Brouwer TM, Witte MD, Verhoek M, Overkleeft HS, Boot RG & Aerts JM (2016) Endo-β-Glucosidase Tag Allows Dual Detection of Fusion Proteins by Fluorescent Mechanism-Based Probes and Activity Measurement. *Chembiochem* **17**, 1698–704.

50 Woeste MA & Wachten D (2018) The Enigmatic Role of GBA2 in Controlling Locomotor Function. *Front Mol Neurosci* **10**, 386.

51 Woeste MA, Stern S, Raju DN, Grahn E, Dittmann D, Gutbrod K, Dörmann P, Hansen JN, Schonauer S, Marx CE, Hamzeh H, Körschen HG, Aerts JMFG, Bönigk W, Endepols H, Sandhoff R, Geyer M, Berger TK, Bradke F & Wachten D (2019) Species-specific differences in nonlysosomal glucosylceramidase GBA2 function underlie locomotor dysfunction arising from loss-of-function mutations. *J Biol Chem* **294**, 3853–3871.

52 Boyd RE, Lee G, Rybczynski P, Benjamin ER, Khanna R, Wustman BA & Valenzano KJ (2013) Pharmacological chaperones as therapeutics for lysosomal storage diseases. *J Med Chem* **56**, 2705–25.

53 Filocamo M & Morrone A (2011) Lysosomal storage disorders: molecular basis and laboratory testing. *Hum*

Genomics **5**, 156–169.

54 Ferraz MJ, Kallemeijn WW, Mirzaian M, Herrera Moro D, Marques A, Wisse P, Boot RG, Willems LI, Overkleeft HS & Aerts JM (2014) Gaucher disease and Fabry disease: New markers and insights in pathophysiology for two distinct glycosphingolipidoses. *BBA Mol Cell Biol Lipids* **1841**, 811–825.

55 Schielen PCJI, Kemper EA & Gelb MH (2017) Newborn Screening for Lysosomal Storage Diseases: A Concise Review of the Literature on Screening Methods, Therapeutic Possibilities and Regional Programs. *Int J Neonatal Screen* **3**, pii: 6.

56 Jiang J, Kuo CL, Wu L, Franke C, Kallemeijn WW, Florea BI, Van Meel E, Van Der Marel GA, Codée JDC, Boot RG, Davies GJ, Overkleeft HS & Aerts JMFG (2016) Detection of active mammalian GH31 α-glucosidases in health and disease using in-class, broad-spectrum activity-based probes. *ACS Cent Sci* **2**, 351–358.

57 Wu L, Jiang J, Jin Y, Kallemeijn WW, Kuo C-L, Artola M, Dai W, van Elk C, van Eijk M, van der Marel GA, Codee JDC, Florea BI, Aerts JMFG, Overkleeft HS & Davies GJ (2017) Activity-based probes for functional interrogation of retaining β-glucuronidases. *Nat Chem Biol* **13**, 867–873.

58 Desnick RJ, Ioannou YA & Eng CM (2001) α galactosidase A deficiency: Fabry disease. In Scriver CR, Beaudet AL, Sly WS & Valle D (Eds) *The metabolic and molecular basis of inherited disease*. (pp 3773–3774). New York, NY: MacGraw-Hill.

59 Linthorst GE, Poorthuis BJ & Hollak CE (2008) Enzyme activity for determination of presence of Fabry disease in women results in 40% false-negative results. *J Am Coll Cardiol* **51**, 2082–2083.

60 Ejiofor A, Robinson D, Wise D, Hamers MN & Tager JM (1978) Hair root analysis in heterozygotes for Fabry's disease. *Adv Exp Med Biol* **101**, 719–725.

61 Willems LA (2014) *Direct and two-step activity-based profiling of proteases and glycosidases* (Doctoral dissertation). Retrieved from Leiden University Repository.

62 Artola M, Wu L, Ferraz MJ, Kuo CL, Raich L, Breen IZ, Offen WA, Codée JDC, van der Marel GA, Rovira C, Aerts JMFG, Davies GJ & Overkleeft HS (2017) 1,6-Cyclophellitol cyclosulfates: a new class of irreversible glycosidase inhibitor. *ACS Cent Sci* **3**, 784–793.

63 Nakagome I, Kato A, Yamaotsu N, Yoshida T, Ozawa SI, Adachi I & Hirono S (2018) Design of a new α- 1-C-alkyl-DAB derivative acting as a pharmacological chaperone for β-glucocerebrosidase using ligand docking and molecular dynamics simulation. *Molecules* **23**, pii: E2683.

64 Verdoes M, Hillaert U, Florea BI, Sae-Heng M, Risseeuw MD, Filippov DV, van der Marel GA & Overkleeft HS (2007) Acetylene functionalized BODIPY dyes and their application in the synthesis of activity based proteasome probes. *Bioorg Med Chem Lett* **17**, 6169–6171.

65 Willems LI, Beenakker TJM, Murray B, Gagestein B, van den Elst H, van Rijssel ER, Codée JDC, Kallemeijn WW, Aerts JMFG, van der Marel GA, & Overkleeft HS (2014) Synthesis of α- and β-galactopyranose-configured isomers of cyclophellitol and cyclophellitol aziridine. *Eur J Org Chem* **2014**, 6044–6056.

66 Crivellato E, Candussio L, Rosati AM, Bartoli-Klugmann F, Mallardi F & Decorti G (2002) The fluorescent probe Bodipy-FL-verapamil is a substrate for both P-glycoprotein and multidrug resistance-related protein (MRP)-1. *J Histochem Cytochem* **50**, 731–734.

67 Strouse JJ, Ivnitski-Steele I, Waller A, Young SM, Perez D, Evangelisti AM, Ursu O, Bologa CG, Carter MB, Salas VM, Tegos G, Larson RS, Oprea TI, Edwards BS & Sklar LA (2013) Fluorescent substrates for flow cytometric evaluation of efflux inhibition in ABCB1, ABCC1, and ABCG2 transporters. *Anal Biochem* **437**, 77–87.

68 Rhee WJ & Bao G (2010) Slow non-specific accumulation of 2'-deoxy and 2'-O-methyl oligonucleotide probes at mitochondria in live cells. *Nucleic Acids Res* **38**, e109.

Summary

This thesis describes the characterization and application of activity-based probes (ABPs) labeling lysosomal glycosidases in mechanism-based manner. Lysosomal glycosidases are acid hydrolases that fragment glycoconjugates in lysosomes. Several inherited diseases in man concern deficiencies of lysosomal glycosidase. The corresponding lysosomal storage disorders (LSDs) show characteristic lysosomal accumulation of undegraded substrates. The **General introduction** provides an overview of lysosomal glycosidases and the associated LSDs. Described are the present therapeutic options and diagnostic methods for these LSDs. Attention is paid to the current knowledge on the life cycle of lysosomal glycosidases and their catalytic mechanisms. Understanding on the latter has allowed generation of activity-based probes (ABPs) irreversibly labeling lysosomal glycosidases in mechanism-based manner. Present ABPs for lysosomal glycosidases are based on cyclophellitol and cyclophellitol aziridine scaffolds. This thesis describes how these probes can be employed as versatile research tools to study the disease-relevant enzymes.

Chapter 1 describes developed protocols for gel-based and microscopy-based detection of lysosomal glycosidases, in particular the β-glucosidase GBA, labeled with ABPs *in vitro* and *in vivo*. Cyclophellitol-type ABPs with an electrophilic epoxide group and a BODIPY fluorophore allow specific and sensitive visualization of active GBA molecules following their irreversible labeling. Labeling can be performed *in vitro* using enzyme preparations or cell and tissue extracts. Due to the amphiphilic nature of the ABP *in situ* labeling and detection of active GBA in cells and organisms is also feasible. The described protocols for ABP labeling of GBA can be extended to other cyclophellitol- and cyclophellitol aziridine-based ABPs labeling distinct glycosidases.

Chapter 2 investigates the *in vivo* target engagement of conduritol B epoxide (CBE) and cyclophellitol (CP) in cells and animals using a gel-based competitive activity-based protein profiling (cABPP) approach. CBE and CP are mechanism-based irreversible GBA inhibitors and have been used to generate Gaucher disease (GD) models in cells and animals. These inhibitors

may have additional *in vivo* glycosidase targets beyond GBA, depending on the applied dose and treatment duration. For the candidate off-target glycosidases a suite of glycosidase ABPs is available which allow the investigation of potential interaction of CBE and CP with these enzymes. Investigation of CBE-treated cells and zebrafish larvae revealed that at higher concentrations CBE also interacts *in vivo* with non-lysosomal glucosylceramidase (GBA2) and lysosomal α-glucosidase (GAA). A tight, but acceptable window for selective inhibition of GBA in the brain of mice is observed. On the other hand, CP was found to inactivate with equal affinity GBA and GBA2 and therefore is not suitable to generate genuine GD-like models.

Chapter 3 reports on novel irreversible inhibitors of GBA that are superior in selectivity towards the enzyme. This class of inhibitors carries a bulky hydrophobic substituent at C8, such as adamantyl and biphenyl-moieties. By gel-based ABPP experiments, it is shown that the novel compounds are far more potent and selective irreversible inhibitors of GBA as compared to CBE *in vitro*, in cultured cells, and in zebrafish larvae. The novel inhibitor with an adamantyl moiety was found to be brain-permeable and to allow generation of an authentic neuropathic GD model in zebrafish by exposure of animals to the compounds via food uptake.

Crucial for optimal application of novel glycosidase ABPs in LSD research is a careful biochemical characterization, which includes mechanism of action, inhibition kinetics, *in vitro* and *in vivo* potency, and identification of off-target glycosidases.

Chapter 4 describes the characterization of newly developed ABPs for the lysosomal glycosidases α-glucosidase (GAA) and β-glucuronidase (GUSB), enzymes deficient in Pompe disease and mucopolysaccharidosis type VII (MPSVII; Sly Syndrome), respectively. The ABPs label their respective target enzymes (GAA and GUSB) potently and in mechanism-based manner. The α-glucose configured ABP labels not only the lysosomal α-glucosidase GAA, but also the ER α-glucosidase II (GANAB). It was observed that at higher concentration GBA is a major off-target of both ABPs. By altering the labeling pH, and by pre-incubating samples with a specific GBA inhibitor, GAA and GUSB can be specifically visualized with the corresponding ABPs *in vitro* as well as in lysates of cells treated *in situ* with the ABPs. The diagnostic potential of the GAA ABP for Pompe disease is demonstrated with cultured fibroblasts from patients.

Chapter 5 documents the development and characterization of α-L-iduronide configured cyclophellitol aziridine ABPs for human α-L-iduronidase, whose inherited deficiency leads to

mucopolysaccharidosis type I (MPS I, Hurler syndrome). The new ABPs are shown to react irreversibly in mechanism-based manner with human recombinant α-L-iduronidase (rIDUA, Aldurazyme®). They exhibit lower potency and slower inhibitory kinetics compared to previous ABPs designed for other retaining glycosidases, possibly due to conformational constraints. The relative low affinity of the present ABP makes detection of endogenous IDUA in cells and tissues challenging. The fluorescent Cy5 ABP can efficiently label therapeutic iduronidase *ex vivo* and the enzyme's delivery to lysosomes—as intended in enzyme replacement therapy (ERT)—was successfully monitored by fluorescence microscopy of MPS I patient fibroblasts.

Chapter 6 reports on the biochemical evaluation of α- and β-mannose configured cyclophellitol aziridine ABPs for the retaining exo-α-mannosidases (Glycoside Hydrolase (GH) family 38) and exo-β-mannosidase (GH family 2)—enzymes implicated in diverse pathologies including cancer and LSDs. The ABPs label their respective target mannosidases in mechanism-based manner and are micromolar inhibitors. All five GH38 α-mannosidases are labeled by α-mannose configured ABPs in mouse testis homogenates. Each enzyme was labeled in lysates of cells upon overexpression. The unique molecular weight and pH optimum of each mannosidase allows their simultaneous activity-based profiling in complex biological samples, and should assist future screening for small molecule inhibitors/activators of these highly relevant enzymes.

β-Mannose configured ABPs label the GH2 β-mannosidase (MANBA) in mouse kidney extracts. Co-labeling of GBA occurs that can be prevented by pre-incubating samples with a GBA inhibitor.

Chapter 7 provides the biochemical characterization of β-galactose configured cyclophellitol aziridine ABPs, intended to label lysosomal β-galactosidase (GLB1) and galactocerebrosidase (GALC). The deficiency of GLB1 causes GM1 gangliosidosis (and Morquio B syndrome) whilst deficiency of GALC leads to Krabbe disease. The examined ABPs inhibit and label in mechanism-based manner recombinant and endogenous GLB1 and GALC. GLB1-like protein 1 and 2 (GLB1L and GLB1L2) are additionally labeled, two putative β-galactosidases with poorly understood physiological roles. The β-glucosidases GBA and GBA2 are major off-targets of β-galactose configured cyclophellitol aziridine ABPs. These new ABPs can assist in future fundamental and applied research on human β-galactosidases.

The thesis is concluded with a discussion summarizing the most important findings and highlighting the current and future applications of cyclophellitol and cyclophellitol aziridine-based ABPs in LSD research.

Appendices

- Summary in Dutch (Nederlandse Samenvatting)
- Summary in Chinese (中文摘要)
- List of Publications
- *Curriculum vitae*

Nederlandse Samenvatting

Dit proefschrift beschrijft de karakterisering en de toepassing van activity-based probes (ABPs) die lysosomale glycosidases labelen op een mechanism-based wijze. Lysosomale glycosidases zijn zure hydrolases die glycoconjugaten fragmenteren in lysosomen. Verschillende erfelijke stofwisselingsziekten worden veroorzaakt door de deficiëntie in een lysosomaal glycosidase. De bijbehorende lysosomale stapelingsziekten (LSDs) vertonen een karakteristieke lysosomale stapeling van niet afgebroken substraten. De algemene inleiding biedt een overzicht betreffende lysosomale glycosidases en hiermee verbonden LSDs. Beschreven worden huidige behandelingsmogelijkheden en diagnostiek. Aandacht wordt besteed aan de kennis over de levenscyclus van lysosomale glycosidases en hun katalytische mechanisme. Begrip van dat laatste heeft mogelijk gemaakt het ontwikkelen van activity-based probes (ABPs), verbindingen die selectief en irreversibel lysosomale glycosidases labelen in mechanism-based manner. De huidige ABPs voor lysosomale glycosidases zijn gebaseerd op cyclophellitol en cyclophellitol aziridine scaffolds. Dit proefschrift beschrijft hoe deze probes worden benut in het onderzoek aan ziekte-relevante enzymen.

Hoofdstuk 1 beschrijft de ontwikkelde protocollen voor detectie in gel en met microscopie van lysosomale glycosidases, in het bijzonder het β-glucosidase GBA, na zijn labeling *in vitro* en *in vivo* met een fluorescent ABP. Cyclophellitol-type ABPs met een electrofiel epoxide en een BODIPY-fluorofoor visualiseren zeer specifiek en sensitief actieve GBA-moleculen middels irreversibele labeling hiervan. Labeling kan worden uitgevoerd *in vitro* met enzympreparaten ofwel cel- en weefsel-extracten. Vanwege de amfifiele aard van de gebruikte ABPs is *in situ* labeling en detectie van actieve GBA-moleculen mogelijk in cellen en organismen. De beschreven protocollen voor ABP labeling van GBA kunnen verbreed worden naar andere cyclophellitol- en cyclophellitol aziridine-based ABPs die andersoortige glycosidases labelen.

Hoofdstuk 2 betreft de bestudering van *in vivo* target engagement van conduritol B-epoxide

(CBE) en cyclophellitol (CP) in cellen en proefdieren. Gebruikt is gemaakt hierbij van een gel-based 'competitive activity-based protein profiling' (cABPP) benadering. CBE en CP zijn mechanism-based irreversibele GBA remmers die volop gebruikt worden voor het genereren van modellen van de ziekte van Gaucher in cellen en proefdieren. Deze remmers kunnen echter *in vivo* additionele glycosidases naast GBA remmen, afhankelijk van gehanteerde dosis en incubatietijd. Voor kandidaat off-target glycosidases is een scala aan glycosidase ABPs beschikbaar waarmee de mogelijke interactie van CBE en CP hiermee kan worden aangetoond. Bestudering van CBE-behandelde cellen en zebravis larven liet zien dat bij hogere concentraties de remmer CBE *in vivo* ook reageert met niet-lysosomaal glucosylceramidase (GBA2) en lysosomaal α-glucosidase (GAA). Een nauw, maar acceptabel, window voor selectieve remming van GBA in het brein van de muis werd waargenomen. Anderzijds, CP bleek met eenzelfde affiniteit GBA en GBA2 te inactiveren en is derhalve ongeschikt voor het genereren van echte modellen van de ziekte van Gaucher.

Hoofdstuk 3 beschrijft nieuwe irreversibele remmers van GBA met een superieure selectiviteit. Deze klasse verbindingen bevat een relatief grote hydrofobe substituent op C8, zoals adamantyl en bifenyl-groepen. Aangetoond kon worden met gel-based ABPP experimenten dat de nieuwe verbindingen veel krachtigere en selectievere irreversibele remmers zijn van GBA dan CBE, zowel *in vitro*, in gekweekte cellen als in zebravis larven. De remmer met een adamantyl groep bleek goed brein-permeabel en maakt het mogelijk om een authentiek model de voor neuropathische vorm van de ziekte van Gaucher te genereren in de zebravis door een blootstelling van de dieren aan de verbinding gedrenkt in het voedsel.

Cruciaal voor het optimale gebruik van nieuwe glycosidase ABPs in LSD onderzoek is een grondige biochemische karakterisering van hun eigenschappen, behelzend onderzoek aan het werkingsmechanisme, bepaling van kinetiek van remming, potentie *in vitro* en *in vivo*, en identificatie van off-target glycosidases.

Hoofstuk 4 beschrijft de karakterisering van nieuw ontwikkelde ABPs voor de lysosomale glycosidases α-glucosidase (GAA) en β-glucuronidase (GUSB), enzymen die deficiënt zijn bij de ziekte van Pompe en mucopolysaccharidosis type VII (MPSVII; Sly Syndroom), respectievelijk. Deze ABPs labelen hun respectievelijke target enzymen (GAA en GUSB) met hoge affiniteit en op mechanism-based wijze. Het α-glucose geconfigureerde ABP labelt niet alleen het lysosomale α-glucosidase GAA, maar ook het ER α-glucosidase II (GANAB). Waargenomen is dat bij een

hogere concentratie van het ABP het β-glucosidase GBA een off-target is. Door het veranderen van de pH tijdens de labeling, en het pre-incuberen van monsters met een specifieke GBA remmer, kan GAA en GUSB specifiek gevisualiseerd worden met ABPs, zowel *in vitro* als in lysaten van cellen die in situ behandeld zijn met de ABPs. De potentieel diagnostische waarde van het GAA ABP voor de biochemische bevestiging van de diagnose van de ziekte van Pompe is aangetoond met gekweekte huidfibroblasten van patiënten.

Hoofdstuk 5 documenteert α-L-iduronide geconfigureerde cyclophellitol aziridine ABPs voor human α-L-iduronidase, waarvan een erfelijke deficiëntie leidt tot mucopolysaccharidosis type I (MPS I, Hurler syndroom). De nieuwe ABPs blijken irreversibel te reageren op mechanism-based wijze met recombinant α-L-iduronidase (rIDUA, Aldurazyme®). De probes vertonen een relatief lage affiniteit van remming in vergelijking tot eerdere ABPs voor andere glycosidases, mogelijk vanwege conformatie beperkingen. The relatief lage affiniteit van de huidige ABP verhindert de detectie van endogeen IDUA in cellen en weefsels. Omdat met de fluorescent Cy5 ABP therapeutisch iduronidase *ex vivo* goed gelabeld kan worden, bleek het mogelijk de afvoer van dit enzym naar lysosomen van fibroblasten van MPS I patiënten goed te vervolgen. Een goede delivery naar lysosomen werd waargenomen, het beoogde doel van enzyme replacement therapy (ERT).

Hoofdstuk 6 rapporteert de biochemische evaluatie van α- en β-mannose geconfigureerde cyclophellitol aziridine ABPs voor retaining exo-α-mannosidases (Glycoside Hydrolase (GH) familie 38) en exo-β-mannosidase (GH familie 2)—enzyme geïmpliceerd in verschillende ziekten zoals kanker en LSDs. De ABPs labelen mannosidases op mechanism-based wijze en zijn micromolaire remmers. Alle vijf GH38 α-mannosidases aanwezig in homogenaten van muis-testis worden gelabeld door de α-mannose geconfigureerde ABPs. Elk van de vijf enzyme werd gelabeld in lysaten van cellen waarin het tot over-expressie was gebracht. Het unieke molecuul gewicht en pH optimum van elk mannosidase staat een simultane activity-based profiling toe in complexe biologische monsters. Die nieuwe probes kunnen van grote dienst zijn bij toekomstige screening voor small molecule inhibitors/activators van deze zeer relevante mannosidases.

β-Mannose geconfigureerde ABPs labelen het GH2 β-mannosidase (MANBA) in muis-nier extracten. Een co-labeling van GBA treedt op met de probes die voorkomen kan worden door pre-incubatie van monsters met een GBA-remmer.

Hoofdstuk 7 biedt een overzicht van de biochemische karakterisering β-galactose geconfigureerde cyclophellitol aziridine ABPs, beoogd om de lysosomale β-galactosidase (GLB1) en galactocerebrosidase (GALC) te labelen. Deficiëntie van GLB1 veroorzaakt GM1 gangliosidosis (en Morquio B syndroom) terwijl deficiëntie van GALC resulteert in de ziekte van Krabbe. De bestudeerde ABPs remmen en labelen in mechanism-based manner recombinant en endogeen GLB1 en GALC. Eveneens gelabeld worden GLB1-like protein 1 en 2 (GLB1L en GLB1L2), twee putatieve β-galactosidases met een grotendeels onbekende fysiologische rol. De β-glucosidases GBA en GBA2 zijn majeure off-targets van β-galactose geconfigureerde cyclophellitol aziridine ABPs. De nieuwe ABPs kunnen toekomstig fundamenteel en toegepast onderzoek aan humane β-galactosidases ondersteunen.

Dit proefschrift eindigt met een beschrijving van de voornaamste resultaten en bevindingen. Tot slot worden de huidige en toekomstige toepassingen voor cyclophellitol en cyclophellitol aziridine-based ABPs in het LSD onderzoek bediscussieerd.

活性探針對糖苷水解酶在生醫研究中的應用

中文摘要

本論文主要研究能依反應機制標記溶小體糖苷水解酶(lysosomal glycosidase)的活性探針（activity-based probes, ABPs）之特性分析及應用。溶小體糖苷水解酶是酸性水解酶，在溶小體中裂解各種糖綴合物(glycoconjugates)。在人體中，有多種遺傳性疾病涉及這些水解酶的缺陷，而其導致的溶小體儲積症（lysosomal storage disorder, LSD）則會顯示在溶小體內積累了特定的未降解底物。**緒論**綜述了溶小體糖苷水解酶和相關的 LSD，以及這些 LSD 現有的治療選項和診斷方法。文中更整理了關於溶小體糖苷水解酶的生命週期及其催化機制的最新知識；對後者的知識尤其使得能固定標記溶小體糖苷水解酶的活性探針（ABPs）因此得以被開發及合成。目前，環殼醇(cyclophellitol)和環殼醇氮丙啶(cyclophellitol aziridine)為建構這些 ABP 的骨架。本論文闡述了在研究這些與疾病相關的水解酶上如何來多元應用這些 ABP。

第 1 章描述了利用凝膠電泳和螢光顯微鏡，在活體外或活體內用 ABP 標記溶小體糖苷水解酶(特別是 β-葡萄糖苷酶(GBA))的實驗步驟。這些環殼醇型 ABP 具有親電環氧基團和 BODIPY 螢光團，能專一及靈敏地顯示被固定標記的活性 GBA 分子。標記能在活體外(*in vitro*)進行，樣本包括酶製劑、細胞提取物、和組織提取物。而由於 ABP 的兩親性質，在細胞和生物體中來原位標記(*in situ* labeling)和檢測活性 GBA 也均可行。另外，本章 ABP 標記 GBA 的實驗步驟還可以進一步擴展到其他環殼醇和六元環多醇氮丙啶(configurational isomers of cyclophellitol aziridine)的 ABP，來標記其他不同類型的糖苷水解酶。

第 2 章使用了凝膠電泳來進行競爭活性蛋白質分析（cABPP），研究了脫水肌醇（CBE）和環殼醇（cyclophellitol, CP）在細胞和動物體內的目標接合作用。CBE 和 CP 是 GBA 依反應機制型的不可逆抑制劑，已被用於在細胞和動物中產生高雪氏症（Gaucher disease, GD）疾病模型。根據施用的劑量和持續時間，這兩種抑制劑在活體內(*in vivo*)則可能有 GBA 以外的糖苷水解酶接合物。對於這些可能的額外接合物，一系列的 ABP 已被開發，並能被用於研究這些酶與 CBE 和 CP 的潛在交互作用。本研究顯示，在對被施以 CBE 的細胞和斑馬魚幼蟲中，較高濃度的 CBE 在體內還會與非溶小體葡糖神經酰胺酶（GBA2）和溶小體 α-葡萄糖苷酶（GAA）交互作用。在小鼠腦中，CBE 對

於 GBA 則有窄小但卻足夠的選擇性抑制窗口。另一方面，本研究還發現 CP 對於 GBA 和 GBA2 的親和力相同，因此其不適合用於產生真正的 GD 相關疾病模型。

第 3 章介紹了對於 GBA 的新型不可逆抑制劑；其對 GBA 的選擇性優異。這類抑制劑在 8 號碳帶有大體積的疏水取代基，包括金剛烷基(adamantyl)和聯苯基(biphenyl)。藉由凝膠電泳 cABPP 實驗，新化合物顯示了與 CBE 相比，其在活體外、培養細胞、和斑馬魚幼蟲中均是更有效和更有選擇性的 GBA 不可逆抑制劑。本研究並發現通過食物攝取，具有金剛烷基的新型抑制劑可滲透至斑馬魚腦部，而這允許了其被用於斑馬魚中產生真正的神經性 GD 疾病模型。

在 LSD 研究中，詳細的生物化學特性分析是來如何最佳應用新型糖苷水解酶 ABP 的關鍵。這些分析包括了 ABP 的作用機制，抑制動力學，活體外和活體內效力，以及非目標糖苷水解酶的鑑定。

第 4 章描述了對新開發的、針對溶小體糖苷水解酶 α-葡萄糖苷酶（GAA）和 β-葡萄醣醛酸酶（GUSB）的 ABP 的特性分析。這兩種水解酶分別在龐貝氏症(Pome disease)和粘多醣貯積症第七型（MPSVII；史萊式症(Sly syndrome)）中缺陷。本研究發現這些 ABP 能依酶反應機制來有效地標記它們各自的目標酶（GAA 和 GUSB）。而 α-葡萄糖構型的 ABP 不僅標記了溶小體 α-葡糖苷酶 GAA，還標記了內質網 α-葡糖苷酶 II（GANAB）。觀察顯示，在較高濃度下，兩種 ABP 的主要非目標酶為 GBA。而通過改變標記時的酸鹼值，和通過使用特定 GBA 抑制劑來預先混入樣品，相應的 ABP 則可用於活體外(包括用 ABP 原位標記的細胞之裂解液)來選擇性地檢測 GAA 和 GUSB。另外，利用體外培養的患者纖維母細胞(fibroblasts)，針對 GAA 的 ABP 顯示了其對龐貝氏症有診斷潛力。

第 5 章詳載了 α-L-艾杜糖苷構型之六元環多醇氮丙啶 ABP 的開發和特性分析。此類 ABP 為針對人類 α-L-艾杜醣醛酸酶而開發，而此酶的遺傳缺陷能導致粘多醣貯積症第一型（MPS I，賀勒式症(Hurler syndrome)）。此新類的 ABP 顯示它們能依酶反應機制與人類重組 α-L-艾杜醣醛酸酶（rIDUA，Aldurazyme®艾德酶）固定接合。不過與先前為其他保留型糖苷水解酶設計的 ABP 相比，它們表現出較低的效力和更慢的抑制動力學—而構象限制為可能的原因。本類 ABP 的相對低親和力使得檢測細胞和組織中的內源性艾杜醣醛酸酶具有挑戰性。不過 Cy5 螢光的 ABP 則可以有效地被用於離體標記用於治療用途的艾杜醣醛酸酶。通過螢光顯微鏡，於 MPS I 患者纖維母細胞中這些酶被成功的監測到被遞送至溶小體—正如酶替代療法（ERT）中所期盼的。

第 6 章介紹了α-和β-甘露糖構型的六元環多醇氮丙啶 ABP 對保留型外切-α-甘露糖苷酶（糖苷水解酶（GH）家族 38）和外切-β-甘露糖苷酶（GH 家族 2）的生化分析。這些酶與不同的疾病相關，包括癌症和 LSD。而甘露糖構型的 ABP 則顯示能以依酶反應機制的方式標記它們各自的目標甘露糖苷酶，並且是微摩爾抑制劑。於小鼠睾丸勻漿中，α-甘露糖構型的 ABP 標記了所有五種 GH38 α-甘露糖苷酶。在過表達時，每種酶在細胞裂解液中也均被此類 ABP 標記。而在復雜的生物樣品中，每種甘露糖苷酶的獨特分子量和最適酸鹼值則允許了它們能同時被進行活性表達譜分析(ABPP)。此結果應有助於將來篩選針對這些和生理及疾病高度相關的酶之小分子抑制劑/活化劑。

β-甘露糖構型的 ABP 則標記了小鼠腎提取物中的 GH2 β-甘露糖苷酶（MANBA）。通過用 GBA 抑制劑預先混入樣品，則可被避免 GBA 被此 ABP 共標記。

第 7 章提供了對β-半乳糖構型的六元環多醇氮丙啶 ABP 的生化特性分析，這些 ABP 旨在用於標記溶酶體β-半乳糖苷酶（GLB1）和半乳糖腦苷脂酶（GALC）。GLB1 的缺陷能導致 GM1 神經節苷脂儲積症（和莫奎歐氏症(Morquio syndrome)B 型），而 GALC 缺陷導致克拉伯病(Krabbe Disease)。據檢測，這些 ABP 是以依酶反應機制的方式抑制和標記重組或內源性的 GLB1 和 GALC。另外，這些 ABP 也同時標記了 GLB1 樣蛋白 1 和 2（GLB1L 和 GLB1L2）。這兩者目前均為假定的β-半乳糖苷酶，而對其生理作用目前則所知甚少。β-葡萄糖苷酶 GBA 和 GBA2 是β-半乳糖購型的六元環多醇氮丙啶 ABP 的主要非目標物。這些新的 ABP 可協助未來對人類β-半乳糖苷酶的基礎和應用研究。

最後，本論文於**討論**部分總結了最重要的發現，並著重探討了在 LSD 研究中帶有環殼醇和六元環多醇氮丙啶結構的 ABP 當前及未來的應用。

APPENDICES

List of Publications

1. Artola M,* **Kuo CL**,* Lelieveld LT, Rowland RJ, van der Marel GA, Codée JDC, Boot RG, Davies GJ, Aerts JMFG & Overkleeft HS (2019) Functionalized cyclophellitols are selective glucocerebrosidase inhibitors and induce a bona fide neuropathic Gaucher model in zebrafish. *J Am Chem Soc* 141, 4214–4218.

2. **Kuo CL**, Kallemeijn WW, Lelieveld LT, Mirzaian M, Zoutendijk I, Vardi A, Futerman AH, Meijer AH, Spaink HP, Overkleeft HS, Aerts JMFG & Artola M (2019) In vivo inactivation of glycosidases by conduritol B epoxide and cyclophellitol as revealed by activity-based protein profiling. *FEBS J* 286, 584–600.

3. Artola M,* **Kuo CL**,* McMahon SA, Oehler V, Hansen T, van der Lienden M, He X, van den Elst H, Florea BI, Kermode AR, van der Marel GA, Gloster TM, Codée JDC, Overkleeft HS & Aerts JMFG (2018) New irreversible α-l-iduronidase inhibitors and activity-based probes. *Chemistry* 24, 19081–19088.

4. **Kuo CL**, van Meel E, Kytidou K, Kallemeijn WW, Witte M, Overkleeft HS, Artola ME & Aerts JM (2018) Activity-based probes for glycosidases: profiling and other applications. *Methods Enzymol* 598, 217–235.

5. Schröder SP, van de Sande JW, Kallemeijn WW, **Kuo CL**, Artola M, van Rooden EJ, Jiang J, Beenakker TJM, Florea BI, Offen WA, Davies GJ, Minnaard AJ, Aerts JMFG, Codée JDC, van der Marel GA & Overkleeft HS (2017) Towards broad spectrum activity-based glycosidase probes: synthesis and evaluation of deoxygenated cyclophellitol aziridines. *Chem Commun (Camb)*. 2017, 53, 12528–12531.

6. Aerts JMFG, Guimaraes Da Lomba Ferraz MJ, Mirzaian M, Gaspar P, Oussoren SV, Wisse P, **Kuo CL**, Lelieveld LT, Kytidou K, Hazeu MD, Boer DEC, Meijer R, Lienden MJC van der, Herrera D, Gabriel TL, Aten J, Overkleeft HS, Eijk MC van, Boot RG & Marques ARA (2017) Lysosomal storage diseases. for better or worse: adapting to defective lysosomal glycosphingolipid breakdown. *eLS*, 1–13.

7. Lahav D, Liu B, van den Berg RJBHN, van den Nieuwendijk AMCH, Wennekes T, Ghisaidoobe AT, Breen I, Ferraz MJ, **Kuo CL**, Wu L, Geurink PP, Ovaa H, van der Marel GA, van der Stelt M, Boot RG, Davies GJ, Aerts JMFG & Overkleeft HS (2017) A fluorescence polarization activity-based protein profiling assay in the discovery of potent, selective inhibitors for human nonlysosomal glucosylceramidase. *J Am Chem Soc* 139, 14192–14197.

8. Artola M, Wu L, Ferraz MJ, **Kuo CL**, Raich L, Breen IZ, Offen WA, Codée JDC, van der Marel GA, Rovira C, Aerts JMFG, Davies GJ & Overkleeft HS (2017) 1,6-Cyclophellitol cyclosulfates: a new class of irreversible glycosidase inhibitor. *ACS Cent Sci* 3, 784–793.

9. Wu L,* Jiang J,* Jin Y,* Kallemeijn WW, **Kuo CL**, Artola M, Dai W, van Elk C, van Eijk M, van der Marel GA, Codée JDC, Florea BI, Aerts JMFG, Overkleeft HS & Davies GJ (2017) Activity-based probes for functional interrogation of retaining β-glucuronidases. *Nat Chem Biol* 13, 867–873.

10. Jiang J, **Kuo CL**, Wu L, Franke C, Kallemeijn WW, Florea BI, van Meel E, van der Marel GA, Codée JD, Boot RG, Davies GJ, Overkleeft HS & Aerts JM (2016) Detection of active mammalian GH31 α-glucosidases in health and disease using in-class, broad-spectrum activity-based probes. *ACS Cent Sci* 2, 351–358.

List of publications prior to PhD research

11. Li N, **Kuo CL**, Paniagua G, van den Elst H, Verdoes M, Willems LI, van der Linden WA, Ruben M, van Genderen E, Gubbens J, van Wezel GP, Overkleeft HS & Florea BI (2013) Relative quantification of proteasome activity by activity-based protein profiling and LC-MS/MS. *Nat Protoc* 8, 1155–1168.

12. Brouwer AJ, Jonker A, Werkhoven P, **Kuo E**, Li N, Gallastegui N, Kemmink J, Florea BI, Groll M, Overkleeft HS & Liskamp RM (2012) Peptido sulfonyl fluorides as new powerful proteasome inhibitors. *J Med Chem* 55, 10995–11003.

13. Geurink PP, Florea BI, Li N, Witte MD, Verasdonck J, **Kuo CL**, van der Marel GA & Overkleeft HS (2010) A cleavable linker based on the levulinoyl ester for activity-based protein profiling. *Angew Chem Int Ed Engl* 49, 6802–6805.

** Shared first-authorship*

Curriculum vitae

Chi-Lin (Ethan) Kuo was born on 11th June 1988 in Kaohsiung, Taiwan. He became interested in biology during an exchange year in high school (2006, West Hartford, Connecticut, USA). In 2007 he graduated from Kaohsiung Senior High School, and commenced his BSc study in Life Sciences at National Cheng-Kung University, Tainan, Taiwan. During this period he sought further experience abroad and was offered an one-year research internship opportunity in Prof. Hermen Overkleeft's group at Leiden University, the Netherlands. Here he studied activity-based probes (ABPs) for proteasomes. Parts of the results were later published in journals including *Nature Protocols* and *Journal of Medicinal Chemistry*.

After his graduation in 2011 (with honor) and a year of military service, he returned to Leiden University in 2012 for his MSc study in Life Science and Technology. Here he studied chemical biology of the ubiquitin-proteasome system under the supervision of Dr. Nan Li and Dr. Bobby Florea in the Overkleeft group. He also studied cellular processes related to autophagy in his minor internship under the supervision of Dr. Ruud Wijdeven in Prof. Sjaak Neefjes' group at the Netherlands Cancer Institute (NKI, Amsterdam). He graduated in 2014 *cum laude.*

In 2015 he began his PhD research in the newly-established Medical Biochemistry group at Leiden University. Under the supervision of Prof. Hans Aerts, the aim was to characterize glycosidase ABPs and their applications, particularly in the lysosomal storage disorders. Parts of the research described in this thesis were presented in the ESGLD conference (Naples, Italy, 2015, poster), the Joint Glycobiology Meeting (Nijmegen, NL, 2016, speech & poster), the annual ABPP conferences (Leiden, NL, 2017, speech; Oxford, UK, 2018, speech), the Molecular Medicines of Sphingolipids conference (Rehovot, Israel, 2018, poster), the Reedijk Symposium (Leiden, NL, 2018, poster), and the EWGGD conference (Clermont-Ferrand, France, 2019, Posters). Chi-Lin is currently a postdoctoral researcher in the Aerts group, while also conducting joint projects with the gene therapy company uniQure NV (Amsterdam, NL).