Activity-based probes for retaining β-glucosidases: Novel tools for research and diagnostics

Kallemeijn, W.W.

Citation for published version (APA):
Introduction, part I

Activity-based inhibitors of glycosidases: Design and applications

Wouter W. Kallemeijn, Martin D. Witte, Tom Wennakes, Johannes M. F. G. Aerts and Hermen S. Overkleeft

Prepared as invited review in Advances in Carbohydrate Chemistry and Biochemistry.

Graphical abstract – An activity-based probe can bind a target enzyme through mimicry of its substrate, form a covalent bond with the target via its catalytic mechanism, and allows detection of the bound enzyme by the implementation of a reporter tag.
Abstract

This introductory chapter, to be published in revised form as invited review, covers recent developments in the design and application of activity-based probes (ABPs) for glycosidases, with emphasis on enzymes involved in glucosylceramide metabolism in humans. Described are the various catalytic reaction mechanisms employed by inverting and retaining glycosidases. Understanding of catalysis at the molecular level has stimulated the design of different types of ABPs for glycosidases. Such compounds range from oxocarbenium ion-like transition state-mimics tagged with reactive moieties, which associate with the target active-site – forming covalent bonds in a rather non-specific manner in or near the catalytic pocket – to probes that (partially) exploit the catalytic mechanism of target retaining glycosidases via the action of fluoroglycosides, epoxides and aziridines, which bind to the catalytic nucleophile. The design of highly specific ABPs for human retaining β-glucosidases, including the lysosomal glucocerebrosidase GBA, non-lysosomal β-glucosidase GBA2, cytosolic β-glucosidase GBA3 and lactase/phlorizin hydrolase LPH, is discussed. A broad spectrum of applications for the ABPs in research on various glycosidases is demonstrated, including specific quantitative visualization of active enzyme molecules in vitro and in vivo, and as tools to unambiguously identify catalytic residues in glycosidases in vitro, and rapid, specific introduction of glycosidase deficiencies in intact cells and animals.

Author affiliations

aDepartment of Medical Biochemistry, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands. bDepartment of Bio-organic Chemistry, Stratingh Institute for Chemistry, University of Groningen, Groningen, The Netherlands. cDepartment of Synthetic Organic Chemistry, Wageningen University, Wageningen, The Netherlands. dDepartment of Bio-organic Synthesis, Leiden Institute of Chemistry, Leiden University, Leiden, The Netherlands.
The human body contains a great diversity of glycoconjugates, located inside cells, their extracellular matrix and in bodily fluids. The essential building blocks of glycoconjugates are carbohydrates, highly variable in composition, stereoisomerism and with the ability to connect to multiple moieties within a single molecule. These factors together result in an astounding level of combinatorial complexity: a simple hexasaccharide has at least $10^{12}$ possible isomers, far greater than achievable with amino acids or nucleotides\(^1\). Linkage of carbohydrates yields oligosaccharides, often named glycans. The number of distinct glycan structures in mammalian glycoconjugates is estimated to exceed 7,000+ structures, a diversity assembled from merely ten different carbohydrates: L-fucose, D-galactose, D-glucose, N-acetyl-D-galactosamine, N-acetyl-D-glucosamine, D-glucuronic acid, L-iduronic acid, D-mannose, N-acetyl-D-neuraminic acid and D-xylose (Fig 1)\(^2\)−\(^6\). The generation of the full mammalian glycan structure diversity involves ~200 glycosyltransferases\(^2\), using nucleotide- or lipid-linked sugars as activated donor substrates\(^7\). Vice versa, glycans are degraded via cleavage of their glycosidic bonds, a reaction catalyzed by specialized glycosylhydrolases (i.e. glycosidases)\(^2\). In total, 1–3% of the genes in the genome is believed to encode carbohydrate-modifying enzymes\(^8, 9\).

**Glycosidases: proficient catalysts of glycoconjugate turnover** – The glycosidic bond is very stable, far more than e.g. phosphodiester bonds in DNA and RNA, and peptide bonds in proteins\(^10\)−\(^13\), glycosidases wield specialized active-sites to accelerate hydrolysis of the glycosidic bond. Glycosidase-catalyzed hydrolysis occurs up to $10^{17}$-fold faster than spontaneous auto-hydrolysis, calculated to occur once per ~5 million years, to mere microseconds\(^14\). To date, a plethora of glycosidases (> 6,300 individual entities)\(^9, 15\) from various organisms have been identified and categorized in 130+ glycosylhydrolase (GH) families, based on their homology in amino-acid sequence similarities, in the Carbohydrate Active enZYme database (CAZy)\(^9\). This classification system correlates with protein-fold and catalytic mechanism more than the enzymatic specificity towards substrates\(^8, 16\)–\(^18\). Due to the conserved nature of glycosidases, the catalytic residues and overall mechanism of newly identified glycosidases can often be predicted\(^8, 9, 16\)–\(^18\).

**Figure 1 | Carbohydrates incorporated in mammalian glycoconjugates.** The ten carbohydrate sugars required for synthesis of the full glycan diversity (estimated > 7,000 structures) in mammals, with α- and β-stereoisomer of C1-hydroxyl coordination depicted by a wavy bond (down or up, respectively).
**Catalytic reaction mechanisms of glycosidases** – Glycosidases hydrolyze the glycosidic bond between two moieties, via a reaction that results in inversion or retention of the anomeric stereochemistry, a mechanism first outlined by Daniel E. Koshland, Jr. in 1953\(^1\). As shown in **Fig. 2a**, inverting glycosidases employ two opposing carboxyl-exposing residues, typically spaced 6–11 Å apart on either side of the active site, to enable entry of both the substrate and a water molecule\(^20\text{-}25\). During the reaction, the carboxylate ion (i.e. the base) residue abstracts a proton from the incoming water molecule, thereby generating a nucleophilic hydroxyl, which in turn attacks the substrate’s anomeric C1 carbon\(^20\text{-}25\). Concomitantly, the opposing carboxylic acid protonates the inter-glycosidic oxygen, thereby assisting its departure from the anomeric center\(^20\text{-}25\). The reaction proceeds through a single oxocarbenium ion-like transition state, stabilized in part through hydrogen-bonding between the C2-hydroxyl and the carboxylate base (see **Fig 2a, finely dotted line**)\(^20\text{-}26\text{-}31\). As the nucleophile can attack the C1 carbon from one side only, the oxocarbenium ion-like transition state causes the thereafter-formed hemi-acetal truncated sugar to exhibit inversed stereochemistry at the anomeric center\(^19\text{-}31\).

In sharp contrast, retaining glycosidases employ a double-displacement mechanism proceeding by way of a glycosyl-enzyme intermediate, flanked by two oxocarbenium ion-like transition states, in order to preserve the initial anomeric oxygen configuration (see **Fig 2b**)\(^19\text{-}31\). This mechanism employs two adjacent carboxyl-exposing residues, spaced ~5.5 Å apart, with one functioning as the direct catalytic nucleophile and one as acid/base residue.

**Figure 2** | Classic Koshland-type catalytic mechanisms of glycosidases. (a) Inverting β-glycosidase mechanism, with oxocarbenium ion-like transition state depicted between brackets (with asterisk; *). Transition state stabilized by hydrogen-bonding between substrate C2-hydroxyl and base residue (finely dotted line). (b) Catalytic mechanism of a retaining β-glycosidase, characterized by a covalent glycosyl-nucleophile adduct, flanked by two oxocarbenium ion-like transition states. Transition states stabilized by hydrogen-bonding between substrate C2-hydroxyl and nucleophile (finely dotted line).
Briefly, the deprotonated carboxylate ion of the nucleophile attacks the substrate’s anomeric C1 carbon, while the protonated carboxylate base of the acid/base donates a proton to the inter-glycosidic oxygen. Proceeding through an oxocarbenium ion-like transition state, the aglycone is expelled while concurrently a glycosyl-enzyme intermediate is formed with the nucleophile, exhibiting an inversed configuration at the anomeric center (Fig 2b). To liberate the nucleophile, the deprotonated carboxylate ion of the acid/base performs an inverting mechanism, characterized by the acid/base abstracting a proton from an incoming water molecule, forming a nucleophilic hydroxyl that attacks the anomeric center (C1) of the glycosyl-enzyme adduct, resulting in a second oxocarbenium ion-like transition state (Fig 2b). In turn, this reverses the changed anomeric configuration of the cleaved hemi-acetal to that present in the initial substrate.

Several variations on the classical Koshland-type mechanisms have been identified. Multiple glycosidases categorized in CAZy GH families 18, 20, 25, 56, 84 and 85 are able to hydrolyze the glycosidic bond at C1 of substrates containing an N-acetyl or N-glycolyl moiety at C2, through a mechanism involving neighboring group participation (Fig 3a, next page). For hydrolysis, a retaining N-acetyl-hexasaminidase does not require a catalytic nucleophile: instead the C2-acetyl moiety executes the role as intra-molecular nucleophile. This leads to the formation of an oxazolinium intermediate, which is not covalently bound to the enzyme, as in retaining enzymes of the classic Koshland-type. Moreover, a carboxylate is generally present to stabilize the charge development during the transition state, thereby partially mimicking the role of the nucleophile in the stabilization of the oxocarbenium ion-like transition state through hydrogen-bonding (Fig 3a).

The retaining mechanism identified in some sialidases (i.e. neuraminidases, in CAZy GH 33, 34) employs a tyrosine instead of a carboxylate as nucleophilic residue (Fig 3b). As the anomeric center of the substrate is already negatively charged, the use of a similarly charged carboxylic acid would cause charge repulsion. A nearby base residue deprotonates the tyrosine to increase its nucleophilic character, thereby enabling the nucleophilic attack to the C1 carbon of sialic acid, forming a covalent glycosyl-enzyme intermediate with inversed anomeric C1 configuration. As the acid/base protonates the inter-glycosidic oxygen, it promotes deglycosylation of the nucleophile-adduct via generating a hydroxyl in a mechanism similar to inverting Koshland-type glycosidases.

Another catalytic mechanism has been described for α-glucan lyases (in GH 31), whom break the glycosidic bond through a two-step mechanism employing elimination (Fig 3c). After a Koshland-type nucleophilic attack and concomitant protonation of the inter-glycosidic oxygen by the acid/base residue, a base residue in close proximity induces the hydrolysis of the formed glycosyl-enzyme intermediate. Withdrawal of the proton at C2 causes rearrangement inside the glycosyl-adduct, causing the formation of a 1,2-anhydro-β-fructose.

Other unusual catalytic mechanisms include redox mechanisms, employing tightly-bound nicotinamide adenine dinucleotide (NAD⁺), which is found in several glycosidases...
Introduction, part I
categorized in CAZy families GH4 and 109, to facilitate glycoside cleavage via an oxidation, elimination, addition, reduction sequence, and myrosinases (members of CAZy GH1), which utilize an exogenous base for hydrolysis. These mechanisms will not be discussed in further detail in this review.

The concept of activity-based probes – Reducing the action of glycosidases through the use of inhibitory sugar-mimics has been used in research on (treatments for) cancer, viral infections, neurological disorders, the metabolic syndrome, and glycosphingolipid storage diseases.

Very potent (reversible) inhibitors for glycosidases can be obtained by the design of compounds that optimally mimic the oxocarbenium ion-like transition state of the target enzyme’s natural substrate during hydrolysis. They offer valuable research tools to study...
Activity-based probes for glycosidases

the physiological function of glycosidases. In addition, such type of reversible inhibitors may act as chemical chaperones promoting and/or stabilizing the correct folding of glycosidases. The value of chemical chaperone therapy is presently investigated for patients suffering from lysosomal glycosidase deficiencies.

This review focuses on another class of inhibitors, which can assist in direct identification, e.g. through visualization, of enzyme molecules by specific covalent linkage in their catalytic pocket. These so-called activity-based probes (ABPs) enable the detection of the targeted glycosidase (or class of glycosidases) in complex mixtures, living cells and animals. In general, activity-based probes feature three structural elements, namely a warhead and a reporter moiety, linked together via a spacer (Fig 4a, left).

Typically, the warhead comprises of a reactive group allowing the covalent attachment to the catalytic pocket of the enzyme. Employed for this are photo-crosslinkers and (latent) electrophiles, which bind to a(ny) nucleophile in the catalytic pocket. Grafting the warhead onto a natural substrate-mimicking molecular core exploits the exhibited specificity of, and affinity towards, the target enzyme’s active site. The composition of the reporter tag and spacer moiety bridging both subunits can further enhance ABP properties, including

![Figure 4](image.png)

**Figure 4 | Concept of activity-based probes.** (a) General activity-based probe (ABP) structure, depicted with reporter (lightbulb), spacer (curved line) and warhead (grey triangle) grafted onto a substrate mimick (white triangle), to gain affinity towards active-site of target enzymes. (b) ABP labeling of targeted enzyme in the complete proteome present in vitro, in intact cells or organisms in vivo, with various read-out possibilities. (c) Types of regularly employed reporter tags. From left to right: fluorescent moieties (BODIPY, fluorescein and rhodamine), affinity tag (biotin) and bio-orthogonal labeling handles (azide-, alkyne- and norbornene-moieties).
augmentation of substrate mimicry and introduction of necessary spacing between the warhead-substrate core, tag and the target enzyme to minimize unfavorable interactions including steric hindrance. Optimally, exposure of the total proteome, either present in a lysate (in vitro), in living cells (in situ) or in intact laboratory animals (in vivo), results in the labeling of only the targeted enzyme or enzyme class (Fig 4b), where after the ABP-labeled proteins can be visualized through various techniques, including fluorescence scanning of SDS-PAGE gels, fluorescence microscopy, fluorescence-assisted cell sorting (FACS), positron emission tomography (PET), but labeling could also enable purification and subsequent MS/MS or NMR techniques.

For detection of ABP-labeled enzyme, numerous reporter tags have been employed (Fig 4c). Generally they include fluorescent moieties (BODIPY, fluorescein, rhodamine), radio-isotopes ($^3$H, $^{19}$F, $^{125}$I) and affinity tags such as biotin. Since the inherent steric bulk of ABP reporter-tags may potentially sever recognition by the targeted enzyme(s), the covalent labeling mechanism or mitigate the cell-permeability of the ABP. An alternative strategy is to replace these reporters with markedly smaller bio-orthogonal ligation handles to enable two-step labeling strategies (Fig 4c).

Bio-orthogonal ligation ideally occurs chemoselectively and without interference of the biological system. Several bio-orthogonal ligation strategies have been developed, these include the Staudinger-Bertozzi ligation, copper(I)-catalyzed click reaction, strain-promoted click reaction and norbornene cycloaddition, which can be performed in the same sample (Fig 5).

**Figure 5 | Bio-orthogonal labeling.** Variants of azide with Staudinger-Bertozzi ligation (top), copper(I)-catalyzed click reaction (second) and strain-promoted click reaction (third), and tetrazine-norbornene cycloaddition (fourth). Presence of reporter tag, e.g. a fluorescent moiety (see also Fig 4c), is illustrated with a lightbulb.
Inhibitors of glycosidases, the road towards true activity-based probes – In recent years, different types of probes have been designed to label the catalytic pockets, or their close environment, in glycosidases. Some merely have affinity for the catalytic pocket and are subsequently covalently linked in rather non-specific manner, whilst true activity-based probes employ the catalytic mechanism of glycosidases for their linkage to the enzyme.

A reversible inhibitor with sufficient affinity for the catalytic pocket of a target glycosidase may be transformed into a labeling probe by addition of a photo-activatable tag, e.g. a benzophenone, diazirine or aryl azide (Fig 6, next page). Optimally, photo-affinity tags are stable in the dark under various pH conditions, bear structural resemblance to the natural substrate of the target enzyme, are sterically non-congested, require a wavelength for photo-activation which does not cause undesired damage to other components in the biological sample, generate highly reactive, short-lived photo-intermediates upon irradiation and form a stable adduct with the target protein.

Benzophenone-type photo-affinity labels generate upon irradiation (λ<sub>EX</sub> 350–360 nm) triplet carbonyl states, which can react with inactive C–H bonds (Fig 6a, next page). While stable, the photo-activatable group is rather bulky which potentially hinders probe-recognition and ultimately also binding to the target protein. However, a bulky tag can also fortuitously improve specificity as is noted for extension of 1-deoxynojiririmycin with a benzophenone-type tag, yielding the potent probe 1 (Fig 6b) for labeling of human acid β-glucosidase (GBA) and non-lysosomal β-glucosidase GBA2.

Compared to benzophenones, aryl azides are relatively small molecules that are readily synthesized, relatively stable and highly reactive upon photolysis (Fig 6c). Irradiation with wavelengths of λ<sub>EX</sub> < 300 nm is required for photolysis, which can induce severe damage to biological systems such as cells. Photolysis generates a singlet nitrene, able to generate a triplet nitrene via intersystem crossing. Importantly, the singlet nitrene can rearrange into bicyclic benzazirine, which generates 1,2-azacycloheptatetraene, which in turn can bind with nucleophiles further away, thereby reducing labeling efficiency and specificity. Overall, aryl azides allow binding to nearest molecules through multiple reaction mechanisms. For instance, development of a photo-affinity probe based on 1-deoxynojiririmycin equipped with an aryl azide (probe 2, see Fig 6d) resulted in the successful labeling of human α-glucosidase.

Diazirines are currently the smallest photo-activatable moiities employed for photo-affinity labeling (Fig 6e), and are relatively inert towards nucleophilic attacks, acidic– and alkaline conditions. Photolysis of diazirines requires irradiation with λ<sub>EX</sub> 350–380 nm, which concomitantly generate a highly reactive species, e.g. a carbene, able to form a covalent bond with the closest target molecule via C–C, C–H, O–H or C/N–H insertion (Fig 6f). As carbenes are quickly quenched by water, the efficiency of labeling with diaziridines is rather poor, but their half-life of nanoseconds enables more specific labeling of target proteins. Nevertheless, diazirine-wielding probe 3, developed by Kuhn et al. to label β-galactosidase from *E.coli* was unsuccessful in labeling its catalytic
Introduction, part I

In general, photo-affinity probes have been only successful to label enzymes such as α-glucosidases, N-acetyl-hexosaminidases and sialidases in a non-specific manner. Photo-affinity probes also have great applicability in the field of lipid research: equipping lipids with these tags allows cross-linking to identify lipid-lipid as well as lipid-protein interactions.

While reversible inhibitors equipped with photo-affinity tags facilitate labeling of both inverting and retaining glycosidases, these probes cannot discriminate between inactive and active target enzyme molecules, and their covalent attachment after photolysis is rather unspecific. For this purpose, focus has shifted from the concept of employing reversible
Activity-based probes for glycosidases

inhibitors, that mimic the oxocarbenium ion-like transition state (see Fig 2), to probes that exploit the catalytic mechanism intrinsic to the targeted glycosidases in order to become covalently attached, whilst mimicking the structure of the natural substrate or the oxocarbenium ion-like transition state.

One such approach encompasses p-nitrophenyltriazene probes, such as probe 4 introduced by Sinnott\textsuperscript{112,113} (Fig 7). After recognition and correct positioning of the probe in the active site, protonation of the aglycone subsequent to cleavage of the glycosidic bond, yields a reactive species that decomposes into nitrogen, arylamine and a highly reactive glycosyl methylcarbenium-like ion, which rapidly reacts with nearby nucleophiles\textsuperscript{68,112,113}. Concomitantly, the probe is immobilized in the active site, thereby inhibiting the enzyme’s activity\textsuperscript{68,112,113}. This approach has been applied successfully to several enzymes to identify the catalytic amino acid residues, such as in E. coli LacZ\textsuperscript{112,113}, and β-galactosidases from other origins\textsuperscript{114−116}.

Another approach is where the glycosidase cleaves the glycosidic bond between a sugar and an inherently non-reactive aglycone\textsuperscript{117}. After hydrolysis, the expelled aglycone rearranges into a highly reactive moiety, which then non-specifically binds to other residues. Examples include glycosides containing a difluoroalkyl or difluorotolyl moiety, for instance probe 5\textsuperscript{118} (see Fig 7b, and c for reaction mechanism). After liberation, these moieties eliminate hydrogen fluoride and a highly reactive acyl fluoride or quinone methide, respectively\textsuperscript{118,119}. This mechanism-based activation allows the labeling of both inverting and retaining glycosidases, but as the reactive group is in the aglycone, i.e. the part of the product lowest affinity for the active site, covalent inactivation of the catalytic pocket is not necessarily very efficient. Accordingly, no studies have been reported on using these probes for active-site identification.

**Figure 7 | Probes with reactive aglycones.** (a) Reaction mechanism of triazene-type probes labeling the catalytic nucleophile (N). (b) Structure of probe 4, a β-о-glucose decorated with a p-nitrophenyltriazene and probe 5, a β-о-glucose decorated with difluorotolyl. (c) Reaction mechanism of difluorotolyl-wielding probe 5. During the retaining β-glucosidase double-displacement mechanism, involving nucleophile (N) and acid/base (A/B), rearrangements inside the released difluorotolyl aglycone result in semi-irreversible, nonspecific labeling of nearby nucleophilic residues (X).
2-Deoxy-2-fluoroglycosides as ABPs for retaining glycosidases – In the 1980s, the idea was conceived that one could temporarily trap retaining glycosidases in the glycosylated stage of their double-displacement mechanism. Initial experiments by Legler revealed that removal of the C2-hydroxyl in p-nitrophenyl β-D-glucoside, resulting in p-nitrophenyl β-D-2-deoxyglucoside 6 (Fig 8a) resulted in an accumulation of the glycosyl-enzyme intermediate of a β-glucosidase, enabling the direct identification of its 2-deoxyglucoside-bound nucleophile \(^{120}\) (see Fig 8b for a similar reaction mechanism).

Accumulation of the glycosylated nucleophile occurred due to the lack of key interactions between the C2-hydroxyl in the oxocarbenium ion-like transition state and the catalytic pocket of the enzyme \(^{120,121}\) (see Fig 2 for proposed regular stabilization). Moreover, the replacement of the C2-hydroxyl group with a hydrogen has only moderate effects on binding, as expressed by the Michaelis constant \(K_m\), but is profoundly detrimental to the rate constant of product formation \(k_{cat}\). Absence of the C2-hydroxyl destabilized the transition state as such that the glycosyl-enzyme intermediate accumulated, due to a \(10^6\)-fold decreased deglycosylation rate \(^{122,123}\). It may well be that interactions with the C2-hydroxyl are not only required for stabilization, but also for the optimal orientation of the substrate to the catalytic residues, in respect to the to-be-cleaved glycosidic bond \(^{122,123}\).

Withers and co-workers subsequently developed activated 2-deoxy-2-fluoroglycosides \(^{121-124}\), where the C2-hydroxyl was substituted for C2-fluorine, for instance in probes 7 and 8 (Fig 8a). Introduction of a fluorine onto the carbohydrate-ring destabilizes the charge development at O5 or C1 in the oxocarbenium ion-like transition state during both the glycosylation and deglycosylation steps, due to the strictly limited hydrogen bonding potential of fluorine (Fig 8b). In addition, fluorine stabilizes the glycosyl-enzyme intermediate through inductive destabilization of the electron-deficient transition states. As fluorine is more electronegative than a hydroxyl-group, it destabilizes the positive charge

---

**Figure 8 | Fluoroglycoside-type probes.** (a) Structures of 4-nitrophenyl 2-deoxyglucose probe 6, 2-deoxy-2-fluoroglucoside probes 7 and 8, 2-deoxy-2,2-difluoroglucoside probe 9 and 5-fluoroglucoside probe 10. (b) Trapping of the catalytic nucleophile (N) in retaining β-glucosidases after the initial step in the double-displacement mechanism, requiring the acid/base (A/B) for protonation of the interglycosidic oxygen linking to the aglycone (R). The oxocarbenium ion-like transition states flanking the glycosyl-nucleophile adduct are depicted between brackets (with asterisk; \(*\)).
developed during the oxocarbenium ion-like transition state. Influence on both steps of the mechanism culminates in 10⁶ to 10⁷-fold reduction of the catalytic rates. However, by introducing aglycones that are good leaving groups, i.e. aglycones with low pKa values, such as 2,4-dinitrophenyl (pKa ~4.1, probe 7) or fluorine (pKa ~3.2, probe 8), the formation of the glycosyl-enzyme intermediate proceeds faster, while the deglycosylation rate remains very poor due to the C2-fluorine destabilizing the subsequent oxocarbenium ion-like transition state. Use of the natural aglycone in fluoroglycosides has also been described. A relatively slowly proceeding inhibition of enzyme activity by fluoroglycosides has been observed, as the result of temporary trapping of the glycosyl-enzyme intermediate. The stability of the intermediate can vary considerably, ranging from 1 to 600 hours, depending on the enzyme and the exact structure of the fluoroglycoside employed.

Retaining α-glycosidases are less effectively inhibited by 2-deoxy-2-fluoroglycosyl fluorides, as these commonly act as slow substrates in which the glycosylation step is rate limiting. This has been attributed to the greater stability of the α-glycosyl enzyme intermediate formed on a β-glycosidase than the inverse situation. The rapid turn-over of the 2-deoxy-2-fluoroglycosyl intermediate in α-glycosidases could be overcome by development of 2,2-difluoroglycosides, where a second fluorine at C2 further slowed both the glycosylation and deglycosylation steps, while further enhancing the reactivity of the leaving group. In the study of Braun et al., 2,4,6-trinitrophenyl 2-deoxy-2,2-difluoro-α-D-glucopyranoside inhibited yeast α-glucosidase, and 2,4,6-trinitrophenyl 2-deoxy-2,2-difluoro-α-D-maltosyl fluoride potently inhibited human α-amylase. Enhanced accumulation of the glycosyl-enzyme intermediate can also be achieved by the introduction of 5-fluoroglycosyl probes such as 10. These 5-fluoroglycosyl fluorides act as slow substrates due to the presence of the C2-hydroxyl. Unfortunately, 5-fluoroglycosides show a very poor $K_M$, caused by the presence of the electron-withdrawing fluorine directly adjacent to the O5, which develops the greatest charge during the oxocarbenium ion-like transition state.

Crystallographic analysis of the glycosyl-enzyme intermediates formed through the reaction of fluoroglycosides has yielded insight in the structure – and roles of – specific residues in the catalytic pocket of the enzymes. Most potent tools to identify the nucleophile of glycosidases are 2-deoxy-2-fluoroglycosyl– and 5-fluoroglycosyl fluorides, and 2,4-dinitrophenyl 2-deoxy-2-fluoroglycosides. The relatively stable, trapped glycosyl-enzyme intermediate also allowed analysis through $^{19}$F-NMR and/or MS/MS peptide sequencing, assisting the identification of the nucleophile residue of various enzymes, including those of multiple α– and β-glycosidases, sialidases and a N-acetyl-β-D-glucosaminidase.

2-Deoxy-2-fluoroglycosides enable the assignment of both the acid/base and nucleophile amino-acid residues, through detailed kinetic studies with mutant enzymes. In this classic approach, the acid/base can be identified by the reduced
catalysis of 2-deoxy-2-fluoroglycosides by enzyme lacking this amino acid residue. Complete absence of activity is seen for enzyme lacking the nucleophile residue. Identification of acid/base and nucleophile residues can be confirmed by co-incubation with an alternative nucleophile such as azide. In case of absence of the acid/base, the configuration of the anomeric carbon is retained upon rescue with azide (Fig 9a), whilst absence of the nucleophile causes subsequent stereochemical inversion (Fig 9b). In a similar approach, fluoroglycosides with inverted stereochemistry may be incubated with enzymes lacking the nucleophile to employ them as glycosynthases forming a glycosidic bond between two (complex) carbohydrate molecules (Fig 9c).

Besides identification of the catalytic residues, fluoroglycosides like on 2-deoxy-2-fluoro-β-D-glucopyranosyl and −β-D-mannopyranosyl compounds injected in rats labeled their target β-glucosidases and β-mannosidases in several organs, including the brain. These data indicate that 18F-labeled fluoroglycosides might be used as imaging probes of

![Diagram](image-url)
glycosidases using positron emission tomography (PET). In animals injected with $^{18}$F-2-deoxy-2-fluorogluco side labeled human acid β-glycosidase, its bio-distribution has been monitored via PET.

**Epoxides as ABPs for retaining glycosidases** – Probes equipped with epoxides have been proven to be true mechanism-based inhibitors for numerous retaining glycosidases, as these ABPs require the target enzyme’s catalytic mechanism to bind covalently to the nucleophile. Typically, upon activation of the epoxide, *i.e.* after specific protonation of the oxirane by the acid/base residue, a reactive species is generated by the concomitant attack by the nucleophile to the anomeric C1 carbon, effectively opening the epoxide and the formation of a stable ester bond with the nucleophile (Fig. 10a, next page).

The reactivity of the epoxide group reduces markedly when conjugated to a glycoside due to the electron-withdrawing effect of hydroxyl groups in the vicinity, consequently increasing both the inherent stability *in solutio* and the enzymatic requirements to open the oxirane. Covalent enzyme inhibition only occurs with epoxides with an appropriate orientation allowing protonation of the epoxide oxygen atom by the acid/base, *i.e.* by a carboxylic acid group in close proximity.

Conduritol epoxides incorporate an endocyclic epoxide grafted to a cyclitol scaffold that mimics the carbohydrate ring of xylose. The most common conduritol epoxide stereoisomer is conduritol β-epoxide (CBE 12, see Fig. 10b), or DL-1,2-anhydro-myoinositol, which was discovered by Legler and meticulously characterized and reviewed. CBE 12 is a potent mechanism-based inhibitor for several retaining glucosidases present in various organisms. It has been employed to identify the catalytic nucleophiles in several glucosidases.

The symmetry in CBE 12 at the C2 axis allows it to irreversibly inhibit both retaining α- and β-glucosidases. Inhibition of retaining α-glucosidases by CBE 12 occurs with reduced reaction kinetics as compared for β-glucosidases. This is due to the alternative orientation the molecule must be in to open the β-epoxide: in β-glucosidases it preferentially opens *trans*-diallyl (Fig. 10a), but in α-glucosidase it opens *trans*-equatorially (Fig. 10c). The *trans*-dialyl opening of the β-epoxide of CBE 12 in β-glucosidases was confirmed by cleaving the $^{14}$C radio-labeled CBE 12-nucleophile adduct with hydroxylamine, which revealed that only D-1,2-anhydro-myoi nositol reacted with the glucosidase, closely mimicking the structure of D-xylose and D-glucose. No β-glucosidase inhibition was observed for L-1,2-anhydro-myoinositol. Alternatively, analogues of CBE 12, including L-1,2-anhydro-myoinositol, conduritol F epoxide 13, conduritol C cis-epoxide 14 and *trans*-epoxide 15 were found to irreversibly inhibit a β-fructosidase, α-mannosidases, β-galactosidases, α-galactosidases and a α-fucosidase, respectively. Symmetry in CBE 12 furthermore permits binding into some catalytic pockets with different orientations, thereby covalently binding catalytic residues other than the nucleophile. This is illustrated by the labeling of alternative residues with conduritol C *cis*-.
epoxide in E. coli LacZ β-galactosidase, and conduritol β-epoxide for both human GBA and almond β-glucosidase. The correct residues were later identified with 2-deoxy-2-fluoroglycosides. The symmetry of CBE permits multiple binding modes, thereby enabling irreversible inhibition of both α- and β-glucosidases through different interactions. Loss of symmetry dramatically restricts the number of orientations of the inhibitor in a catalytic pocket. Introduction of a structural group in the cyclitol ring of CBE, which breaks its symmetry, results in superior target specificity and inhibitory potential towards specific retaining β-glucosidases.

The asymmetric cyclophellitol β-epoxide, isolated from the Phellinus sp. mushroom, indeed inhibits β-glucosidases several orders of magnitude more potently. Its covalent binding to the catalytic nucleophile was demonstrated firstly through crystallography of a β-glucosidase from Thermotoga maritima. The substitution of the C5-hydroxyl group of CBE with a 5-hydroxymethyl group makes the inhibitor resemble β-D-glucose more closely than β-D-xylene. Importantly, this β-epoxide has a reduced affinity for α-glucosidases. Specific targeting of α-glucosidases can be achieved by synthesis of the
diastereomer 1,6-epi-cyclophellitol \(17^{192}\), which potently inactivated \(\alpha\)-glucosidases in cancer cells. Diastereomers with \(\alpha\)-glucose \(17\), \(\beta\)-mannose \(18\) and \(\alpha\)-mannose \(19\) configurations were expectedly potent inhibitors of the corresponding \(\alpha\)-glucosidases, \(\beta\)-mannosidases and \(\alpha\)-mannosidases\(^{190-193}\). Exo-alkyl epoxide-glycosides, \textit{i.e.} compounds with the epoxide-group attached \textit{via} an alkyl-spacer to a sugar (see \textit{e.g.} probe \(11\), in Fig \(10a\)) deserve also discussion. These compounds lack the structural rigidity of conduritol– and cyclophellitol structures and, due to their flexibility, can also label catalytic carboxylates of several glycosidases\(^{194-206}\), including exoglucosidases and endoglycosidases such as cellulases\(^{194,199-202}\). Interestingly, 2,3-epoxypropyl \(\beta\)-D-xyloside probe \(11\) labeled the catalytic nucleophile of a retaining 1,4-xylanase, while the 3,4-epoxybutyl form covalently bound to the enzyme’s acid/base\(^{204}\). Apparently, the length of alkyl chains causes in this case alternative orientations of the epoxide in the catalytic pocket\(^{206}\), a flexibility that can be exploited. By variation of the alkyl chain length and configurations of the epoxide, specific ABPs can be tailored for a particular enzyme\(^{201,207}\).

Another class of irreversible, mechanism-based inhibitors encompasses compounds equipped with an aziridine as warhead (Fig \(11\)). Glycosyl aziridines have received far less attention than epoxide-containing inhibitors. Typically, the nitrogen in the aziridine head-group is protonated by the acid/base residue. The protonated nitrogen thereby becomes an optimal leaving group, allowing aziridine-ring opening following a nucleophilic attack, resulting in the formation of a stable ester-bond with the nucleophile, similar as observed/proposed for epoxide-containing probes (Fig \(11a\)). Theoretically, aziridines have a high affinity for the rather negatively charged catalytic pockets of glycosidases due to the positive charge of the protonated nitrogen atom in the aziridine moiety. Aziridines with particular configurations are highly specific inhibitors of corresponding retaining glycosidases, similarly as observed for epoxide-containing inhibitors. Of note, conduritol \(\beta\)-aziridine \(20\) (Fig \(11b\)) inhibits both \(\beta\)-glucosidases and \(\alpha\)-glucosidases due to its symmetry at the C2 axis, which is similar as in CBE \(12\)^{208,209}.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure11}
\caption{Catalytic nucleophile-binding by aziridine-wielding probes. (a) Reaction mechanism of (conduritol) \(\beta\)-azidirine labeling the nucleophile (N) of a retaining \(\beta\)-glucosidase (through \textit{trans}-diaxial ring opening), after activation of the warhead by the acid/base (A/B). The oxocarbenium ion-like transition state is depicted between brackets (with asterisk; e). (b) Structure of conduritol \(\beta\)-aziridine \(20\).}
\end{figure}

**ABPs for the human glucosylceramide-metabolizing glycosidases** – Conduritol \(\beta\)-epoxide \(12\) and cyclophellitol \(\beta\)-epoxide \(16\) are irreversible inhibitors of the human
lysosomal acid β-glucosidase, named glucocerebrosidase (GBA)\textsuperscript{210–213}. GBA is responsible for the breakdown of glucosylceramide, by catalyzing the hydrolytic cleavage of the glycosidic bond between β-D-glucose and ceramide, which is the final step in the degradation of glycosphingolipids.

This important class of glycoconjugates was discovered in 1884 by Johan L. W. Thudichum, who coined their name ‘glycosphingolipids’, after the Sphinx and its riddle in his book\textsuperscript{214} entitled ‘A Treatise on the Chemical Constitution of Brain’. Glycosphingolipids are currently known to be involved in a myriad of biological processes, including the formation of membrane domains\textsuperscript{215}, function as bioactive molecules involved in regulation of the cell-cycle\textsuperscript{216}, differentiation\textsuperscript{217}, senescence\textsuperscript{218}, necrosis\textsuperscript{219}, proliferation\textsuperscript{220}, apoptosis\textsuperscript{221}, and act as structural components required for cell adhesion\textsuperscript{222}. As a consequence, glycosphingolipids are implicated in multiple pathological processes\textsuperscript{223}, including cancer\textsuperscript{224–228}, the metabolic syndrome\textsuperscript{227–229} and glycosphingolipid storage disorders\textsuperscript{230–233}. Glycolipids make up a predicted 80% of all glycoconjugates, with the class of glycosphingolipids estimated to be the largest (> 9,000 species)\textsuperscript{234}. In human plasma alone, more than 200 distinct glycosphingolipids have been identified recently\textsuperscript{235}.

**Glycosphingolipid metabolism** – The structure of glycosphingolipids is comprised of several elements (Fig 12, next page), first of which is a long-chain sphingoid base, represented by sphingosine, (2S,3R,4E)-2-amino-octadec-4-ene-1,3-diol, also referred to as (E)-sphing-4-enine, the major sphingoid base found in mammals\textsuperscript{234, 236, 237}. The majority of sphingoid bases are amide-linked with a long- or very-long-chain fatty acid (second structural feature) to form ceramides, which (third structural feature) can be further derivatized by addition of a head-group to form (glyco)sphingolipids, including sphingomyelin, glucosylceramide, galactosylceramide and more complex glycosphingolipids containing numerous sugar residues\textsuperscript{234, 237–239}. See for excellent overviews of (glyco)sphingolipid chemistry references\textsuperscript{234, 237, 239}.

The biosynthesis and degradation of glycosphingolipids is highly complex, as it requires the expression and correct localization of an array of enzymes\textsuperscript{234, 237}. Namely, after synthesis of dihydroceramide from L-serine and Coenzyme A-activated fatty acids at the cytosolic leaflet of the ER membrane\textsuperscript{240–242}, dihydroceramide desaturase produces ceramide. Additionally, ceramide can be produced by acylation of sphingosine, freed during lysosomal degradation of (glyco)sphingolipids\textsuperscript{234}. Subsequent transport of ceramide to the membrane of the cis-Golgi apparatus\textsuperscript{243–245} allows glucosylceramide synthase (GCS) to glycasylate ceramide with UDP-activated α-D-glucose to produce β-D-glucosylceramide, one of the simplest known glycosphingolipids (Fig 12a)\textsuperscript{246–249}. Additionally, ceramide can be equipped with a α-D-galactose, resulting in β-D-galactosylceramide, which serves as the basis for different glycosphingolipid species, or ceramide is employed to synthesize sphingomyelin\textsuperscript{234}. Creation of more complex glucosphingolipids occurs upon flipping of glucosylceramide to the luminal side of the Golgi apparatus\textsuperscript{250–253}, where other UDP-
Activity-based probes for glycosidases

activated sugars can be added to through a step-wise process. As a first step, glucosylceramide is extended with UDP-activated galactose by GalT1 to form lactosylceramide (Fig 12a)\(^{234}\). Next, highly complex glycosphingolipids containing multiple branches and comprising of different sugar moieties are formed\(^{234, 237}\), with each biosynthetic step tightly regulated through localization of the glycosyltransferase, availability of appropriate glycosphingolipid substrate and the presence of activated glycosyl donor substrate.

**Vice versa**, during the catabolism of complex glycosphingolipids, each glycoside is removed in a stepwise fashion, requiring an array of specialized glycosidases (Fig 12b)\(^{234, 237}\). Breakdown occurs predominantly in the endosomal–lyosomal compartment, of which the membrane is protected against glycosidase-executed degradation through the presence of a glycosalyx consisting of heavily glycosylated membrane proteins\(^{234}\). Glycosphingolipids destined for degradation may enter the endosomal/lysosomal apparatus through receptor-mediated endocytosis of lipoproteins, endocytosis of the plasma membrane or phagocytosis of membrane structures and/or senescent– and apoptotic cells. The latter process occurs predominantly in macrophages, which, amongst others, phagocytose senescent erythrocytes containing high levels of (blood-group)
glycosphingolipids. In the lumen of lysosomes, glycosidases require specific (glyco)sphingolipid activator proteins (SAPs) to assist in the interaction with the membrane-bound glycosphingolipid substrate, in order to remove sequentially the carbohydrates from the glycosphingolipids, starting at the non-reducing end of the molecule.

Inability of cells to successfully catalyze the cleavage of all glycosidic bonds encountered in glycosphingolipids inevitably results in the accumulation of the retained, uncleavable glycosphingolipid, thereby eventually causing lysosomal dysfunction and a glycosphingolipid storage disorder. The most common lysosomal glycosphingolipid storage disorder is Gaucher disease, named after the French clinician Philippe C. E. Gaucher, who firstly authored a detailed case report in his doctoral thesis – ‘De l’Epithelioma Primitif de la Rate, Hypertrophie Idiopathique de la Rate sans Leucemie’ – in 1882. This thesis included a description of the clinical manifestations of a 32-year old woman presenting an unexplained hepatosplenomegalgy. After it became clear that macrophages in Gaucher disease accumulated a lipoid structure, the French chemist Aghion identified this chemical structure as glucosylceramide, one of the simplest glycosphingolipids. Roscoe O. Brady and co-workers discovered in the 1960s that the activity of the lysosomal glucosylhydrolase glucocerebrosidase (GBA) was deficient in Gaucher disease, causing the hallmark accumulation of glucosylceramide. The gene GBA encoding the GBA enzyme was independently cloned in the early 1980s by the research groups of Barranger, Beutler and Horowitz. Almost three decades later the existence of another, non-lysosomal, β-glucosidase (GBA2) was firstly described. The gene encoding this membrane-bound enzyme was independently cloned by Yildiz and Boot in 2007. Other human enzymes that (may) also be able to hydrolyze glucosylceramide are the intestinal lactase/phlorizin hydrolase (LPH) and the cytosolic non-specific β-glucosidase GBA3, albeit the latter with very low affinity.

Design of ABPs for glucocerebrosidase – Recently, a 2-deoxy-2-fluoro-D-glucoside was designed with an imidate as optimized leaving group (Fig 13). This probe was found to avidly label GBA in vitro and in living cells. As discussed previously, the binding of 2-deoxy-2-fluoro-D-glucoside to glutamate E340 is not irreversible. Reactivation of GBA, and loss of labeling, occurs spontaneously also at a low rate (Chapter 3).

Figure 13 | Fluorescent 2-deoxy-2-fluoro-D-glucoside. Structure of fluorescent β-2-deoxy-2-fluoro-D-glucoside probe 21, containing imidate as leaving group for optimal in situ labeling of GBA. See also Chapter 3, or ref.
In parallel, CBE 12 and cyclophellitol β-epoxide 16 were utilized as leads for the development of specific fluorescent activity-based probes for GBA. An azide-group was grafted to the C6 position of cyclophellitol β-epoxide 16, yielding azido-cyclophellitol probe 22 (Fig 14a). Azido-cyclophellitol 22 was 100-fold more potent than CBE 12. It was also observed that when 22 was covalently extended at C6 with BODIPY fluorophore moieties, these ABPs became even more potent inhibitors for GBA280, 281. These fluorescent ABPs, green fluorescent β-epoxide 23 and red fluorescent β-epoxide 24, proved superior probes to azido-cyclophellitol 22, with an approximately 5,000-fold improved inhibitory potency when compared to CBE 12 (1.24 and 1.94 nM versus 9,497 nM). The inhibition of GBA with β-epoxide 23 occurred most efficiently at its pH optimum, indicating that only active enzyme molecules are labeled with the probe (Fig 14b). This was expected since the β-epoxide warhead should only open after activation by the acid/base and concomitant nucleophilic attack (see Fig 10 for reaction mechanism). Analysis of labeled GBA with MS/MS indeed validated that azido-cyclophellitol 22 bound to the nucleophile E340281. Moreover, direct attachment of the BODIPY tag allows for immediate fluorescence scanning of β-epoxide 23-labeled proteins on for instance SDS-PAGE gels, and in turn revealed that direct labeling can be ablated by pre-incubation with known active-site inhibitors or denaturing conditions (Fig 14c).

Moreover, the detection of the BODIPY-emitted fluorescence proved ultra-sensitive, with visualization achievable of 20 attomoles of GBA molecules (20 × 10⁻¹⁸ mol, or ~12

![Figure 14](image)

Figure 14 | GBA specific cyclophellitol epoxide-type ABPs. (a) Structures of cyclophellitol β-epoxide 16, azido-cyclophellitol 22 (KY170), and fluorescent β-epoxide 23 (MDW933, green fluorescence) and β-epoxide 24 (MDW941, red). (b) Effect of pH on relative glucocerebrosidase (GBA) activity towards artificial 4MU-β-D-Glc substrate (open circles; o–o) compared to the IC₅₀ of β-epoxide 23 (closed squares; ••••) across the pH-range. (c) Labeling of GBA was blocked by incubating with irreversible inhibitor CBE 12, competitive reversible inhibitor AMP-DNM or by denaturing conditions (1% (w/v) SDS for 4 min at 100 °C) prior to labeling with β-epoxide 23. (d) Sensitivity of detection and labeling. GBA was incubated with an excess of β-epoxide 23 and dilutions applied on gel. (e) Fluorescent labeling of murine liver and duodenum lysates exposed to β-epoxide 23. Arrows indicate molecular weight of recombinant GBA (~59 kDa). Figure is a combination of excerpts from Chapter 1, see also ref281.
millions of molecules), or less (Fig 14d). Incubation of complex lysates such as homogenates from murine liver (Fig 14e) revealed that β-epoxide 23 labeled GBA highly specifically. This was observed in numerous lysates, except in those of small intestine, where β-epoxide 23 additionally bound to LPH, and fragments thereof. Importantly, cell-types as including hepatocytes, present in the liver, contain besides GBA also the aforementioned retaining β-glucosidases GBA2 and GBA3. However, these remained evidently unlabeled by β-epoxide 23. This was confirmed by separate enzyme assays, showing β-epoxide 23 very poorly inhibits GBA2 and GBA3 (IC50 > 100 μM, for both enzymes). The amphiphilic nature of β-epoxide 23 also allows passage of the ABP across membranes and results in efficient in situ labeling of active GBA molecules in living cells and mice.

Intriguingly, the attachment of the bulky BODIPY moiety at the C6 position in cyclophellitol β-epoxide 16 further enhanced and skewed the specificity of the inhibitor for GBA and not to GBA2, GBA3 or LPH. Albeit this finding was unexpected, this effect can be explained by the degree of acceptance toward substrates exhibited by the different retaining β-glucosidases expressed in man and mice. While GBA, GBA2, GBA3 and LPH are able to catalyze the hydrolysis of the glycosidic bond between β-D-glucose and an aglycone (e.g. ceramide or 4-methylumbelliferone), they are categorized into different CAZy glycosyl hydrolase families: GBA in GH30, GBA2 in GH116 and GBA3/LPH in GH1. The specificity of β-epoxide 23 for GBA is likely due to its partial resemblance of the structure of β-1,6-glucan, as can be seen in Fig 15a. Along with GBA, several retaining β-1,6-glucanases are categorized in GH30, indicating still some resemblance between GBA and this class of glucosidases is remaining in its structure and therefore bulky groups at C6 are readily accepted.

To create potent ABPs for all retaining β-glucosidases, an ABP structure was envisioned that resembled significantly more that of glucosylceramide (Fig 15a). For this purpose, conduritol β-aziridine 20 was identified as putative scaffold. The nitrogen of the aziridine can be acylated to yield an ABP structure containing a large bulky BODIPY group close to the C1 position, resembling the positioning of the ceramide aglycone in glucosylceramide. Moreover, the C6-position of conduritol β-aziridine 20 was changed from an exocyclic hydroxyl to an exocyclic hydroxymethyl, yielding cyclophellitol β-aziridine 25, which improved specificity similar as observed for the cyclophellitol β-epoxides 16 and 22–24. Acylation of the nitrogen in the aziridine with an alkyl-group allowed extension of the β-aziridine (probe 26), to yield both the green fluorescent β-aziridine 27 and a biotin-containing β-aziridine 28. Incubation of GBA in vitro with β-aziridine 27 revealed that it could still bind GBA at high pH values where the enzymatic activity was lost (Fig 15b). This indicated β-aziridine 27 can bind regardless of the action/state of the acid/base, in contrast to β-epoxides such as 23. The acid/base-independent binding mechanism of β-aziridine 27 was verified by investigations with GBA with mutated acid/base (E235) or nucleophile (E340). In absence of the acid/base, β-aziridine 27 still binds GBA, whilst β-epoxide 23 cannot (Fig 15c). The acid/base-independent binding mechanism of β-aziridines 25–28 is
Activity-based probes for glycosidases

Figure 15 | Broad-spectrum aziridine-type ABPs for retaining β-glucosidases. (a) Structure of fluorescent cyclorrhizol β-epoxide 23 aligned with those of β-1,6-glucan, glucosylceramide, cyclorrhizol β-aziridine 25 and novel β-aziridine probes 26–28. (b) Quantified labeling β-epoxide 23 (open circles; ○–○) and β-aziridine 27 (closed squares; ●–●), compared to GBA β-glucosidase activity profile towards artificial 4MU-β-D-Glc substrate across the pH-range (open triangles; △–△). Data are average of triplicates ± SD, normalized to conditions at pH 5.0. (c) Labeling of GBA active-site mutants. Over-expressed myc/His-tagged GBA, either as wild-type or lacking the acid/base (E235G or E235Q), or nucleophile (E340G or E340Q) labeled with β-aziridine 27 in total lysate (top), or isolated from labeling with β-epoxide 23 or β-aziridine 27 in total lysate after pull-down (middle), with α-myc as loading control (bottom). (d) Labeling of endogenously expressed retaining β-glucosidases in homogenates of murine brain and using β-epoxide 23 and β-aziridine 27. Molecular weights are marked of recombinant H. sapiens GBA (59 kDa, imiglucerase, asterisk; ●), endogenous murine GBA (circle; ●), GBA2 (open diamond; ○), GBA3 (square; △), LPH and its degradation fragments (triangle; ▲). Figure is a combination of excerpts from Chapter 2, see also ref282.
extraordinarily fast, hampering the determination of kinetic rate constants. As discussed below, β-aziridines 27 and 28 prove to be broad-spectrum ABPs, labeling also the retaining β-glucosidases GBA2, GBA3 and LPH (see Fig 15d)282 and from various other organisms, such as bacteria and fungi282. Other retaining glycosidases, including fungal cellulases or xylosidases, were not labeled by either ABP, indicating cross-species selectivity for retaining β-glucosidases282.

**Use of ABPs to investigate retaining β-glucosidases** – For some decades, GBA properties have been studied with irreversible inhibitors. For example, through labeling of GBA with β-epoxide 23 and subsequent MS/MS analysis, the catalytic nucleophile could be correctly identified as E340. Using the novel fluorescent β-epoxide- and β-aziridine ABPs in parallel, the acid/base and nucleophile residues of GBA can be easily mapped. GBA lacking the catalytic nucleophile E340 cannot bind fluorescent β-epoxide 23 or β-aziridine 27. Where labeling of GBA by β-epoxide ABPs requires activation through protonation by the acid/base that by β-aziridine 27 does not282. As a result, GBA with an inactivated acid/base residue E235 labels avidly with β-aziridine 27 but not with β-epoxide 23, as can be conveniently detected with SDS-PAGE and fluorescence scanning. In the case other retaining β-glucosidases than GBA are used, the catalytic nucleophile can be mapped by identification of the amino acid to which an epoxide or aziridine ABP is covalently bound, either directly by mass spectrometry or indirectly using site-directed mutagenesis of putative nucleophile residues. The acid/base residues of these enzymes can be identified using external nucleophiles such as azide, able to substitute a lacking acid/base and thus rescue activity and labeling by a β-epoxide ABP148−151. The latter can be very sensitively detected using SDS-PAGE and fluorescence scanning. In Chapter 4, this approach is applied to identify the key catalytic residues of all known retaining β-glucosidases in man. Another use of CBE 12 and other epoxide ABPs is to differentiate retaining β-glucosidases in complex (tissue) homogenates283, 284. Selective inhibitors have for example aided in the discovery of non-lysosomal β-glucosidase GBA272.

An important application for β-epoxide 23 and β-aziridine 27 is laid in the visualization and quantification of (active) GBA and other retaining β-glucosidases. Incubation of wild-type fibroblasts with fluorescent β-epoxide 25 results in *in vivo* labeling of active GBA and allows *in situ* detection with (time-lapse) confocal fluorescence microscopy (Fig 16a, next page). After harvesting of ABP-labeled cells and fixation, the pattern of *in vivo* labeled enzyme overlaps with GBA detected with monoclonal α-GBA antibody 8E4, albeit the latter is markedly less sensitive. In fibroblasts of Gaucher patients completely lacking (active) GBA, no labeling with β-epoxide 24 occurs281. β-Epoxide 23 selectively labeling active GBA can be used for diagnosis of Gaucher disease. Following incubation of cells or lysates thereof with 23, GBA molecules can be visualized by fluorescence scanning after SDS-PAGE. A clear inverse relationship between the cellular amount of labeled GBA and the severity of Gaucher disease has been observed for fibroblasts (Fig 16b).
The fact that GBA is labeled specifically by $\beta$-epoxides $23$ and $24$ can be exploited for determination of the half-life of GBA in living cells (Fig 16c). An overnight pulse with red $\beta$-epoxide $24$ followed by a continuous chase with green $\beta$-aziridine $27$, allows analysis of the rate of formation of de novo red $24$-labeled GBA. Similar results were obtained with the combination of green and red $\beta$-epoxides $23$ and $24$. In addition, one can follow the modifications in newly formed enzyme: newly synthesized GBA in the ER contains four high mannose-type N-linked glycans (apparent molecular weight ~62 kDa)$^{285-289}$, in the Golgi apparatus these are modified to sialylated complex-type glycans resulting in an increased molecular weight (62 to 66 kDa); finally, in the lysosome the glycans of GBA are trimmed by the sequential action of exoglycosidases resulting in a decrease in molecular weight (66 kDa reduced back to the bare 58 kDa peptide)$^{285-289}$.
The selectivity of CBE 12 for GBA has been employed to specifically ablate its activity in intact cells and animals, causing the accumulation of glucosylceramide and glucosylsphingosine. Incubation of macrophages with CBE 12 results in altered morphology and formation of glucosylceramide-accumulating Gaucher-like cells. Loading of CBE-treated macrophages with lysed red blood cells has been argued to closely mimic the formation of Gaucher cells. Inhibition of GBA by CBE has also been used to obtain insight in the role of glycosphingolipids in the epidermis, most particularly the importance of hydrolysis glucosylceramide to ceramide in the stratum corneum. Based on such studies, amongst others, GBA is thought to play a crucial role in maintaining the barrier function of the skin, e.g. prevent excessive water loss and absorption of exogenous molecules. The most severe Gaucher patients, so-called collodion babies, suffer from a compromised barrier function of the skin due to a (near) total absence of GBA activity.

CBE 12 has also been used to study the role of GBA in neuronal cells. Inhibition of GBA was found to result in significantly increased neuronal axon length and branching. Finally, CBE 12 has been employed to induce GBA deficiency in wild-type mice to attempt to generate a genuine Gaucher disease model. Indeed, repeated administration of CBE 12 results in accumulation of glucosylceramide, albeit much more moderate than observed in Gaucher patients. CBE 12-treated mice are short-lived as the result of lethal neurological complications. Nevertheless, CBE 12-treated mice have been used as a short-lived animal model to test gene therapy as treatment for Gaucher disease. Titration studies with CBE 12 revealed that GBA could be inhibited about 85% in mice, without causing Gaucher symptoms. This suggests that therapies who elevate GBA activity to at least ~15% of levels observed in control individuals, might render beneficial clinical responses. Such a partial increase in GBA activity is aimed at in current studies with small molecular chaperones, which may stabilize the ‘native’ fold of mutant GBA in the endoplasmic reticulum (ER), thereby reduce degradation by the ER-associated degradation system and thereby enhance GBA delivery to lysosomes.

A major drawback of the use of CBE 12 to induce Gaucher disease in animal models, is its ability to permeate into the brain and irreversibly inhibit GBA there, and as a consequence, induce fatal neurological complications. Cyclophellitol β-epoxide 16 has also been used to inactivate GBA in living cells and mice. Again, 16 also inhibits GBA activity in the brain of animals treated in vivo, and thus does not allow generation of a long-lived Gaucher mouse model. Neither a complete GBA knockout renders a completely genuine Gaucher model in mice, since it is embryonically lethal. More sophisticated inducible and/or cell-type specific deficiency of GBA and knock-ins of specific mutations occurring in Gaucher patients, have been employed to generate a variety of mouse models.

The newly developed β-epoxide 23 offers an exciting new possibility to generate a long-lived Gaucher mouse model with prominent visceral symptomatology. This is based on the
ability of β-epoxide 23 to inhibit GBA irreversibly in all tissues except for the brain, skin and eye. Intravenous administration of the β-epoxide 23 is well tolerated to dosages up to at least 1 nmol/mouse (20 μg kg⁻¹)282. Intravenous administration of wild-type mice with β-epoxide 23 results in specific, dose-dependent labeling of GBA in various organs (Fig 16d). Besides GBA, no other retaining β-glucosidases are labeled except LPH in the small intestine. Repeated intravenous administration to mice of β-epoxide 23 generates a Gaucher-like phenotype, exhibiting classic hallmarks such as splenomegaly, infiltration of tissue with macrophages expressing markers of Gaucher cells such as gpNMB, and increased tissue and plasma concentrations of glucosylceramide and glucosylsphingosine (see also Chapter 7). In the case of β-aziridine 27, intravenous administration results not only in prominent in vivo ABP-labeling, and thus inactivation, of GBA in various organs but also in that of GBA2 and GBA3, as well as LPH in the small intestine (Fig 16d). Again, β-aziridine 27 does not label β-glucosidases in brain and eye, most likely through its active removal by P-glycoproteins present in the blood brain barrier282.

Finally, ABPs such as β-epoxide 23 and β-aziridine 27 can be used to elegantly, fluorescently label therapeutic enzymes in vitro. The fate of pre-labeled enzymes following administration to cultured cells or after intravenous infusion into mice can be sensitively monitored. Enzyme therapy is presently the first choice treatment for type I Gaucher disease. This non-neuropathic variant of the disorder is thought to be a macrophage disorder in which glucosylceramide-laden phagocytes (Gaucher cells) underlie the visceral symptoms327–329. Prevention and/or removal of these lipid-laden macrophages is the rational treatment goal for type I Gaucher disease330,331. Brady and colleagues developed an enzyme therapy approach to specifically supplement visceral macrophages with exogenous GBA332,333, exploiting the presence of mannose-lectin(s) in this cell-type334–339. To date, several GBA preparations have been produced with exposed terminal mannose-moieties mediating recognition and uptake by mannose-lectin(s). Chronic two-weekly intravenous administration of therapeutic GBAs, such as alglucerase340, imiglucerase341, velaglucerase342 and taliglucerase343, results in impressive clinical responses of type I Gaucher patients344,345. Direct comparisons of the different therapeutic GBAs, including bio-distribution studies, are scarce342,346–351. Differential labeling of different therapeutic GBAs with ABPs with distinguishable tags, for instance with green β-epoxide 23 and red β-epoxide 24, and their equimolar administration to cell- and animal models enables a head-to-head analysis of the binding and uptake by cells and bodily distribution of these therapeutir GBAs, visualized by methods as SDS-PAGE, FACS, or fluorescence microscopy (see also Chapter 6). An exciting future development is the development of ABPs carrying specialized tags that allow non-invasive imaging and bodily distribution in animal models and individual Gaucher patients, when trace-labeled onto therapeutic enzymes.

Another potential future application of ABPs may be laid in the development of more potent chemical chaperones to stabilize the ‘native’ fold(ing) and lysosomal delivery of mutant GBA enzymes in Gaucher patients. Semi-irreversible ABPs such as 2-deoxy-2-
fluoro-D-glucosides, equipped with optimized leaving groups as discussed previously appear attractive. These types of ABPs should only transiently bind to the nucleophile of GBA, thereby promoting the critical process of folding in the ER and sorting to lysosomes. Inside lysosomes the rescued enzyme should be reactivated for glucosylceramide hydrolysis, following the slow irreversible release of 2-deoxy-2-fluoro-D-glucose (see Chapter 5). Obviously, the use of appropriate concentrations of such probes is of eminent importance: too little probe will be without effect and too much will backfire by inhibiting all residual enzyme activity. Animal studies will have to reveal whether an optimal concentration of such chemical chaperons can be reached to in order to result in a sustained clinical benefit.

**Concluding remarks**

Recently designed activity-based probes for retaining β-glucosidases offer extraordinarily versatile tools that can be broadly applied. Applications are found in fundamental characterization of the catalytic mechanism of these enzymes, diagnosis of patients suffering from enzyme deficiencies, determination of the in situ half-life of individual enzymes, their processing and subcellular trafficking, induction of long-lived Gaucher-disease in animal models and the in vivo monitoring of enzyme therapy efficacy in individual patients by imaging of therapeutic enzyme trace-labeled with ABPs. The potential application of semi-irreversible activity-based probes as transient chemical chaperones promoting successful folding in the ER of mutant GBA and consequently improving cellular enzymatic activity warrants further investigation.
References


Introduction, part I


Introduction, part I

5719–5722.
Activity-based probes for glycosidases


Introduction, part I


144. Withers SG, et al. (1990) Unequivocal demonstration of the involvement of a glutamate residue as a


Biochim Biophys Acta. 1758, 1864–1884.


