PERSPECTIVE

Irreversible inhibitors and activity-based probes as research tools in chemical glycobiology

Martin D. Witte,[a] Gijsbert A. van der Marel,[b] Johannes M. F. G. Aerts[c] and Herman S. Overkleeft*[b]

In this review, we will discuss the enzymes that are involved in the synthesis and degradation of glycoconjugates and we will give an overview of the inhibitors and activity-based probes (ABPs) that have been used to study these. Following discussion of some general aspects of the biosynthesis and degradation of *N*-linked glycoproteins, attention is focused on the enzymes that hydrolyze the protein–carbohydrate linkage, peptide *N*-glycanase and glycosylasparaginase and their mechanism. We then focus on the biosynthesis of *O*-linked glycoproteins and glycolipids and in particular on the enzymes that hydrolyze the interglycosidic linkages in these, the glycosidases. Some important mechanism-based glycosidase inhibitors that form a covalent bond with the targeted enzyme(s), their corresponding ABPs and their application to study this class of enzymes are highlighted. Finally, alternative pathways for degradation of glycoconjugates and an ABP-based strategy to study these will be discussed.

Glycoconjugates are a highly diverse class of biomolecules that partake in many biological processes.[1] Glycoconjugates can be

aWhitehead Institute For Biomedical Research, 9 Cambridge Center, Cambridge, MA, 02142, USA
bLeiden Institute of Chemistry and the Netherlands Proteomics Centre, Leiden University, P.O. Box 9502, 2300 RA, Leiden, The Netherlands. E-mail: h.s.overkleeft@lic.leidenuniv.nl; Fax: (+31) 71 527 4307
cDepartment of Medical Biochemistry, Academic Medical Center, Meibergdreef 15, 1105 AZ, Amsterdam, The Netherlands

divided into three major types, namely *N*-linked glycoproteins, *O*-linked glycoproteins and glycolipids. The β-glycosyl–asparagine amide bond, formed by post-translational modification of the side chain of asparagine, is the major type of glycosidic linkage in *N*-linked glycoproteins.[2] The linkages found in *O*-glycoproteins and glycolipids are much more diverse, and threonine, serine, tyrosine, hydroxylysine and hydroxyproline are subject to modification with monosaccharides or oligosaccharides of various natures.[3] Glycolipids in turn differ in both the lipid part (for instance

Martin Witte

Martin Witte (14-09-1981) received his MSc in chemistry from Leiden University, the Netherlands, in 2004. During his MSc studies, he performed a research internship in the Bio-organic synthesis group of Prof. Jacques van Boom (Leiden University) and in the group of Prof. Matthew Bogyo (Department of Pathology, Stanford University). In 2004, he started his graduate research under the supervision of Prof. Gijsbert van der Marel and Prof. Herman Overkleeft (Leiden University) and in 2009 he obtained his PhD in chemistry cum laude. After a postdoctoral period at Leiden University (Prof. Herman Overkleeft, 2009–2010), he recently moved to the Whitehead Institute for biomedical research to do postdoctoral research in the group of Prof. Hidde Ploegh.

Gijsbert van der Marel

Gijsbert van der Marel (07-04-1952) performed his PhD research in the group of the late professor Jacques van Boom on the subject of synthetic nucleic acid chemistry. Upon graduating in 1982 he pursued his independent career at Leiden University where he was appointed Professor in Organic Chemistry in 2004. His current research interests are in organic synthesis with a focus on nucleic acids and carbohydrates.

cholesterol, ceramide) and the nature of the (oligo)saccharide fragment. Aberrations in either the nature of glycoconjugates or their metabolism are at the basis of a variety of diseases, and thus insight in glycoconjugate processing and functioning is of utmost importance in understanding and combating these diseases. For instance, altered glycoconjugate patterns are often found on human tumors and provide an entry for anti-tumor drug development.[4] Aberrations in glycolipid metabolism or the degradation of *N*-glycosylated proteins are at the basis of a wide number of human genetic disorders called lysosomal storage disorders.[5]

In the study of glycoconjugate processing enzymes in biological samples, mechanism-based inhibitors and activity-based probes (ABPs) are increasingly applied as research tools. In this review, we will discuss the enzymes that are involved in the synthesis and degradation of glycoconjugates and we will give an overview of the inhibitors and ABPs that have been used to study these. Following a short introduction on the concept of activity-based protein profiling some general aspects on the biosynthesis and degradation of *N*-linked glycoproteins are discussed. Attention is focused on the enzymes that hydrolyze the protein–carbohydrate linkage, peptide *N*-glycanase and glycosylasparaginase and their mechanism. The next part of this review focuses on the biosynthesis of *O*-linked glycoproteins and glycolipids and in particular on the enzymes that hydrolyze the interglycosidic linkages in these, the glycosidases. Some important mechanism-based glycosidase inhibitors that form a covalent bond with the targeted enzyme(s), their corresponding ABPs and their application to study this class of enzymes are highlighted. Finally, alternative pathways for degradation of glycoconjugates and ABP-based strategies to study these will be discussed.

Activity-based protein profiling

In activity-based protein profiling (ABPP), activity-based probes (ABPs) are employed to specifically modify an enzyme or a class of enzymes in a complex sample (cell extract, living cells or animal models). In the next step, the tagged enzymes are identified and/or analyzed by fluorescence read out or mass spectrometry-based proteomics techniques (Fig. 1A).[6] An ABP generally consists of three structural elements, being a reactive group (also known as the warhead), a linker and a tag/ligation handle, or label (Fig. 1B). The warhead covalently attaches the probe to the enzyme (or enzyme family) of interest and should only react with active enzymes; hence the name activity-based probe. The design of a reactive group is thus guided by the mechanism of the enzyme. Often an electrophilic group is used as a warhead, which selectively reacts with the nucleophilic residue in the active site. The linker connects the warhead to the label and thereby introduces additional spacing between both functional groups. This spacing minimizes the steric hindrance between the reporter group and active site of the enzyme in order to disturb as little as possible the binding of the probe. Additionally, the linker moiety of the probe can be adapted to provide selectivity for the enzyme (or enzymes) of interest, to increase the hydrophilicity of the ABP and for quantification of the labeled proteins by mass spectrometry. For example, introduction of structural elements in the linker part, such as peptides, resulted in specific protease probes.[7] PEG-spacers have been used to increase the solubility of ABPs. For quantification purposes, isotope coded linkers may be incorporated in ABPs.[8] Such probes contain either H/D or ^{12}C/^{13}C and share almost the same physical and chemical properties. After labeling and digestion of two different protein samples, the labeled peptides will therefore co-elute in LCMS experiments. By comparing the relative signals of these peptides, the relative abundance of the protein in the two different samples can be determined. The tag is used to visualize and/or purify the labeled enzymes. Both affinity tags, such as biotin, and fluorophores, such as a BODIPY,[9] a fluorescein or a rhodamine moiety, are frequently applied in activity-based proteomics. Incorporation of a fluorophore tag is an attractive strategy to allow rapid in-gel detection of labeled proteins (Fig. 1A, path i). Furthermore, such ABPs can be used to study enzymes in living cells by means of fluorescence imaging microscopy and fluorescence activated cell sorting. Biotin is often used to enrich labeled proteins by means of streptavidin pull-down after which the enriched proteins are identified by mass spectrometry (Fig. 1A, path ii).[10] Alternatively, the biotin moiety allows in-gel detection by streptavidin-horseradish peroxidase

Hans Aerts

Hans Aerts (30-10-1956) received his PhD degree on the molecular basis of Gaucher disease at the University of Amsterdam supervised by Professor J. M. Tager and Professor J. A. Branger (USC) in 1989. Since then he has been employed at the Academic Medical Centre, University of Amsterdam, where he is now Professor in Medical Biochemistry and Head of the Biochemistry Department. His current research interests focus on glycosphingolipids in health and disease and cover inherited lysosomal storage disorders and the acquired metabolic syndrome.

Herman Overkleeft

Herman Overkleeft (12-04-1969) performed his graduate research at the University of Amsterdam under the guidance of Prof. Upendra Pandit. After postdoctoral research at Leiden University (Prof. Jacques van Boom, 1997–1999) and Harvard Medical School (Prof. Hidde Ploegh, 1999–2001) he was appointed Professor in Bioorganic Chemistry at Leiden University. He currently pursues chemical biology research in relation to glycolipid metabolism and antigen processing.

This journal is © The Royal Society of Chemistry 2011
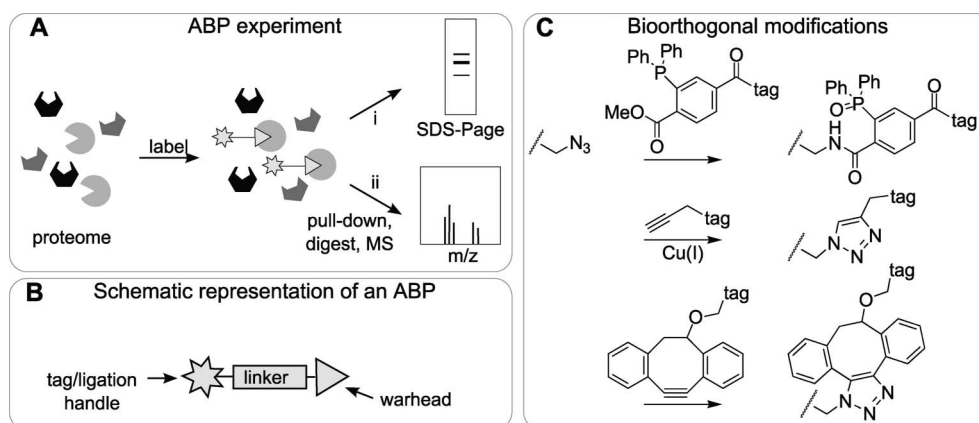
Org. Biomol. Chem., 2011, **9**, 5908–5926 | 5909

Fig. 1 (A) In a typical ABP labeling experiment, a proteome is treated with an activity-based probe. The labeled enzymes are either directly visualized with SDS-PAGE followed by fluorescent imaging (i) or are purified by streptavidin pull-down, digested with trypsin after which the peptides are analyzed by mass spectrometry (ii). In a two-step labeling experiment, a bioorthogonal ligation is performed after treating the proteome with the ABP. (B) Schematic representation of an ABP. (C) Bioorthogonal reactions used to modify azide containing ABPs. Top: Staudinger–Bertozzi ligation, middle: copper-catalyzed click reaction, bottom: strain-promoted click reaction.
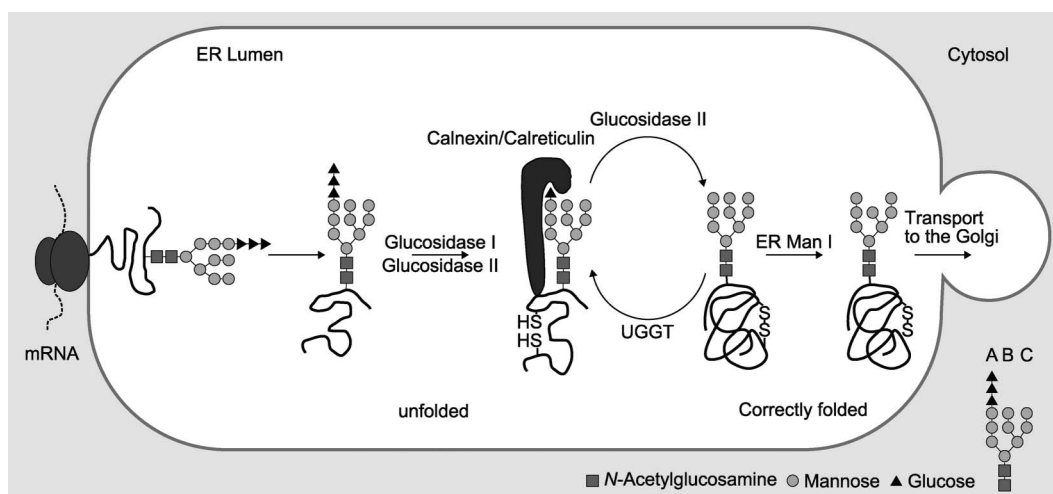

Fig. 2 The synthesis of glycoproteins.

blotting. On the downside of the ABPs described so far, the inherent steric bulk of a tag (biotin/fluorophore) may obstruct binding of the ABP to the targeted enzyme(s). Next to this the tag, especially biotin, may have a detrimental effect on the cell-permeability of the probe. Obviously, biotin or fluorescent tags are normally not part of the natural substrate of the targeted enzyme(s), and ABPs equipped with these may be structurally rather removed from the actual enzyme substrates. This poses some difficulties on the translation of data on enzyme–ABP binding to substrate preferences as exerted by the enzyme or enzymes at hand. With the specific aim to overcome these caveats, two-step labeling strategies have come to the fore in recent years. In these strategies, the tag in an ABP is replaced by a small bioorthogonal ligation handle, most commonly an azide or alkyne. After labeling, these ligation handles can be conjugated to the tag using the Staudinger–Bertozzi ligation, the copper-catalyzed click reaction or the strain promoted click reaction (Fig. 1C).[11] The implementation of ABP tools has met with considerable success in the areas of esterases and proteases, both in the study of known entities in physiological processes and in unearthing new hydrolytic activities belonging

to these classes. Recent years have witnessed the first successful development of ABP tools and techniques aimed at enzyme classes other than esterases/proteases, in particular also the subject of this review, glycosidase activities.

The biosynthesis of *N*-linked glycoproteins

In eukaryotes, the majority of *N*-linked glycoproteins are synthesized by membrane-bound ribosomes on the endoplasmic reticulum (ER) (Fig. 2). The newly synthesized proteins are inserted into the ER lumen *via* a specialized structure referred to as the translocon.[12] In a co-translational process, oligosaccharyl transferase may glycosylate asparagine residues within the Asn–Xxx–Ser/Thr (Xxx = any amino acid excepting Pro) consensus sequence forming an *N*-linked glycan.[13] Glycosylation of the growing peptide chain increases the hydrophilicity of the unfolded peptide and thereby prevents its aggregation. The outermost glucose residue of the $Glc_3Man_9GlcNAc_2$ *N*-linked oligosaccharide is rapidly removed by glucosidase I, which is followed by hydrolysis of the second glucose residue by glucosidase

II. The chaperones calnexin and calreticulin bind to the formed GlcMan$_9$GlcNAc$_2$ and aid in folding of the peptide chain.[14] Protein disulfide isomerase binds to these chaperones and catalyzes the formation of disulfide bonds.[15] Upon release of the protein from calnexin and calreticulin, glucosidase II cleaves the inner most glucose. In mammalian cells, the folding state of the protein is monitored by UDP-glucose:glycoprotein glucosyltransferase (UGGT), which recognizes hydrophobic patches.[16] Misfolded proteins are re-glucosylated and subsequently re-enter the calnexin/calreticulin cycle. Properly folded proteins progress through the ER and Golgi. A series of deglycosylation/glycosylation events transform the high mannose N-glycans into complex-type N-glycans. These glycans in turn help in guiding the glycoprotein to its final destination, such as the cell surface or the endocytic pathway. There are two main degradation pathways for N-linked glycoproteins. The folding state and location of the glycoprotein determines the pathway by which it is degraded. Newly synthesized N-glycoproteins that are persistently misfolded in the ER are degraded by the endoplasmic reticulum associated degradation pathway (ERAD). Mature proteins are, after performing their function, finally degraded in the lysosome.

Endoplasmic reticulum associated degradation pathway

Up to 10% of the newly synthesized glycoproteins in the ER are persistently misfolded and are destined for degradation.[17] Depending on region that is misfolded, the proteins will be recognized and extracted by a different subset of proteins. For glycoproteins, it was shown that the N-glycans play an important role in this ingenious process. In contrast to normally folded glycoproteins, the N-glycans on misfolded proteins are extensively trimmed (Fig. 3) in the ER and it was therefore long thought that the N-glycans acted as a folding timer. Hydrolysis of the α-1-2 linked mannose of branch A blocks reglucosylation of the glycan by UGGT[18] and thereby prevents futile folding attempts. Furthermore, the α-1-6 linked mannose residue formed upon trimming of branch C either in combination with trimming of the branch B[19] or as sole modification[20] serves as a signal for degradation. The enzyme that removes the mannose residue of branch B was identified as ER α-mannosidase I.[21] Which mannosidases are involved in the trimming of the branches A and C is still a matter of debate. The ER degradation enhancing α-mannosidase I like proteins 1–3 (EDEMs 1–3) have been suggested as potential candidates. Despite their resemblance with ER α-mannosidase I, it was originally thought that the EDEMs act as lectins which specifically recognize trimmed N-glycans.[22] Olivari et al. however showed that overexpression of EDEM1 resulted in the removal of the α-1,2-linked mannose residue of branch A in vivo thereby accelerating degradation of the glycoproteins[18] and similar results were obtained for overexpression Htm1p, the yeast homologue of EDEMs.[19] Additionally, overexpression of EDEM3 leads to the formation of Man$_7$GlcNAc$_2$ and Man$_6$GlcNAc$_2$ in vivo.[23] The influence of EDEM1 and EDEM3 on trimming was abolished by mutating the putative catalytic residues. This data points to the EDEMs as being active mannosidases. However data of another study using human EDEM1 suggest that instead of being a mannosidase, EDEM1 downregulates the proteolytic degradation of ER α-mannosidase I and that the resulting increased levels of ER α-mannosidase I account for the extensive trimming of N-glycans.[24] The lectin, human OS-9, specifically binds to the α 1–6 linked mannose residue of branch C of the trimmed glycan.[25] This lectin forms a large protein complex that includes an ubiquitin ligase.[26] This complex is thought to be responsible for ubiquitination and retrotranslocation of the glycoproteins to the cytosol. Inside the cytosol, the glycoproteins are deglycosylated by cleavage of the β-aspartyl-glucosamine bond by cytoplasmic peptide N-glycanase (PNGase) and degraded by the proteasome and aminopeptidases.

PNGase belongs to the transglutaminase family and was first detected in yeast.[27] After identification of the gene encoding PNGase in yeast, the function and structure of this enzyme has received a lot of attention.[28] X-ray crystallography showed that the catalytic residue of PNGase is located in the middle of a long deep cleft (8 Å wide and 30 Å long).[29] The carbohydrate part of a glycoprotein binds to one end of the cleft and the peptide part binds at the other side of it. To bind efficiently to PNGase, the folding state of the substrate is of importance as was demonstrated in 2004 by Hirsch et al. in an activity-based assay. They found that RNAse B has to be denatured to enable deglycosylation by PNGase.[30] Computer models of native glycosylated proteins bound to the Röntgen diffraction structure of PNGase illustrate why proteins need to be unfolded before they can bind to PNGase. The walls of the deep cleft in which the active site is located obstruct binding of native N-linked glycoproteins. The globular structure of unfolded proteins, on the other hand, does fit with the requirements of the active site and can therefore bind to PNGase. Upon binding of the substrate, the carbohydrate-protein linkage is hydrolyzed as depicted in Fig. 4. Cysteine 191 of the catalytic Cys, His, Asp triad, characteristic for many cysteine proteases/transglutaminases, attacks the amide bond forming tetrahedral adduct **1**. Thio-ester **2** is formed upon collapse of oxyanion **1** and the oligosaccharide is liberated. Subsequently, thio-ester **2** is hydrolyzed, regenerating the active enzyme. The Röntgen diffraction structure of yeast PNGase
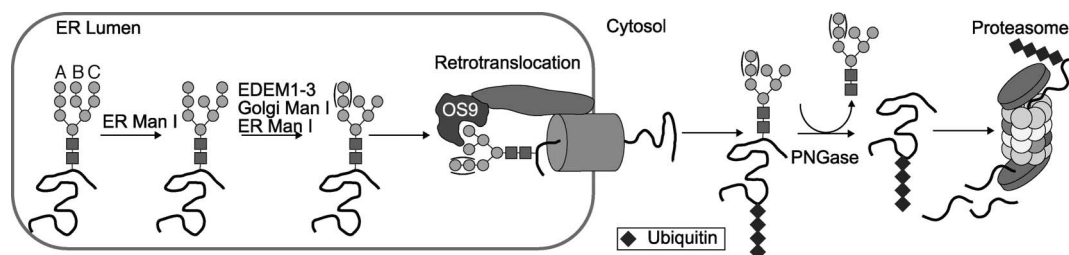

Fig. 3 Schematic representation of ERAD.

This journal is © The Royal Society of Chemistry 2011

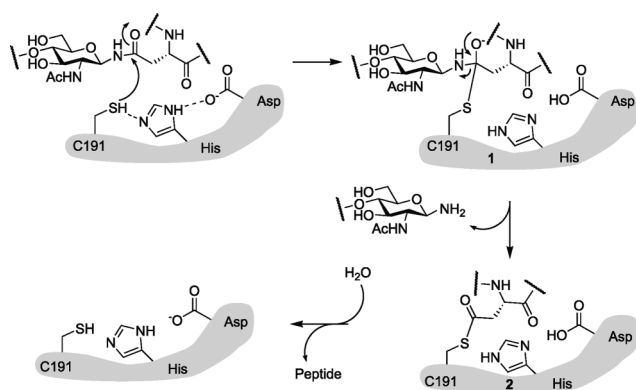*Org. Biomol. Chem.*, 2011, **9**, 5908–5926 | 5911

Fig. 4 Mechanism of peptide *N*-glycanase.

furthermore reveals that the N-terminal and C-terminal parts of the yeast protein form a hydrophobic patch which interacts with the ubiquitin receptor Rad23, coupling yeast PNGase to the proteasomal pathway.[31] In higher organisms, PNGase contains an N-terminal putative protein–protein interaction domain (PUB-domain).[32] Next to Rad23 homologues, various other proteins have been reported to form a complex with these patches, including the ATPase Associated with diverse cellular Activities which is possibly involved in the extraction of misfolded proteins from the ER and ubiquitin ligases which are responsible for ubiquitination of misfolded proteins

PNGase inhibitors and activity-based probes

Broad-spectrum caspase inhibitor Z-VAD-Fmk **3** (Fig. 5) was the first inhibitor reported for PNGase and was discovered by library screening.[33] Misaghi *et al.* revealed that this peptidyl fluoromethyl ketone covalently and irreversibly modified PNGase, as the activity of PNGase inhibited with **3** could not be restored by dialysis. To study the site of binding of Z-VAD-Fmk **3**, an active site mutant with the nucleophilic cysteine mutated to an alanine residue was used. According to MALDI-MS, enzymes lacking this active site cysteine residue were, in contrast to wild-type enzyme, not labeled by Z-VAD-Fmk. These results indicate that the inhibition of PNGase is caused by the selective modification of the thiol of cysteine 191. Later, the binding-site of **3** was validated by the X-ray structure of **3** with PNGase.[25] To determine which structural parts of the inhibitor are important for binding, Misaghi *et al.* synthesized diastereomerically pure Z-VAD-Fmk **3a** and **3b** and Z-VAD-Fmk analogues **4–6**. The stereochemistry of the aspartic

acid residue in Z-VAD-Fmk (commercial available Z-VAD-Fmk is a diastereomeric mixture) proved not important for binding, and both isomers **3a** and **3b** inhibit the enzyme with equal efficiency. The position of the electrophilic trap and the aspartyl side chain at position 1 however plays a key role in inhibition of PNGase. Analogues bearing a fluoromethyl ketone on their side chain (**4** and **5**) did not inhibit the enzyme. Compounds lacking the carbonyl at their side chain such as Z-VAA-Fmk **6** or compounds in which the aspartyl side chain is replaced by a glutamyl side chain such as Z-VAE(OMe)-Fmk **7** were also inactive. A main and obvious disadvantage of Z-VAD-Fmk as PNGase inhibitor is its intrinsic reactivity towards caspases.

In the search for selective inhibitors, attention has been focused on the design of PNGase inhibitors based on the carbohydrate part of *N*-glycoproteins. Initially, Ito and co-workers designed a set of five inhibitors (**8–12**, Fig. 6).[34] The high-mannose-type oligosaccharides **8** and **9** equipped with an iodoacetamide trap are mimics of the high mannose *N*-glycans of the natural substrate. Inhibitors **10–12** contain the core chitobiose as recognition element. In a substrate-based assay, it appeared that oligosaccharides **8** and **9** are very potent inhibitors of PNGase (IC$_{50}$ ~1.6 μM). To evaluate whether **8** and **9** covalently modified PNGase, Ito and co-workers performed a labeling experiment. SDS-PAGE revealed a distinct shift in molecular weight of PNGase treated with **8** and **9**. Furthermore, they could visualize the glycosyl–PNGase adduct using lectin blotting. Labeling proved to be very selective and caspases 2, 3 and 7 and bovine serum albumin were not modified with **8** and **9**. Also in *E. coli* extracts, exclusive labeling of PNGase was observed. Mass spectrometry analysis clearly revealed Cys191 of the catalytic triad as the target of the inhibitors. Interestingly, truncated analogues such as disaccharides **10**, **11** and **12** also inactivated PNGase efficiently (IC$_{50}$ 4, 0.1 and 0.1 μM), despite their reduced binding affinity. Shortly after these studies, we published two new chitobiose-based inhibitors of PNGase, **13** and **14**.[35] Of these, epoxysuccinate **13** is a potent inhibitor of PNGase and the potency of this compound is comparable to chloroacetamide **10**. Fluoromethylketone **14** however proved to be a poor inhibitor of PNGase that blocks activity at high micromolar concentrations. These results prompted the study towards structural requirements essential for binding to PNGase. To establish the influence of the reactive group on the binding affinity, we synthesized a set of fifteen chitobiose-based inhibitors containing a variety of warheads, including methylketones/acetamides equipped with a leaving group, epoxides, aziridines and Michael acceptors (**15–29**, Fig. 6).[36] After biological evaluation of this set of inhibitors, it
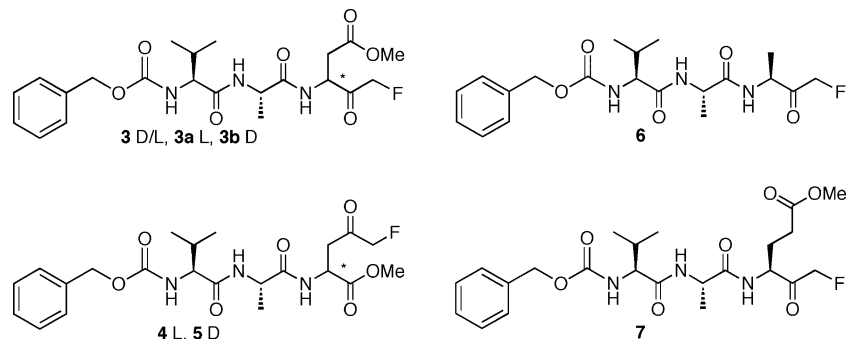
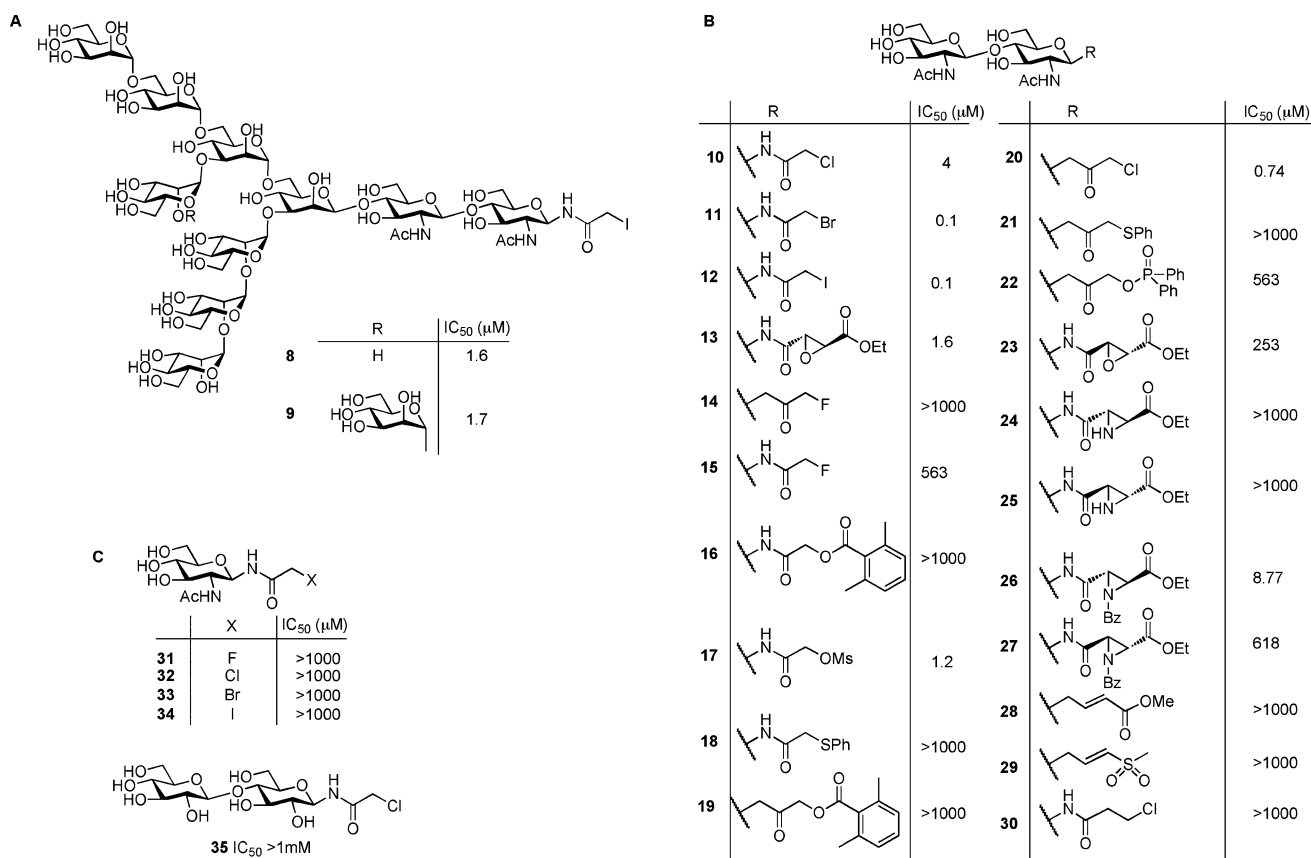

Fig. 5 Peptide based inhibitors of PNGase.

Fig. 6 A) High mannose based inhibitors of PNGase. (B) Chitobiose-based inhibitors. (C) Monosaccharide and cellobiose based inhibitors.

became evident that the nature of the warhead is crucial for the potency of the inhibitor. In general, three trends can be observed. (1) Acetamides/methylketones inhibitors bearing a good leaving group (**17** and **20**) are superior to inhibitors containing a poor leaving group (**15**, **16**, **18**, **19**, **21** and **22**). (2) The potency of *S,S*-configured epoxysuccinate/aziridine dicarboxylate warhead containing compounds such as **13** and **26** is comparable to that of chloroacetamides/chloromethylketones and these compounds are approximately 100-fold more active then compounds containing the corresponding *R,R*-configured warheads (**23** and **27**). Furthermore, acylation of the aziridine warhead is a prerequisite. (3) Finally Michael acceptors, such as vinyl sulfone **28** and vinyl ester **29** appeared to be poor inhibitors of PNGase. Next to the nature of the warhead also the location thereof plays a vital role as was shown by Ito and co-workers.[37] In contrast to good inhibitor **10**, chloropropionamide **30** in which the leaving group is shifted one carbon atom do not inhibit PNGase. In the same paper,

Ito and co-workers investigated the influence of the carbohydrate part on binding with monosaccharide inhibitors (**31–34**, Fig. 6) and cellobiose inhibitor **35**. At least two glucosamine residues are required for binding, since monosaccharides **31–34** do not inhibit the enzyme, this in contrast to chitobiose inhibitors **10–12** and **14**. Furthermore, the importance of the *N*-acetyl group was demonstrated. Compounds lacking the *N*-acetyl, such as disaccharide **35**, are poor inhibitors.

The peptide and carbohydrate-based inhibitors described above have been used as lead for the design of two ABPs for PNGase. ABP **36** was obtained by replacing the benzyloxycarbonyl group in broad-spectrum caspase inhibitor Z-VAD-Fmk **3** for a BODIPY TMR fluorophore (Fig. 7).[30] Chitobiose-based inhibitor **10** has been converted to activity-based PNGase probe **37** by the introduction of a BODIPY fluorophore at the 4-OH of the non-reducing GlcNAc.[38] Both peptide-based ABP **36** and carbohydrate based ABP **37** label purified PNGase and labeling
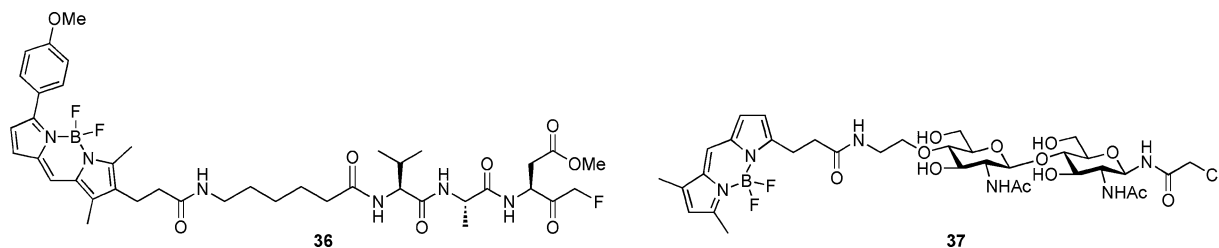

Fig. 7 Activity-based probes for peptide *N*-glycanase.

This journal is © The Royal Society of Chemistry 2011

Org. Biomol. Chem., 2011, **9**, 5908–5926 | 5913

can be completely abolished by pre-incubation with Z-VAD-Fmk. Labeling of *E. coli* cell-extracts overexpressing yeast PNGase with either probe resulted in a single band in the fluorescence image. Furthermore, ABP **37** is cell-permeable and inhibits PNGase *in vivo* as was evidenced in a substrate-based assay.

Lysosomal degradation of *N*-linked glycoproteins: glycosylasparaginase and its inhibitors

Mature *N*-linked glycoproteins are degraded in the lysosome in a bidirectional process. The non-reducing end carbohydrates of complex *N*-glycans are removed stepwise by the exo-glycosidases present in the lysosome (Fig. 8). Classification of glycosidases is based on their substrate, mode of action and mechanism. These enzymes will be further discussed in the second part of this review.
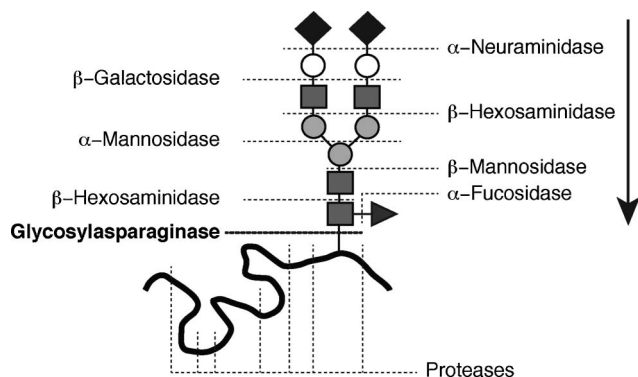

Fig. 8 Bidirectional degradation of *N*-linked glycoproteins. The glycan is degraded starting from the non-reducing (direction denoted by the arrow). In the final step, the glycosyl–asparagine linkage is hydrolyzed by glycosylasparaginase (bold). Carbohydrates are indicated as follows: D-sialic acid (black diamond), D-galactose (clear circle), D-*N*-acetylglucosamine (grey square), D-mannose (grey circle) and L-fucose (grey triangle).

Simultaneous to the degradation of the carbohydrate part, the protein is disassembled by proteinases such as cathepsins A, B, C, D, H, and L.[39] The final step in the degradation is hydrolysis of the protein–carbohydrate linkage by glycosylasparaginase (GA).

This amidase belongs to the N-terminal nucleophile hydrolase superfamily and is produced as zymogen. Autoproteolitic cleavage of a distorted scissile bond in the precursor liberates the nucleophilic N-terminal threonine residue located near the top of the funnel-shaped active site.[40] The binding-pockets for the α-amino and α-carboxyl group of the asparagine residue are deep in the active site of GA and binding of (*N*-GlcNAc)Asn and analogues thereof to these pockets has been studied in detail. Prerequisite for binding of substrates to GA is complete degradation of the peptide part as was revealed by Risely *et al.*[41] Substrates wherein the carboxylic acid was altered were not hydrolyzed by GA, clearly indicating the importance of the carboxylic acid in binding. Furthermore, it was shown that the α-amine group acts as an anchor and could be replaced by non-polar groups with a similar size. Modification of the carbohydrate part is well tolerated by the funnel shaped active site. A wide variety of asparagine analogues have been synthesized in which the GlcNAc moiety is replaced by various moieties, including carbohydrates (other than GlcNAc), amino acids and methylcoumarine. All these substrate analogues are hydrolyzed by GA. Removal of remaining α-1,6-fucose residues by fucosidases however is essential, since these residues obstruct binding of the substrate to GA.[42] The mechanism of GA was unravelled with a combination of crystallographic studies[43] using mutated enzyme and kinetic studies[44] employing analogues of (*N*-GlcNAc)Asn in which the carbohydrate part is substituted with a 4′-substituted-aniline. The following mechanism was established (Fig. 9). The α-amino group of the catalytic residue acts as a base and polarizes the nucleophilic β-hydroxyl through the side chain of a threonine residue 170 (GA from *F. meningosepticum*). Nucleophilic attack on
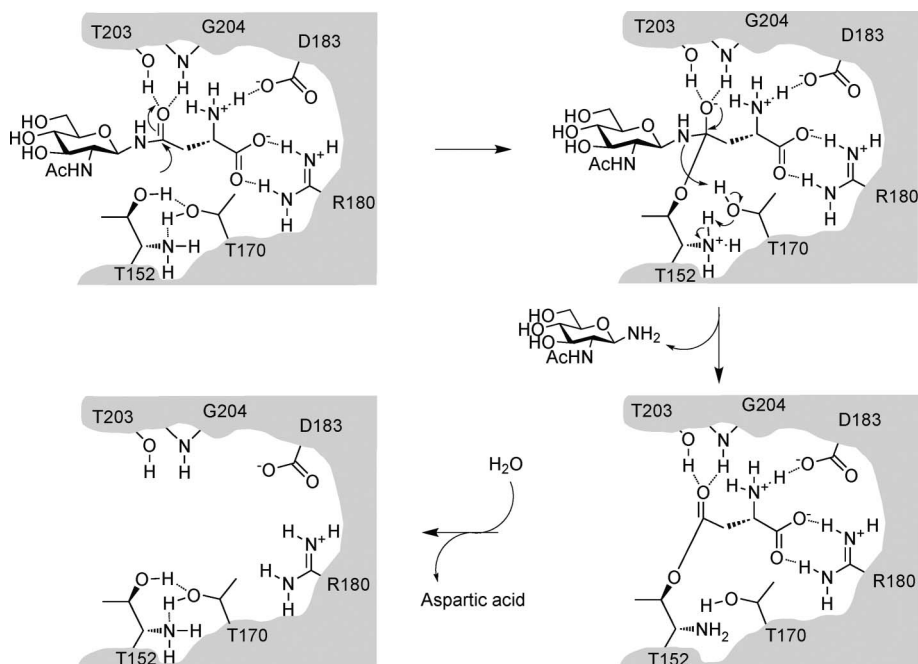

Fig. 9 The hydrolysis of glycosyl-asparagine by *F. meningosepticum* glycosyl-asparaginase.

the amide bond causes the formation of a tetrahedral intermediate, which is stabilized by the oxyanion hole. Breakdown of the tetrahedral adduct liberates *N*-acetyl glucosamine and forms β-aspartic acid ester, which in kinetic studies turned out to be the rate limiting step. Hydrolysis of the ester regenerates the active enzyme. The only reported mechanism based inhibitor of GA so far, 5-diazo-4-oxo-L-norvaline (**38**), covalently modified the active site residue forming an ether-bond (Fig. 10).[45] With the mechanism of the enzyme known in detail however the design of ABPs should be feasible. For instance, the mechanism of action of the enzyme rather resembles that of another N-terminal threonine hydrolase, the proteasome. Proteasomes in turn have been subject to numerous ABPP studies and a variety of electrophilic traps with which proteasome active sites are efficiently modified are known. These include vinyl sulfones and epoxyketones and grafting these moieties onto GlcNAc-Asp may be an entrance to the design of GA ABPs.
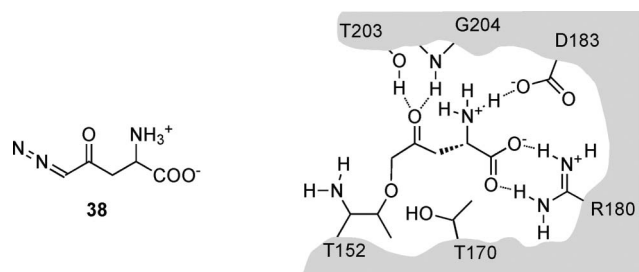

Fig. 10 Mechanism based inhibitor of glycosyl asparaginase.

Biosynthesis of *O*-glycoconjugates

O-glycoproteins and their structurally related lipid counterparts, the glycolipids constitute a large and diverse class of glyco-conjugates. Glycosyltransferases catalyze the synthesis of these glycoconjugates by transferring the carbohydrate from a glycosyl donor to their substrate, the acceptor.[46] In mammals, two main folds have been observed for glycosyltransferases, GT-A and GT-B, which are both Rossmann type.[47] This fold is characteristic for nucleotide binding enzymes and it accommodates the nucleotide diphosphate leaving group of the donor, the most common leaving group in mammals. Examples of donors containing this leaving group are UDP-α-Glc **39**, UDP-α-Gal **40**, UDP-α-GlcNAc **41** and GDP-α-Man **42**. Sialyl transferases employ

nucleotide monophosphates such as CMP-β-Neu5Ac **43** (Fig. 11). Departure of the leaving group, which is often facilitated by a divalent cation located at the active site is followed by the transfer of the carbohydrate. Depending on the enzyme, glycosylation can happen either with inversion or retention of the anomeric centre. In inverting transferases, the leaving group is replaced in a direct S_N2 like displacement. The mechanism of retaining glycosyltransferases is less clear. A double displacement mechanism has been proposed. First, a nucleophilic residue of the enzyme would replace the leaving group forming a covalent glycosyl–enzyme complex. In the second step, the acceptor reacts with the formed adduct. Both the lack of a generally conserved nucleophilic residue at the active site of glycosyl transferases and the fact that no covalently linked glycosyl–enzyme complex has been observed to date led to the postulation of a second mechanism. In this mechanism, the leaving group is replaced in an S_N1 like fashion, with the reactive species being a short-lived oxocarbenium ion. Recently, glycosyl transferases that use donors containing an alternative leaving group (lipid phosphates and unsubstituted phosphates) have been reported and some of these enzymes do not have the typical Rossmann fold.[48] The acceptor is often a hydroxyl of another carbohydrate but it can also be a lipid, protein, nucleic acid and a whole variety of small molecules. In enzymes with the GT-A fold, the C-terminus is highly variable and associated with the accommodation of the acceptor.[49] Within the GT-B fold, the N-terminus is variable and therefore believed to be in involved with the recognition of substrate.[50]

Degradation of *O*-glycoconjugates

The glycosidic bond in *O*-glyconjugates is hydrolyzed by glycoside hydrolases, also known as glycosidases.[51] This large class of enzymes comprises up to 1% of the genome. Various attempts have been made to classify glycosidases. A first classification is based on the substrate specificity of the enzyme. For example, β-glucosides are the optimal β-glucosidase substrates. This specificity allows differentiation in various classes. Although being the simplest classification, it has some disadvantages. Some glycosidases are capable of hydrolyzing several substrates and, furthermore, structurally unrelated enzymes can have an identical classification. A second classification has been made on the mechanism of the enzymes. Based on the stereochemical outcome of the anomeric center of the product, there are two main mechanisms for hydrolysis of glycosidic bonds, namely inverting and retaining.
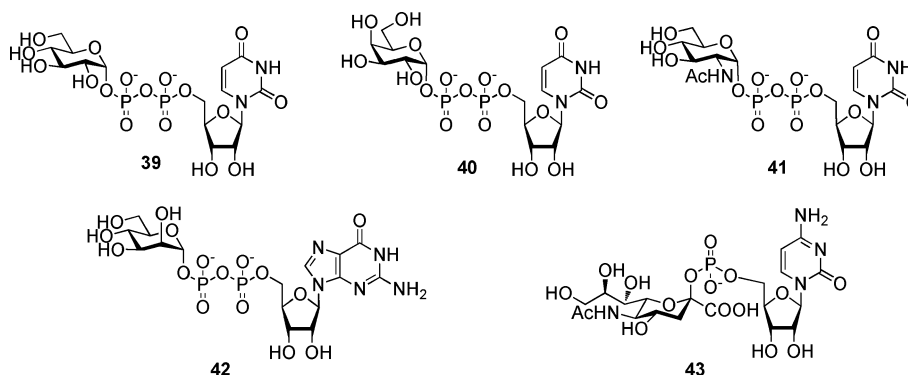

Fig. 11 Nucleotide based glycosyl donors.

This journal is © The Royal Society of Chemistry 2011

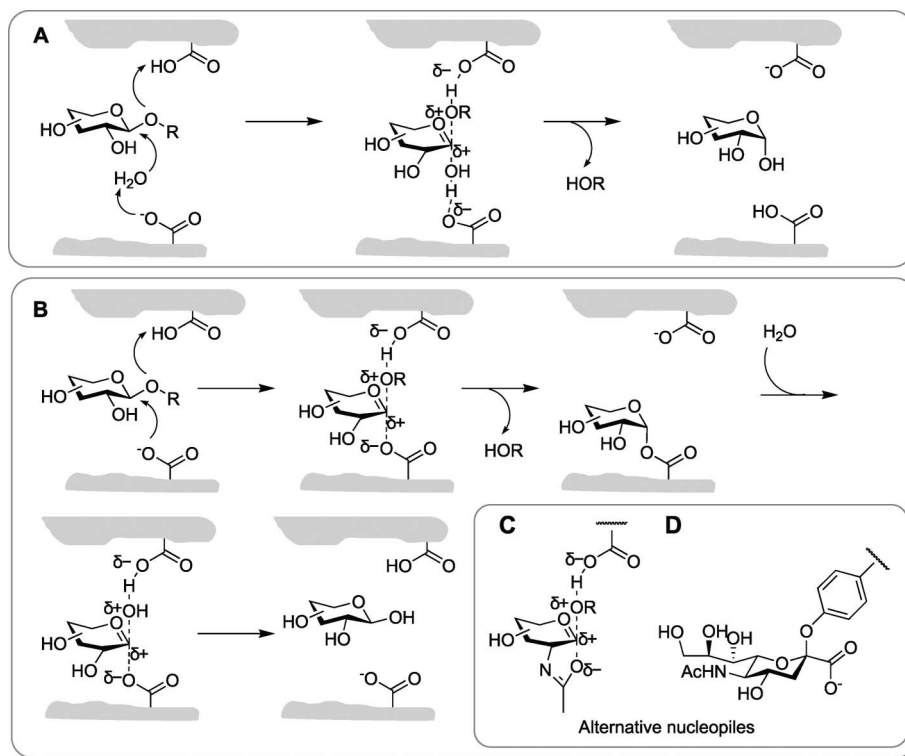*Org. Biomol. Chem.*, 2011, **9**, 5908–5926 | 5915

Fig. 12 General mechanism of glycosidases (A) inverting, (B) retaining. Alternative nucleophiles are depicted in (C) and (D).

Hydrolysis by inverting glycosidases results in inversion of the configuration of the anomeric center (for instance, α becomes β), whereas the configuration at the anomeric center is not changed by retaining glycosidases. Koshland was the first to recognize this in 1953 and postulated that inversion of the anomeric center was caused by general acid activation of the glycosidic bond followed by S$_N$2 substitution of the anomeric center with a water molecule.[52] He also proposed a double displacement mechanism for the retaining glycosidases, in which first an enzyme–glycoside complex is formed, which is hydrolyzed in the second step. The two postulated mechanisms are well established today. In inverting glycosidases, two carboxylic acid residues are located at the opposing sites of the active site, and are at least 6 Å apart. One of the carboxylic acids is protonated. This carboxylic acid acts as an acid catalyst to activate the glycosidic bond. The other carboxylic acid is deprotonated and is responsible for deprotonation of the water molecule. Hydrolysis proceeds through a single oxocarbenium ion-like transition state. No covalent glycosyl–enzyme adducts are formed during hydrolysis (Fig. 12A). The active site of retaining glycosidases also contains two carboxylic acids which are generally separated by ~5.5 Å. Similar to the inverting glycosidase, one of the carboxylic residues acts as a general acid catalyst. Now, however, the other residue does not act as base, but instead performs a nucleophilic attack forming a covalent glycosyl–enzyme adduct. In the next step, the formed covalent adduct is hydrolyzed *via* the reversed pathway. The residue that acted as an acid in the first step of hydrolysis now acts as a base abstracting a proton from the incoming water molecule (Fig. 12B). Besides carboxylic acids, several other nucleophiles have been observed in retaining glycosidase. The *N*-acetyl group adjacent to the anomeric center of the substrate acts as the nucleophilic residue in hexosaminidases

forming an oxazoline intermediate (Fig. 12C),[53] and in retaining sialidases a tyrosine residue functions as nucleophile (Fig. 12D).[54] Although the mechanism gives useful information about the enzyme, on its own this classification is not suitable.

Yet another classification has been made on the mode of action of the enzymes. Depending on the hydrolysis of polysaccharides, glycosidases can be divided in endo- and exo-glycosidases.[55] Exo-glycosidases selectively remove the terminal residues of the reducing end of polysaccharides. In the active sites, extensive interactions are made with the substrate enabling recognition of a specific glycoside. Endo-glycosidases hydrolyze glycosidic bonds within polysaccharides and have a cleft/tunnel shaped active site. Classification by the mode of action can be confusing too, since many glycosidases possess intermediate activity. Finally, in 1991 Henrissat came up with another proposal.[56] In this work, glycosidases were classified on their amino acid sequence and predicted structural relationship. Using this strategy, the glycosidases have been divided into 100 families, which can be found on http://www.cazy.org. Since the structure of the enzymes is related in these families, the mechanism of the hydrolysis of the glycosidic bond is also conserved. In some cases, the members from different families share an equal mechanism and these enzymes form a so-called glycoside hydrolase clan.

Glycosidase inhibitors and activity-based probes

The development of inhibitors and activity-based probes for glycosidases has received considerable attention, due to their potential therapeutic value as well as their usefulness in studies towards the mechanism and active site residues of glycosidases. Two types of inhibitors, namely non-covalent and covalent, have
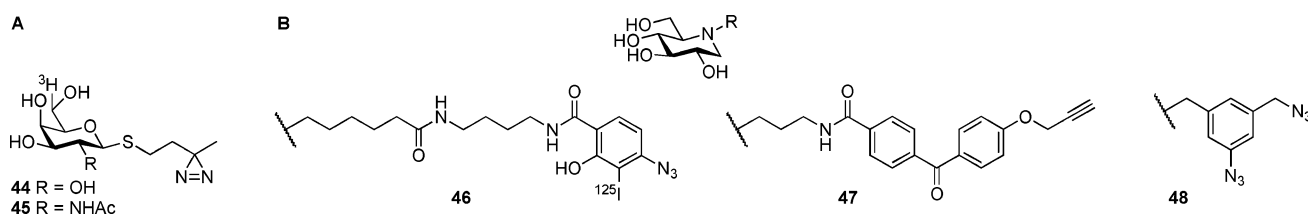
Fig. 13 (A) substrate-based photoaffinity labels. (B) Deoxynojirimycin based photoaffinity labels.

been described in literature. The non-covalent inhibitors, the largest class, have extensively been reviewed.[57] Here, the focus will be on the covalent inhibitors and their conversion to ABPs. The use of non-covalent inhibitors as leads for ABPs will be discussed briefly.

Probes based on non-covalent inhibitors

Photoaffinity probes have been applied for many enzymes. Such molecules are quite useful to label enzymes that do not form a co-valent substrate–enzyme intermediate such as metalloproteases.[58] The general structure of these probes is comparable to that of activity-based probes. Both consist of a label/reporter group for visualization purposes and a recognition element (in the case of a photoprobe, a non-covalent inhibitor), which facilitates selective and strong binding to the enzyme. In contrast to ABPs, photoactivatable probes do not have a warhead that reacts in a mechanism based fashion with the enzyme. Instead, they are equipped with a photoactivatable group. After irradiation with light of the appropriate wavelength, this group is converted into a highly reactive intermediate, which reacts with the enzyme. In two separate papers, Khun et al. utilized this approach to label a galactosidase and a hexosaminidase respectively with tritium containing photo-probes **44** and **45** (Fig. 13).[59] Upon photolysis of the diazirine in **44** and **45**, a carbene is formed which undergoes an insertion reaction with a nearby residue within the enzyme. Two active site fragments of the β-galactosidase and one of hexosaminidase were identified by proteolysis and Edmann degradation of the radioactive fragments of enzyme labeled with **44** or **45** respectively. For the β-galactosidase an additionally a peptide which was located remote from the active site was detected.[60]

Recently, three photoaffinity probes based on the broad-spectrum glucosidase inhibitor deoxynojirimycin have been re-ported. Alkylation of deoxynojirimycin with a linker equipped with a *p*-azidosalicyl amide, as photocrosslinker, and ^{125}I, as radioactive label (Fig. 13) afforded a probe for glucosidase I (**46**).[61] Radioactive labeling of a single 85 kDa protein corresponding to the size of glucosidase I was observed when a microsomal protein fraction was subjected to **46** and this signal could be abolished by incubation with reversible inhibitors showing the selectivity of **46**. Van Scherpenzeel et al. used a similar approach for the synthesis of a probe for lysosomal glucocerebrosidase and non-lysosomal glucoceramidase.[62] By alkylating deoxynojirimycin with a benzophenone photocrosslinker and an alkyne as ligation handle probe **47** was obtained which potently inhibits and after irradiation with UV light and copper catalyzed click reaction with a fluorophore can be used to visualize these enzymes. Very recently,

Gandy et al. capitalized on the same idea for the synthesis of an exo-α glucosidase probe **48**.[63]

Quinone methide probes

In the early nineties, the search for selective glycosidase inhibitors let to the introduction of quinone methide inhibitors.[64] These compounds are solely activated after enzymatic cleavage and should therefore have enhanced selectivity for the targeted enzyme. Their design was guided by the successful inhibition of proteases and esterases with chloromethylaryl esters and amides.[65] The corresponding glycoside analogues synthesized by Halazy et al. contain an *ortho* or *para*-difluoromethylaryl group **49** as the latent reactive group. Hydrolysis of the glycosidic bond by a glycosidases liberates difluoromethyl phenolate **50** as is shown in Fig. 14. Fluorine rapidly eliminates forming reactive quinone methide intermediate **51**. Any nucleophile present in the active site can perform a Michael reaction with **51** after which a covalent adduct is formed (Fig. 14).

Although this method, together with the photoaffinity labeling approach, is currently the only method that can be used to label inverting glycosidases, it has a major disadvantage and that is that the affinity for the enzyme is lost by cleavage of the glycosidic bond. Diffusion of the reactive quinone methide from the active site of the enzyme leads to cross-reactivity and labeling of the targeted enzyme at multiple sites, limiting the use quinone methide probes. This limitation notwithstanding, the approach has found some successful applications. In 1997, Janda and co-workers used the (difluoromethyl)aryl-β-glycosides as lead for the design of activity-based probe **52**.[66] The biotin in **52** allowed detection of antibodies showing galactosidase activity via a facile streptavidin-based ELISA assay (Fig. 15). Ichikawa elaborated on this paper and designed a set of probes (**53**–**55**), which were used to label partially purified O-GlcNAcase.[67] Western-blotting showed that several proteins were tagged with biotin by **53**–**55**. However, affinity purification of the formed covalent adduct proved in the case of the difluorine probes **53** and **54** troublesome. Elimination of the remaining fluorine facilitates the regeneration of a quinone methide. Attack of water on the resulting methide is followed by conversion to the aldehyde liberating the enzyme (lower part of Fig. 14). Release of the enzyme could be prevented by reacting O-GlcNAcase with monofluoromethyl probe **55** and a single protein was obtained after enrichment. A similar probe (**56**) was synthesized for β-glucosidases by Lo and co-workers.[68] This probe, containing a fluoromethyl at the *para*-position, worked well on purified β-glucosidases. However, in complex protein samples severe cross-reactivity was observed. In later studies, various purified glycosidases such as galactosidases, xylanases and neuroaminidases have successfully been labeled with
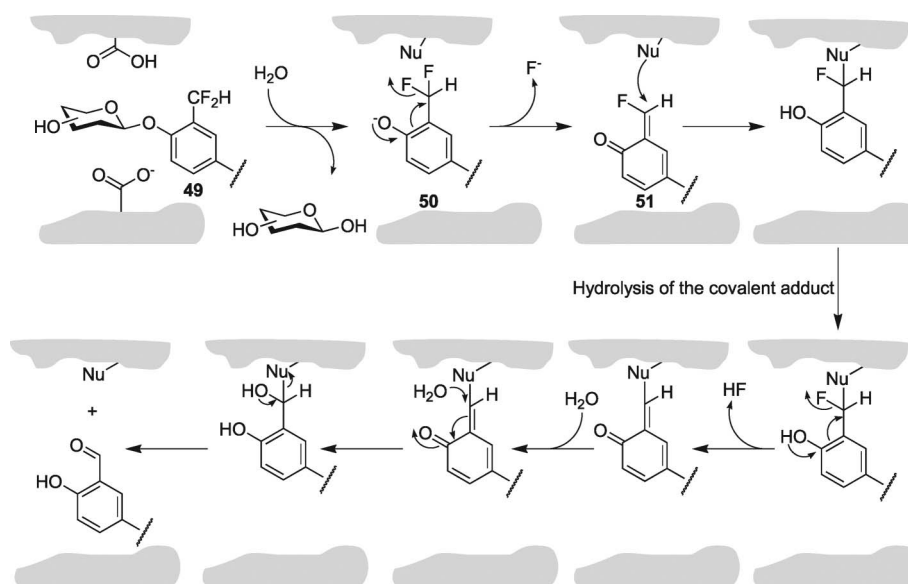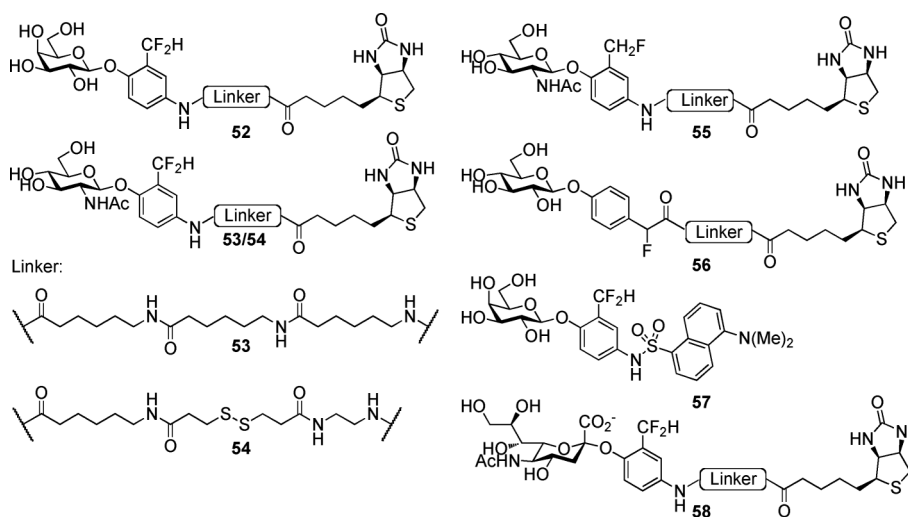
Fig. 14 Mechanism of quinone methides.

Fig. 15 Quinone methide activity-based probes. A range of linkers was used in these probes.

quinone methide probes. Especially noteworthy are the papers of Kurogohchi *et al.* and Lu *et al.* in which glycosidases are labeled in biologically relevant samples.[69] In these papers, probe **57** is used to label galactosidases with a dansyl group and probe **58** is used to biotinylate neuroaminidases on the surface membrane of influenza.

Finally, in a very recent contribution, Withers and co-workers developed novel histological and cell-labeling reagents that in contrast to fluorogenic substrates (which have previously been used for this purpose) do not diffuse away from the site of cleavage.[70] Their elegant strategy capitalizes on the intrinsic disadvantage of quinone methide probes, namely that upon release of the Michael acceptor, reaction with a nearby nucleophile is slow and therefore in all likelihood the reactive group will to a large extent diffuse from the active site. They reasoned that although diffusion would out-compete the formation of an enzyme adduct, the reactivity of the Michael acceptor would be sufficient enough to link the reporter to a nearby entity. In this fashion, the

reporter would at least be trapped (and accumulate) in the same subcellular compartment as the target glycosidase. They proved the validity of this concept by the design of a series of coumarin glycosides **59–61** (Fig. 16) targeting glucuronidase, glucosidase and galactosidase. Histological staining of glucuronidase with **59** in *Arabidopsis* plants was readily achieved whereas labeling with methylumbelliferyl glucuronide did not show any labeling at all. Cell-labeling experiments with *E. coli* cells expressing *Argobacterium* sp. β-glucosidase/galactosidase and *P. pastoris* expressing Thai rosewood β-glucosidase revealed that cells labeled with **60** and **61** remained fluorescent even after extensive washing and could be sorted using fluorescence assisted cell sorting.

2-Deoxy-2-fluoroglycoside probes

Activated 2-deoxy-2-fluoroglycosides were introduced as inhibitors of retaining β-glycosidases in 1987 by Withers and co-workers.[71] The 2-fluoride group destabilizes the
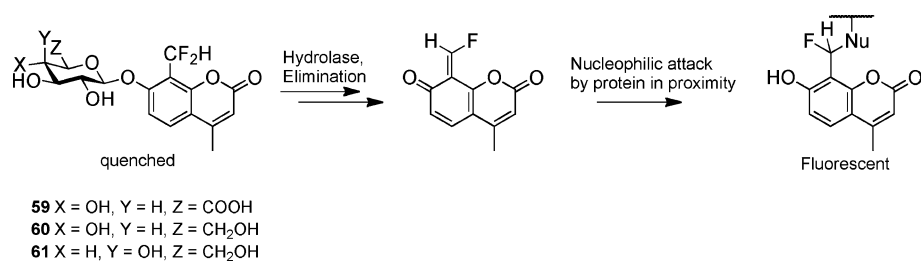
Fig. 16 Quinone methide probes that can be used for histological staining.

59 X = OH, Y = H, Z = COOH
60 X = OH, Y = H, Z = CH₂OH
61 X = H, Y = OH, Z = CH₂OH

oxocarbenium-like transition state formed during the glycosyla-tion and deglycosylation step. Consequently, the rate of formation and the rate of hydrolysis of the glycosyl-enzyme adduct is decreased. An activated anomeric leaving group increases the glycosylation rate leading to accumulation of the glycosyl–enzyme adduct (Fig. 17). Later it was reported that activated 5-fluoroglycosides inhibit glycosidases in a similar fashion.[72] Interestingly, these compounds inhibit both α- and β-glycosidases. Lifetimes of the formed glycosyl–enzyme adducts are sufficient to allow their isolation for ensuing sequence analysis. The application of fluoroglycosides led to the identification of the nucleophilic residue of various enzymes, including the nucleophilic residues of various α- and β-glycosidases,[73] sialidases[74] and a glucosaminidase.[75] In an interesting recent application, human acid glucosylceramidase, the enzyme deficient in Gaucher patients, was covalently and irreversibly labeled with a ^{18}F-2-deoxy-2-fluoroglucoside and subsequently administered to animal models. In this fashion, distribution of the enzyme can be monitored *via* positron electron tomography (PET), which might be of use in the assessment of the efficacy of enzyme replacement therapy for Gaucher patients.[76]
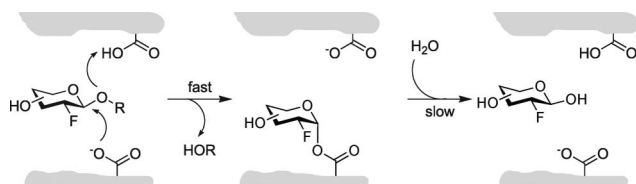

Fig. 17 Mechanism based inhibition of glycosidases with fluorosugar.

In the past decade, 2-deoxy-2-fluoroglycosides have served as inspiration for the development of ABPs aimed at retaining glycosidases. Bertozzi and Vocadlo developed an elegant strategy to profile exo-glycosidases. They realized that binding of an ABP containing a large reporter group might be hampered by the pocket shaped active site of exo-glycosidases, that the introduction of a small ligation handle would be tolerated by the enzyme and that this handle could be elaborated with a reporter group after labeling. To this end, 2-deoxy-2-fluorogalactosyl fluoride was converted into ABP **62** by incorporation of an azido group (Fig. 18).[77] Kinetics studies with *E. coli* β-galactosidase (LacZ) revealed that the modification of the C-6 position of galactosides was tolerated and therefore the azido group was introduced at this position. Although being time- and concentration-dependent, inactivation by **62** was rather slow with a second-order rate constant of 0.2 M⁻¹ min⁻¹. Both purified LacZ and LacZ in lysate of *E. coli* induced with IPTG could be inhibited with probe **62**.
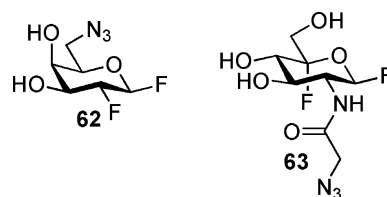

Fig. 18 Exo glycosidase probes.

The formed covalent adduct was visualized by modification of the azido group with FLAG-tag using the Staudinger–Bertozzi ligation followed by Western-blotting. In this fashion, not only LacZ was labeled with **62** but six different retaining β-glycosidases from the families 1, 2 and 35 were labeled as well, demonstrating the versatility of this approach.

Recently, Stubbs *et al.* reported a probe for retaining β-glucosaminidases based on the idea of Vocadlo and Bertozzi.[78] In the X-ray structure of *Vibrio cholerae* NagZ (VCNagZ), they observed a large pocket around the 2-acetamido binding site. They envisioned that replacing the 2-acetamido-group by an azidoacetyl would minimize loss in carbohydrate-enzyme interactions. Indeed, the resulting probe **63** proved a good glucosaminidase inhibitor. Despite the inherent instability of the *O*-acylal linkage (hydrolysis of the linkage was observed during gel-electrophoresis), probe **63** was used to successfully label purified VCNagZ in combination with FLAG-tagged phosphine. From this point of view, the assay was remarkably sensitive and as little as 80 ng could be visualized. Stubbs took advantage of the inherent instability to increase the sensitivity of glucosaminidase labeling by the development of the following procedure. Cell-lysate was reacted with probe **63** and modified with a biotin using the copper catalyzed click reaction. The biotinylated proteins were immobilized on avidine resin. Unlabeled proteins were washed away after which the resin was boiled in SDS-PAGE loading buffer to hydrolyze the acylal linkage. The liberated enzymes were resolved by gel-electrophoresis and stained using general protein staining. A putative glucosaminidase of *P. aeruginosa* was captured using this method.

Fluoroglycoside probes have also found use in protein profiling of endo-glycosidases (Fig. 19). In contrast to the above-described exo-glycosidase, modification of the glycoside with large tags is well tolerated by the canyon like active site of these enzymes. Withers and co-workers anticipated this and synthesized xylanase probes **64** and **65** containing a biotin as reporter group.[79] After confirming that compounds **64** and **65** still inhibited β-glycanases and that they could be used as probes, Withers and co-workers used them to study the proteome secreted by *Cellulomonas fimi*. These soil bacteria degrade cellulose and xylan from plant
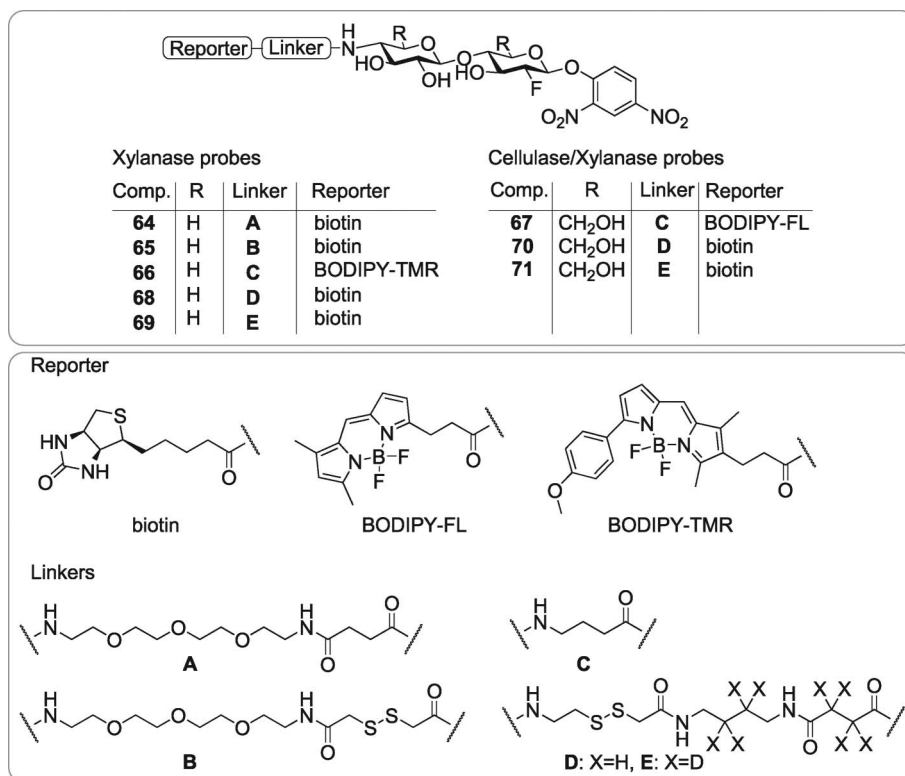
This journal is © The Royal Society of Chemistry 2011

Org. Biomol. Chem., 2011, **9**, 5908–5926 | 5919

Fig. 19   Endoglycosidases probes.

sources, for which they produce and secrete an array of xylanases and cellulases. Using probe **65**, a new β-glycanase activity was discovered in the extracellular proteome of *C. fimi*.[80] The excreted glycanases often show mixed substrate specificity (that is endoxylanase/cellulase specificity). Xylanases from family 10 are able to degrade both cellulose and xylan. The family 11 xylanases are "true" xylanases and exclusively degrade xylan. To examine the specificity of the secreted enzymes in greater detail, Withers and co-workers presented a well-designed approach to distinguish endoxylanase/cellulases from true xylanases.[81] Enzymes from family 10 and 11 were treated with two fluorescent probes, a 2-deoxy-2-fluoro-β-xylobioside condensed to a red fluorophore (**66**) and a 2-deoxy-2-fluoro-β-cellobioside conjugated to a green fluorophore (**67**). As anticipated, xylanases from family 10 could be labeled with either of the probes and xylanases from family 11 could only be labeled by **66**. This difference in selectivity was exploited to investigate the influence of extracellular surroundings on the proteome secreted by bacteria *C. fimi*. Whereas induction with xylan led to secretion of a mixture of endoxylanase/cellulases and some specific xylanases, induction with cellulose led to secretion of enzymes with a mixed specificity and one specific cellulase. Isotope-coded affinity tagging (ICAT) analogues **68–71** were published as tools to quantify the labeled proteins by mass spectrometry.[82]

Cyclitol epoxides and exocyclic epoxides

Conduritol B epoxide (CBE) or DL-1,2-anhydro-*myo*-inositol (**72**) is a potent irreversible inhibitor of β-glucosidases (Fig. 20). After its discovery in 1966 by Legler,[83] it has been used to

study β-glucosidases of many sources such as *Aspergillus*, yeast, sweet almonds and mammals.[84] Interaction of the hydroxyls in **72** with the substrate-binding pockets of the enzyme ensures specific binding of the inhibitor to glucosidases. Activation of the epoxide by a carboxylic acid in the active site is required for inhibition. Upon *trans*-diaxial opening of the epoxide by the nucleophilic residue, a stable ester bond is formed (Fig. 20A). This mechanism was confirmed with ^{14}C-labeled CBE.[85] Cleavage of the covalent adduct by reacting it with hydroxylamine and isolation of the released radioactive compound exclusively afforded 1-D-*chiro*-inositol, which is formed by diaxial opening of D-1,2-anhydro-*myo*-inositol. It was therefore reasoned that only the D-isomer of CBE reacts with glucosidases. This makes sense since it resembles the natural substrate, D-glucose. A later study employing chirally pure L-1,2-anhydro-*myo*-inositol showed that the L-isomer was indeed inactive to β-glucosidases.[86] Interestingly, Quaroni *et al.* found that CBE also inhibits the sucrase-isomaltase complex, an α-glucosidase.[87] Binding to α-glucosidases can be explained by the C-2 symmetry axis in CBE (Fig. 20B). This axis allows the molecule to orient itself in the active site such that the epoxide is activated and opened *trans*-equatorially by the nucleophilic residue, albeit with reduced reaction rates.[88] Later it was shown that other α-glucosidases of various sources could be inhibited by CBE, including yeast α-glucosidase, human lysosomal α-glucosidase and plant α-glucosidases.[89] The scope of glycosidases that could be inhibited was broadened by the synthesis of L-1,2-anhydro-*myo*-inositol (the L-isomer of CBE), conduritol F epoxide **73**, conduritol C *cis*-epoxide **74** and *trans*-epoxide **75** (Fig. 20C). These analogues of CBE inactivate yeast β-fructosidase, β-mannosidases,
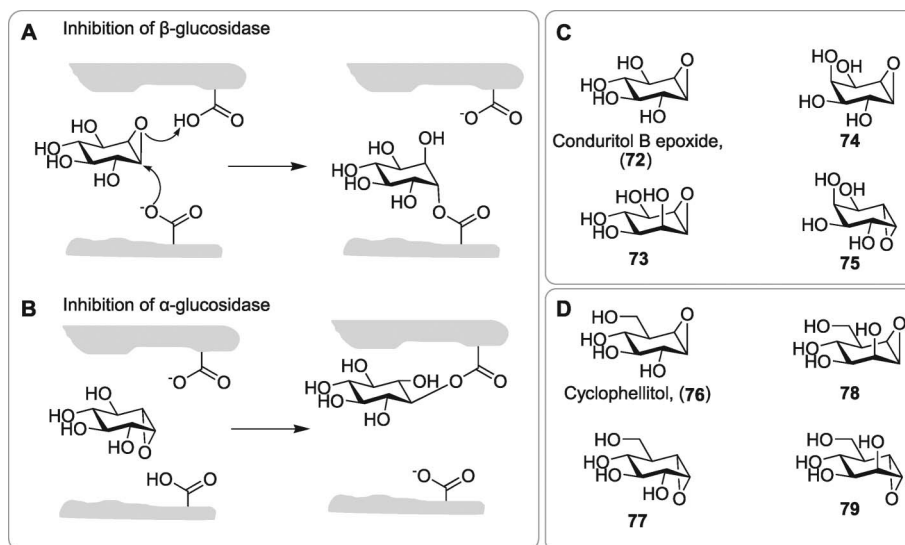
Fig. 20 Mechanism of inhibition by CBE (A) of β-glucosidases, (B) α-glucosidases. (C) Structures of conduritol B epoxide and its analogues, (D) structure of cyclophellitol and its analogues.

β-galactosidases, α-galactosidases and a α-fucosidase respectively.[90] Radiolabeled versions of **72–75** have been used to determine the active site residues of α-and β-glycosidases. An analogue of conduritol B epoxide containing an exocyclic methylene, cyclophellitol **76**, was isolated from the mushroom strain *Phellinus sp.* (Fig. 20D).[91] Prior to its discovery this analogue had already been proposed as a more specific and potent inhibitor of retaining β-glucosidases. The idea was that the introduction of a methylene would not only increase the binding-affinity but also break the symmetry of the molecule and thereby prevent binding to α-glucosidases. Indeed this proved the case, and cyclophellitol **76** potently inhibits β-glucosidases (inhibition of almond β-glucosidase is 92 fold more effective than CBE)[92] and leaves other glycosidases practically untouched (partial inhibition of β-xylosidase and α-glucosidase activities has been observed).[93] Soon after the discovery of cyclophellitol, unnatural diastereomers with the α-gluco **77**, β-manno **78** and α-manno **79** configuration were synthesized and it was shown that these compounds inhibit the corresponding α-glucosidases, β-mannosidases and α-mannosidases.[94]

Exocyclic epoxides **80—85** (Fig. 21) have also been explored as glycosidase inhibitors. In general these compounds consist of a carbohydrate tailored at the reducing end with an epoxy-alkyl chain. A diminished activity was observed for exo-glycosidases compared to CBE,[95] but these compounds proved to be excellent inhibitors of endo-glycosidases.[96] By changing the spacer length and stereochemistry of the warhead of these compounds, specificity for a certain enzyme could be generated.[97]
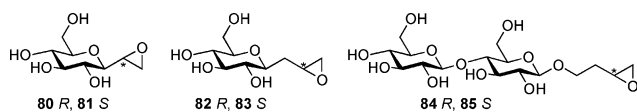

Fig. 21 Exocyclic epoxides.

Of the epoxide based inhibitors, conduritol B epoxide especially, and to a lesser extent cyclophellitol, have found application in

glycobiology studies. Both compounds selectively inhibit human acid β-glucosidase, also known as glucocerebrosidase or GBA-1, in mammals and have therefore been widely applied in the study towards Gaucher disease. Glucocerebrosidase is deficient in Gaucher patients resulting in accumulation of its substrate, glucosylceramide. By treating cells and mice with **72** and **76** a phenotype resembling Gaucher disease could be induced.[98] The selectivity of CBE was also exploited for the discovery of unknown mammalian β-glucosidases. Selective inhibition of glucocerebrosidase by CBE or cyclophellitol followed by identification of the enzymes responsible for the residual activity led to the discovery of acid β-glucosidase 2 (GBA-2) and the broad specificity β-glucosidase (GBA-3).[99] Recently, the crystal structure of CBE covalently bound to glucocerebrosidase was published.[100]

Although widely applied as inhibitors to study glycosidase activities, CBE and cyclophellitol were only recognized as leads for the development of glycosidase ABPs. This is rather surprising given that their mode of action towards retaining beta-glucosidase, which involves the formation of a covalent adduct, invites the design of ABPs by grafting a reporter group onto the cyclitol core. As outlined above, this strategy has met with considerable success starting from 2-deoxy-2-fluoroglycosides and the available literature point towards cyclitol epoxides as the superior inhibitor of the two classes. We therefore turned our attention to this, and in a first study compared the efficacy of both classes of compounds in inactivating two retaining beta-glucosidases, namely human acid glucosylceramidase (GBA) and almond beta-glucosidase (ABG).[101] In our studies we included an evaluation of the merits of direct labeling (grafting a BODIPY fluorophore onto the ABP core) and two-step bioorthogonal labeling (in which case an azide was installed as the ligation handle). This led to the design of the panel of glycosidase ABPs **86–94** (Fig. 22). As expected, we found that the cyclophellitol derivatives (**91–94**) vastly outperform their 2-deoxy-2-fluoro counterparts (**86–90**) in inhibiting both enzymes. A rather more unexpected finding is that cyclophellitol probes **93** and **94**, with the BODIPY installed at C6, proved to be the most effective GBA ABP of the series. Apparently, and as opposed
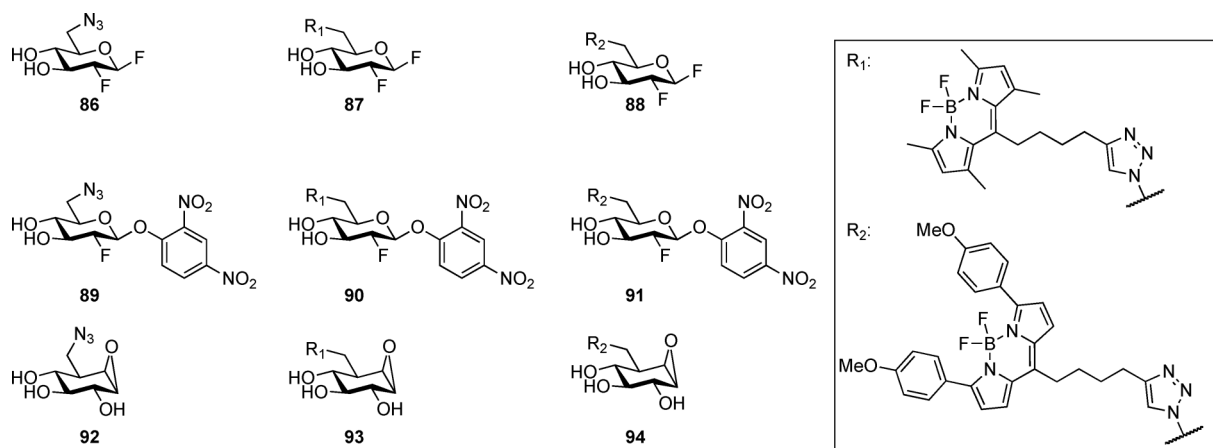
This journal is © The Royal Society of Chemistry 2011

Org. Biomol. Chem., 2011, **9**, 5908–5926 | 5921

Fig. 22 Epoxide based glucosidase probes and their corresponding fluoroglycoside probes.

to ABG, GBA is not averse, and in fact welcomes the presence of a large hydrophobic group at this position, making the probe both very sensitive and highly specific (labeling of tissue from various origin gave only a single band, corresponding to the – low abundance – protein, GBA). We capitalized on this finding by monitoring *in situ* in tissue and animal models the activity of GBA in the context of Gaucher disease and found, amongst other results, that mutated GBA can be rescued to some extent by administering isofagomine as a chemical chaperone.[102] The ability to monitor GBA activity inside cells, thanks to the ABP probe, sheds new light on such active site directed chemical chaperones and we found that the efficacy of isofagomine is quite a bit lower when looking in the appropriate environment than what is alluded to in the literature. Returning to the bioorthogonal chemistry, we found that detecting glycosidase activities in cell extracts through click chemistry using azide probe **91** proved cumbersome. Here optimization of the ligation protocols, or indeed the development of new strategies is needed.[103]

Alternative degradation pathways

For the degradation of glycoproteins, alternative pathways have been described such as excretion and autophagy of the ER.[104] Moreover, ABP-like studies revealed that the protein part of these glycoconjugates can be degraded prior to hydrolysis of the glycosidic bonds. Until recently, it was generally believed that deglycosylation of glycoproteins was required prior to degradation by the proteasome. It was reasoned that the steric bulk opposed by the glycan would prevent entry of the glycoprotein to the catalytic core of the proteasome and that it as such prevents breakdown of the glycoprotein. Interestingly, Ploegh *et al.* observed that inhibition of PNGase with Z-VAD-Fmk did not inhibit proteasomal degradation of *N*-linked glycoproteins. Ito and co-workers used fluorescent glycopeptides **95–98** to unambiguously establish that *N*-linked glycoproteins could be degraded by proteasome (Fig. 23).[105] Peptides **95–98** consist of a proteasomal recognition element, Z-Leu–Leu–Leu, a *N*-glycosylation sequence, Asn–Gly–Thr, and rhodamine as a fluorescent tag. Cleavage of the proteasomal recognition element could be monitored by HPLC. Hydrolysis of the peptide bond results in the formation of a novel fluorescent glycopeptide. Quantification of the fluorescence of the
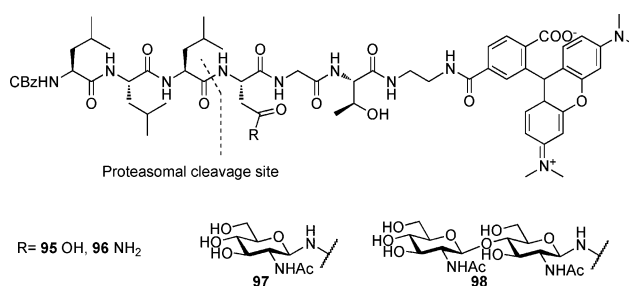

Fig. 23 Fluorescent glycosylated proteasome substrates.

newly formed peptide revealed that the proteasome is capable of degrading *N*-linked glycopeptides, although with decreased effectiveness. These results were corroborated by Navon and co-workers.[106] Glycosylated proteins containing one or multiple *N*-linked glycans were completely degraded by the proteasome. The global degradation pattern was not changed by *N*-linked glycans. They did however have a local effect. Reduced expression of epitopes near a glycosylation site was observed in *in vivo* experiments. Analysis of the degradation products *in vitro* supported these results. N-terminally extended peptides were observed near *N*-linked glycans.

Next to the degradation of *N*-linked glycoproteins also the degradation of *O*-linked glycoproteins by the proteasome has been studied. Protein *O*-GlcNAcylation is another major post-translational modification in the nucleus and cytosol and it was therefore reasoned that the proteasome will encounter *O*-GlcNAcylated substrates.[107] We used an ABP-based strategy to study whether the proteasome can handle such substrates.[108] The selective proteasome inhibitor epoxomicin was N-terminally elongated with *O*-GlcNAc-Ser derivative and equipped with an azidoacetyl group which could be modified with biotin for visualization purposes (probes **99–102**, Fig. 24). Incorporating the azidoacetyl group either at N-terminus of peptide (**99** and **100**) or at GlcNAc residue (**101** and **102**) allowed monitoring deglycosylation of probe by comparing the labeling signal of the probes (Deglycosylation of probes **101** and **102** leads to removal of the azido group whereas it will leave the azido group untouched in probes **99** and **100**). In a competition experiment,[109] these probes not only blocked labeling, but they did so with only slightly diminished potency compared to the parent compound epoxomicin.
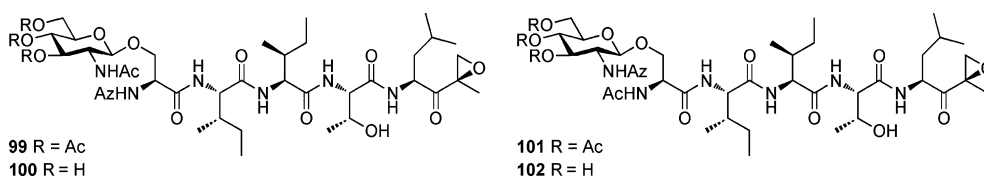
Fig. 24 *O*-GlcNAcylated proteasome probes.

Table 1 Overview of the enzymes discussed in this review, examples of the reported probes and their advantages/disadvantages

Enzyme	Active site residues	Covalent adduct	Probes		Advantages of the probe	Disadvantages of the probe
PNGase	Cys, His, Asp triad	Yes	Peptide based		- Peptide synthesis - Potent	- Labeling of proteases (*i.e.* caspases) - Less selective than carbohydrate-based probes
			Carbohydrate based:		- More selective than peptide probes - Potent	- Probe may label *endo*-glycosidases - Synthetically challenging
Glycosyl-asparaginase	N-terminal threonine	Yes	—		—	—
Glycosidase	Inverting: Glu, H₂O, Glu	No	Photoactivatable:		- Retaining and inverting glycosidases are labeled	- Labeling is affinity-based, not mechanism-based (not only active enzyme may be labeled)
			Quinone methide:		- Retaining and inverting glycosidases are labeled	- Non-specific labeling (reactive species diffuses away from the active site)
					- Mechanism-based	- Only suitable to stain compartment where the glycosidase resides
					- Fluorescently quenched probes	
	Retaining: Glu, Glu	Yes	2-deoxy-2-fluoroglycoside:		- Very selective - Mechanism-based	- Less potent than cyclitol epoxide probes - Covalent adduct in some cases unstable
						- Only retaining enzymes are labeled
			Cyclitol epoxides		- Selective - Mechanism-based	- Less selective than 2-deoxy-2-fluoroglycosides - Direct labeling strategy can only be used for GBA
					- Very potent - Highly stable adduct	- Only retaining enzymes are labeled

The covalent adduct could be visualized by Staudinger–Bertozzi modification, clearly demonstrating that deglycosylation prior to proteasome binding is not a prerequisite. Furthermore, there is not a significant difference in the labeling pattern which indicates that the probes are not subject to deglycosylation. It is therefore valid to ascertain that *O*-GlcNAcylated proteins are degraded by the proteasome.

Concluding remarks

Activity-based profiling of glycoconjugate processing enzymes has come a long way in the past decade and ABPs have been used to study both the glycoconjugation enzymes and glycoconjugate processing. Although generally considered difficult to target, ABPs for various enzymes involved in degradation of glyconjugates have been reported and with the current knowledge of the mechanisms of the enzymes it should be possible to design probes for most of the untargeted enzymes, such as glucosylasparaginase (Table 1). Despite this, it is clear that there are still issues that should be addressed to obtain successful labeling. All reported probes so far have their advantages but also disadvantages. Major disadvantages of PNGase probes may be their selectivity and by combining peptidic and glycosidic structural elements in a single probe this issue may be addressed. For glycosidase ABPs, several problems need to be addressed. First of all, inverting glycosidases do not form a covalent intermediate and can therefore only be labeled with photoaffinity probes and quinone methide probes, but their use is some what limited due to their moderate selectivity. Another problem in the development of glycosidase probes is the exquisite selectivity of exo-glycosidases. Incorporation of a reporter group decreases in most cases the activity and the selectivity and sensitivity of two-step labeling strategies currently employed is not sufficient. Optimization of the two-step labeling or alternative strategies (such as the incorporation of the reporter group at the aglycone site) are required to obtain broad spectrum probes for this class of enzymes. When these issues are addressed, new tools for glycoconjugate processing enzymes will be useful both for fundamental studies as well as diagnostics.

References

1 For Reviews see: (a) L. L. Kiessling and C. R. Bertozzi, *Science*, 2001, **291**, 2357; (b) R. A. Dwek, *Chem. Rev.*, 1996, **96**, 683; (c) P. Gagneux and A. Varki, *Glycobiology*, 1999, **9**, 747.
2 A. Larkin and B. Imperiali, *Biochemistry*, 2011, **50**, 4411.
3 R. G. Spiro, *Glycobiology*, 2002, **12**, 43R.
4 H. J. Gabius, H. C. Siebert, S. Andre., J. Jimenez-Barbedo and H. Rudigerm, *ChemBioChem*, 2004, **5**, 740.
5 (a) T. Wennekes, R. J. B. H. N. van den Berg, R. G. Boot, G. A. van der Marel, H. S. Overkleeft and J. M. F. G. Aerts, *Angew. Chem., Int. Ed.*, 2009, **48**, 8848; (b) N. N. Aronson Jr., *Biochim. Biophys. Acta*, 1999, **1455**, 139.
6 For reviews see: (a) M. Fonović and M. Bogyo, *Expert Rev. Proteomics*, 2008, **5**, 721; (b) B. F. Cravatt, A. T. Wright and J. W. Kozarich, *Annu. Rev. Biochem.*, 2008, **77**, 383.
7 D. C. Greenbaum, W. D. Arnold, F. Lu, L. Hayrapetian, A. Baruch, J. Krumrine, S. Toba, K. Chehade, D. Brömme, I. D. Kuntz and M. Bogyo, *Chem. Biol.*, 2002, **9**, 1085.
8 P. F. van Swieten, R. Maehr, A. M. C. H. van den Nieuwendijk, B. M. Kessler, M. Reich, C.-S. Wong, H. Kalbacher, M. A. Leeuwenburgh, C. Driessen, G. A. van der Marel, H. L. Ploegh and H. S. Overkleeft, *Bioorg. Med. Chem. Lett.*, 2004, **14**, 3131.
9 (a) G. Ulrich, R. Ziessel and A. Harriman, *Angew. Chem., Int. Ed.*, 2008, **47**, 1184; (b) A. Loudet and K. Burgess, *Chem. Rev.*, 2007, **107**, 4891.
10 N. Jessani, S. Niessen, B. Q. Wei, M. Nicolau, M. Humphrey, Y. Ji, W. Han, D.-Y. Noh, J. R. Yates III, S. S. Jeffrey and B. F. Cravatt, *Nat. Methods*, 2005, **2**, 691.
11 (a) E. Saxon and C. R. Berozzi, *Science*, 2000, **287**, 2007; (b) H. Ovaa, P. F. van Swieten, B. M. Kessler, M. A. Leeuwenburgh, E. Fiebinger, A. M. C. H. van den Nieuwendijk, P. J. Galardy, G. A. van der Marel, H. L. Ploegh and H. S. Overkleeft, *Angew. Chem., Int. Ed.*, 2003, **42**, 3626; (c) A. E. Speers, G. C. Adam and B. F. Cravatt, *J. Am. Chem. Soc.*, 2003, **125**, 4686; (d) J. M. Baskin, J. A. Prescher, S. T. Lauglin, N. J. Agard, P. V. Chang, I. A. Miller, A. Lo, J. A. Codelli and C. R. Bertozzi, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 16793; (e) X. Ning, M. A. Wolfert and G.-J. Boons, *Angew. Chem., Int. Ed.*, 2008, **47**, 2253. For an example in which bioorthogonal reactions are used to label carbohydrate-binding proteins, see: (f) L Ballell, K. J. Alink, M. Slijper, C. Versluis, R. M. J. Liskamp and R. J. Pieters, *ChemBioChem*, 2005, **6**, 291.
12 (a) A. E. Johnson and M. A. van Waes, *Annu. Rev. Cell Dev. Biol.*, 1999, **15**, 799; (b) T. A Rapoport, *Nature*, 2007, **450**, 663.
13 B. Imperiali and T. L. Hendrickson, *Bioorg. Med. Chem.*, 1995, **3**, 1565.
14 (a) A. R. Rodan, J. F. Simons, E. S. Trombetta and A. Helenius, *EMBO J.*, 1996, **15**, 6921; (b) A. Zapun, S. M. Petrescu, P. M. Rudd, R. A. Dwek and D. Y. Thomas, *Cell*, 1997, **88**, 29–38.
15 J. D. Oliver, H. Llewelyn Roderick, D. H. Llewellyn and S. High, *Mol. Bio. Cell*, 1999, **10**, 2573.
16 (a) J. J. Caramelo, O. A. Castro, L. G. Alonso, G. de Prat-Gay and A. J. Parodi, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 86; (b) S. C. Taylor, P. Thibault, D. C. Tessier, J. J. M. Bergeron and D. Y. Thomas, *EMBO Rep.*, 2003, **4**, 405.
17 R. M. Vabulas and F. U. Hartl, *Science*, 2005, **310**, 1960.
18 S. Olivari, T. Cali, K. E. H. Salo, P. Paganetti and L. W. Ruddock, *Biochem. Biophys. Res. Commun.*, 2006, **349**, 1278.
19 S. Clerc, C. Hirsch, D. M. Oggier, P. Deprez, C. Jakob, T. Sommer and M. Aebi, *J. Cell Biol.*, 2009, **184**, 159.
20 A. Hosomi, K. Tanabe, H. Hirayama, I. Kim, H. Rao and T. Suzuki, *J. Biol. Chem.*, 2010, **285**, 24324.
21 A. Herscovics, *Biochimie*, 2001, **83**, 757.
22 N. Hosokawa, I. Wada, K. Hasegawa, T. Yorihuzi, L. O. Tremblay, A. Herscovics and K. Nagata, *EMBO Rep.*, 2001, **2**, 415.
23 K. Hirao, Y. Natsuka, T. Tamura, I. Wada, D. Morito, S. Natsuka, P. Romero, B. Sleno, L. O. Tremblay and A. Herscovics, *J. Biol. Chem.*, 2006, **281**, 9650.
24 D. J. Termine, K. W. Moremen and R. N. Sifers, *J. Cell Sci.*, 2009, **122**, 976.
25 N. Hosokawa, Y. Kamiya, D. Kamiya, K. Kato and K. Nagata, *J. Biol. Chem.*, 2009, **284**, 17061.
26 J. C. Christianson, T. A. Shaler, R. E. Tyler and R. R. Kopito, *Nat. Cell Biol.*, 2008, **10**, 272.
27 T. Suzuki, A. Seko, K. Kitajima, Y. Inoue and S. Inoue, *Biochem. Biophys. Res. Commun.*, 1993, **194**, 1124.
28 T. Suzuki, H. Park, N. M. Hollingsworth, R. Sternglanz and W. J. Lennarz, *J. Cell Biol.*, 2000, **149**, 1039.
29 J.-H. Lee, J. M. Choi, C. Lee, K. J. Yi and Y. Cho, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 9144.
30 C. Hirsch, S. Misaghi, D. Blom, M. E. Pacold and H. L. Ploegh, *EMBO Rep.*, 2004, **5**, 201.
31 N. P. Dantuma, C. Heinen and D. Hoogstraten, *DNA Repair*, 2009, **8**, 449.
32 (a) G. Li, G. Zhao, X. Zhou, H. Schindelin and W. J. Lennarz, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 15809; (b) G. Li, G. Zhao, X. Zhou, H. Schindelin and W. J. Lennarz, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 8348.
33 S. Misaghi, M. E. Pacold, D. Blom, H. L. Ploegh and G. A. Korbel, *Chem. Biol.*, 2004, **11**, 1677.
34 T. Suzuki, I. Hara, M. Nakano, G. Zhao, W. J. Lennarz, H. Schindelin, N. Taniguchi, K. Totani, I. Matsuo and Y. Ito, *J. Biol. Chem.*, 2006, **281**, 22152.
35 M. D. Witte, C. V. Descals, S. V. De Lavoir, B. I. Florea, G. A. Van der Marel and H. S. Overkleeft, *Org. Biomol. Chem.*, 2007, **5**, 3690.
36 M. D. Witte, D. Horst, E. J. Wiertz, G. A. Van der Marel and H. S. Overkleeft, *J. Org. Chem.*, 2009, **74**, 605.
37 A. Miyazaki, I. Matsuo, S. Hagihara, A. Kakegawa, T. Suzuki and Y. Ito, *Glycoconjugate J.*, 2009, **26**, 133.
38 S. Hagihara, A. Miyazaki, I. Matsuo, A. Tatami, T. Suzuki and Y. Ito, *Glycobiology*, 2007, **17**, 1070.

5924 | *Org. Biomol. Chem.*, 2011, **9**, 5908–5926

This journal is © The Royal Society of Chemistry 2011

39 (a) N. N. Aronson Jr. and M. J. Kuranda, *FASEB J.*, 1989, **3**, 2615; (b) B. Winchester, *Glycobiology*, 2005, **15**, 1R.

40 (a) C. Guan, T. Cui, V. Rao, W. Liao, J. Benner, C.-L. Lin and D. Comb, *J. Biol. Chem.*, 1996, **271**, 1732; (b) Q. Xu, D. Buckley, C. Guan and H.-C. Guo, *Cell*, 1999, **98**, 651; (c) Y Wang and H.-C. Guo, *J. Mol. Biol.*, 2010, **403**, 120.

41 J. M. Risely, D. H. Huang, J. J. Kaylor, J. J. Malik, Y.-Q. Xia and W. M. York, *Arch. Biochem. Biophys.*, 2001, **391**, 165.

42 T. Noronkoski and I. Mononen, *Glycobiology*, 1997, **7**, 217.

43 Y. Wang and H.-C. Guo, *J. Mol. Biol.*, 2007, **366**, 82.

44 W. Du and J. M. Risely, *Org. Biomol. Chem.*, 2003, **1**, 1900.

45 V. Kaartinen, J. C. Williams, J. Tomich, J. R. III Yates, L. E. Hood and I Mononen, *J. Biol. Chem.*, 1991, **266**, 5860.

46 L. L. Lairson, B. Henrissat, G. J. Davies and S. G. Withers, *Annu. Rev. Biochem.*, 2008, **77**, 521.

47 A. M. Lesk, *Curr. Opin. Struct. Biol.*, 1995, **5**, 775.

48 A. L. Lovering, L. H. de Castro, D. Lim and N. C. Strynadka, *Science*, 2007, **315**, 1402.

49 S. J. Charnock and G. J. Davies, *Biochemistry*, 1999, **38**, 6380.

50 D. Hoffmeister, K. Ichinose and A. Bechthold, *Chem. Biol.*, 2001, **8**, 557.

51 D. J. Vocadlo and G. J. Davies, *Curr. Opin. Chem. Biol.*, 2008, **12**, 539.

52 D. E. Koshland, *Biol. Rev.*, 1953, **28**, 416.

53 (a) B. L. Mark, D. J. Vocadlo, S. Knapp, B. L. Triggs-Raine, S. G. Withers and M. N. G. James, *J. Biol. Chem.*, 2001, **276**, 10330; (b) I. Tews, A. Perrakis, A. Oppenheim, Z. Danter, K. S. Wilson and C. E. Vorgias, *Nat. Struct. Biol.*, 1996, **3**, 638.

54 (a) A. G. Watts, I. Damager, M. L. Amaya, A. Buschiazzo, P. Alzari, A. C. Frasch and S. G. Withers, *J. Am. Chem. Soc.*, 2003, **125**, 7532; (b) J. N. Watson, V. Dookhun, T. J. Borgford and A. J. Bennet, *Biochemistry*, 2003, **42**, 12682.

55 G. Davies and B. Henrissat, *Structure*, 1995, **3**, 853.

56 B. Henrissat, *Biochem. J.*, 1991, **280**, 309.

57 (a) E. Borges de Melo, A. da Silveira Gomes and I. Carvalho, *Tetrahedron*, 2006, **62**, 10277; (b) B. G. Winchester, *Tetrahedron: Asymmetry*, 2009, **20**, 645; (c) B. G. Davis, *Tetrahedron: Asymmetry*, 2009, **20**, 652.

58 (a) S. A. Sieber, S. Niessen, H. S. Hoover and B. F. Cravatt, *Nat. Chem. Biol.*, 2006, **2**, 274; (b) E. W. S. Chan, S. Chattopadhaya, R. C. Panicker, X. Huang and S. Q. Yao, *J. Am. Chem. Soc.*, 2004, **126**, 14435.

59 (a) C.-S. Kuhn, J. Lehmann, G. Jung and S. Stevanović, *Carbohydr. Res.*, 1992, **232**, 227; (b) C.-S. Kuhn, J. Lehmann and K. Sandhoff, *Bioconjugate Chem.*, 1992, **3**, 230.

60 B. Liessem, G. J. Glombitza, F. Knoll, J. Lehmann, J. Kellermann, F. Lottspeich and K. Sandhoff, *J. Biol. Chem.*, 1995, **270**, 23693.

61 A. V. Romaniouk, A. Silva, J. Feng and I. K. Vijay, *Glycobiology*, 2004, **14**, 301.

62 M. Van Scherpenzeel, R. J. H. B. N. Van der Berg, W. E. Donker-Koopman, R. M. J. Liskamp, J. M. F. G. Aerts, H. S. Overkleeft and R. J. Pieters, *Bioorg. Med. Chem.*, 2010, **18**, 267.

63 M. N. Gandy, A. W. Debowski and K. A. Stubbs, *Chem. Commun.*, 2011, **47**, 5037.

64 S. Halazy, V. Berges, A Ehrhard and C. Danzin, *Bioorg. Chem.*, 1990, **18**, 330.

65 M. Wakselman, *Nouveau Journal de Chimie*, 1983, **7**, 439.

66 K. D. Janda, L.-C. Lo, C.-H. L. Lo, M.-M Sim, R. Wang, C. H. Wong and R. A. Lerner, *Science*, 1997, **275**, 945.

67 M. Ichikawa and Y. Ichikawa, *Bioorg. Med. Chem. Lett.*, 2001, **11**, 1769.

68 C.-S. Tsai, Y.-K. Li and L.-C. Lo, *Org. Lett.*, 2002, **4**, 3607.

69 (a) M. Kurogochi, S.-I. Nishimura and Y. C. Lee, *J. Biol. Chem.*, 2004, **279**, 44704; (b) C. P. Lu, C.-T Ren, Y.-N. Lai, S.-H. Wu, W.-M. Wang, J.-Y. Chen and L.-C. Lo, *Angew. Chem., Int. Ed.*, 2005, **44**, 6888.

70 D. H. Kwan, H.-M. Chen, K. Ratananikom, S. M. Hancock, Y. Watanabe, P. T. Kongsaeree, A. L. Samuels and S. G. Withers, *Angew. Chem., Int. Ed.*, 2011, **50**, 300.

71 S. G. Withers, I. P. Street, P. Bird and D. H. Dolphin, *J. Am. Chem. Soc.*, 1987, **109**, 7530.

72 J. D. McCarter and S. G. Withers, *J. Am. Chem. Soc.*, 1996, **118**, 241.

73 (a) S. G. Withers, R. A. J. Warren, I. A. P. Street, K. Rupitz, J. B. Kempton and R. Aebersold, *J. Am. Chem. Soc.*, 1990, **112**, 5887; (b) H. D. Ly, S. Howard, K. Shum, S. He, A. Zhu and S. G. Withers, *Carbohydr. Res.*, 2000, **329**, 539.

74 M. F. Amaya, A. G. Watts, I. Damager, A. Wehenkel, T. Nguyen, A. Buschiazzo, G. Paris, A. C. C. Frasch, S. G. Withers and P. M. Alzari, *Structure*, 2004, **12**, 775.

75 D. J. Vocadlo, C. Mayer, S. He and S. G. Withers, *Biochemistry*, 2000, **39**, 117.

76 C. P. Phenix, B. P. Rempel, K. Colobong, D. J. Doudet, M. J. Adam, L. A. Clarke and S. G. Withers, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 10842.

77 D. J. Vocadlo and C. R. Bertozzi, *Angew. Chem., Int. Ed.*, 2004, **43**, 5338.

78 K. Stubbs, A. Scaffidi, A. W. Debowski, B. L. Mark, R. V Stick and D. J. Vocadlo, *J. Am. Chem. Soc.*, 2008, **130**, 327.

79 S. J. Williams, O. Hekmat and S. G. Withers, *ChemBioChem*, 2006, **7**, 116.

80 O. Hekmat, Y.-W. Kim, S. J. Williams, S. He and S. G. Withers, *J. Biol. Chem.*, 2005, **280**, 35126.

81 O. Hekmat, C. Florizone, Y.-W. Kim, L. D. Eltis, R. A. J. Warren and S. G Withers, *ChemBioChem*, 2007, **8**, 2125.

82 O. Hekmat, S. He, R. A. Warren and S. G. Withers, *J. Proteome Res.*, 2008, **7**, 3282.

83 G. Legler, *Z. Physiol. Chem.*, 1966, **345**, 197.

84 (a) G. Legler, *Hoppe-Seyler's Z. Physiol. Chem.*, 1968, **349**, 767; (b) G. Legler and S. N. Hasnain, *Hoppe-Seyler's Z. Physiol. Chem.*, 1970, **351**, 25; (c) N. Kanfer, G. Legler, J. Sullivan, S. S. Raghavan and R. A. Mumford, *Biochem. Biophys. Res. Commun.*, 1975, **67**, 85.

85 G. Legler, *Biochim. Biophys. Acta.*, 1968, **151**, 728.

86 D. Mercier, A. Olesker, S. D. Gero and J. E. G. Barnett, *Carbohydr. Res.*, 1971, **18**, 227.

87 A. Quaroni, E. Gershon and G. Semenza, *J. Biol. Chem.*, 1974, **249**, 6424.

88 H. Braun, G. Legler, J. Deshusses and G. Semenza, *Biochim. Biophys. Acta*, 1977, **483**, 135.

89 (a) M. M. P. Hermans, M. A. Kroos, J. van Beeumen, B. A. Oostra and A. J. J. Reuser, *J. Biol. Chem.*, 1991, **266**, 13507; (b) Y. Shou-Jun, G. Su-guo, Z. Yu-cheng and Z. Su-zheng, *Biochim. Biophys. Acta*, 1985, **828**, 236; (c) S. Iwanami, H. Matsui, A. Kimura, H. Ito, H. Mori, M. Honma and S. Chiba, *Biosci., Biotechnol., Biochem.*, 1995, **59**, 459.

90 (a) H. Braun, *Biochim. Biophys. Acta*, 1976, **452**, 452; (b) K. T. Cavanagh, R. A. Fischer, G. Legler, M. Herrchen, M. Z. Jones, E. Julich, R. P. Sewell-Alger, M. L. Sinnot and F. E. Wilkinson, *Enzyme*, 1985, **34**, 75; (c) G. Legler and M. Herrchen, *FEBS Lett.*, 1981, **135**, 139; (d) W. J. Jr. White, K. J. Schray, G. Legler and J. A. Alhadeff, *Biochim. Biophys. Acta*, 1986, **873**, 198.

91 S. Atsumi, K. Umezawa, H. Iinuma, H. Naganawa, H. Nakamura, Y. Iitaka and T. Takeuchi, *J. Antibiot.*, 1990, **43**, 49.

92 S. G. Withers and K. Umezawa, *Biochem. Biophys. Res. Commun.*, 1991, **177**, 532.

93 S. Atsumi, H. Iinuma, C. Nosaka and K. Umezawa, *J. Antibiot.*, 1990, **43**, 1579.

94 (a) K. Tatsuta, Y. Niwata, K. Umezawa, K. Tashima and M. Nakata, *J. Antibiot.*, 1991, **44**, 456; (b) K. Tatsuta, Y. Niwata, K. Umezawa, K. Tashima and M. Nakata, *J. Antibiot.*, 1991, **44**, 912.

95 A, S. D. Shiyan and A. Y. Khorlin, *Biochim. Biophys. Acta*, 1976, **445**, 169.

96 (a) E. W. Thomas, J. F. McKelvy and N. Sharon, *Nature*, 1969, **222**, 485–486; (b) G. Legler and E. Bause, *Carbohydr. Res.*, 1973, **28**, 45.

97 P. B. Hoj, E. B. Rodriguez, J. R. Iser, R. V. Stick and B. A. Stone, *J. Biol. Chem.*, 1991, **266**, 11628.

98 (a) J. N. Kanfer, G. Legler, J. Sullivan, S. S. Raghavan and R. A. Mumford, *Biochem. Biophys. Res. Commun.*, 1975, **67**, 85; (b) S. Atsumi, C. Nosaka, H. Iinuma and K. Umezawa, *Arch. Biochem. Biophys.*, 1992, **297**, 362; (c) S. Atsumi, C. Nosaka, H. Iinuma and K. Umezawa, *Arch. Biochem. Biophys.*, 1993, **304**, 302.

99 (a) R. G. Boot, M. Verhoek, W. Donker-Koopman, A. Strijland, J. van Marle, H. S. Overkleeft, T. Wennekes and J. M. F. G. Aerts, *J. Biol. Chem.*, 2007, **282**, 1305; (b) Y. Hayashi, N. Okino, Y. Kakuta, T. Shikanai, M. Tani, H. Narimatsu and M. Ito, *J. Biol. Chem.*, 2007, **282**, 30889.

100 L. Premkumar, A. R. Sawkar, S. Boldin-Adamsky, L. Toker, I. Silman, J. W. Kelly, A. H. Futerman and J. L. Sussman, *J. Biol. Chem.*, 2005, **280**, 23815.

101 M. D. Witte, M. T. C Walvoort, K.-Y Li, W. W. Kallemeijn, W. E. Donker-Koopman, R. G. Boot, J. M. F. G. Aerts, J. D. C Codee, G. A. Van der Marel and H. S. Overkleeft, *ChemBioChem*, 2011, **12**, 1263.

This journal is © The Royal Society of Chemistry 2011

Org. Biomol. Chem., 2011, **9**, 5908–5926 | 5925

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

103 (*a*) M. F. Debets, C. W. J. van der Doelen, F. P. J. T. Rutjes and F. L. van Delft, *ChemBioChem*, 2010, **11**, 1168; (*b*) N. Devaraj, S. Hilderbrand, R. Upadhyay, R. Mazitschek and R. Weissleder, *Angew. Chem., Int. Ed.*, 2010, **49**, 2869.

104 U. Nair and D. J. Klionsky, *J. Biol. Chem.*, 2005, **280**, 41785.

105 S. Hagihara, K. Goda, I. Matsuo and Y. Ito, *Biochem. Biophys. Res. Commun.*, 2007, **360**, 357.

106 E. Kario, B. Tirosh, H. L. Ploegh and A. Navon, *J. Biol. Chem.*, 2008, **283**, 244.

107 L. Wells, K. Vosseller and G. W. Hart, *Science*, 2001, **291**, 2376.

108 M. D. Witte, B. I. Florea, M. Verdoes, O. Adeyanju, G. A. Van der Marel and H. S. Overkleeft, *J. Am. Chem. Soc.*, 2009, **131**, 12064.

109 M. Verdoes, B. I. Florea, V. Menendez-Benito, C. J. Maynard, M. D. Witte, W. A. van der Linden, A. M. C. H. van den Nieuwendijk, T. Hofmann, C. R. Berkers, F. W. van Leeuwen, T. A. Groothuis, M. A. Leeuwenburgh, H. Ovaa, J. J. Neefjes, D. V. Filippov, G. A. van der Marel, N. P. Dantuma and H. S. Overkleeft, *Chem. Biol.*, 2006, **13**, 1217.