ACS central science

http://pubs.acs.org/journal/acscii

Research Article

Azido Groups Hamper Glycan Acceptance by Carbohydrate Processing Enzymes

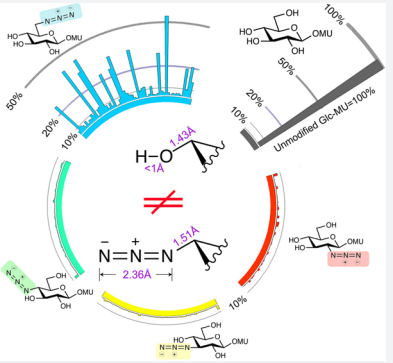
Feng Liu, Hong-Ming Chen, Zachary Armstrong, and Stephen G. Withers*

Cite This: *ACS Cent. Sci.* 2022, 8, 656−662

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Azido sugars have found frequent use as probes of biological systems in approaches ranging from cell surface metabolic labeling to activity-based proteomic profiling of glycosidases. However, little attention is typically paid to how well azide-substituted sugars represent the parent molecule, despite the substantial difference in size and structure of an azide compared to a hydroxyl. To quantitatively assess how well azides are accommodated, we have used glycosidases as tractable model enzyme systems reflecting what would also be expected for glycosyltransferases and other sugar binding/modifying proteins. In this vein, specificity constants have been measured for the hydrolysis of a series of azidodeoxy glucosides and *N*-acetylhexosaminides by a large number of glycosidases produced from expressed synthetic gene and metagenomic libraries. Azides at secondary carbons are not significantly accommodated, and thus, associated substrates are not processed, while those at primary carbons are productively recognized by only a small subset of the enzymes and often then only very poorly. Accordingly, in the absence of careful controls, results obtained with azide-modified sugars may not be representative of the situation with the natural sugar and should be interpreted with considerable caution. Azide incorporation can indeed provide a useful tool to monitor and detect glycosylation, but careful consideration should go into the selection of sites of azide substitution; such studies should not be used to quantitate glycosylation or to infer the absence of glycosylation activity.

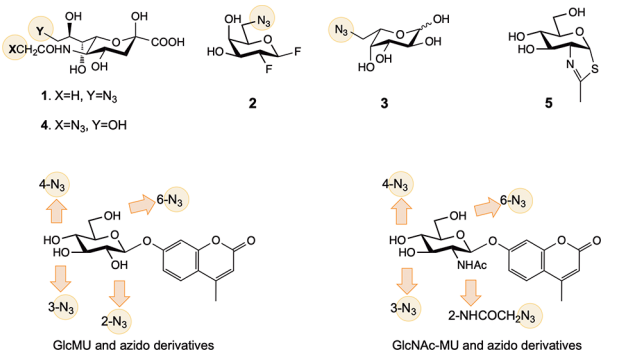


INTRODUCTION

The azide functionality vaulted into the repertoire of chemical biologists in the early 2000s based on the remarkable utility of the Huisgen azide/alkyne cycloaddition reaction popularized by Sharpless and Meldal, and through the groundbreaking work of Bertozzi in developing bio-orthogonal reactions such as the Staudinger ligation.[1−3] Such reactions have found particular application in glycobiology as a means to label glycans on cell surfaces and, when done properly, have led to valuable insights into cell biology.[4−10]

Frequently, an azide is substituted in place of a hydroxyl group as in the commonly employed 9-azido-9-deoxy sialic acid (**1**, or its azido-ManNAc precursor) and as in the 6-azido-2-fluoro-2,6-dideoxy-β-galactosyl fluoride (**2**) used to tag β-galactosidases in proteomes.[7,9,11] In other cases, it may replace a hydrogen atom, as in 6-azidofucose derivatives (**3**) and 5-azidoacetamido-sialic acid (**4**).[12] Studies of such types have provided useful insights in many cases, but care must be taken in the choice of position of substitution and in the interpretation of resultant data since the substitution is far from isosteric or isoelectronic. Therefore, the enzymes and binding partners involved in its incorporation, recognition, or turnover may well function substantially differently on the modified substrate or, in many cases, not at all. Typically, azido sugars are used directly in cell biology studies, and the presence or absence of an azide moiety in the treated biological sample is assessed by clicking on a fluorophore or a tag such as biotin to locate, identify, and/or quantitate the product. (See Chart 1.) In the past, it was rare for any assessment to be made of how well the azide functionality was accommodated by the individual enzymes involved in order to

Chart 1. Sample Azido Sugars and Compounds Used in Determination of Relative Reactivities



Received: September 27, 2021
Published: May 10, 2022

https://doi.org/10.1021/acscentsci.1c01172
ACS Cent. Sci. 2022, 8, 656−662

gauge how well the result obtained reflected the behavior of the parent sugar. However, several recent studies have probed the tolerance of specific enzymes *in vitro* for a variety of substitutions and, in some cases, have followed this up with parallel cell-based studies.[13−15] Taking one step further in a particularly illuminating mass spectrometry-based profiling study, Shajahan et al. quantitatively analyzed the classes of azido sugar glycans formed, on a temporal basis.[16] They observed substantial differences in outcome, thus the glycome, between cell types and specific sugar precursors, which they attributed to differential acceptance by metabolic enzymes. Thus, knowledge on how best to minimize differential recognition is important: this study aims to address that issue.

The azide functionality is a rigid rod of 2.36 Å in length, attached via a C−N bond that is itself 1.51 Å long.[17] The substituent thus requires a substantially larger binding pocket than that ordinarily occupied by a well-accommodated hydroxyl of C−O bond length 1.43 Å and an oxygen atomic radius of 0.48 Å, much less a hydrogen. Further, the azide is unlikely to satisfy the hydrogen bonding partners evolved in response to a hydroxyl at the position in question. The concern then is that, unless the position of substitution is well-chosen, the biological result derived might be misleading since hydroxyl groups are typically tightly coordinated in enzyme active sites leaving very little room for additional bulk. Since significant steric clashes are typically energetically prohibitive, this may well mean that the azido sugar is not bound/processed, yielding results that are not representative of the true situation.

In order to gain some measure of the extent of this concern in a reasonably unbiased manner, we wanted to assay a large number of enzymes to assess how well a series of azido sugars of the *Gluco* and *GlcNAc* series that are monosubstituted at primary and secondary carbons are processed. Gaining access to, and assaying, hundreds of different enzymes with a series of azide-substituted substrates is a nontrivial task. We propose that the use of a single class of enzymes (glycosidases) to assess how well azides are accommodated in protein binding sites that evolved to recognize a hydroxyl group or a hydrogen is a valid approach since there is unlikely to be any fundamental difference in the way that a hydroxyl group (or a hydrogen) is accommodated by one class of sugar binding/processing protein as compared to another: the same laws of physics and evolutionary trajectories apply. Glycosyltransferases might seem to be the most appropriate enzymes on which to perform such a study, but this is simply not feasible on scale, due to the challenges of expression of hundreds of such enzymes, synthesis of azide-substituted nucleotide sugars, and then determination of accurate kinetic parameters. Glycosidases are much more viable as a model system since they typically express well, and diverse libraries are available through metagenomic studies as well as a synthetic gene library. In addition, assays are much simpler and amenable to high-throughput approaches. Importantly, glycosidases and glycosyltransferases catalyze hydrolysis via very similar oxocarbenium ion-like transition states; thus, impacts of azide substitution upon the stabilities of such transition states are likely to be globally similar in the two cases. In addition, specific information on the effects of azide substitution on the turnover of the modified glycan may be just as important for cell biological studies as that on incorporation by providing specific insights into how the glycosidase-catalyzed degradation of azidoglycoconjugates might be impacted *in cellulo* when performing metabolic labeling experiments.

All glycosidase libraries were expressed in a host that itself lacked the enzyme activity of interest. This allowed a comparative assay of the enzyme with the parent and modified substrate within crude cell extracts. The useful, and meaningful, relative values of k_{cat}/K_M, the kinetic parameter best defining substrate specificity, can be measured by means of the substrate depletion method in which reaction time courses are monitored at low concentrations of substrate and fit to a first-order equation.[18,19] By performing this measurement with both the parent (glucoside or *N*-acetylglucosaminide) and the azide-substituted substrates with identical or known relative amounts of enzyme, the *relative* values of k_{cat}/K_M can be determined for each substitution with each enzyme. It is this rate constant ratio that is of interest in this study since it is the best measure of the effect of substrate substitution upon enzyme performance. This study was performed with a set of approximately 100 enzymes from a single CAZy[20] glycoside hydrolase (GH) family as well as across a large number of bacterial *β*-glucosidases and *β*-*N*-acetylhexosaminidases from a range of sources, derived from a large metagenomic library.

RESULTS

Accommodation of Azide Substituents within a Single Structural Fold. We started out by assessing the tolerance for azide substitution at primary and secondary carbon centers of the substrate 4-methylumbelliferyl *β*-D-glucopyranoside (GlcMU) using a set of enzymes within a single sequence-related CAZy family, GH1, following up on qualitative studies reported as part of a previous study.[21,22] All enzymes within this family have the same basic fold and effect catalysis via the same double displacement mechanism. The substrates most commonly cleaved by GH1 enzymes are *β*-glucosides and *β*-galactosides, but also cleaved are 6-phospho-*β*-glucosides, *β*-xylosides, *β*-mannosides, and *β*-glucuronides.[20] These enzymes were expressed from *Escherichia coli* bearing a synthetic gene library that is phylogenetically representative of the family and covers members of the known subfamilies therein.[21] Since the enzymes were expressed in an *E. coli* host that is devoid of *β*-glucosidase activity under the conditions used, it was possible to evaluate the relative abilities of their cell lysates to cleave the set of azidoglucosides without the need for purification.

A full deep well plate of 96 GH1-bearing clones was grown, harvested, and lysed, and the supernatants were assayed. A total of 58 of the 96 clones expressed enzyme that could cleave GlcMU. Some of the 38 inactive enzymes are 6-phosphoglucosidases, which have an absolute requirement for the phosphate moiety. The rest of the clones either do not express active enzyme or are specific for other substrates.[21] The 58 clones that were able to hydrolyze GlcMU were then assayed with the full set of 2-, 3-, 4-, and 6- GlcMU analogues (each at 40 μM). Using defined dilutions of lysate for each sample, reaction progress curves were followed to completion and relative values of k_{cat}/K_M extracted by fitting the data to a first-order equation for each since, at substrate concentrations $\ll K_M$, the observed rate is approximated by $v = (k_{cat}/K_M)[E][S]$. Results for the 58 active clones are shown in Figure 1 and Table S1 as a relative % k_{cat}/K_M value, with that of GlcMU set at 100% for each.

As is readily seen, only the 6-position is tolerated to any extent, and even then, only 11 of the enzymes had k_{cat}/K_M values >20% of the values corresponding to the native substrate, GlcMU, while 38 of these had k_{cat}/K_M values under 5%. Thus, just 19% of the active library members could cleave the 6-azido substrate with a catalytic efficiency within 20% of the parent

https://doi.org/10.1021/acscentsci.1c01172
ACS Cent. Sci. 2022, 8, 656−662

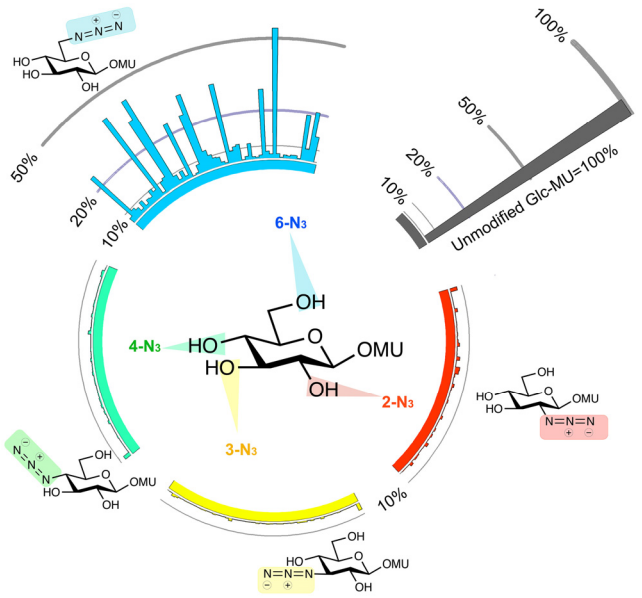


Figure 1. Relative k_{cat}/K_M value (%) of 2-, 3-, 4-, and 6-Az-GlcMU compared to GlcMU (set as 100%). Red histogram, 2-Az; yellow, 3-Az; green, 4-Az; blue, 6-Az.

substrate while the activities shown by the majority (~66%) were less than 5% that of GlcMU. For many experiments, such low activity levels are probably not going to be detected or have any real relevance in a biological setting, especially when competing with a much more efficient natural substrate.

Substitution at C2, C3, or C4 was considerably more deleterious. Such substrates are not cleaved in most cases, while a few clones have relative k_{cat}/K_M values of 1−2% of the parent. Clearly, at least in GH1 enzymes, azides at secondary carbons are essentially not accommodated. Consequently, any conclusions drawn about the presence of specific glycosidases based upon experiments with secondary azides would be erroneous while screening methodologies based upon such reagents would miss essentially everything. The fact that a small subset of the enzymes were able to accommodate an azide at C-6, albeit with activities substantially lower than that of the parent substrate, likely reflects the greater conformational flexibility at C-6, thus a much greater choice of specific binding poses to accommodate the bulky substituent.

Broader Analysis of Azide Acceptance by Screening of Metagenomic Libraries. Since it is possible that the GH1 fold imposes unusually tight constraints on substrate binding, we carried out a second series of assays on glycosidases within metagenomic libraries. This allowed us to explore more widely how well azides are accommodated, in a structurally agnostic manner. The library employed in the study was a sublibrary derived by first assaying very large libraries of metagenome-derived clones with a parent *Gluco-* or *GlcNAc*-based substrate and picking those that display activity above a defined threshold, as described previously.[23,24] Combined hits from that screen of multiple different metagenomic libraries of terrestrial, marine, and mammalian gut origin were assembled into a "hit library" which was then assayed with both the parent substrate and the azide-modified versions.

Glucosides. We previously generated a large (>300 000 clones) metagenomic library sourced from a variety of natural and engineered ecosystems, with each clone containing a ~35 kb segment of metagenomic DNA and possibly expressing several different enzymes. Screening of this large library with 4-methylumbelliferyl β-cellobioside allowed detection of cellulases, cellobiohydrolases, and β-glucosidases. 891 active clones were shown to cleave MU-glucoside or MU-cellobioside. The known GH families present on some of these fosmids (164 out of 891) were identified by bioinformatics, and the ratios of the numbers of such genes in each GH family are shown as a pie chart in Figure 2A. Particularly prominent GH families that are

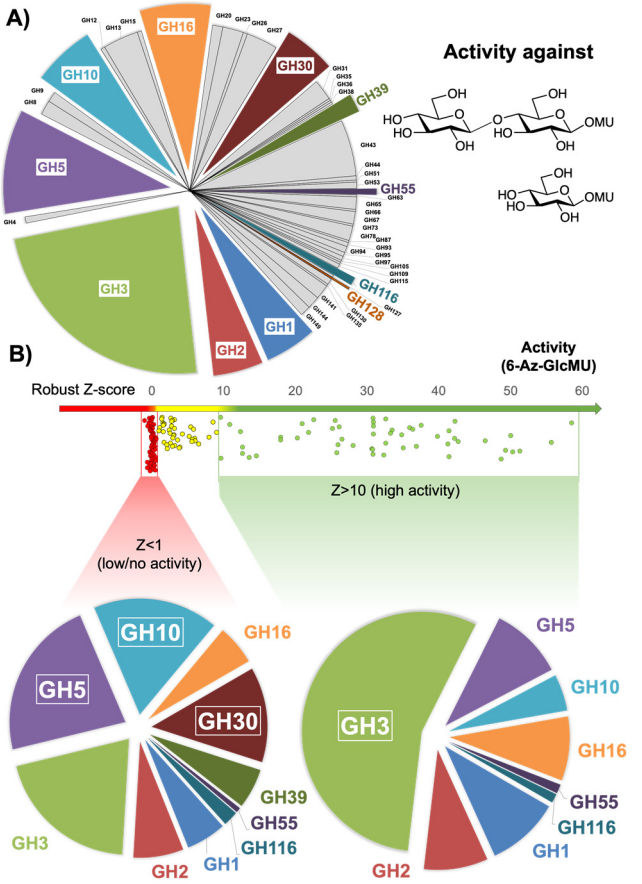


Figure 2. (A) Ratio of GH families identified in 164 fully sequenced clones that cleaved MU-cellobioside or glucoside. Families relevant to exo-acting β-glucosidase activity are colored; other families are in gray. (B) Ratios of relevant GH families contained in the 164 clones with high ($Z > 10$) or low/no ($Z < 1$) activities against 6-Az-GlcU, respectively.

known to contain enzymes with exo-acting β-glucosidase activity include GH1, GH2, GH3, GH5, GH10, GH16, GH30, GH39, GH55, GH116, and GH128, with the first four families present at relatively high abundance.[23−25] This subset of 891 glucosidases was screened for their ability to hydrolyze the 3-azido, 4-azido, and 6-azido MU-glucosides. Results very much mirrored those found with the GH1 library, even though only 15% of the active clones contained a GH1 gene, as estimated from the 164 sequenced clones. The majority (85%) contained one or more GH3 genes, with smaller numbers (<20%) containing genes from GH2, GH5, GH10, GH16, and GH30. A total of 157 of the 891 cleaved the 6-azido substrate with a robust *Z*-score >10, while only one and eight, respectively, cleaved the 3-azido and 4-azido substrates, and most of these with only *very* low activity relative to the parent (Figure S1).[23−25] Clearly, azides are again not significantly accom-

https://doi.org/10.1021/acscentsci.1c01172
ACS Cent. Sci. 2022, 8, 656−662

modated at secondary positions, especially knowing that some of the 4-azidoglucoside cleaving enzymes are likely cellobiohydrolases that happen to also cleave glucosides. Such enzymes can accommodate a bulky substituent at the 4-position of the *gluco* substrate, where the second sugar is normally bound.[26] It was of interest to see if there was differential recognition of the 6-azide substituent among enzymes of different GH families. The 164 fully sequenced fosmids are therefore given a closer look. Figure 2B shows the ratios of relevant (exoacting beta glucosidase) GH families in the 164 sequenced fosmids of high ($Z > 10$) and low/no ($Z < 1$) activity for cleavage of the 6-azido substrates. Clearly, in general, GH3 members better accommodate the 6-azide substituent while members of families such as GH5, GH10, and GH30 do not.

N-Acetylhexosaminides. In order to search beyond glucosidases, we also carried out separate metagenomic screens for enzymes that hydrolyze *N*-acetyl glucosaminides and created a sublibrary of 42 clones carrying enzymes (HexNAcases) that cleave 4-methylumbelliferyl β-*N*-acetylglucosaminide (GlcNAc-MU). Using that library, we measured relative k_{cat}/K_M values with four different azido substrates: the 3-, 4-, and 6-azidodeoxy β-*N*-acetylglucosaminides, as well as the 2-azidoacetamido glycoside. This latter substrate has an important position of substitution since many studies with azido HexNAc glycosides are performed with sugars bearing azidoacetyl groups. Further, although some HexNAcases such as those in GH3 follow a standard Koshland mechanism via a covalent glycosyl enzyme intermediate, enzymes within the principal HexNAcase family (GH20) employ a mechanism in which the acetamide functions as an intramolecular nucleophile forming an oxazoline intermediate.[27,28] This, and the fact that the azide replaces a hydrogen of the methyl group rather than a hydroxyl, may well place additional constraints on the ability to accommodate an azide at that position. Results are shown in Figure 3 and Table S2.

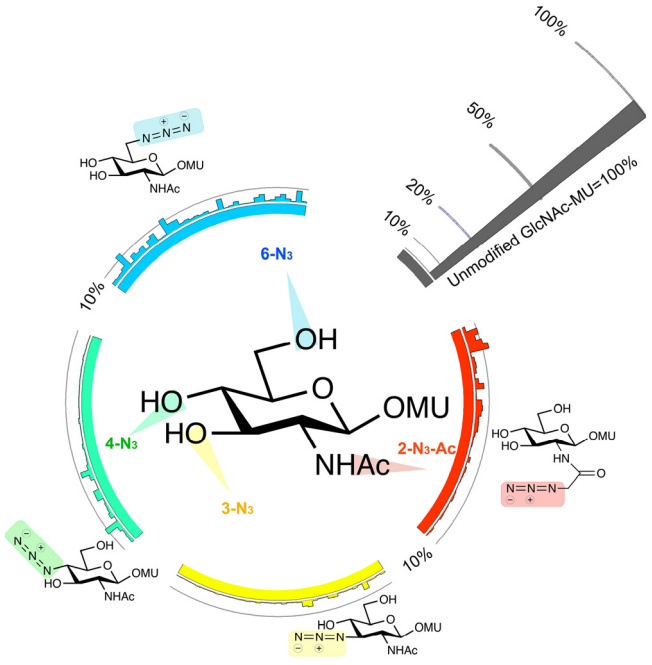


Figure 3. Relative k_{cat}/K_M values (%) of 2-AzAc, 3-Az, 4-Az, and 6-Az-GlcNAc-MU compared to GlcNAc-MU (set as 100%). Red histogram, 2-AzAc; yellow, 3-Az; green, 4-Az; blue, 6-Az.

As is readily apparent, the *N*-acetylhexosaminidases studied are even less accommodating of azide substitution at the 6-position than are the β-glucosidases. Only two of the enzymes cleave 6-azido GlcNAc substrates with k_{cat}/K_M values above 5% of the parent and only 23 with values over 1%. As with the β-glucosidases, the 3- and 4-positions are less accommodating than the 6-position, but with somewhat less stringency than the GH1 enzymes: one enzyme has activities over 5% for the 4-position and eight over 1%. Only three cleave the 3-azido substrates with k_{cat}/K_M values above 1% of the parent and none over 5%. Most interestingly, azide incorporation within the acetamide moiety is not well-tolerated. Only 3 of the 42 processed the 2-azidoacetamide substrate at a rate above 5% of the parent, while an additional 11 cleave at rates between 1% and 5% of the parent. The remaining 28 do cleave the substrate, but *extremely* slowly. This is particularly relevant given the common usage of azidoacetamide substitutions.

To assess which of these enzymes use such an oxazoline-based mechanism, we screened the full set of HexNAcases for inhibition by 25 μM GlcNAc thiazoline (**5**). This hydrolytically stable sulfur analogue of the oxazoline is bound relatively tightly by enzymes using such a mechanism. As seen in Figure 4, all of the enzymes were significantly inhibited, suggesting that all utilize an oxazoline intermediate.

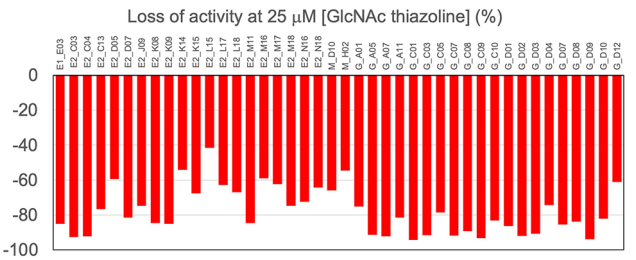


Figure 4. Inhibition of hexosaminidase activity by 25 μM GlcNAc thiazoline using 100 μM GlcNAc-MU.

DISCUSSION

The lesson emerging from these data is that care must be taken in the choice of azide-substituted sugars used in attempts to probe complex biological systems, and suitable control experiments should be performed. Key among these would be to avoid sugars bearing azides at secondary carbons unless the relevant enzymes are known to be very permissive at that position. In some contexts, azides are often (but not always) accommodated. For example, a 6-azidofucose-containing mechanism-based inactivator proved to function well on the specific fucosidases tested in recent studies, though its broader use as a probe in cases where accommodation cannot be tested must be in question.[29] In other studies of fucosidases, the C-6 azide was accepted to a degree while substitution at all other positions led to inactive reagents.[12] Due diligence was thus fittingly applied. Most notably, many sialyltransferases accommodate azides at the 5- and 9-positions quite well. Indeed, in some cases, much bigger substituents such as PEG polymers or large oligosaccharides can be appended.[30,31] This fortunate reality underlies much of the success attained in the use of 5- and 9-azido sialic acids (or their azidoManNAc precursors) in cell surface metabolic labeling studies. This is further bolstered by the fact that the four human sialidases that could be responsible for removal of these modifications all appear to accommodate an

https://doi.org/10.1021/acscentsci.1c01172
ACS Cent. Sci. 2022, 8, 656−662

azide at C-5 relatively well. However, bulkier substituents at C-9 are not as broadly tolerated.[32,33] This could result in "over-accumulation" of such modified analogues on the cell relative to the normal scenario, which would be helpful in observing incorporation but may, for some studies, be misleading.

Experiments in which azido sugars are used to probe potential pathways for incorporation of the parent sugar through pull-down or imaging studies after fluorophore conjugation can certainly be informative. These can yield valuable information on labeling sites since even a small amount of labeling can be detected if suitably sensitive methods are applied. However, such experiments cannot be relied upon to identify *all* sites of labeling or to provide relative quantitation, and results are not reliably transportable between cell types since the abilities of the processing enzymes to accommodate the azide could differ substantially between systems. This was particularly apparent in the high-throughput/high-content mass spectrometry-based analysis of azido sugar incorporation into glycoconjugates of different classes where quite different glycan profiles were observed for the same azido sugars between different cell types.[16]

However, the use of azido sugars to profile the full range of a certain class of enzymes either through enzyme assay/screen or through activity-based proteomics is likely to be misleading at best.[34] Likewise, the use of azido sugars in microbiome level studies as direct proxies for the parent sugar in studies to identify which bacteria incorporate that sugar on their surface runs the risk of only identifying the subset that can handle the azide modification, not the full set.[35]

We believe that our findings are representative of the general limitations on azide acceptance by biological systems. Indeed, in most metabolic labeling experiments, where multiple enzymes may be involved in processing the azido sugar precursor prior to incorporation, the limitations imposed by the use of azide tags are much greater since lack of azide accommodation by any one of the enzymes in the pathway would shut down incorporation. We therefore suggest that suitable care be taken in the choice of azido sugar and the design and interpretation of studies and that suitable controls should be performed. When biosynthesis/sugar incorporation is being probed, where possible, the study should include an independent assessment of the levels of processing of the modified sugar by the relevant enzymes of the nucleotide sugar processing pathway as well as by the GT in question, preferably in comparison to natural or isotopically labeled substrates. Similarly reductive experiments should likewise be applied to pathways associated with glycan processing and degradation. This can represent a major body of work if several different cell lines are being investigated but would be necessary to ensure that the cellular studies are meaningful.

MATERIALS AND METHODS

The GH1 synthetic gene library was generously provided by the Joint Genome Institute.[21] Metagenomic library construction and primary screening have been described previously.[36] Substrates and GlcNAc thiazoline were synthesized according to published procedures.[37,38]

Depletion Curve Analysis. Relative values of k_{cat}/K_M are extracted from depletion curves by fitting the data to a first-order equation for each, since at substrate concentrations $\ll K_M$, the observed rate is approximated by $v = (k_{cat}/K_M)[E][S]$. By then taking the ratio (and correcting for any differences in amount of enzyme used), the relative value of k_{cat}/K_M is obtained. Data from early regions of curves from substrate concentrations above this concentration regime do not fit a first-order curve[18] and so are discarded. For those reactions for which the progress is too slow to reasonably monitor until substrate depletion (mostly 2-azidodeoxy/azidoacetimido, 3-azidodeoxy, and 4-azidodeoxy substituted GlcMUs and GlcNAcMUs), initial velocities were used to calculate approximate values of relative k_{cat}/K_M, assuming $[S] \ll K_M$.

Determination of Relative Activities (GH 1 Synthetic Library). 7.5 mL of ZYP-5052 media containing 50 mg/L carbenicillin was inoculated with 5 μL of glycerol stock, incubated at 37 °C for 16 h, and shaken at 225 rpm. Cells were harvested by centrifugation at 3900 rpm for 20 min. After supernatant was decanted, cell pellets were resuspended in 1.2 mL of lysis buffer (50 mM Na-HEPES, 1% Triton X-100, 0.25 mg/mL lysozyme, benzonase, cOmplete Protease Inhibitor EDTA-free, pH 7.0) and incubated for 2 h with stirring at room temperature before centrifugation (3900 rpm, 30 min). Supernatants, with appropriate dilutions, were used in the next step.

In 96-well plates (Corning 3694 half area, black) was added 2× reaction buffer (50 mM Na-HEPES, pH 7.0, 80 μM of substrate, 40 μL). Reactions were initiated by the addition of 40 μL of supernatant of 1× lysate (all substrates) or 1/10−1/1000× diluted lysate (GlcMU) into each well, at 22 °C. Reactions were analyzed by fluorescence spectroscopy on a Beckman Coulter DTX-880 multimode detector (λ_{ex} = 360 nm, λ_{em} = 465 nm). Dilution factors were multiplied back when calculating relative rate constants.

Determination of Relative Activities (Metagenomic Libraries). A 1 mL portion of Luria broth (LB) containing 12.5 mg/L chloramphenicol and 100 mg/L arabinose was inoculated and incubated at 37 °C, 225 rpm, for 20 h. Cells were harvested by centrifugation at 3000 rpm for 30 min. After supernatant was decanted, cell pellets were resuspended in 400 μL of lysis buffer (50 mM Na-HEPES, 1% Triton X-100, 0.25 mg/mL lysozyme, Benzonase nuclease, cOmplete Protease Inhibitor EDTA-free, pH 7.0) and incubated for 2 h, with shaking at 60 rpm with glass beads, at room temperature before centrifugation (3000 rpm, 30 min). Supernatants, with appropriate dilutions, were used in the next step.

In 96-well plates (Corning 3694 half area, black) was added 2× reaction buffer (50 mM Na-HEPES, pH 7.0, 80 μM of substrate, 50 μL). Reactions were initiated by the addition of 50 μL of supernatant of 1× lysate (all substrates) or 1/10× diluted lysate (GlcNAcMU) into each well, at 22 °C. Reactions were analyzed by fluorescence spectroscopy on a Beckman Coulter DTX-880 multimode detector (λ_{ex} = 360 nm, λ_{em} = 465 nm). Dilution factors were multiplied back when calculating relative rates.

Inhibition of N-Acetylhexosaminidases by GlcNAc Thiazoline. Lysates of clones showing hexosaminidase activity were diluted 10×. 0 or 25 μM GlcNAc thiazoline was added to each lysate which was then incubated at 25 °C for 30 min. Reactions were initiated by the addition of 100 μM (final) GlcNAc-MU in Na-HEPES buffer (50 mM, pH = 7) and monitored by fluorescence spectroscopy on a Beckman Coulter DTX-880 multimode detector (λ_{ex} = 360 nm, λ_{em} = 465 nm).

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acscentsci.1c01172.

https://doi.org/10.1021/acscentsci.1c01172
ACS Cent. Sci. 2022, 8, 656−662

Relative k_{cat}/K_M and Z scores (PDF)

AUTHOR INFORMATION

Corresponding Author

Stephen G. Withers − *Department of Chemistry, University of British Columbia, Vancouver, British Columbia V6T1Z1, Canada;* orcid.org/0000-0002-6722-5701; Email: withers@chem.ubc.ca

Authors

Feng Liu − *Department of Chemistry, University of British Columbia, Vancouver, British Columbia V6T1Z1, Canada*

Hong-Ming Chen − *Department of Chemistry, University of British Columbia, Vancouver, British Columbia V6T1Z1, Canada*

Zachary Armstrong − *Department of Chemistry, University of British Columbia, Vancouver, British Columbia V6T1Z1, Canada;* Present Address: Z.A.: Leiden Institute of Chemistry, Gorlaeus Building, Einsteinweg 552333 CC Leiden, Netherlands

Complete contact information is available at: https://pubs.acs.org/10.1021/acscentsci.1c01172

Author Contributions

F.L. and Z.A. performed all screening and enzyme assays. H.-M.C. performed all syntheses. S.G.W. and F.L. wrote the manuscript with editing from the other authors. S.G.W. conceived and supervised the project.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Professor S.J. Hallam for introducing us to metagenomics and advising on our screening efforts. We also thank the Joint Genome Institute of the U.S. Department of Defense for generous provision of the GH1 synthetic gene library. We thank the Natural Sciences and Engineering Research Council of Canada (NSERC), the Canadian Institutes of Health Research (CIHR), and the Canadian Glycomics Network (GlycoNet) for financial support of this work.

ABBREVIATIONS

Az, azido; GH, glycoside hydrolase; Glc, glucose; GlcNAc, *N*-acetylglucosamine, HEPES, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; ManNAc, *N*-acetylmannosamine; MU, 4-methylumbelliferyl

REFERENCES

(1) Kolb, H. C.; Finn, M. G.; Sharpless, K. B. Click Chemistry: Diverse Chemical Function from a Few Good Reactions. *Angew. Chem., Int. Ed.* **2001**, *40* (11), 2004−2021.

(2) Meldal, M.; Tornoe, C. W. Cu-catalyzed azide-alkyne cycloaddition. *Chem. Rev.* **2008**, *108* (8), 2952−3015.

(3) Saxon, E.; Armstrong, J. I.; Bertozzi, C. R. A Traceless Staudinger Ligation for the Chemoselective Synthesis of Amide Bonds. *Org. Lett.* **2000**, *2* (14), 2141−2143.

(4) Agard, N. J.; Baskin, J. M.; Prescher, J. A.; Lo, A.; Bertozzi, C. R. A comparative study of bioorthogonal reactions with azides. *ACS Chem. Biol.* **2006**, *1* (10), 644−648.

(5) Shieh, P.; Bertozzi, C. R. Design strategies for bioorthogonal smart probes. *Organic & Biomolecular Chemistry* **2014**, *12* (46), 9307−9320.

(6) Laughlin, S. T.; Bertozzi, C. R. Imaging the glycome. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (1), 12−17.

(7) Luchansky, S. J.; Goon, S.; Bertozzi, C. R. Expanding the diversity of unnatural cell-surface sialic acids. *Chembiochem* **2004**, *5* (3), 371−374.

(8) Caneque, T.; Mueller, S.; Rodriguez, R. Visualizing biologically active small molecules in cells using click chemistry. *Nature Reviews Chemistry* **2018**, *2* (9), 202−215.

(9) Pedowitz, N. J.; Pratt, M. R. Design and synthesis of metabolic chemical reporters for the visualization and identification of glycoproteins. *Rsc Chemical Biology* **2021**, *2* (2), 306−321.

(10) Row, R. D.; Prescher, J. A. Constructing New Bioorthogonal Reagents and Reactions. *Acc. Chem. Res.* **2018**, *51* (5), 1073−1081.

(11) Vocadlo, D. J.; Bertozzi, C. R. A strategy for functional proteomic analysis of glycosidase activity from cell lysates. *Angew. Chem., Int. Ed.* **2004**, *43* (40), 5338−5342.

(12) Rabuka, D.; Hubbard, S. C.; Laughlin, S. T.; Argade, S. P.; Bertozzi, C. R. A chemical reporter strategy to probe glycoprotein fucosylation. *J. Am. Chem. Soc.* **2006**, *128* (37), 12078−12079.

(13) Jackson, E. G.; Cutolo, G.; Yang, B.; Yarravarapu, N.; Burns, M. W. N.; Bineva-Todd, G.; Roustan, C.; Thoden, J. B.; Lin-Jones, H. M.; van Kuppevelt, T. H.; Holden, H. M.; Schumann, B.; Kohler, J. J.; Woo, C. M.; Pratt, M. R. 4-Deoxy-4-fluoro-GalNAz (4FGalNAz) Is a Metabolic Chemical Reporter of O-GlcNAc Modifications, Highlighting the Notable Substrate Flexibility of O-GlcNAc Transferase. *ACS Chem. Biol.* **2022**, *17*, 159.

(14) Ma, X.; Liu, P.; Yan, H.; Sun, H.; Liu, X.; Zhou, F.; Li, L.; Chen, Y.; Muthana, M. M.; Chen, X.; Wang, P. G.; Zhang, L. Substrate Specificity Provides Insights into the Sugar Donor Recognition Mechanism of O-GlcNAc Transferase (OGT). *PLoS One* **2013**, *8* (5), e63452.

(15) Wen, L.; Gadi, M. R.; Zheng, Y.; Gibbons, C.; Kondengaden, S. M.; Zhang, J.; Wang, P. G. Chemoenzymatic Synthesis of Unnatural Nucleotide Sugars for Enzymatic Bioorthogonal Labeling. *ACS Catal.* **2018**, *8* (8), 7659−7666.

(16) Shajahan, A.; Supekar, N. T.; Wu, H.; Wands, A. M.; Bhat, G.; Kalimurthy, A.; Matsubara, M.; Ranzinger, R.; Kohler, J. J.; Azadi, P. Mass Spectrometric Method for the Unambiguous Profiling of Cellular Dynamic Glycosylation. *ACS Chem. Biol.* **2020**, *15* (10), 2692−2701.

(17) Birnbaum, G. I.; Giziewicz, J.; Gabe, E. J.; Lin, T. S.; Prusoff, W. H. Structure and conformation of 3′-azido-3-deoxythymidine (AZT), an inhibitor of the HIV (AIDS) virus. *Canadian Journal of Chemistry-Revue Canadienne De Chimie* **1987**, *65* (9), 2135−2139.

(18) Crompton, I. E.; Waley, S. G. The determination of specificity constants in enzyme-catalyzed reactions. *Biochem. J.* **1986**, *239* (1), 221−224.

(19) Vocadlo, D. J.; Wicki, J.; Rupitz, K.; Withers, S. G. Mechanism of Thermoanaerobacterium saccharolyticum ss-xylosidase: Kinetic studies. *Biochemistry* **2002**, *41* (31), 9727−9735.

(20) Lombard, V.; Ramulu, H. G.; Drula, E.; Coutinho, P. M.; Henrissat, B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* **2014**, *42* (D1), D490−D495.

(21) Heins, R. A.; Cheng, X. L.; Nath, S.; Deng, K.; Bowen, B. P.; Chivian, D. C.; Datta, S.; Friedland, G. D.; D'Haeseleer, P.; Wu, D. Y.; Tran-Gyamfi, M.; Scullin, C. S.; Singh, S.; Shi, W. B.; Hamilton, M. G.; Bendall, M. L.; Sczyrba, A.; Thompson, J.; Feldman, T.; Guenther, J. M.; Gladden, J. M.; Cheng, J. F.; Adams, P. D.; Rubin, E. M.; Simmons, B. A.; Sale, K. L.; Northen, T. R.; Deutsch, S. Phylogenomically Guided Identification of Industrially Relevant GH1 beta-Glucosidases through DNA Synthesis and Nanostructure-Initiator Mass Spectrometry. *ACS Chem. Biol.* **2014**, *9* (9), 2082−2091.

(22) Armstrong, Z.; Liu, F.; Chen, H.-M.; Hallam, S. J.; Withers, S. G. Systematic Screening of Synthetic Gene-Encoded Enzymes for Synthesis of Modified Glycosides. *ACS Catal.* **2019**, *9* (4), 3219−3227.

(23) Armstrong, Z. *Harnessing Natural Diversity for the Discovery of Glycoside Hydrolases and Design of New Glycosynthases*; University of British Columbia: Vancouver, 2018.

(24) Armstrong, Z.; Liu, F.; Kheirandish, S.; Chen, H.-M.; Mewis, K.; Duo, T.; Morgan-Lang, C.; Hallam, S. J.; Withers, S. G. High-Throughput Recovery and Characterization of Metagenome-Derived

Glycoside Hydrolase-Containing Clones as a Resource for Biocatalyst Development. *Msystems* **2019**, *4* (4), e00082.
(25) Armstrong, Z.; Mewis, K.; Liu, F.; Morgan-Lang, C.; Scofield, M.; Durno, E.; Chen, H. M.; Mehr, K.; Withers, S. G.; Hallam, S. J. Metagenomics reveals functional synergy and novel polysaccharide utilization loci in the Castor canadensis fecal microbiome. *ISME J.* **2018**, *12* (11), 2757−2769.
(26) Williams, S. J.; Hekmat, O.; Withers, S. G. Synthesis and testing of mechanism-based protein-profiling probes for retaining endo-glycosidases. *Chembiochem* **2006**, *7* (1), 116−124.
(27) Jones, C. S.; Kosman, D. J. Purification, properties, kinetics and mechanism of beta-N-acetylglucosaminidase from Aspergillus niger. *J. Biol. Chem.* **1980**, *255* (24), 11861−11869.
(28) Zechel, D. L.; Withers, S. G. Glycosidase mechanisms: Anatomy of a finely tuned catalyst. *Acc. Chem. Res.* **2000**, *33* (1), 11−18.
(29) Luijkx, Y. M. C. A.; Jongkees, S.; Strijbis, K.; Wennekes, T. Development of a 1,2-difluorofucoside activity-based probe for profiling GH29 fucosidases. *Organic & Biomolecular Chemistry* **2021**, *19* (13), 2968−2977.
(30) Capicciotti, C. J.; Zong, C.; Sheikh, M. O.; Sun, T.; Wells, L.; Boons, G.-J. Cell-Surface Glyco-Engineering by Exogenous Enzymatic Transfer Using a Bifunctional CMP-Neu5Ac Derivative. *J. Am. Chem. Soc.* **2017**, *139* (38), 13342−13348.
(31) DeFrees, S.; Wang, Z.-G.; Xing, R.; Scott, A. E.; Wang, J.; Zopf, D.; Gouty, D. L.; Sjoberg, E. R.; Panneerselvam, K.; Brinkman-Van der Linden, E. C. M.; Bayer, R. J.; Tarp, M. A.; Clausen, H. GlycoPEGylation of recombinant therapeutic proteins produced in Escherichia coli. *Glycobiology* **2006**, *16* (9), 833−843.
(32) Guo, T.; Heon-Roberts, R.; Zou, C.; Zheng, R.; Pshezhetsky, A. V.; Cairo, C. W. Selective Inhibitors of Human Neuraminidase 1 (NEU1). *J. Med. Chem.* **2018**, *61* (24), 11261−11279.
(33) Albohy, A.; Zhang, Y.; Smutova, V.; Pshezhetsky, A. V.; Cairo, C. W. Identification of Selective Nanomolar Inhibitors of the Human Neuraminidase, NEU4. *ACS Med. Chem. Lett.* **2013**, *4* (6), 532−537.
(34) Ruprecht, C.; Bartetzko, M. P.; Senf, D.; Lakhina, A.; Smith, P. J.; Soto, M. J.; Oh, H.; Yang, J.-Y.; Chapla, D.; Silva, D. V.; Clausen, M. H.; Hahn, M. G.; Moremen, K. W.; Urbanowicz, B. R.; Pfrengle, F. A Glycan Array-Based Assay for the Identification and Characterization of Plant Glycosyltransferases. *Angew. Chem., Int. Ed.* **2020**, *59* (30), 12493−12498.
(35) Weiss, G. A.; Grabinger, T.; Glaus Garzon, J.; Hasler, T.; Greppi, A.; Lacroix, C.; Khanzhin, N.; Hennet, T. Intestinal inflammation alters mucosal carbohydrate foraging and monosaccharide incorporation into microbial glycans. *Cellular Microbiology* **2021**, *23* (1), e13269.
(36) Armstrong, Z.; Liu, F.; Kheirandish, S.; Chen, H.-M.; Mewis, K.; Duo, T.; Morgan-Lang, C.; Hallam, S. J.; Withers, S. G. High-Throughput Recovery and Characterization of Metagenome-Derived Glycoside Hydrolase-Containing Clones as a Resource for Biocatalyst Development. *Msystems* **2019**, *4* (4), e00082.
(37) Chen, H.-M.; Withers, S. G. Synthesis of azido-deoxy and amino-deoxy glycosides and glycosyl fluorides for screening of glycosidase libraries and assembly of substituted glycosides. *Carbohydr. Res.* **2018**, *467*, 33−44.
(38) Knapp, S.; Vocadlo, D.; Gao, Z. N.; Kirk, B.; Lou, J. P.; Withers, S. G. NAG-thiazoline, an N-acetyl-beta-hexosaminidase inhibitor that implicates acetamido participation. *J. Am. Chem. Soc.* **1996**, *118* (28), 6804−6805.

https://doi.org/10.1021/acscentsci.1c01172
ACS Cent. Sci. 2022, 8, 656−662