Studies Towards Finding Selective Substrates for Human β-Glucosidases

by

Jeongmee Hailey Kim

B.Sc., University of the Fraser Valley, 2017

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Master of Science

in the
Department of Chemistry
Faculty of Science

SIMON FRASER UNIVERSITY
Fall 2021

Declaration of Committee

Name:	**Jeongmee Hailey Kim**
Degree:	**Master of Science (Chemistry)**
Title:	**Studies Towards Finding Selective Substrates for Human β-Glucosidases**
Committee:	**Chair:** **Hogan Yu** Professor, Chemistry
	David J. Vocadlo Supervisor Professor, Chemistry
	Andrew J. Bennet Committee Member Professor, Chemistry
	Robert A. Britton Committee Member Professor, Chemistry
	Robert N. Young Examiner Professor, Chemistry

Abstract

Humans have three types of β-glucosidases (GBA) enzymes named GBA1, GBA2, and GBA3 that break down glucosylceramide (GlcCer). Studies have suggested that GBA2 is involved in Gaucher disease via feedback mechanism with GBA1, and lack of GBA2 activity leads to other neurodegenerative diseases. However, currently with no selective substrate probe developed for GBA2, making it difficult to study the enzyme activity, the role of GBA2 in pathogenesis of the diseases remains unclear. To find GBA2 selective substrates, the substrate specificity of GBA enzymes was investigated by preparing a series of mono-substituted deoxyfluoro and deoxy analogs of the substrate and testing the compounds with GBA enzymes to obtain Michaelis-Menten parameters, which allowed the evaluation of the affinity of each compound against each enzyme. This approach led to a compound that shows selectivity towards GBA2, which can be further developed into a cell lysate probe for brain cells.

Keywords: Gaucher disease; glucosidase; selective substrate; Michaelis-Menten kinetics; deoxyfluoro analogs; deoxy analogs

Acknowledgements

I would like to thank all the great people who have helped me achieve this thesis project and supported me throughout my Master's degree journey. First, I would like to begin by thanking Dr. David Vocadlo for providing me the opportunity to do research work in his lab with his amazingly smart team. I am forever thankful for his invaluable mentorship, encouragement and belief in my abilities. I would like to thank Dr. Andrew Bennet for his guidance and support by his insightful feedbacks during group meetings, which pushed me to think deeper. I would also like to thank Dr. Robert Britton for his valuable insights and challenging questions during committee meetings, which helped me to shape my project.

I want to extend my thanks to all members of the Vocadlo lab and Bennet lab whom I had great pleasure to work with. Among all incredible team members, I would like to give special thanks to Dr. Sandeep Bhosale for his guidance in chemical synthesis, Dr. Femi Akintola for helping me out tremendously with enzyme kinetics, Dr. Marco Farren-Dai for drawing me Pymol diagrams, Matthew Deen for sharing his knowledge on GBA enzymes as well as his creative ideas on my project. I also want to thank Matthew Deen and Dr. Roger Ashmus for proofreading this thesis work and help me to shape it better.

Last, but not least, I would like to express my deep gratitude to my family and friends as I would not be where I am today without their care, trust and support. I am truly indebted to my mom and my sister Gina for their love, encouragement and patience. I am so fortunate to have my partner Alec by my side who always pushes me to be better and to help me see the bigger picture in life. I would like to also express my thanks to my friend Kristen for helping me out with writing and also making my time at Simon Fraser University enjoyable and fun.

Table of Contents

List of Tables

List of Figures

List of Acronyms

CBE	Conduritol B Epoxide
CNS	Central Nervous System
ER	Endoplasmic Reticulum
ES	Enzyme-Substrate
ESI	Electron Spray Ionization
FRET	Forster Resonance Energy Transfer
GD	Gaucher Disease
GH	Glycoside Hydrolase
HRMS	High Resolution Mass Spectrometry
LBD	Lewy Body Dementia
LC/MS	Liquid Chromatography/Mass Spectrometry
M-M	Michaelis-Menten
MS	Mass Spectrometry
NMR	Nuclear Magnetic Resonance
PD	Parkinson's Disease
ppm	parts per million
RFU	Relative Fluorescence Unit
TLC	Thin Layer Chromatography
TS	Transition State
UV	Ultra Violet

Chapter 1.

Introduction

1.1. Carbohydrates

1.1.1. Significance of carbohydrate molecules in living organisms

Comprised of carbon, hydrogen, and oxygen, carbohydrates ($C_m(H_2O)_n$) are originally thought to be only as one of the four fundamental macromolecule building blocks of life alongside proteins, lipids, and nucleic acids. Carbohydrates are commonly thought to be fundamental for their structural and energy roles. An example of carbohydrate as an energy storage molecule for living cells is glycogen, which is comprised of multibranched glucose molecules. Another well-known example of a carbohydrate molecule that plays a structural role is cellulose, a component of plant cell walls that maintains cell shape and integrity. However, the roles of carbohydrate molecules in humans are much more diverse than just as an energy source or as a structural component; in fact, carbohydrates have broad significance in living organisms.

Following their biosynthesis, carbohydrates can be covalently attached to lipid molecules to form glycolipids through glycosylation. During glycosylation, carbohydrates attach to amine and hydroxyl functional groups of lipids to form *N*-and *O*-glycosidic bonds, respectively. Each carbohydrate unit is attached to the lipid sequentially by a corresponding glycosyltransferase, the enzyme that catalyzes the transfer of an activated carbohydrate unit to a nucleophilic glycosyl acceptor. Glycolipids are present on all eukaryotic cell membranes, with the lipid residue embedded in the phospholipid bilayer and the sugar moiety facing the extracellular space. The glycan moiety that faces outside the cell is involved in various vital biological processes. For example, glycans bind to carbohydrate-binding proteins on neighbouring cells where cell to cell recognition occurs and processes such as cell growth, regulation or apoptosis proceeds as required.[1] Cell surface glycans are also involved in immune responses. Viruses recognize and bind to cell surface glycan receptors composed of glycoproteins or glycolipids. The bound viruses are internalized into the cell, and the cell's immune mechanism works to eradicate the virus and stop it from replicating.

Carbohydrates not only attach to lipids to increase their functionality but the same process also occurs with proteins. To increase the functional diversity of the proteome, carbohydrates are also covalently attached to proteins through a process of a post-translational modification. Post-translational modification often occurs to the side chains of serine, threonine and tyrosine to form *O*-linked glycosidic bonds or lysine, arginine, histidine, asparagine and glutamine to form *N*-linked glycosidic bonds. The physiological roles of glycoproteins are diverse and range from structural, such as collagen, to immunological, such as immunoglobulin (antibody). Glycosylation on collagens allows the protein to form structural fibrils, which act as a scaffold for structural integrity and strengthen connective tissues.[2] Glycans on immunoglobulins not only play a structural role but also are involved in antigen binding, such as bacteria and viruses, and proceeds to destruction.

Glycans are ubiquitous in living systems, as described above. However, relative to the advances made in fields like genomics and proteomics, glycoscience has progressed more slowly. However, for the past several decades, researchers in glycoscience have established that glycans play direct roles in human physiology and the etiology of diseases, and countless studies have been done that have enabled advances in our understanding of human diseases.[3] Continued advances in elucidating the physiological roles of glycans in the human body as well as enzymes and other factors that affect the glycan functions will allow further understanding of glycans and should contribute to the development of therapeutic drugs.

1.1.2. Nomenclature of glucose

Glucose is a hexose with a molecular formula of $C_6H_{12}O_6$, and it is the most abundant carbohydrate found in living systems. Like most sugar molecules, glucose mutarotates in an aqueous solution, and it exists in equilibrium as a complex mixture of pyranose, furanose, and acyclic forms.[4] The two major forms are the α-pyranose and the β-pyranose that interconvert via the acyclic aldehyde (**Figure 1.1**).

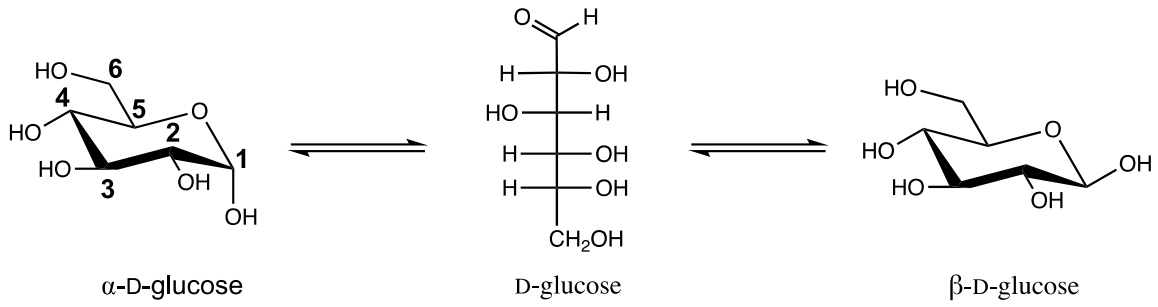
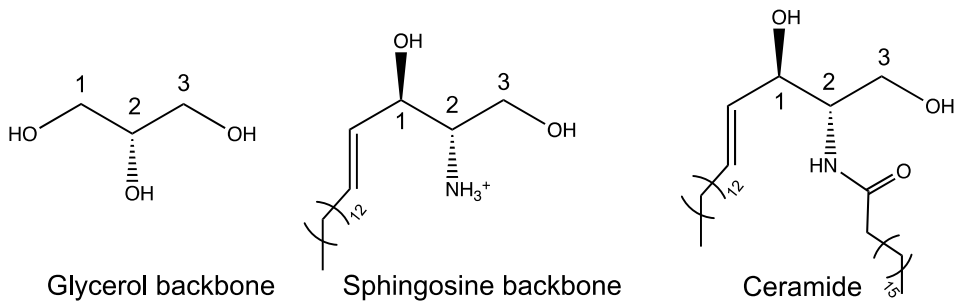


Figure 1.1. Mutarotation of glucose

The carbon atoms on glucose are numbered starting with C_1 at the hemiacetal and continuing around the ring ending at C_6 with primary alcohol (**Figure 1.1**). The C_1 carbon is referred to as the anomeric carbon. When a group is attached to the anomeric oxygen via a glycosidic bond, the sugar is called a glucoside. The absolute configurations of the two enantiomers of a sugar molecule are annotated as L or D depending on the stereochemistry of the hydroxyl group attached to the highest numbered chiral center. When drawn in Fisher projection, L- and D-configurations have the OH-group on the highest numbered center of chirality on the left and right, respectively.

1.2. Glycosphingolipids

The term glycolipid encompasses all lipid molecules with covalently attached carbohydrates. There are two major types of lipid backbone molecules: glycerol and sphingosine. Both the glycerol and sphingosine backbones consist of three carbons. The glycerol backbone has a hydroxyl group on each carbon whereas the sphingosine backbone has a long alkyl chain attached to C1 and an amine group at C2. The amine group at C2 of sphingosine can be acylated with a long-chain fatty acid, generating ceramide.



Glycerol backbone Sphingosine backbone Ceramide

Figure 1.2. Glycerol and sphingosine backbone of glycolipids and ceramide

Glycosphingolipids (GSL) are a sub-class of glycolipids that contain ceramide as the backbone and a glycan connected to the C3 hydroxyl via a β-glycosidic bond. GSLs are found in many species, from bacteria and fungi to all animals, and as mentioned earlier, they are predominantly located on cell membranes and function in cell recognition and communication as well as in immune response. The discovery of GSLs stemmed from observations of accumulated GSLs in patient tissue samples diagnosed with lysosomal storage diseases as well as those with genetic disorders that result in a lack of enzymatic activity that degrades certain GSLs.[5] Glucosylceramide (GlcCer), a type of GSL, became a molecule of interest when researchers observed its accumulation in pathogenic levels within spleens of Gaucher disease patients.[6] From there, countless studies were done on GlcCer accumulation in relation to lysosomal storage disease[6] and neurodegenerative diseases[7], a topic that will be further discussed in Section 1.6.

1.3. Glucosylceramide in the glycosphingolipid metabolic pathway

GlcCer is one of the products of the glycosphingolipid metabolic pathway. The synthesis of GSLs occurs in a step-wise fashion, and it starts in the endoplasmic reticulum (ER), where the condensation of sphingosine with acyl-Coenzyme A to produce ceramide is catalyzed by ceramide synthase.[8] Ceramide then follows one of the two pathways to the Golgi apparatus. It can be transferred to the luminal side of the ER then trafficked to the *trans*-Golgi network by ceramide transfer protein, where it is primarily used to synthesize sphingomyelins. Ceramide can also end up in the *cis*-Golgi via vesicular transport. On the cytosolic leaflet of the *cis*-Golgi, ceramide is glycosylated by glucosylceramide synthase (GCS) to form GlcCer. GlcCer flips into the lumen of the Golgi by flippase, where it is being used downstream as a major precursor for various

types of complex GSLs via step-wise glycosylation catalyzed by glycosyltransferases. Then GSLs are transferred onto the cell membrane, where they are incorporated into lipid rafts, a rigid and ordered mobile domain on the outer leaflet of the plasma membrane enriched with lipids and GSLs. The glycan moieties of GSLs are involved in multiple different processes, including regulation of cell growth, differentiation and cell to cell communication. They also have been found to regulate cell apoptosis.[9]

GSLs go through the endocytic pathway from the plasma membrane to lysosomes for degradation.[8] The GSLs are degraded by internalization into endocytic vesicles followed by delivery to lysosomes, where specific glycosyl hydrolases along with co-factor proteins sequentially cleave off each sugar unit. After hydrolysis of all glycan units, ceramide remains as the product at the end of the GSL catabolic pathway.

1.4. β-Glucosidases: enzymes that act on glucosylceramide

Lysosomes are the digestive compartment of the cell containing many enzymes allowing for the degradation of many different macromolecules. Glycoside hydrolases (GH) present in the lysosomes degrade glycans of GSLs to generate GlcCer. The final glucose unit of GlcCer is cleaved by lysosomal β-glucosidase (GBA1). In fact, there are three types of β-glucosidases in human cells that degrade GlcCer, distinguished by their different cellular locations. Along with the lysosomal β-glucosidase, there is a non-lysosomal β-glucosidase (GBA2) that is responsible for degrading GlcCer localized on the cytosolic face of the membrane of ER and Golgi and cytosolic β-glucosidase (GBA3), which possesses significant β-glucosidase activity, but its natural substrate is unknown.[10,11]

All carbohydrate-active enzymes, including glycoside hydrolases, are classified into different families in the CAZy (Carbohydrate-Active Enzymes) database system according to their protein sequence and structural relevance.[12] While all three human β-glucosidases catalyze the degradation of GlcCer, each enzyme is categorized into different families: lysosomal GBA1, membrane-associated GBA2, and cytosolic GBA3 belong to family GH 30, GH 116, and GH 1, respectively. The fact that each β-glucosidase belongs to a different family suggests that although they catalyze the same

reaction, they are structurally different at the active site where the substrate binds and the reaction occurs.

1.4.1. Lysosomal β-glucosidase: GBA1

The primary degradation of GlcCer is by the lysosomal β-glucosidase, GBA1, also named GCase. GBA1 is encoded by the *GBA* gene, which consists of a 7.6 kb sequence with 11 exons and 10 introns and is located on chromosome 1.[13] GBA1 has optimum activity around pH 5, the pH of the lysosomal compartment, and it is associated at the lysosomal membrane.[14] The enzyme is ubiquitously distributed throughout various tissues.[15] *In vivo*, degradation of GlcCer by GBA1 is assisted by a cofactor peptide called saposin C (Sap C).[16] It was proposed by Romero et al. that Sap C enhances GBA1 function by stabilizing conformational change at the GBA1 substrate-binding site.[17] They have also shown that a deficiency of Sap C leads to a decrease in GBA1 activity, and humans with a Sap C deficiency exhibits symptoms of Gaucher disease (GD), a type of neurodegenerative disease. A substrate docking model of GBA1 showed the glucose of GlcCer fitting into the substrate-enzyme binding pocket and the lipid moiety sticking out of the pocket.[18] It is proposed that the ceramide moiety could be interacting with Sap C or may assist in the interaction of GBA1 with the lysosomal phospholipid bilayer during the catalysis.[16]

The X-ray structure of human GBA1 was first reported by Dvir et al. in 2003.[18] GBA1 consists of 497 amino acids and had a molecular weight of 59 kDa. This X-ray crystal structure at 2.0 Å resolution revealed the configuration of its active site. The catalytic residues E235 and E340, which act as an acid/base catalyst and a nucleophile, respectively, are found in domain III.[18] The distance between the carboxylic oxygens of the two catalytic residues is about 5-5.5 Å, a distance that is highly conserved within the GH30 family of enzymes.[12] The hydrogen-bonding interactions of the inhibitor, *N*-nonyldeoxynojirimycin (NN-DNJ), with amino acid residues at the GBA1 active site were elucidated (**Figure 1.4**). The 2OH-group was H-bonded with two glutamate residues (Glu 340, Glu 235) and one asparagine residue (Asn 234). 3OH and 4OH-groups were H-bonded with an aspartate residue (Asp 127), and each OH-group was also interacting with two different tryptophan residues (Trp 179, Trp 381). The 6OH-group formed only one H-bond interaction with Asp 396.[19]

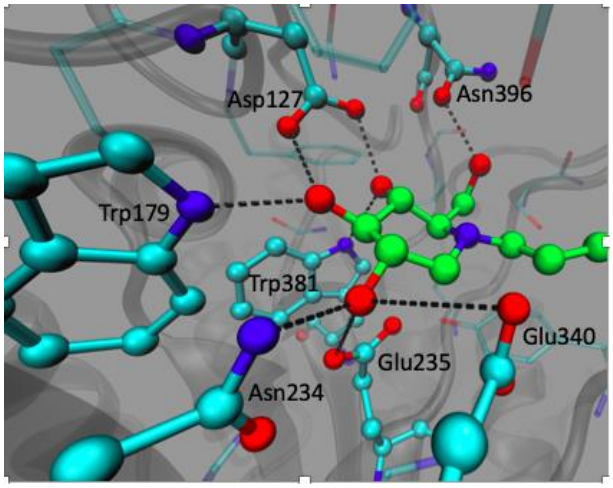


Figure 1.3. Interactions of GBA1 active site residues with NN-DNJ[18]

It is known that substrates with a bulky substituents at the 6O-position are tolerated by GBA1 and exhibit selectivity toward GBA1 over other β-glucosidases.[20] Witte et al. designed a fluorescent probe for visualizing the activity of GBA1 *in situ* by taking a cyclophellitol, a potent inhibitor of β-glucosidases, and attaching a fluorophore onto the pseudo-6C-position, which gave the probe selectivity towards GBA1.[21] Appending a morpholine derivative on the 6O through a triazole linker gave a non-selective resorufin β-glucoside selectivity towards GBA1.[20] Conduritol B epoxide (CBE) does not contain 6C and thus, this structural feature contributes to CBE being a selective inhibitor of GBA1 over GBA2 and GBA3. The above studies proved that a 6-hydroxy group is not crucial for the activity of GBA1.

1.4.2. Membrane-associated β-glucosidase: GBA2

Non-lysosomal β-glucosidase (GBA2) was first reported by Matern and coworkers in 1992 when they noticed that liver cells contained an enzyme that acts on bile acid β-glucosides.[22] Soon after, in 2003, Van Weely et al. described a β-glucosidase with distinct features from the lysosomal β-glucosidase GBA1. This new β-glucosidase was insensitive to CBE, an irreversible inhibitor of GBA1, and did not require Sap C for its activity.[23] Also, it had optimal activity at pH 5.5 – 6.0, a pH slightly less acidic than the optimal pH of GBA1.[24] Unlike the lysosomal GBA1, the newly identified β-glucosidase was not deficient in patients with GD.[23] The identity of this new β-glucosidase was further confirmed later around 2006 as a non-lysosomal β-glucocerebrosidase, involved

in extra-lysosomal metabolism of GlcCer.[25,26] The non-lysosomal β-glucocerebrosidase is encoded by the *GBA2* gene, which is localized on chromosome 9. The enzyme itself is often referred to as GBA2. The GBA2 protein has 927 amino acids in length with a molecular weight of 104 kDa.[27] GBA2 is ubiquitously distributed but primarily expressed in the brain and testes and expressed in lower levels in the liver, spleen, and small intestine.[27] The primary structure of GBA2 was first elucidated in 2001 by Matern et al., where they found that GBA2 possesses a transmembrane domain leading to the proposal that GBA2 is a transmembrane protein.[27] However, in 2013, it was established that GBA2 is actually not a protein embedded on the membrane but instead is tightly associated on the cytosolic face of the membrane of ER and Golgi.[10] Because of the locations of GBA2, researchers suggested that GBA2 works to degrade GlcCer that somehow ended up near ER and Golgi.

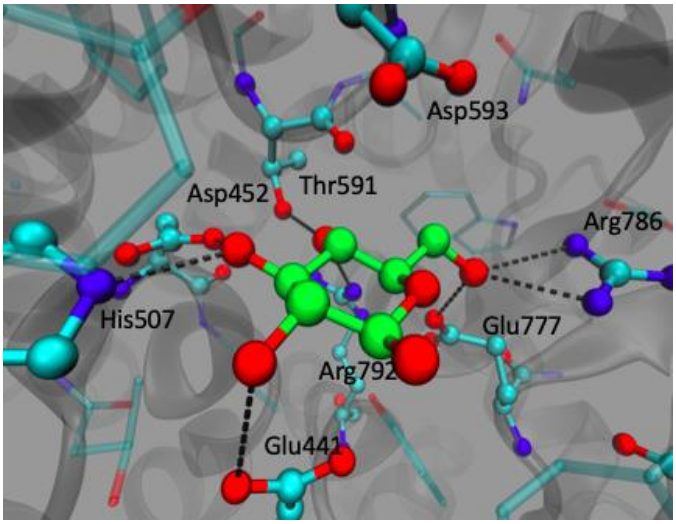


Figure 1.4. Interactions of GBA2 active site residues with glucose[28]

Currently, no 3D structure of human GBA2 has been elucidated, but an X-ray structure of a bacterial homolog of human GBA2 is available.[28] TxGH116 from *Thermoanaerobacterium xylanolyticum* shares about 40% sequence homology with human GBA2 at the catalytic domain and possess the same substrate stabilizing residues at the active site. Human GBA2 and TxGH116 have a similar inhibition profile. TxGH116 is not inhibited by CBE, an inhibitor that distinguishes GBA1 activity (CBE sensitive) from GBA2 activity (CBE insensitive). TxGH116, like the human GBA2, is inactivated by cyclophellitol. These suggest that the study of the TxGH116 active site would give insights into the human GBA2 active site. The catalytic domain of TxGH116

is located in the C-terminus. The nucleophilic residue (Glu 441) and the catalytic acid/base residue (Asp 593) are oriented with their carboxylate oxygens about 9 Å apart, which is greater than a separation of 4.5-6.5 Å for most β-glucosidases from various sources. The interaction between the TxGH116 active site residues and glucose was proposed by Charoenwattanasatien et al. and is shown in **Figure 1.5**.[28] All of the OH-groups of glucose are stabilized via H-bonding interactions with the active site residues except for the 2-hydroxyl group. The 3-hydroxyl group has one H-bonding interaction with Glu 441. The 4-hydroxyl and the 5-hydroxyl groups are stabilized by two H-bonding interactions, each with His 507, Asp 452, and Thr 591, Arg 792, respectively. The 6-hydroxyl group has the greatest number of H-bonding interactions of three, which may indicate that GBA2 does not tolerate modifications at C6 and also it could explain the differences in the inhibition produced by CBE and cyclophellitol.[28]

1.4.3. Neutral β-glucosidase: GBA3

The third and the least studied human enzyme that hydrolyzes GlcCer is a cytosolic β-glucosidase (GBA3). GBA3 was first discovered as a Klotho-related protein that possesses one GH motif.[29] GBA3 is predominantly expressed in liver tissues followed by spleen, kidney and small intestine and in negligible amounts in the brain. GBA3 is found in the cytosol, and it has optimum activity around pH 6.0.[14,30]

Based on the X-ray crystal structure of human GBA3 at 1.6 Å resolution, its catalytic nucleophile (Glu 373) and the acid/base (Glu 165) are located near the carboxyl-terminus.[11] The carboxyl oxygens of the catalytic residues are estimated to be 5.3 Å apart. It is proposed that during the catalysis, the glucose portion of GlcCer is tucked in a small tight inner pocket of the active site where it is stabilized via H-bonding interactions with the active site residues, and the ceramide moiety lies in a larger pocket closer to the surface where the lipid interacts with the hydrophobic amino acid residues.[14]

GBA3 has shown broad substrate specificity *in vitro*, with the ability to hydrolyze a wide range of substrates, including β-D-galactosides, β-D-fucosides, β-D-xylosides α-L-arabinosides and other β-D-glucosides.[30] The studies performed to elucidate the biological significance of GBA3 show conflicting results. Hayashi et al. established a crystal structure of GBA3 with GlcCer and galactosylceramide (GalCer), and GlcCer was

shown to fit better in the active site than GalCer.[14] Their results also showed that GBA3 had higher neutral glucosylceramidase activity than galactosylceramidase activity, which led to the proposal that GBA3 may be involved in a novel non-lysosomal pathway of GlcCer degradation. However, in the study done by Degraaf et al., GBA3 showed higher β-D-galactosidase activity than β-D-glucosidase activity.[30] Currently, the identity of the GBA3 endogenous substrate remains unknown.

1.5. Mechanism of GlcCer hydrolysis by β-glucosidases: Koshland double-displacement mechanism

GHs catalyze the hydrolysis of glycosidic bonds of glycoconjugates. GHs commonly function through one of the two mechanisms: either through retention or inversion of the stereochemistry at the anomeric site. β-Glucosidases are β-retaining enzymes, meaning that both their substrates and products have β-configurations at the anomeric carbon. β-Glucosidases employ a Koshland double displacement mechanism where the hydrolysis occurs in two steps (**Figure 1.3**).[31] The catalysis occurs with the assistance of the active site acid/base and nucleophilic residues, typically glutamic and/or aspartic acid residues. In the first step, the enzyme and the substrate come into contact to form the Michaelis-Menten complex (ES), resulting in a change in the conformation of the enzyme. Then, simultaneously, the nucleophilic residue attacks the anomeric carbon of the substrate while the general acid protonates the aglycon leaving group. (**1**) The enzyme-substrate complex goes through an oxocarbenium ion-like transition state (TS) (**2**) to form a covalent glycosyl-enzyme intermediate where glucose is covalently attached to the nucleophile of the enzyme. (**3**) In the second step, an activated water molecule, with assistance from the general base, attacks the anomeric carbon of the covalent glycosyl-enzyme intermediate. After passing through another oxocarbenium ion-like TS (**4**), the hydrolyzed product is formed and released from the active site pocket.

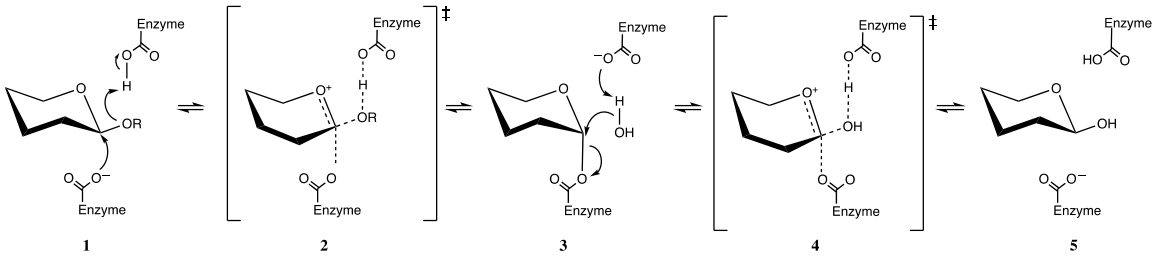


Figure 1.5. Koshland double displacement mechanism of retaining β-glucosidases

1.6. GBA1 relevance to neurodegenerative diseases

1.6.1. GBA1 mutation is the cause of Gaucher disease

GBA1 is the first identified human β-glucosidase and has been studied extensively since the finding that a deficiency in GBA1 is associated with neurodegenerative diseases.[32] GD is a rare genetic disease that causes neurodegeneration that is progressive and potentially life-threatening. It is a lysosomal storage disease resulting from homozygous mutation of the *GBA* gene, which causes lack, or in some cases, absence of GBA1 enzymatic activity followed by accumulation of GlcCer in the lysosomes, particularly in liver and spleen cells.[6] This causes enlargement of the liver and spleen, and the GlcCer buildup in the nervous system interferes with its normal function, which can lead to brain damage. GD is classified into three types: type 1, type 2, and type 3.[6] Symptoms of type 1 GD include abnormal swelling of the liver and spleen, although brain development is unaffected. Both type 2 and type 3 GD are neuronopathic forms of the disorder, each to a differing extent, which involve complications in the central nervous system (CNS) that cause neurological deteriorations and seizures.

1.6.2. GBA1 mutation associated with synucleinopathy

GBA1 mutations are associated with synucleinopathy, a group of neurodegenerative disorders caused by abnormal aggregation of α-synuclein proteins in nerve cells in the substantia nigra and cortex parts of the brain. One synucleinopathy is Parkinson's disease (PD), one of the most common neurodegenerative diseases that affect about 1 in 500 people in Canada and in 2020, more than 10 million people worldwide were living with PD.[33] PD is caused by degeneration of nerve cells in a part of

the brain called substantia nigra where dopamine is produced. Dopamine plays a crucial role in enabling precision movements. The degeneration of nerve cells is caused by the aggregation of α-synucleins, which are neuronal proteins abundant in dopamine neurons of the brain, and the formation of toxic clumps called Lewy bodies. This causes a decrease in dopamine levels, which causes mobility complications. Genetic studies have revealed a strong association of mutations in the gene encoding GBA1 with PD. The GBA1 mutations L444P and N370S were found in 5% to 7% of PD patients, which makes it the most significant genetic risk factor for developing PD over other genetic mutations.[34] Carriers with heterozygous and homozygous mutations were at the same risk of developing PD.

GBA1 mutations are also observed in patients with Lewy body dementia (LBD), another synucleinopathy similar to PD. In patients diagnosed with LBD, the GBA1 mutations were observed in 23% of patients, which is about four times more frequent than in PD patients. This suggests that GBA1 mutations may play a greater role in the development of LBD than in PD.[35]

1.7. Involvement of GBA2 in neurodegenerative diseases

The generation of GBA2 knockout mice (*GBA2*$^{-/-}$) in 2006 allowed the discovery of GBA2 involvement in the pathophysiology of diseases in mice.[25] In *GBA2*$^{-/-}$ mice, abnormal accumulation of GlcCer was observed primarily in tissues of the brain, liver and testes. In mice, this led to male infertility, delay of liver regeneration, and cytoskeletal defects. Soon after, in 2007, a pharmacological GBA2-deficient mice model using the GBA2 selective inhibitors, *N*-butylgalactonojirimycin (NB-DGJ) and *N*-butyldeoxynojirimycin (NB-DNJ), was established, which showed the same results as those obtained *in GBA2*$^{-/-}$ mice.[36] Unlike in these mice models, in humans, the loss of function mutation of GBA2 affected the CNS and resulted in cognitive impairments as well as led to neurodegenerative diseases such as cerebellar ataxia and spastic paraplegia.[37]

1.7.1. Contribution of GBA2 to the pathogenesis of Gaucher disease

Multiples studies have reported a possible contribution of GBA2 to the pathogenesis observed in GD through a feedback mechanism with GBA1 in regulating

GlcCer levels in GD cells. Marques et al. experimented with Niemann-Pick type C (*Npc*$^{-/-}$) mice displaying secondary deficiency in GBA1 activity.[38] They labelled the GBA1 and GBA2 enzymes with an activity-based probe, which covalently binds to the active sites of both enzymes, followed by a quantitative analysis of the protein content using western blot. In these mice, an increase in the expression of GBA2 was observed. Similarly, Burke et al. assessed GBA2 levels in GBA1 deficient mice and from leucocytes of GD patients.[39] They used 4-methylumbeliferyl β-D-glucopyranoside to measure the activity of β-glucosidases by detecting the released fluorescence derived from the 4-methylumbeliferone (4-MU), and the isolated GBA2 activity was determined by a subtractive assay using the GBA2 selective inhibitor, deoxynojirimycin (DNJ). The authors found elevated GBA2 activity in the brains of GBA1-deficient mice leading to the proposal of GBA2 up-regulation as a compensatory mechanism to limit GlcCer accumulation. By comparison, Mistry et al. reported a beneficial effect of eliminating GBA2 activity.[40] They suggested that the compensatory effect of GBA2 negatively affects GD cells. In GD cells with GBA1 deficiency, accumulated GlcCer in lysosomes leaks out into the cytoplasm, thus allowing GBA2 to hydrolyze cytosolic GlcCer, leading to increased ceramide levels in the cytoplasm. This ceramide then gets degraded into another type of lipid called sphingosine by the cytosolic enzyme, ceramidase. High levels of ceramide and/or sphingosine in the cytoplasm are active agents responsible for symptoms of type 1 GD. Their proposal was confirmed by the deletion of the *GBA2* gene, which removed the enzyme activity and alleviated the symptoms of type1 GD.

A conflicting result was reported by Schonauer et al., who reported that GBA1-dependent down-regulation of GBA2 was observed in Gaucher cells that resulted from a negative feedback loop of GBA2.[41] They proposed that when GBA1 activity was low, GBA2 activity increased, leading to accumulated ceramide in the cytosol, which then is converted to sphingosine by neutral ceramidase. Sphingosine, a GBA2 metabolite, directly binds to GBA2 and inhibits its activity.

There is plenty of evidence that suggests GBA2 is involved in GD; however, the results on the role of GBA2 in the pathogenesis of this disease are conflicting. Also, although there are many suggestions about the intracellular localization of GBA2 and its possible roles, the physiological role of GBA2 in living cells is not fully defined. Therefore, this information on GBA2 must be clarified to understand its role in GD and

other neurodegenerative diseases, which can aid in the development of its therapeutic treatments.

1.8. Limitations of current methods for detecting GBA2 activity

The discrepancy in the results of GBA2 activity in the literature published above by Marques et al., Burke et al., Mistery et al., and Schonauer et al. may be due to differences in the methods of measuring enzyme activity. In most studies, the GBA2 activity was determined by subtractive measurement by using a substrate with an incorporated fluorescence reporter group, commonly 4-MU β-D-glucopyranoside[39] with a combination of selective inhibitors of GBA1 and GBA2. Unfortunately, this type of measurement requires multiple experiments to measure GBA2 activity and errors associated with each experiment are additive and accumulate in the final value, which decreases the reliability and accuracy of the experiment.

It is pretty clear from the precedent studies that GBA2 is involved in the pathogenesis of GD and that GBA2, possibly through crosstalk with GBA1, is involved in GlcCer level regulation. However, the subtractive measurement method presents limitations as described above, therefore there is a great need for a better chemical tool to explore the function of GBA2.

1.9. Development of a selective chemical tool to study GBA2

Developments in the field of chemical biology have been achieved through the use of chemical tools. Chemical tools are small molecules used as probes for biochemical processes, which provide new insights into the target molecule, such as its role or the pathway it is involved in. Substrate probes are the least developed type of chemical tool for the study of β-glucosidases.

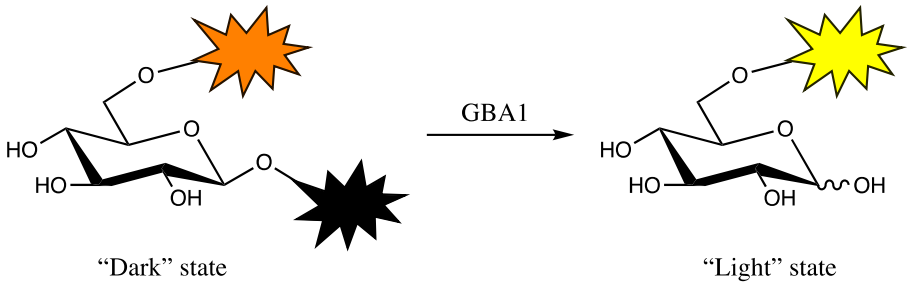


Figure 1.6. GBA1 selective substrate probe for live cell imaging[42]

Currently, a selective substrate probe for biological imaging of live cells has only been developed for one β-glucosidase, GBA1. Yadav et al. synthesized a fluorescence-quenched substrate probe for GBA1 consisting of a glucose molecule covalently attached with a fluorophore on 6O and a quencher on the anomeric position.[42] Before the hydrolysis of the probe by GBA1, the fluorophore and the quencher are in close proximity and go through Forster resonance energy transfer (FRET), which keeps the probe in the "dark" state. FRET is a distance-dependent process of non-radiative energy transfer from a donor chromophore to an acceptor chromophore. Hydrolysis of the probe allows the pair to move away from each other, and the probe enters the "light" state. This probe allows the localization and quantification of GBA1 enzymatic activity in live cells.

In an effort to understand the biological function of GBA2, many efforts have been made to develop a selective and potent inhibitor for GBA2. Most of the known GBA2 inhibitors are iminosugars that are derivatives of miglustat, a compound sold under the drug name, Zavasca (**Figure 1.7**). Miglustat was marketed as an inhibitor of GCS, the enzyme that synthesizes GlcCer, to treat GD type 1. However, since the inhibitor is a conformational and charge mimicry of the transition state of enzymatic glycosylation, it displays "off-target" inhibition of GBA1 and GBA2 in with IC_{50} of 74μM and 140 nM, respectively.[26,24] These off-target effects of the compound caused gastrointestinal symptoms such as diarrhea and weight loss in patients treated with miglustat.[43]

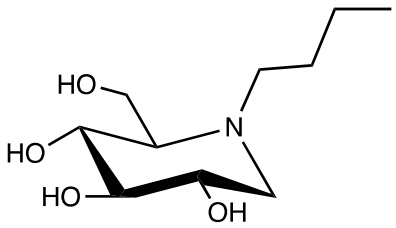


Figure 1.7. Miglustat (NB-DNJ or Zavesca)

Efforts to develop a potent GBA2 inhibitor were undertaken through structure-activity studies by exploring different substituents on the nitrogen of miglustat.[44,45,46] Nevertheless, it has been challenging to develop an inhibitor with high potency towards GBA2 while maintaining high selectivity over other β-glucosidases. The off-target effects from using inhibitors with poor selectivity add uncertainty to the conclusion.

While the work towards identifying selective inhibitors of GBA2 has been undertaken, research on the development of a GBA2 selective substrate probe is limited. Currently, there is no selective substrate probe for GBA2 to directly and selectively measure its activity in live cells and tissue and cell lysates. A GBA2 selective substrate that is potent and also highly selective towards GBA2 will simplify the measurement method and eliminate the errors associated with the subtractive assay for tissue and cell lysate studies.

1.10. Aims of the thesis

The focus of this research is to explore the substrate specificity of the non-lysosomal β-glucosidase, GBA2, with the goal being to develop a substrate that can be used to selectively monitor GBA2 activity over the functionally-related glucosidases, GBA1 and GBA3. This thesis reports the design and preparation of a panel of 4MU β-D-glucopyranoside substrate analogs in Chapter 2 and the tests of these analogs against GBA1, GBA2, and GBA3 through kinetic analysis to generate Michaelis-Menten plots as described in Chapter 3.

Chapter 2. Synthesis of the synthetic substrates

A fluorogenic reporter substrate commonly consists of a substrate that is recognized by the enzyme of interest, which contains a leaving group that becomes fluorescent once liberated by the action of the enzyme. The aim is to find a selective substrate for GBA2. To do this, a series of synthetic substrates were prepared in which glucose analogs were used to generate β-glycosides with a 4-MU leaving group.

2.1. Use of deoxy- and deoxyfluoro-substituted substrates for probing enzyme selectivity

Fluorinated and deoxyfluorinated substrate analogs are widely used in studying glycan-protein interactions.[47,48] Substitution of a hydroxyl group with fluorine or generating a deoxy-analog alters the physiochemical properties of the compound, which can often lead to enhanced pharmacological properties.[49,50] Such changes to carbohydrates are a useful strategy for studying the field effects in ligand-protein interactions and can provide information on the significance of each substrate hydroxyl group for binding and catalysis.[51,52] Both fluorine and hydrogen atoms are smaller in size than a hydroxyl group but differ greatly in their electronic properties, which makes these substitutions a good strategy for probing electronic effects without having to be unduly concerned about steric effects of the functional group substitution. Two principle types of interactions are involved in sugar-protein binding. The major interaction is the hydrogen bonding of the hydroxyl groups of the sugar molecule with heteroatoms of amino acids within the protein binding pocket. Important but less favorable interactions are the CH-π interactions of the carbohydrate with the aromatic amino acids within the active site pocket.[53] The C-H bonds on the sugar ring are electropositive due to the polarization by their geminal hydroxyl groups. Functional groups of aromatic amino acids are partially negatively charged due to the delocalized π-electrons. This allows weak non-covalent interactions to occur between C-H bonds of carbohydrate and π-systems of aromatic amino acids. That being so, the substitution of a hydroxyl group of carbohydrate can lead to disruption of key interactions driving the protein-carbohydrate interactions.

2.1.1. Impact of deoxyfluorination on substrate binding and catalysis

Oxygen and fluorine are the two most electronegative atoms, and they are similar in size. However, there are, of course, differences between the two atoms in their intrinsic properties and the substitution of a hydroxyl group with fluorine can affect the affinity and enzymatic process of ligand-protein interactions.[51,54]

Fluorine substitution on a substrate can have three direct effects. Firstly, the effect of the fluorine substitution is (i) the decreased H-bond acceptor capability. The lone pairs of fluorine have a higher ionization potential and are less polarizable than those on oxygen. This property makes fluorine less able to participate in H-bonding interactions than oxygen.[55] Secondly, the substitution leads to (ii) loss of H-bond donor capability. Fluorine is a monovalent atom and thus cannot act as a H-bond donor.[56] And lastly, the fluorine substitution affects the interactions by (iii) presenting a sterically smaller (1.35 Å)[57] substituent than a hydroxyl group (1.43 Å).[58]

There are several ways a fluorine substitution of a hydroxyl group within a monosaccharide can affect the binding of a ligand to its protein receptor by altering the overall characteristics of the carbohydrate ring. The substitution can alter several properties, including; (i) polarizability of the molecule, (ii) overall lipophilicity, (iii) capacity for engaging in CH-π interactions, and (iv) ability to interact with water molecules. Firstly, the greater electronegativity of fluorine leads to the C-F bond being more polarized than the C-OH bond, affecting the electron density of the adjacent functional groups.[56] This can lead to structural and conformational changes of the ring; however, substitution with fluorine generally does not cause significant distortion of the conformations adopted by a monosaccharide.[59] Secondly, fluorination affects the molecule's lipophilicity. Lipophilicity describes the ability of a compound to dissolve in fats, lipids, or non-polar solvents, and it is the measure of membrane permeability used in the pharmaceutical industry. Lipophilicity is measured as partition coefficient (P) between 1-octanol and water, and it is usually written as log P.[60] Naturally, carbohydrate molecules are quite hydrophilic due to the multiple hydroxyl groups present. The substitution of a hydroxyl group with a fluorine atom increases the lipophilicity of the molecule. Testing lipophilicities of a series of mono-fluorinated glucose analogs revealed that the fluorination leads to about an order of magnitude higher log P value comparison to glucose. This change in lipophilicity affects the desolvation energy (energy required to

displace water molecules surrounding the compound and the enzyme) involved in the binding of the compound with an enzyme, thus affecting the binding affinity of the compound with the protein.[56] Thirdly, the substitution also affects CH-π interaction occurring during the binding. A fluorine atom is more electronegative than a hydroxyl group (oxygen bonded with hydrogen makes it less electronegative); therefore, the F substitution affects the CH-π interactions by making the C-H bond more polarized.[53,56] Lastly, incorporating F in the substrate can cause changes in the interactions with water molecules during binding and catalysis.[61] It could introduce new water molecules or change the H_2O-mediated H-bond orientations.

2.1.2. Impact of deoxygenation on substrate binding and catalysis

Deoxygenation of carbohydrates typically affects key interactions involved in the binding of carbohydrate ligands and substrates to proteins, including glycoside hydrolases.[52,54] The principal effect is the loss of the ability to form H-bonding interactions used to bind the natural ligand. Deoxygenation can also affect the binding of a ligand with the protein by altering the characteristics of the sugar ring. As described for deoxyfluorination just above, deoxygenation also changes properties of the ligand, including its: (i) polarizability, (ii) lipophilicity, (iii) ability to form CH-π interactions, and (iv) ability to form interactions with the surrounding water molecules. To illustrate, removing the highly polarized hydroxyl group affects the electron density of the pyranose ring and adjacent functional groups.[54] Depending on the location of the deoxygenation, it causes a conformational change in the pyranose ring.[62] Also, we expect that removing a hydroxyl group will increase the compound's lipophilicity, and the effects are more pronounced than deoxyfluorination.[56] The resulting change in lipophilicity affects the desolvation energy of the ligand and can, accordingly, sometimes lead to enhanced binding or partly counteract the effect of losing a hydrogen bonding interaction. In natural carbohydrate ligands, the C-H bonds of the pyranose ring are polarized by its geminal hydroxyl groups. These polarized C-H bonds are often involved in CH-π interactions formed with aromatic amino acids at the ligand binding site.[53] Loss of the hydroxyl group leaves the geminal C-H less polarized and generally leads to decreased strength of the CH-π polar interaction. Lastly, in a similar manner as with deoxyfluorinated sugars, deoxygenation of a ligand can lead to changes in the orientations of H_2O molecules bound proximally in the active site by eliminating hydrogen bonds and, by creating a

space, may favor binding of H_2O molecules to the site previously occupied by the hydroxyl group.[56] This could create new water-mediated H-bonds, which affect the stabilization of the protein with the bound substrate.

2.2. Target compounds for probing GBA2 selectivity

Eight synthetic glucoside substrate analogs were selected as candidates that might yield a GBA2 selective substrate (**Figure 2.1**). The target substrate analogs were prepared with the same leaving group. Then these analogs were tested against each GBA enzyme while monitoring the resulting fluorescence arising from the enzyme-catalyzed hydrolysis of the substrates.

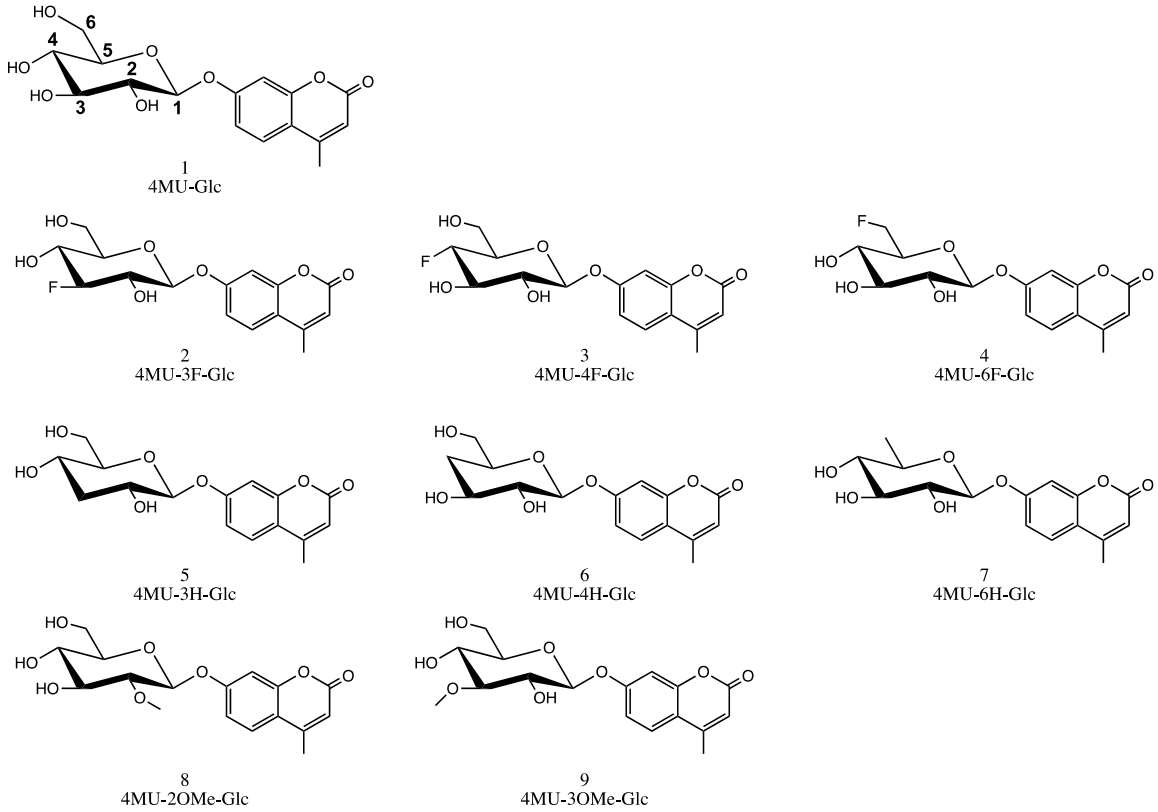


Figure 2.1. Targeted synthetic substrate analogs for investigating substrate selectivity of GBA enzymes. The carbon atoms on glucose of 4MU-Glc are numbered.

The substrates analogs as glycosides of 4-MU linked via a β-glycosidic bond were synthesized. 4-MU is coumarin, which is one of the most commonly used leaving groups for fluorogenic substrates. Its popularity is due to its several advantages. 4-MU is

a low molecular weight compound with fair water solubility and low toxicity, making it useful for live cell studies. In water, 4-MU has excitation and emission bands with maxima at 380 nm and 454 nm, respectively,[63] and it is intensely fluorescent when deprotonated in its phenolate form but comparatively dark when protonated or as a glycoside. Accordingly, its fluorescence intensity is highly pH-dependent. Given its pK_a value of 7.79, it is weakly fluorescent at acidic pH values but approaches its maximum fluorescence intensity as the pH of a solution approaches 10.[64,65] With its pK_a value of 7.79, 4-MU is considered a weak acid and therefore presents an excellent leaving group under physiological pH.[66] For simplicity, substrates using 4-MU as a reporter are often used in stopped assays, where a base is added at the end of the enzymatic assay to increase the pH of the solution to terminate the enzymatic reaction and generate an increased amount of the phenolate form of 4-MU.[20,67] Fluorogenic substrates containing 4-MU have been widely used to monitor the activities of many β-D-glycosidases, including GBA1, GBA2, and GBA3.[68] Burke et al. measured GBA1 activity by measuring the total GBA activity using 4-MU β-D-glucoside, then subtracting the same assay containing additional GBA1 inhibitor.[39] Similarly, Sardi et al. determined the brain total glucocerebrosidase activity using 4-MU β-D-glucoside as an artificial substrate.[69] Pritsch et al. used a 4-MU labelled fluorogenic substrate analog to measure β-glucosidase activities from various fungi.[67]

2.2.1. Deoxyfluoro- and deoxy-glucose analogs as chemical probes for GBA2 selectivity

A series of substrates with modifications at each position of the pyranose ring of glucose, including C3, C4, and C6, were synthesized. Each hydroxyl group at these carbons is involved in hydrogen bonding interactions within the active sites of each GBA enzyme. The 2-hydroxyl group was not chosen for modification since it is well known that deoxyfluorination of this position leads to inductive destabilization of the oxocarbenium ion-like transition state as the electronegative fluorine atom withdraws electron density away from the anomeric carbon.[70] Accordingly, the turnover of 2-deoxy-2-fluoro glycosides by glycoside hydrolases is extremely slow, making such substrates poorly suited to monitoring enzymatic activities.[70]

Regarding literature reports on the influence of deoxy- and deoxyfluoro-modifications of glycosides, previous studies examining the spontaneous acid-catalyzed

hydrolysis of substrate analogs, as I propose to make found that the strong inductive effect of fluorine destabilizes the oxocarbenium ion-like transition state regardless of the position. The observed relative order of rate constants for hydrolysis rates of a series of deoxyfluorinated β-D-glucopyranosides from fastest to slowest were found to be 6-fluoro > 3-fluoro > 4-fluoro >> 2-fluoro.[54] The further away the F is from the anomeric carbon, the more efficient enzymatic catalysis occurs.[52] The rate constants governing hydrolysis of various deoxy-β-D-glucopyranosides had the following order from fastest to slowest: 2-deoxy > 4-deoxy > 3-deoxy > 6-deoxy. Notably, the opposite trend reported with the corresponding deoxy-sugar analog was reasoned to be caused by a more electron-rich pyranose ring compared to glucose, thus stabilizing the oxocarbenium ion-like TS. When the deoxygenated position is further away from the anomeric carbon, the oxocarbenium ion-like TS is stabilized to a lesser extent.[52,54] These effects on spontaneous hydrolysis at these positions, excluding the 2-hydroxyl, which, as noted above, lead to profound effects, are relatively small compared to the effects of such changes on glycoside hydrolase catalyzed turnover of glycosides.

With regards to the effects of such changes on enzyme-catalyzed processes, the effects of such variations on glycosides vary from enzyme to enzyme. This phenomenon was demonstrated in other carbohydrate-processing enzymes. In a recent paper by Keenan et al., a similar study was done with galactokinases (GalK).[71] GalK catalyzes *in vitro* ATP-dependent phosphorylation of α-D-galactose into its sugar-1-phosphate. The authors investigated the effects of mono-deoxyfluorinated α-D-galactose at C2, C3, C4, and C6-positions on substrate conversion by seven different GalK enzymes isolated from different organisms. The Conversion rate of a synthetic substrate varied between seven different GalK enzymes, allowing the elucidation of GalK enzymes with narrow substrate specificities and GalK enzymes with broader substrate specificities. Another example is shown by Slamova et al. using β-*N*-acetylhexoaminidases, a type of glycosyl hydrolase that catalyzes the hydrolysis of the terminal *N*-acetylglucosamine (GlcNAc) and *N*-acetylgalactosamine (GalNAc) residues from oligosaccharides and glycoconjugates.[72] To investigate the differences in substrate tolerances of β-*N*-acetylhexoaminidases, they tested 4-hydroxy analog of the substrate with the enzymes from various origins, including fungal, animal, bacterial and human. The β-*N*-acetylhexoaminidases from different origins processed the substrate of 4-deoxy-analog at different extents. These enzymes catalyze the same biochemical reactions; however,

the degree of processing of the deoxy-modified substrate varied from enzyme to enzyme.

In general, these reports underscore how such changes affect different enzymes to differing extents. The importance of the hydroxyl groups varies from enzyme to enzyme because of differences in the roles of hydrogen bonding between substrate and enzyme in the ground state as well as between the enzyme and transition state.

Given these observations and the fact that the three GBA enzymes are from different glycoside hydrolase families and therefore have different active site architectures, w reasoned that the effects of these deoxy- and deoxyfluoro-substitutions would differ between the three GBA enzymes. Given the importance of the hydroxyl groups, both of the deoxy- and the deoxyfluoro-sugar analogs were expected to be worse substrates than the parent substrate (4MU-Glc). Therefore, the goal was to identify from among these compounds the one that had only a slight change in the rate constant for one enzyme, whereas the other enzymes would be more profoundly affected.

2.2.2. *O*-Methylated substrates for probing GBA2 selectivity

Interesting results were found when various substrates were tested with a bacterial homolog of human GBA2 (*Thermoanaerobacterium xylanolyticum, Tx*GH116).[28] 4-nitrophenyl 2-acetamido-2-deoxy-β-D-glucopyranoside was turned over with 0.45% relative activity as compared to 4-nitrophenyl β-D-glucopyranoside. Although the turnover of this compound is relatively much lower than the parent compound, the fact that this GBA2 homolog turned over the compound with a large substituent at C2 was interesting to us. Moreover, a coworker showed that 2-deoxy-2-fluoro-β-D-glucopyranoside was turned over very slowly by human GBA2, whereas it was not turned over by GBA1. These findings led to considering probing the effect of sterics at C2-position to determine if this might lend selectivity for GBA2. Therefore, a small structural variant was prepared, namely the 4-MU 2-*O*-methyl-β-D-glucopyranoside (4MU-2OMe-Glc). It reasoned that the small methyl group is not as bulky as an acetamide and therefore increasing the chance GBA2 may tolerate it, yet it is much larger than a proton to possibly hinder turnover by GBA1. Notably, the 4MU-2-OMe-Glc will have lost the ability to donate a hydrogen bond but will have increased hydrogen

bonding acceptor capability due to the electron-donating effect of the methyl unit. In addition, this modification will increase the lipophilicity of the compound, which may also affect its binding affinity for these enzymes. As will be discussed below, during the synthesis of the target compound, 4-MU 3-*O*-methyl-β-D-glucopyranoside (4MU-3OMe-Glc) and 4-MU 2,3-di-*O*-methyl-β-D-glucopyranoside were formed as by-products. These by-products were isolated, and 4MU-3OMe-Glc was also assessed as a substrate for the target enzymes.

2.3. Synthesis of fluoro- and deoxyfluoro-glucose substrate analogs

2.3.1. General synthetic route

Synthesis of the substrate analogues was carried out either starting from their corresponding deprotected deoxy- or deoxyfluoro-monosaccharide precursors purchased (D-glucose, 4-deoxy-4-fluoro-D-glucose, 6-deoxy-6-fluoro-D-glucose, 3-deoxy-D-glucose, 4-deoxy-D-glucose, 6-deoxy-D-glucose) or synthesized (3-deoxy-3-fluoro-D-glucose) from 1,2:5,6-di-*O*-isopropylidene-α-D-allofuranose. The general synthetic route is illustrated here (**Figure 2.2**) using the parent compound, glucose, as an example. The route involves *O*-acetylation, glucosyl bromide donor sugar formation, followed by glycosylation using 4-MU, and finally deacetylation.

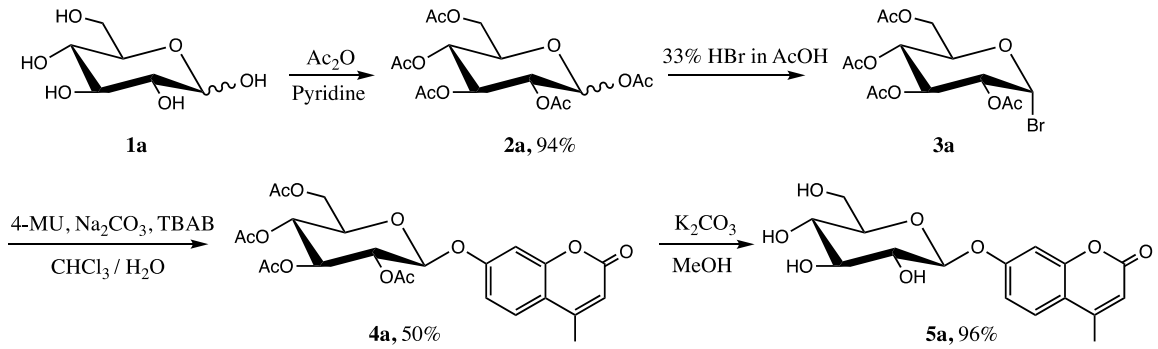


Figure 2.2. Synthetic route for the preparation of 4MU-Glc

2.3.2. *O*-Acetyl protection and deprotection of carbohydrates

Acetylation of carbohydrates is a widely used strategy in carbohydrate chemistry used to protect hydroxyl groups. Per-*O*-acetylated monosaccharides are often the

starting point for preparing a wide range of glycoconjugates. Methods to acetylate monosaccharides can yield products that are exclusively the α-anomer or the β-anomer or may, in some cases, deliver a mixture of the two anomers. In general, the α-anomer is the most thermodynamically stable product, whereas the β-anomer is the kinetically favoured product.[73] Different reaction conditions will lead to different compositions of products. Typically, acetylation can be done under three different conditions; acidic or basic conditions or using nucleophilic catalysis such as pyridine.[73]

Acetylation under acidic conditions with polar, aprotic solvents predominantly yields the α-anomer.[73] The oxocarbenium ion formed during this reaction can have the acetate attack from the top or the bottom face at the anomeric carbon; however, the α-anomer is the more thermodynamically stable product due to the anomeric effect. The anomeric effect is observed when an electronegative substituent is at the anomeric carbon. The α-anomer is favored due to the stabilization arising from hyperconjugation involving an antiperiplanar arrangement of the lone pair of the endocyclic oxygen with the anti-bonding sigma orbital (σ^*) of α-oriented C-X glycosidic bond. Because the acetylation reaction is reversible, the thermodynamically favored α-product accumulates over time. However, acetylation of the glucose analogs done under basic conditions predominantly yields the β-anomer. The use of sodium acetate allows a rapid mutarotation to occur at which β-anomer is deprotonated and forms β-oxyanion. β-oxyanion is more reactive than the α-anomer due to its unfavorable dipole-dipole interactions between O5 and O1. Therefore, more reactive β-anomer gets acetylated to yield the kinetic β-product.[73,74] Also, the neighboring participating effect of the 2-*O*-acetyl group on the sugar contributes to the β-selectivity. The 2-*O*-acetyl group on the sugar ring attacks the anomeric carbon from the axial position and forms a 5-membered ring. This 5-membered ring blocks the acetate from attacking from the axial position; therefore, acetylation on anomeric carbon displays β-selectivity.[73]

Here, global acetylation of glucose and its fluoro- and deoxyfluoro-analogs were performed using acetic anhydride as the electrophilic source of the acetyl groups in pyridine at room temperature. Under these conditions, nucleophilic catalysis occurs by pyridine acting as a nucleophile for the carbonyl groups of acetic anhydride. Acetylation in pyridine generally yields products with the same ratio of α- and β-anomers present in the starting sugars.[73] Literature on acetylation of monosaccharides using pyridine as a

solvent often report using 4-dimethylaminopyridine (DMAP) as a catalyst.[75] However, no such catalyst was needed here, and the reactions gave excellent yields (>90%) of the desired products.

The acetyl group deprotection was done via base-catalyzed solvolysis using K_2CO_3 in MeOH.[76] K_2CO_3 acts as a base catalyst and deprotonates the solvent methanol. Methanol as a nucleophile attacks the carbonyl of the acetate group. This releases methyl acetate and yields deacetylated product.

2.3.3. Phase-transfer glycosylation

Before glycosylation reactions, per-*O*-acetylated monosaccharides were converted into suitable glycosyl donors by generating the corresponding glycosyl bromides. Glycosyl bromides are convenient donor sugars because they are simple to make, readily isolated, and relatively stable.[77] Per-*O*-acetylated glucose and its analogs to the corresponding glycosyl bromides were reliably converted using 33% HBr in acetic acid. This reaction was a quick and simple process that yielded one product. However, it was found that the glycosyl bromides were not stable enough to be purified as the products turned black during drying under a high vacuum at room temperature for several hours. Therefore, due to the instability of the glycosyl bromides, the yields of the products were not measured, and the compounds were used for the subsequent glycosylation reaction without purification.

Phase-transfer glycosylation is performed in a biphasic reaction in which one phase is organic and the other is aqueous. Commonly, quaternary alkylammonium salts, such as tetrabutylammonium bromide (TBAB), are used as catalysts that aid in the transfer of an anion from the aqueous phase to the organic phase by acting as a detergent.[78] In the syntheses here, a 1:1 ratio of chloroform to water was used as the solvent, with Na_2CO_3 and TBAB as the phase transfer catalyst. In the basic aqueous phase, 4-MU is deprotonated, and TBAB facilitates the phase transfer of the anionic form of 4-MU to the organic phase where the glycosylation reaction occurs. This reaction is well precedented and was described by Wu et al..[78] Here, the glycosylated products were obtained in around 50% yields over two steps from the corresponding per-*O*-acetylated monosaccharides.

2.3.4. Fluorination of carbohydrate

There are several different synthetic methods for the deoxyfluorination of monosaccharides. The most popular fluorinating agent is probably diethylaminosulfur trifluoride (DAST), which has been used in myriad examples to convert alcohols to alkyl fluorides as well as ketones, sulfides, epoxides.[79]

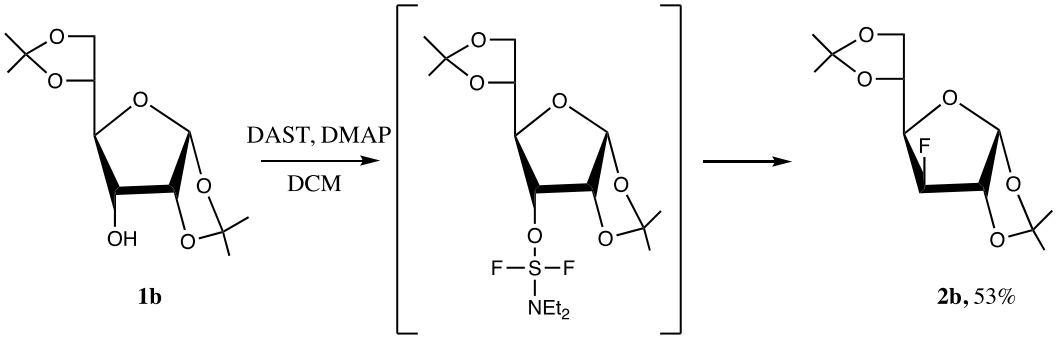


Figure 2.3. Fluorine substitution on 1,2;5,6-di-*O*-isopropylidene-α-D-allofuranose using DAST as fluorinating agent

3-Deoxy-3-fluoro-D-glucose was synthesized from 1,2:5,6-di-*O*-isopropylidene α-D-allofuranose (**1b**). Fluorine was introduced to **1b** using DAST in DCM. DMAP was used as a nucleophilic catalyst instead of more commonly used pyridine. This nucleophilic catalyst was used because the solid-state of DMAP allows the reaction flask with the starting material and DMAP to be dried under a high vacuum prior to the synthesis. The fluorination reaction occurs by the free alcohol attacking the sulfur center of DAST. The lone pair of the amine enables the loss of fluoride. An S_N2 reaction occurs with fluoride ion attacking the activated carbon, leading to displacement with inversion of configuration forming **2b**.

2.4. Synthesis of *O*-alkylated glucoside substrates

4-MU β-D-glucopyranoside (4MU-Glc), prepared using the general synthetic route described previously (**Figure 2.2**), was used as the starting point for the synthesis of *O*-alkylated compounds. The 4 and 6 hydroxyl groups on 4MU-Glc were masked with a benzylidene protecting group. The resulting compound, **3h**, which has only 2 and 3 hydroxyl groups free, was treated with methyl iodide (MeI) for non-selective methylation of each hydroxyl groups.

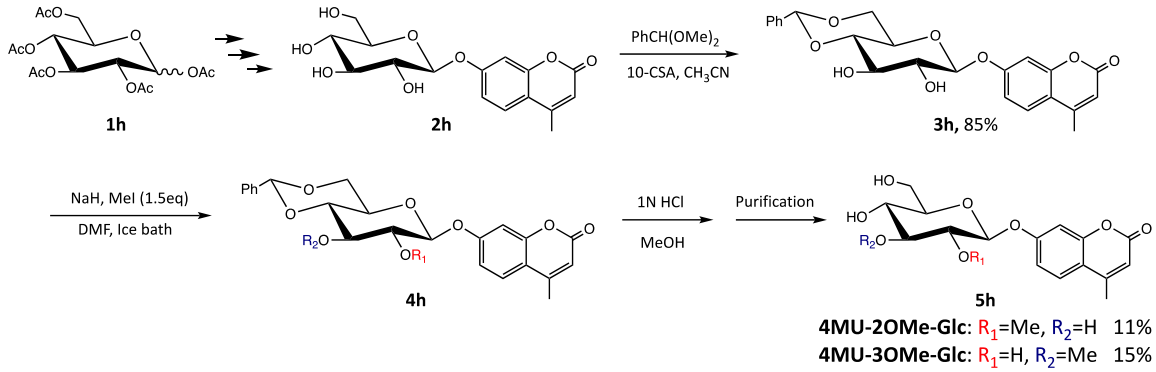


Figure 2.4. Synthetic route for the preparation of mono-*O*-methylated products: 4MU-2OMe-Glc and 4MU-3OMe-Glc

2.4.1. Benzylidene group protection and deprotection

Benzylidene acetal protecting groups are commonly used in carbohydrate chemistry to selectively mask the 1,3-diols driven by the thermodynamically favored the formation of 6-membered rings.[80] In the case with compound **2h, the** benzylidene acetal forms at the 4- and 6-hydroxyl groups with them being incorporated into the acetal functionality. This strategy leaves the 2- and 3-hydroxyl groups free for subsequent modification, including *O*-alkylation. Benzylidene protecting group of diols is commonly done under acidic conditions.[73,81] I installed the benzylidene group using benzaldehyde dimethyl acetal with 10-camphorsulfonic acid (10-CSA) as the acid catalyst in acetonitrile. The acid catalyst is required since the alcohol is a weak nucleophile. The acid protonates the benzylidene acetal allowing the attack of the hydroxyl group. Usually, this reaction is reversible as water molecules can reverse the reaction. However, in this case, the product **3h** forms a more stable bicyclic structure than the reactants, which helps the reaction go to completion.[82] The removal of the benzylidene protecting group was done by acidic hydrolysis using HCl.

2.4.2. Non- selective *O*-alkylation reaction of 4-MU 4,6-*O*-benzylidene-β-D-glucopyranoside

The strategy for synthesizing 4-methylumbelliferyl 2-*O*-methyl-β-D-glucopyranoside (4MU-2OMe-Glc) product involved non-selective *O*-alkylation of compound **3h** (**Figure 2.4**), which has the 2- and 3-hydroxyl groups free. The approach used here uses a hydride base (NaH) with MeI as the alkylation agent in DMF. On a

simple glucose, the 2-hydroxyl would be considered slightly more acidic than the 3-hydroxyl because the 2-position is closer to the anomeric carbon, which lowers the electron density. Therefore, in such a case, different acidities of the 2- and 3-hydroxyl groups could determine which *O*-methylation site is favored. However, with compound **3h**, the benzylidene group likely influences the relative acidities of these substituents, making it difficult to predict the outcome of the reaction. Moreover, using a strong base such as MeI could make the acidity differences of the hydroxyl groups relatively insignificant.

In my initial attempts, MeI and NaH were added at the same time as described in the literature for similar reactions.[83] However, the reaction did not proceed. Later, different attempt was made by adding NaH with stirring before adding MeI. A product was formed in 90% yield. However, the NMR spectra of the formed product were not consistent with the desired products. Structure elucidation of the formed product using 1D and 2D NMR spectra and MS allowed the proposal of a possible side reaction and that led to the unwanted product formation (**Figure 2.5**). Treatment with NaH may have been allowed to proceed for too long, which may have generated hydroxide ions from residual water that attack the ester of 4-MU, causing the opening of the lactone ring. The carboxylate moiety of the compound, formed after the ring-opening, likely is methylated and yielded the unwanted compound as the major product. The molecular weight and the structure of the side product were consistent with the MS and NMR data, respectively.

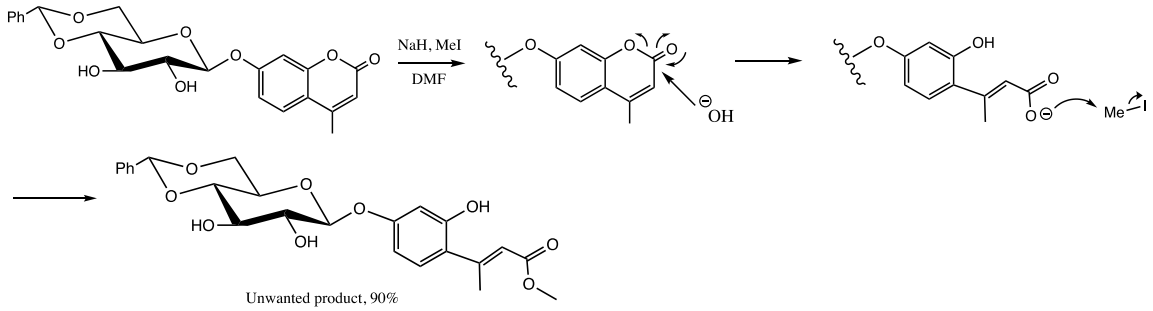


Figure 2.5. Proposed mechanism of unwanted 4-MU ring-opening while the substrate was reacted with NaH for longer than an hour

Ultimately, by not letting the substrate react with NaH for more than an hour, the *O*-methylation reaction yielded three *O*-methylated products; two mono-*O*-methylated compounds at O2 and O3, and a compound di-*O*-methylated at those positions. One

hour of running the reaction after the addition of MeI gave an average 35% yield of all three methylated products combined and 25% of unreacted starting material. The unreacted material was isolated and used for additional rounds of *O*-methylation. The di-*O*-methylated product was isolated in a 7% yield. The two mono-*O*-methylated products had very close Rf values, and multiple rounds of purification were needed to isolate each compound.

2.5. Experimental Procedures

2.5.1. Materials and general synthetic procedures

The glucose analog sugars, chemical reagents, solvents and salts for buffers were purchased from Carbosynth, Sigma-Aldrich or Alfa Aesar.

Prior to performing anhydrous reactions, solid reagents were placed in a flame-dried round-bottom flask and placed under a high vacuum overnight. During anhydrous reactions, reaction flasks were capped with septa, and liquid reagents were added using needles and syringes. The reaction progress was monitored by thin-layer chromatography using Merck pre-coated silica gel aluminum plates in combination with solvents of either EtOAc/Hex or MeOH/DCM. UV-active compounds were visualized under a short-wave UV lamp (254 nm). Sugar molecules were visualized by Seebach staining (2.5% w/v phosphomolybdic acid, 1% cerium sulfate w/v in 1 M H_2SO_4) followed by charring with a heat gun.

^{1}H NMR, ^{13}C NMR, COSY, HSQC, HMBC spectra of the prepared compounds were recorded using Bruker Avance 400 (400 MHz/101 MHz), Bruker Avance 500 (500 MHz/126 MHz) or Bruker Avance 600 (600 MHz/151 MHz). NMR samples were dissolved in either deuterated chloroform ($CDCl_3$), methanol (MeOD), or water (D_2O). The reported chemical shifts (ppm) were relative to the residual signals of deuterated solvents. The final compounds were analyzed for their purity with Agilent 1100 HPLC using Eclipse XDB C18 semi-preparatory column (solvents used: A = H_2O, B = MeOH) with a gradient of 40 - 70% B over 20 mins then 90% B over 5 mins at a flow rate of 2 mL/min. High-resolution mass spectra (HRMS) of the final compounds were recorded with a positive electrospray ionization (ESI) ion source on an Agilent™ Time-of-flight LC/MS mass spectrometer (Model 6210) using direct infusion.

2.5.2. Synthesis of 4-methylumbelliferyl β-D-glucopyranoside (1)

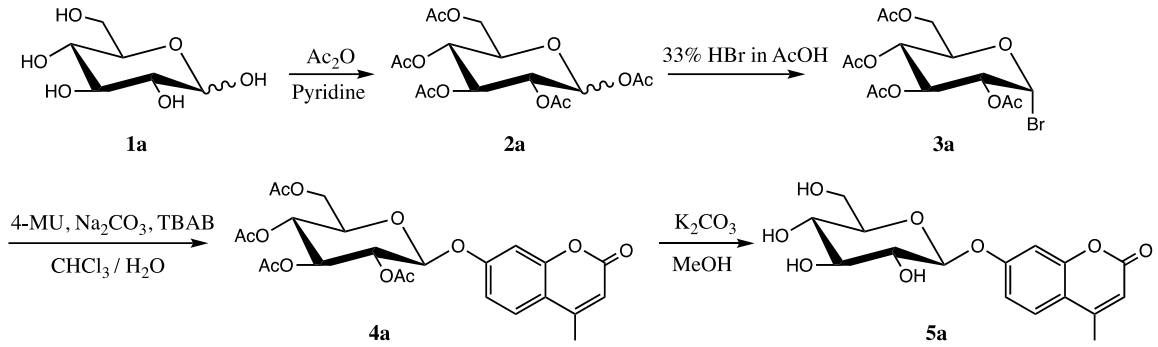


Figure 2.6. Synthetic scheme of 4MU-Glc

1,2,3,4,6-Penta-*O*-acetyl-D-glucopyranose[84] (2a): general acetylation reaction

The general acetylation conditions were adapted from the method used by Withers et al.[54]. To compound **1a** (1.00 g, 5.55 mmol) in a 50 mL round bottom flask under an atmosphere of argon, pyridine (20 mL) was added. The solution was cooled to 0 °C, then acetyl anhydride (20 mL, 211 mmol) was added dropwise over 10 mins using a syringe placed through a septum covering the reaction. The reaction was then allowed to gradually warm to room temperature and stirred overnight under an atmosphere of argon. The reaction progress was monitored by TLC by taking small samples of the reaction and performing a small-scale workup by washing the reaction mixture with 1 M HCl and then extracting the product with ethyl acetate. After confirming that no more starting material was present, the reaction was worked up. By first diluting the reaction mixture with DCM (50 mL), then washing the organic layer with a 1 M HCl solution (5 × 40 mL) until no more pyridine was detected using TLC. The organic layer was then washed with a saturated sodium bicarbonate solution (2 × 30 mL) followed by water (1 × 30 mL). The organic layer was dried over Na_2SO_4, filtered, concentrated *in vacuo,* and dried under a high vacuum overnight to give product **2a** (0.94 g, 94%) as a white powder. Rf: 0.6 (EtOAc:Hex, 1:1 V:V)

^{1}H NMR (400MHz, Chloroform-*d*) δ 5.77 (d, *J* = 7.2 Hz, 1H, H-1), 5.31 (m, 1H, H-4), 5.12 (t, *J* = 1.25 Hz, 1H, H-2), 5.10 (dd, *J* = 1.27 Hz, 1H, H-3), 4.20 (m, 2H, H-6a, H-6b), 3.84 (ddd, *J* = 6.7, 5.1, 3.2 Hz, 1H, H-5), 2.15 (s, 3H, H-OAc), 2.08 (s, 3H, H-OAc), 2.06 (s, 6H, H-OAc), 2.02 (s, 3H, OAc).

2,3,4,6-Tetra-*O*-acetyl-α-D-glucopyranosyl bromide[85] (3a): general bromination reaction

The general bromination conditions were adapted from the method used by Goto et al.[86] To a round bottom flask containing compound **2a** (0.94 g, 2.41 mmol) in DCM (10 mL) cooled to 0 °C and under argon was added 33% w/w hydrobromic acid in acetic acid (2.5 mL, 46 mmol) dropwise using a syringe placed through a septum covering the reaction. The reaction was allowed to gradually warm to room temperature then stirred for 2 hours, at which point no starting material was observed on TLC. The reaction mixture was poured into a separatory funnel containing ice-cold water (25 mL). The product was extracted with DCM (3 × 30 mL). The organic layers were combined and then washed with water (2 × 30 mL), a saturated $NaHCO_3$ solution (2 × 30 mL), followed by brine (2 × 30 mL). The organic layer was dried over Na_2SO_4, filtered, and concentrated *in vacuo* to yield product **3a** (Crude 1.10 g) as thick yellowish-clear oil, which was used immediately without further purification. Rf: 0.6 (EtOAc:Hex, 1:1 V:V)

4-Methylumbelliferyl 2,3,4,6-tetra-*O*-acetyl-β-D-glucopyranoside[76] (4a): general glycosylation reaction

The general glycosylation conditions were adapted from method used by Wu et al.[78] To a round bottom flask containing a 1:1 ratio of $CHCl_3$:H_2O (220 mL) was added 4-MU (0.47 g, 2.67 mmol, 1 eq), Na_2CO_3•H_2O (0.66 g, 5.34 mmol, 2 eq), and tetrabutylammonium bromide (TBAB) (0.86 g, 2.67 mmol, 1 eq). The solution appeared fluorescent yellow. The reaction was heated to 50-55 °C, then compound **3a** (1.10 g, 2.67 mmol) dissolved in minimal $CHCl_3$ was added dropwise over 5 mins and stirred overnight. After confirming the formation of the product by TLC, the reaction was proceeded to work up. The organic layer of the reaction mixture was separated, and the aqueous layer was extracted with DCM (3 × 30 mL). The combined organic layer was washed with water (3 × 100 mL), brine (3 × 100 mL), then dried over Na_2SO_4 before being filtered and concentrated *in vacuo*. Pure product **4a** (0.68 g, 50%) was isolated as a white powder using silica gel flash column chromatography. Rf: 0.2 (EtOAc:Hex, 1:1 V:V)

^{1}H NMR (400MHz, Chloroform-*d*) δ 7.52 (d, *J* = 8.66 Hz, 1H, H-5'), 6.94 (m, 2H, H-6', H-8'), 6.20 (t, *J* = 1.27 Hz, 1H, H-3'), 5.32 (m, 2H, H-1, H-3), 5.17 (m, 2H, H-2, H-4), 4.30

(dd, *J* = 12.36, 5.72 Hz, 1H, H-6a), 4.18 (dd, *J* = 12.31, 2.34 Hz, 1H, H-6b), 3.91 (ddd, *J* = 10.03, 5.72, 2.37 Hz, 1H, H-5), 2.41 (d, *J* = 1.27, 3H, H-Me'), 2.12 (s, 3H, OAc), 2.07 (s, 3H, OAc), 2.06 (s, 3H, OAc), 2.04 (s, 3H, OAc).

^{13}C NMR (101MHz, Chloroform-*d*) δ 170.73, 170.30, 169.53, 169.38 (4× OAc), 160.86 (C-2'), 159.35, 155.02 (C-7'), 152.24, 125.86 (C-5'), 115.71 (C-4'), 114.16 (C-8'), 113.41 (C-3'), 104.15 (C-6'), 98.57 (C-2), 72.74 (C-5), 72.63 (C-1), 71.13 (C-3), 68.26 (C-4), 61.98 (C-6), 20.85, 20.76, 20.74, 20.73 (4× OAc), 18.83 (C-Me').

4-Methylumbelliferyl β-D-glucopyranoside[76] (5a): General deacetylation reaction

The general deacetylation conditions were adapted from the method used by Wei et al.[76] To a flask containing Compound **4a** (0.68 g, 1.34 mmol) in anhydrous MeOH (15 mL), a catalytic amount of K_2CO_3 (0.033 g, 0.24 mmol) was added. As the reaction solution was stirred at room temperature for 2 hours, the solids dissolved then precipitated out of the solution. The reaction solution was diluted with a few millilitres of DCM to dissolve the product into the solution. Pure product **5a** (0.43 g, 96%) as a white powder was obtained using silica gel flash column chromatography. Rf: 0.2 (DCM:MeOH, 9:1 V:V)

^{1}H NMR (400 MHz, Methanol-d_4) δ 7.73 (d, *J* = 8.8 Hz, 1H), 7.19 – 7.05 (m, 2H), 6.22 (q, *J* = 1.2 Hz, 1H), 5.06 (dd, *J* = 5.4, 2.2 Hz, 1H), 3.93 (dd, *J* = 12.1, 2.3 Hz, 1H), 3.73 (dd, *J* = 12.1, 5.7 Hz, 1H), 3.57 – 3.47 (m, 3H), 3.46 – 3.39 (m, 1H), 2.48 (d, *J* = 1.2 Hz, 3H).

^{13}C NMR (101 MHz, Methanol-d_4) δ 161.93, 160.68, 154.67, 154.06, 125.88, 114.62, 113.58, 111.48, 103.63, 100.52, 76.96, 76.48, 73.37, 69.87, 61.03, 17.22.

HRMS (ESI$^+$) *m/z* [M+H]$^+$: Calc'd for $C_{16}H_{19}O_8$ 339.1080, found 339.1034.

2.5.3. Synthesis of 4-methylumbelliferyl 3-deoxy-3-fluoro-β-D-glucopyranoside (2)

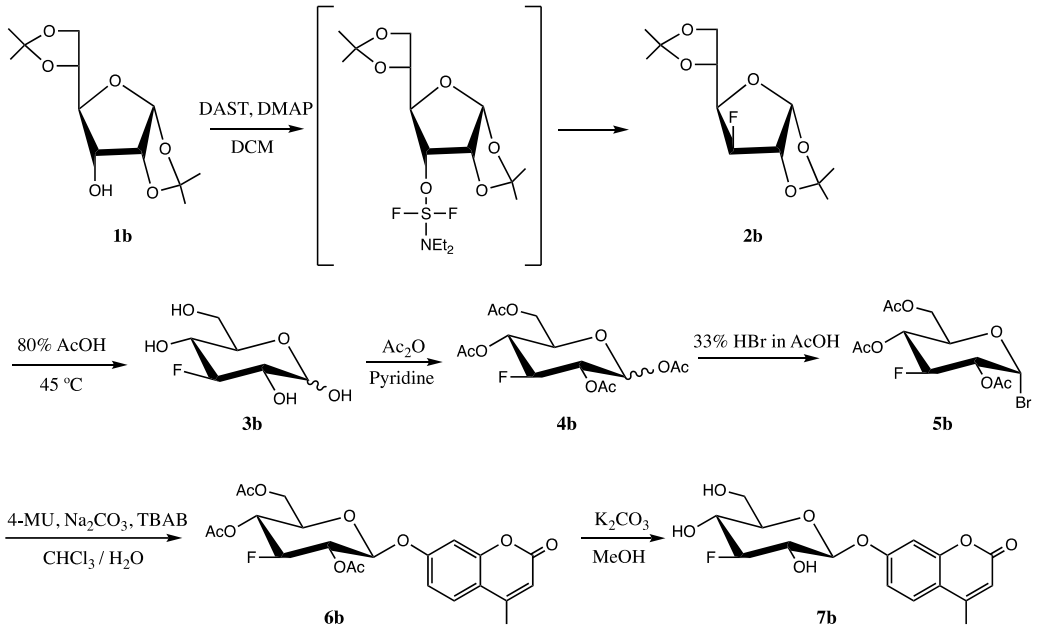


Figure 2.7. Synthetic scheme of 4MU-3F-Glc

3-Deoxy-3-fluoro-1,2:5,6-di-*O*-isopropylidene-α-D-glucofuranose[87] (2b)

The fluorination conditions were adapted from methods used by Withers et al.[54] A flask containing 1,2:5,6-di-*O*-isopropylidene-α-D-allofuranose **1b** (0.50 g, 1.92 mmol) and 4-(dimethylamino)pyridine (DMAP) (0.5 g, 4.09 mmol) was dried under high vacuum overnight. The compounds in the reaction flask were dissolved in anhydrous DCM (20 mL). The reaction was cooled to -20 °C, and diethylaminosulfer trifluoride (DAST) (0.5 mL, 3.78 mmol) was added dropwise via syringe over 5 mins. The reaction was allowed to warm to room temperature and stirred overnight. Once the reaction was completed as judged by TLC, MeOH (5 mL) was added under ice-bath to quench the remaining DAST. The reaction solution was washed with a saturated $NaHCO_3$ solution (3 × 15 mL), water (3 × 15 mL), and brine (3 × 15 mL). The organic layer was dried over Na_2SO_4, filtered, then condensed *in vacuo*. The pure product **2b** (0.266 g, 53%) was isolated as thick colorless oil by silica gel flash column chromatography. Rf: 0.4 (EtOAc:Hex, 1:8 V:V)

^{1}H NMR (500 MHz, Chloroform-*d*) δ 5.93 (d, *J* = 3.7 Hz, 1H, H-1), 4.98 (dd, *J* = 49.8, 2.3 Hz, 1H, H-3), 4.67 (dd, *J* = 10.7, 3.7 Hz, 1H, H-2), 4.26 (ddd, *J* = 8.3, 6.2, 4.8 Hz, 1H, H-

5), 4.14 – 3.97 (m, 4H, H-4, H-6a, H-6b), 1.48 (s, 3H, H-Me), 1.42 (s, 3H, H-Me), 1.34 (s, 3H, H-Me), 1.30 (s, 3H, H-Me).

3-Deoxy-3-fluoro-D-glucose[87] (3b)

The deprotection conditions of isopropylidene groups were adapted from the method used by Egli et al.[87] To a flask containing compound **2b** (0.266 g, 1.014 mmol) in 50% 1,4-dioxane in water (5.6 mL), Amberlite IR120 H^+ resin was added. The reaction was heated to 80 °C and stirred overnight. Once no more starting material was seen on TLC, the beads were filtered through sintered glass funnel and the beads were washed with 80% acetonitrile in water (60 mL). The filtrate was collected and condensed *in vacuo* to give product **3b** as a thick clear oil (0.166 g, 90%). Rf: 0.5 (EtOAc:MeOH:water, 9:3:1 V:V:V)

^{1}H NMR (500 MHz, D_2O) δ 5.27 (t, *J* = 3.9 Hz, 1H, H-1), 4.44 (m, 1H, H-3), 3.79 (m, 4H, H-2, H-4, H-6a, H-6b), 3.51 (m, 1H, H-5).

1,2,4,6-Tetra-*O*-acetyl-3-deoxy-3-fluoro-D-glucopyranose[88] (4b)

Compound **4b** was prepared according to the general procedure for acetylation reported by the synthesis of **2a**. Compound **3b** (0.166 g, 0.911 mmol), acetic anhydride (4 mL), and anhydrous pyridine (4 mL) were used to give thick clear oil (0.287 g, 90%). Rf: 0.3 (EtOAc:Hex, 1:1 V:V)

^{1}H NMR (400 MHz, Chloroform-*d*) δ 5.61 (d, *J* = 8.3 Hz, 1H, H-1), 5.20 – 5.13 (m, 3H, H-2, H-3, H-4), 4.25 (dd, *J* = 12.7, 4.5 Hz, 1H, H-6a), 4.21 (dd, *J* = 12.4, 2.5 Hz, 1H, H-6b), 3.84 (ddd, *J* = 10.1, 4.4, 2.2 Hz, 1H, H-5), 2.13 (s, 3H, H-OAc), 2.11 (s, 3H, H-OAc), 2.07 (s, 3H, H-OAc), 2.04 (s, 3H, H-OAc).

2,4,6-Tri-O-acetyl-3-Deoxy-3-fluoro-α-D-glucopyranosyl bromide (5b)

Compound **5b** was prepared according to general procedure for bromination reaction used to the synthesis of **3a**. Compound **4b** (0.319 g, 0.911 mmol) with 33% HBr in AcOH (2 mL) and anhydrous DCM (3 mL) were used to give crude product **5b** (0.0.392 g) of thick yellow oil. Rf: 0.7 (EtOAc:Hex, 1:1 V:V)

4-Methylumbeliferyl 2,4,6-tri-*O*-acetyl-3-deoxy-3-fluoro-β-D-glucopyranoside (6b)

Compound **6b** was prepared according to general procedure for glycosylation reaction reported by the synthesis of **4a**. The reaction was done with compound **5b** (0.392 g, 1.057 mmol) along with the reagents 4-methylumbeliferone (0.186 g, 1.057 mmol), $Na_2CO_3 \cdot H_2O$ (0.262 g, 2.114 mmol), and TBAB (0.341 g, 1.507 mmol) in 75 mL solvent of CH_3Cl:H_2O (1:1 V:V). Pure product **6b** (0.142 g, 29%) was isolated as white powder after purification with silica gel flash column chromatography. Rf: 0.25 (EtOAc:Hex, 1:1 V:V)

^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.55 (d, *J* = 8.8 Hz, 1H, H-5'), 7.01 – 6.89 (m, 2H, H-6', H-8'), 6.20 (q, *J* = 1.4 Hz, 1H, H-3'), 5.49 (d, *J* = 9.0 Hz, 1H, H-1), 5.25 (m, 2H, H-2, H-4), 4.49 (dd, *J* = 50.5, 8.7 Hz, 1H, H-3), 4.42 (dd, *J* = 12.5, 4.8 Hz, 1H, H-6a), 4.38 (dd, *J* = 12.3, 2.3 Hz, 1H, H-6b), 3.95 (ddd, *J* = 10.9, 4.7, 2.3 Hz, 1H, H-5), 2.38 (d, *J* = 1.2 Hz, 3H, H-Me), 2.15 (s, 3H, H-OAc), 2.11 (s, 3H, H-OAc), 2.07 (s, 3H, H-OAc).

^{13}C NMR (126 MHz, Chloroform-*d*) δ 171.10 (C-OAc), 170.02 (C-OAc), 169.80 (C-OAc), 163.11 (C-2'), 161.13 (C-7'), 155.21, 153.10, 125. 01 (C-5'), 115.39 (C-4'), 114.69 (C-3'), 113.41 (C-6'), 104.01 (C-8'), 98.50 (d, *J* = 12.5 Hz, C-1), 89.87 (d, *J* = 183.1 Hz, C-3), 79.61 (d, *J* = 18.3 Hz, C-2), 71.78 (d, *J* = 18.5 Hz, C-4), 67.98 (d, *J* = 8.0 Hz, C-5), 60.15 (C-6), 21.05 (C-Me), 20.89 (C-OAc), 20.78 (C-OAc), 20.51 (C-OAc).

HRMS (ESI$^+$) *m/z* [M+H]$^+$: Calc'd for $C_{22}H_{24}FO_{10}$ 467.1354, found 467.1252.

4-Methylumbeliferyl 3-deoxy-3-fluoro-β-D-glucopyranoside (7b)

Compound **7b** was prepared according to general procedure for deacetylation reaction reported as synthesis of **5a**. Compound **6b** (0.142 g, 0.303 mmol) was reacted with catalytic K_2CO_3 (0.0073 g, 0.175 mmol) in MeOH (2.3 mL). The crude product was purified with silica gel flash column chromatography to give pure compound **7b** (0.102 g, 99%) as white powder. Rf; 0.45 (DCM:MeOH, 9:1 V:V)

^{1}H NMR (600 MHz, Methanol-d_4) δ 7.74 (d, *J* = 8.8 Hz, 1H, H-5'), 7.13 (dd, *J* = 8.9, 2.3 Hz, 1H, H-6'), 7.10 (d, *J* = 2.5 Hz, 1H, H-8'), 6.24 (q, *J* = 1.4 Hz, 1H, H-3'), 5.12 (d, *J* = 7.8 Hz, 1H, H-1), 4.41 (dt, *J* = 52.4, 8.8 Hz, 1H, H-3), 3.94 (dt, *J* = 11.9, 1.9 Hz, 1H, H-6a), 3.74 (m, 3H, H-2, H-4, H-6b), 3.57 (dd, *J* = 10.4, 5.0 Hz, 1H, H-5), 2.48 (d, *J* = 1.4 Hz, 3H, H-Me).

^{13}C NMR (151 MHz, Methanol-d_4) δ 163.27 (C-2'), 161.83, 156.02, 155.42 (C-7'), 127.33 (C-5'), 116.20 (C-4'), 114.91 (C-6'), 112.99 (C-3'), 105.00 (C-8'), 101.03 (d, *J* = 12.6 Hz, C-1), 98.14 (d, *J* = 183.4 Hz, C-3), 76.99 (d, *J* = 8.0 Hz, C-5), 73.21 (d, *J* = 18.6 Hz, C-4), 69.30 (d, *J* = 18.4 Hz, C-2), 61.86 (C-6), 18.66 (C-Me).

HRMS (ESI$^+$) *m/z* [M+H]$^+$: Calc'd for $C_{16}H_{18}FO_7$ 341.1037, found 341.1001.

2.5.4. Synthesis of 4-methylumbelliferyl 4-deoxy-4-fluoro-β-D-glucopyranoside (3)

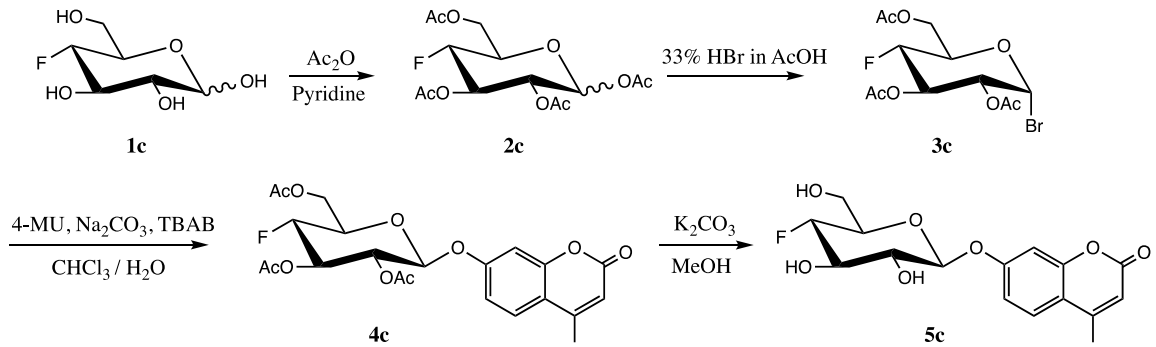


Figure 2.8. Synthetic scheme of 4MU-4F-Glc

1,2,3,6-Tetra-*O*-acetyl-4-deoxy-4-fluoro-D-glucopyranose[89] (2c)

Compound **2c** was synthesized according to the general procedure for acetylation reported by the synthesis of **2a**. Compound **1c** (0.50 g, 2.7 mmol), acetic anhydride (12 mL), and anhydrous pyridine (12 mL) were used to give product **2c** as a thick clear oil (0.49 g, 52%). Rf: 0.4 (EtOAc:Hex, 2:3 V:V)

^{1}H NMR (400 MHz, Chloroform-*d*) δ 5.71 (d, *J* = 8.3 Hz, 1H, H-1), 5.29 – 5.07 (m, 3H, H-2, H-3, H-4), 4.29 (dd, *J* = 12.5, 4.6 Hz, 1H, H-6a), 4.11 (dd, *J* = 12.5, 2.2 Hz, 1H, H-6b), 3.84 (ddd, *J* = 10.0, 4.6, 2.2 Hz, 1H, H-5), 2.11 (s, 3H, H-OAc), 2.08 (s, 3H, H-OAc), 2.03 (s, 3H, H-OAc), 2.01 (s, 3H, H-OAc).

2,3,6-Tri-*O*-acetyl-4-deoxy-4-fluoro-α-D-glucopyranosyl bromide[89] (3c)

Compound **3c** was prepared according to general procedure for bromination reaction used to the synthesis of **3a**. Compound **2c** (0.216 g, 0.617 mmol) with 33% HBr

in AcOH (2 mL) and anhydrous DCM (3 mL) were used to give crude product **3c** (0.066 g) of thick yellow oil. Rf: 0.4 (EtOAc:Hex, 3:7 V:V)

4-Methylumbeliferyl 2,3,6-tri-*O*-acetyl-4-deoxy-4-fluoro-β-D-glucopyranoside (4c)

Compound **4c** was prepared according to general procedure for glycosylation reaction reported as the synthesis of **4a**. The reaction was done with compound **3c** (0.066 g, 0.177 mmol) along with the reagents 4-methylumbeliferone (0.0625 g, 0.3546 mmol), $Na_2CO_3 \cdot H_2O$ (0.044 g, 0.355 mmol), and TBAB (0.057 g, 0.177 mmol) in 13 mL solvent of CH_3Cl:H_2O (1:1 V:V). Pure product **4c** (0.019 g, 23%) was isolated as white powder after purification with silica gel flash column chromatography. Rf: 0.35 ($CHCl_3$:EtOAc, 4:1 V:V)

^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.52 (d, *J* = 8.7 Hz, 1H, H-5'), 6.98 – 6.88 (m, 2H, H-6', H-8'), 6.19 (q, *J* = 1.4 Hz, 1H, H-3'), 5.75 (d, *J* = 8.3 Hz, 1H, H-1), 5.40 (m, 1H, H-3), 5.10 (dd, *J* = 9.7, 8.3 Hz, 1H, H-2), 4.55 (ddd, *J* = 50.2, 9.9, 8.9 Hz, 1H, H-4), 4.43 (dt, *J* = 12.4, 2.1 Hz, 1H, H-6a), 4.29 (ddd, *J* = 12.4, 4.8, 1.7 Hz, 1H, H-6b), 3.91 (ddd, *J* = 9.9, 5.0, 2.6 Hz, 1H, H-5), 2.40 (d, *J* = 1.3 Hz, 3H, H-Me), 2.18 – 2.04 (m, 9H, H-OAc × 3).

^{13}C NMR (126 MHz, Chloroform-*d*) δ 170.47 (C-OAc), 169.83 (C-OAc), 169.40 (C-OAc), 162.31 (C-2'), 160.01 (C-7'), 154.98, 153.07, 126.04 (C-5'), 115.40 (C-4'), 113.70 (C-3'), 113.58 (C-6'), 103.63 (C-8'), 91.69 (C-1), 86.42 (d, *J* = 188.5 Hz, C-4), 72.66 (d, *J* = 19.8 Hz, C-3), 72.20 (d, *J* = 24.1 Hz, C-5), 70.20 (d, *J* = 7.8 Hz, C-2), 61.82 (C-6), 20.64 (C-OAc), 20.73 (C-OAc), 20.77 (C-OAc), 20.50 (C-Me).

HRMS (ESI$^+$) *m/z* $[M+H]^+$: Calc'd for $C_{22}H_{24}FO_{10}$ 467.1354, found 467.1217.

4-Methylumbeliferyl 4-deoxy-4-fluoro-β-D-glucopyranoside (5c)

Compound **5c** was prepared according to general procedure for deacetylation reaction reported as the synthesis of **5a**. Compound **4c** (0.019 g, 0.040 mmol) was reacted with catalytic K_2CO_3 (0.001 g, 0.007 mmol) in MeOH (1 mL). The crude product was purified with silica gel flash column chromatography to give pure compound **5c** (0.0094 g, 69%) as a white powder. Rf: 0.30 (DCM:MeOH, 9:1 V:V)

^{1}H NMR (500 MHz, Methanol-d_4) δ 7.71 (d, *J* = 8.8 Hz, 1H, H-5'), 7.11 (dd, *J* = 8.8, 2.4 Hz, 1H, H-6'), 7.07 (d, *J* = 2.4 Hz, 1H, H-8'), 6.20 (q, *J* = 1.2 Hz, 1H, H-3'), 5.12 (d, *J* = 7.8 Hz, 1H, H-1), 4.46 – 4.27 (m, 1H, H-4), 3.93 – 3.85 (m, 1H, H-6a), 3.84 – 3.70 (m, 3H, H-3, H-5, H-6b), 3.54 (dd, *J* = 9.3, 7.8 Hz, 1H, H-2), 2.45 (d, *J* = 1.2 Hz, 3H, H-Me).

^{13}C NMR (126 MHz, MeOD) δ 163.26 (C-2'), 161.84, 156.02, 155.42 (C-7'), 127.31 (C-5'), 116.18 (C-4'), 114.97 (C-6'), 112.98 (C-3'), 104.99 (C-8'), 101.59 (C-1), 90.17 (d, *J* = 180.9 Hz, C-4), 75.68 (d, *J* = 12.3 Hz, C-5), 75.51 (d, *J* = 18.6 Hz, C-3), 74.50 (d, *J* = 8.8 Hz, C-2), 61.46 (C-6), 18.63 (C-Me).

HRMS (ESI$^+$) *m/z* [M+H]$^+$: Calc'd for $C_{16}H_{18}FO_7$ 341.1037, found 341.0989.

2.5.5. Synthesis of 4-methylumbelliferyl 6-deoxy-6-fluoro-β-D-glucopyranoside (4)

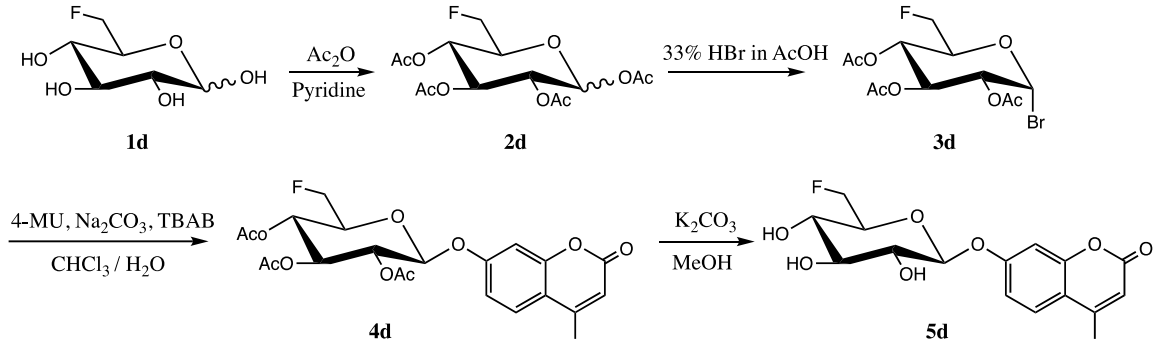


Figure 2.9. Synthetic scheme of 4MU-6F-Glc

1,2,3,4-Tetra-*O*-acetyl-6-deoxy-6-fluoro-D-glucopyranose[54] (2d)

Compound **2d** was synthesized according to the general procedure for acetylation reported by the synthesis of **2a**. Compound **1d** (0.20 g, 1.098 mmol), acetic anhydride (4 mL), and anhydrous pyridine (4 mL) were used to give product **3d** of a thick clear oil (0.377 g, 98%). Rf: 0.5 (EtOAc:Hex, 1:1)

^{1}H NMR (400 MHz, Chloroform-*d*) δ 5.73 (d, *J* = 8.1 Hz, 1H, H-1), 5.28 – 5.10 (m, 3H, H-2, H-3, H-4), 4.50 (ddd, *J* = 49.1, 10.7, 2.2 Hz, 1H, H-6a), 4.45 (ddd, *J* = 47.5, 10.6, 4.0 Hz, 1H, H-6b), 3.99 (m, 1H, H-5), 2.11 (s, 3H, H-OAc), 2.07 (s, 3H, H-OAc), 2.05 (s, 3H, H-OAc), 2.03 (s, 3H, H-OAc).

2,3,4-Tri-*O*-acetyl-6-deoxy-6-fluoro-α-D-glucopyranosyl bromide (3d)

Compound **3d** was synthesized according to the general procedure for bromination reaction reported by the synthesis of **3a**. Compound **2d** (0.377 g, 1.076 mmol) with 33% HBr in AcOH (3 mL) and anhydrous DCM (4 mL) were used to give crude product **3d** (0.361 g) as light-yellow solid. Rf: 0.6 (EtOAc:Hex, 1:1 V:V)

4-Methylumbeliferyl 2,3,4-tri-*O*-acetyl-6-deoxy-6-fluoro-β-D-glucopyranoside (4d)

Compound **4d** was synthesized according to general procedure for glycosylation reaction reported as synthesis of **4a**. The reaction was done with compound **3d** (0.361 g, 0.972 mmol) along with the reagents 4-methylumbeliferone (0.171 g, 0.972 mmol), $Na_2CO_3 \cdot H_2O$ (0.241 g, 1.943 mmol), and TBAB (0.313 g, 0.972 mmol) in 70 mL solvent of CH_3Cl:H_2O (1:1 V:V). Pure product **4d** (0.139 g, 31%) was isolated as white powder after purification with silica gel flash column chromatography. Rf: 0.29 (EtOAc:Hex, 1:1 V:V)

^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.60 (d, *J* = 8.8 Hz, 1H, H-5'), 7.10 – 6.89 (m, 2H, H-6', H-8'), 6.20 (q, *J* = 1.5 Hz, 1H, H-3'), 5.50 (d, *J* = 8.8 Hz, 1H, H-1), 5.30 – 5.15 (m, 3H, H-2, H-3, H-4), 4.47 (ddd, *J* = 50.0, 10.1, 2.1 Hz, 1H, H-6a), 4.41 (ddd, *J* = 49.8, 10.5, 4.3 Hz, 1H, H-6b), 3.97 (m, 1H, H-5), 2.41 (d, *J* = 1.4 Hz, 3H, H-Me), 2.25 – 2.08 (m, 9H, H-OAc × 3).

^{13}C NMR (126 MHz, Chloroform-*d*) δ 170.57 (C-OAc), 169.24 (C-OAc), 168.68 (C-OAc), 161.01 (C-2'), 156.56 (C-7'), 155.29, 154.12, 125.50 (C-5'), 115.23 (C-4'), 114.10 (C-3'), 111.81 (C-8'), 110.15 (C-6'), 87.11 (C-1), 72.43 (d, *J* = 167.5 Hz, C-6), 71.81 (C-2), 70.01 (d, *J* = 18.4 Hz,C-5), 69.18 (C-3), 40.05 (d, 12.0 Hz, C-4), 21.05 (C-OAc), 20.84 (C-OAc), 20.54 (C-OAc), 19.59 (C-Me).

HRMS (ESI$^+$) *m/z* $[M+H]^+$: Calc'd for $C_{22}H_{24}FO_{10}$ 467.1354, found 467.1266.

4-Methylumbeliferyl 6-deoxy-6-fluoro-β-D-glucopyranoside (5d)

Compound **5d** was synthesized according to general procedure for deacetylation reaction reported by the synthesis of **5a**. Compound **4d** (0.139 g, 0.299 mmol) was reacted with catalytic K_2CO_3 (0.007 g, 0.052 mmol) in MeOH (2.2 mL). The crude

product was purified with silica gel flash column chromatography to give pure compound **5d** (0.072 g, 71%) as white powder. Rf: 0.40 (DCM:MeOH, 9:1 V:V)

^{1}H NMR (600 MHz, Methanol-d_4) δ 7.72 (d, *J* = 8.8 Hz, 1H, H-5'), 7.11 (dd, *J* = 8.8, 2.4 Hz, 1H, H-6'), 7.06 (d, *J* = 2.4 Hz, 1H, H-8'), 6.21 (q, *J* = 1.2 Hz, 1H, H-3'), 5.08 (d, *J* = 7.2 Hz, 1H, H-1), 4.65 (m, 2H, H-6a, H-6b), 3.72 (dddd, *J* = 24.4, 10.0, 4.9, 1.8 Hz, 1H, H-5), 3.49 (m, 2H, H-2, H-3), 3.43 (m, 1H, H-4), 2.46 (d, *J* = 1.2 Hz, 3H, H-Me).

^{13}C NMR (151 MHz, Methanol-d_4) δ 163.32 (C-2'), 161.88, 156.02, 155.48 (C-7'), 127.33 (C-5'), 116.14 (C-4'), 114.90 (C-6'), 112.93 (C-3'), 104.87 (C-8'), 101.62 (H-1), 83.11 (d, *J* = 171.9 Hz, C-6), 77.65 (C-2), 76.62 (d, *J* = 18.0 Hz, C-5), 74.60 (C-3), 70.08 (d, *J* = 7.1 Hz, C-4), 18.66 (C-Me).

HRMS (ESI$^+$) *m/z* [M+H]$^+$: Calc'd for $C_{16}H_{18}FO_7$ 341.1037, found 341.1036.

2.5.6. Synthesis of 4-methylumbelliferyl 3-deoxy-β-D-ribo-hexopyranoside (5)

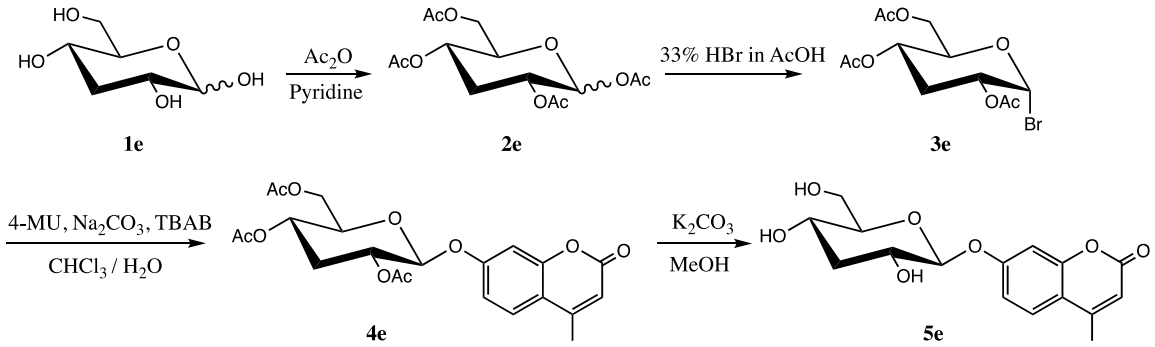


Figure 2.10. Synthetic scheme of 4MU-3H-Glc

1,2,4,6-Tetra-*O*-acetyl-3-deoxy-D-ribo-hexopyranose[90] (2e)

Compound **2e** was synthesized according to the general procedure for acetylation reported by the synthesis of **2a**. Compound **1e** (0.100 g, 0.609 mmol), acetic anhydride (2.2 mL), and anhydrous pyridine (2.2 mL) were used to give product **2e** as thick clear oil (0.190 g, 97%). Rf: 0.5 (EtOAc:Hex, 1:1 V:V)

^{1}H NMR (400 MHz, Chloroform-*d*) δ 5.69 (d, *J* = 7.9 Hz, 1H, H-1), 4.87 (dddd, *J* = 14.5, 10.8, 8.9, 5.1 Hz, 2H, H-2, H-5), 4.23 (dt, *J* = 12.3, 4.2 Hz, 1H, H-4), 4.14 (m, 1H, H-6a), 3.82 (ddd, *J* = 9.5, 5.1, 2.6 Hz, 1H, H-6b), 2.63 (dt, *J* = 12.4, 5.0 Hz, 1H, H-3a), 2.11 (s, 3H, H-OAc), 2.06 (s, 3H, H-OAc), 2.05 (s, 3H, H-OAc), 2.04 (s, 3H, H-OAc), 1.66 (dt, *J* = 12.5, 10.9 Hz, 1H, H-3b).

2,4,6-Tri-*O*-acetyl-3-deoxy-α-D-ribo-hexopyranosyl bromide[90] (3e)

Compound **3e** was prepared according to general procedure for bromination reaction used to the synthesis of **3a**. Compound **2e** (0.190 g, 0.572 mmol) with 33% HBr in AcOH (1 mL) and anhydrous DCM (1.6 mL) were used to give crude product **3e** (0.176 g) as light-yellow oil. Rf: 0.7 (EtOAc:Hex, 1:1 V:V)

4-Methylumbeliferyl 2,4,6-tri-*O*-acetyl-3-deoxy-β-D-ribo-hexopyranoside (4e)

Compound **4e** was prepared according to the general procedure for glycosylation reaction reported as the synthesis of **4a**. The reaction was done with compound **3e** (0.176 g, 0.498 mmol) along with the reagents 4-methylumbeliferone (0.088 g, 0.498 mmol), $Na_2CO_3 \cdot H_2O$ (0.124 g, 0.996 mmol), and TBAB (0.161 g, 0.498 mmol) in 36 mL solvent of CH_3Cl:H_2O (1:1). Pure product **4e** (0.131 g, 59%) was isolated as white powder after purification with silica gel flash column chromatography. Rf: 0.24 (EtOAc:Hex, 1:1 V:V)

^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.50 (d, *J* = 8.2 Hz, 1H, H-5'), 7.07 (dd, *J* = 8.2, 2.4 Hz, 1H, H-6'), 6.97 (d *J* = 2.4 Hz, 1H, H-8'), 6.20 (q, *J* = 1.4 Hz, 1H, H-3'), 5.40 (d, *J* = 7.8 Hz, 1H, H-1), 4.91 (m, 2H, H-2, H-5), 4.84 (m, 1H, H-4), 4.10 – 3.91 (m, 2H, H-6a, H-6b), 2.58 (m, 1H, H-3a), 2.42 (d, *J* = 1.5 Hz, 3H, H-Me), 2.21 (s, 3H, H-OAc), 2.16 (s, 3H, H-OAc), 2.10 (s, 3H, H-OAc), 1.50 (dt, *J* = 12.5, 10.5 Hz, 1H, H-3b).

^{13}C NMR (126 MHz, Chloroform-*d*) δ 171.13 (C-OAc), 170.80 (C-OAc), 169.41 (C-OAc), 163.08 (C-2'), 161.44 (C-7'), 155.90, 155.12, 124.98 (C-5'), 115.50 (C-4'), 114.01 (C-3'), 113.14 (C-6'), 103.91 (C-8'), 81.09 (C-1), 71.48 (C-4), 70.50 (C-5), 70.12 (C-2), 65.14 (C-6), 31.40 (C-3), 20.54 (C-OAc), 20.21 (C-OAc), 20.10 (C-OAc), 19.05 (C-Me).

HRMS (ESI$^+$) *m/z* [M+H]$^+$: Calc'd for $C_{22}H_{25}O_{10}$ 449.1448, found 449.1318.

4-Methylumbeliferyl 3-deoxy-β-D-ribo-hexopyranoside (5e)

Compound **5e** was prepared according to the general procedure for deacetylation reaction reported as the synthesis of **5a**. Compound **4e** (0.131 g, 0.291 mmol) was reacted with catalytic K_2CO_3 (0.007 g, 0.051 mmol) in MeOH (2.1 mL). The crude product was purified with silica gel flash column chromatography to give pure compound **5e** (0.074 g, 80%) as a white powder. Rf: 0.31 (DCM:MeOH, 9:1 V:V)

^{1}H NMR (600 MHz, Methanol-d_4) δ 7.69 (d, *J* = 8.8 Hz, 1H, H-5'), 7.11 (dd, *J* = 8.8, 2.4 Hz, 1H, H-6'), 7.07 (d, *J* = 2.4 Hz, 1H, H-8'), 6.20 (q, *J* = 1.3 Hz, 1H, H-3'), 4.96 (d, *J* = 7.6 Hz, 1H, H-1), 3.89 (dd, *J* = 12.1, 2.5 Hz, 1H, H-3), 3.67 (m, 3H, H-2, H-3, H-5), 3.48 (ddd, *J* = 9.3, 5.8, 2.5 Hz, 1H, H-4), 2.45 (d, *J* = 1.2 Hz, 3H, H-Me), 2.41 (dt, *J* = 12.3, 4.9 Hz, 1H, H-6a), 1.62 (q, *J* = 11.8 Hz, 1H, H-6b).

^{13}C NMR (151 MHz, Methanol-d_4) δ 163.30 (C-2'), 161.89, 155.87, 155.30 (C-7'), 127.09 (C-5'), 115.85 (C-4'), 114.85 (C-6'), 112.75 (C-3'), 104.90 (C-8'), 103.70 (C-1), 81.89 (C-4), 68.76 (C-2), 65.51 (C-5), 62.31 (C-3), 40.21 (C-6), 18.72 (C-Me).

HRMS (ESI$^+$) *m/z* $[M+H]^+$: Calc'd for $C_{16}H_{19}O_7$ 323.1131, found 323.1091.

2.5.7. Synthesis of 4-methylumbelliferyl 4-deoxy-β-D-glucopyranoside (6)

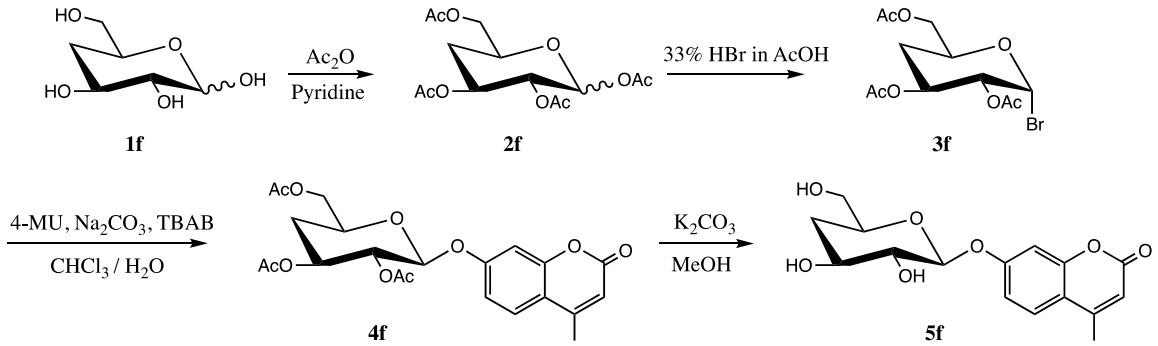


Figure 2.11. Synthetic scheme of 4MU-4H-Glc

1,2,3,6-Tetra-*O*-acetyl-4-deoxy-D-glucopyranose[91] (2f)

Compound **2f** was synthesized according to the general procedure for acetylation reported by the synthesis of **2a**. Compound **1f** (0.150 g, 0.914 mmol), acetic anhydride (3.3 mL), and anhydrous pyridine (3.3 mL) were used to give product **2f** as a thick clear oil (0.283 g, 94%). Rf: 0.4 (EtOAc:Hex, 1:1 V:V)

^{1}H NMR (400 MHz, Chloroform-*d*) δ 5.70 (d, *J* = 7.5 Hz, 1H, H-1), 5.15 (t, *J* = 1.5 Hz, 1H, H-2), 4.55 (m, 2H, H-3, H-5), 4.12 (m, 2H, H-6a, H-6b), 2.85 (dt, *J* = 12.5, 5.5 Hz, 1H, H-4a), 2.10 (s, 3H, H-OAc), 2.07 (s, 3H, H-OAc), 2.06 (s, 3H, H-OAc), 2.05 (s, 3H, H-OAc), 1.66 (dt, *J* = 12.3, 5.3 Hz, 1H, H-4b).

2,3,6-Tri-*O*-acetyl-4-deoxy-α-D-glucopyranosyl bromide (3f)

Compound **3f** was prepared according to the general procedure for bromination reaction reported by the synthesis of **3a**. Compound **2f** (0.283 g, 0.851 mmol) with 33% HBr in AcOH (2 mL) and anhydrous DCM (2 mL) were used to give crude product **3f** (0.262 g) of clear colorless oil. Rf: 0.6 (EtOAc:Hex, 1:1 V:V)

4-Methylumbeliferyl 2,3,6-tri-*O*-acetyl-4-deoxy-β-D-glucopyranoside (4f)

Compound **4f** was prepared according to the general procedure for glycosylation reaction reported by the synthesis of **4a**. The reaction was done with compound **3f** (0.262 g, 0.742 mmol) along with the reagents 4-methylumbeliferone (0.131 g, 0.742 mmol), $Na_2CO_3 \cdot H_2O$ (0.184 g, 1.484 mmol), and TBAB (0.239 g, 0.742 mmol) in 53 mL solvent of CH_3Cl:H_2O (1:1). Pure product **4f** (0.183 g, 55%) was isolated as a white powder after purification with silica gel flash column chromatography. Rf: 0.30 (EtOAc:Hex, 1:1 V:V)

^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.51 (d, *J* = 8.7 Hz, 1H, H-5'), 6.96 (d, *J* = 2.4 Hz, 1H, H-8'), 6.92 (dd, *J* = 8.7, 2.4 Hz, 1H, H-6'), 6.19 (t, *J* = 1.3 Hz, 1H, H-3'), 5.21 (dd, *J* = 9.5, 7.7 Hz, 1H, H-1), 5.11 (ddd, *J* = 11.4, 9.5, 5.3 Hz, 1H, H-3), 5.07 (d, *J* = 7.7 Hz, 1H, H-2), 4.25 – 4.14 (m, 2H, H-6a, H-6b), 3.97 (dddd, *J* = 12.3, 6.2, 3.7, 1.9 Hz, 1H, H-5), 2.41 (d, *J* = 1.2 Hz, 3H, H-Me), 2.27 – 2.18 (m, 1H, H-4a), 2.13 (s, 3H, H-OAc), 2.08 (s, 3H, H-OAc), 2.07 (s, 3H, H-OAc), 1.72 (dt, *J* = 12.9, 11.6 Hz, 1H, H-4b).

^{13}C NMR (126 MHz, Chloroform-*d*) δ 171.63 (C-OAc), 168.50 (C-OAc), 167.23 (C-OAc), 163.01 (C-2'), 161.07 (C-7'), 155.58, 154.47, 125.67 (C-5'), 116.05 (C-4'), 113.96 (C-6'), 113.16 (C-3'), 104.16 (C-8'), 98.86 (C-2), 71.62 (C-1), 70.31 (C-3), 70.25 (C-5), 64.93 (C-6), 32.21 (C-4), 20.92 (C-OAc), 20.80 (C-OAc), 20.77 (C-OAc), 18.70 (C-Me).

HRMS (ESI^{+}) *m/z* $[M+H]^{+}$: Calc'd for $C_{22}H_{25}O_{10}$ 449.1448, found 449.1375.

4-Methylumbeliferyl 4-deoxy-β-D-glucopyranoside (5f)

Compound **5f** was prepared according to the general procedure for deacetylation reaction reported by the synthesis of **5a**. Compound **4f** (0.183 g, 0.408 mmol) was reacted with catalytic K_2CO_3 (0.010 g, 0.071 mmol) in MeOH (3 mL). The crude product was purified with silica gel flash column chromatography to give pure compound **5f** (0.119 g, 91%) as a white powder. Rf: 0.23 (DCM:MeOH, 9:1 V:V)

^{1}H NMR (500 MHz, Methanol-d_4) δ 7.70 (d, *J* = 8.8 Hz, 1H, H-5'), 7.11 (dd, *J* = 8.8, 2.4 Hz, 1H, H-6'), 7.06 (d, *J* = 2.4 Hz, 1H, H-8'), 6.20 (q, *J* = 1.2 Hz, 1H, H-3'), 4.98 (d, *J* = 7.7 Hz, 1H, H-1), 3.83 – 3.70 (m, 2H, H-3, H-5), 3.66 – 3.56 (m, 2H, H-6a, H-6b), 3.39 (dd, *J* = 9.0, 7.7 Hz, 1H, H-2), 2.45 (d, *J* = 1.2 Hz, 3H, H-Me), 2.00 (ddd, *J* = 12.9, 5.2, 1.9 Hz, 1H, H-4a), 1.49 (dt, *J* = 12.9, 11.6 Hz, 1H, H-4b).

^{13}C NMR (126 MHz, MeOD) δ 163.36 (C-2'), 162.21 (C-7'), 156.07, 155.50, 127.25 (C-5'), 115.99 (C-4'), 115.04 (C-6'), 112.83 (C-3'), 104.96 (C-8'), 102.30 (C-1), 76.47 (C-2), 74.48 (C-5), 71.95 (C-3), 65.29 (C-6), 36.07 (C-4), 18.63 (C-Me).

HRMS (ESI$^+$) *m/z* [M+H]$^+$: Calc'd for $C_{16}H_{19}O_7$ 323.1131, found 323.1070.

2.5.8. Synthesis of 4-methylumbelliferyl 6-deoxy-β-D-glucopyranoside (7)

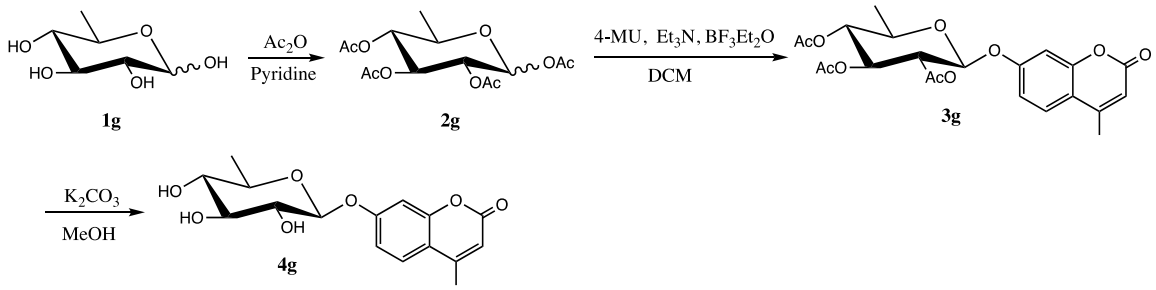


Figure 2.12. Synthetic scheme of 4MU-6H-Glc

1,2,3,4-Tetra-*O*-acetyl-6-deoxy-D-glucopyranose[92] (2g)

Compound **2g** was synthesized according to the general procedure for acetylation reported by the synthesis of **2a**. Compound **1g** (0.189 g, 1.15 mmol), acetic anhydride (1 mL), and anhydrous pyridine (9 mL) were used to give product **2g** as a white solid (0.247 g, 65%). Rf: 0.6 (EtOAc:Hex, 1:1 V:V)

^{1}H NMR (400MHz, Chloroform-*d*) δ 6.24 (d, *J* = 4.0 Hz, 1H, H-1), 5.35 (t, *J* = 10.1 Hz 1H, H-3), 5.05 (dd, *J* = 10.1, 3.9 Hz, 1H, H-2), 4.83 (dd, *J* = 10.2 Hz, 1H, H-4), 4.01 (m, 1H, H-5), 2.14 (s, 3H, H-OAc), 2.05 (s, 3H, H-OAc), 2.03 (s, 6H, H-OAc), 1.98 (s, 3H, OAc), 1.20 (d, *J* = 6.3 Hz, 3H, H-6).

4-Methylumbeliferyl 2,3,4-tri-*O*-acetyl-6-deoxy-β-D-glucopyranoside (3g)

The glycosylation conditions were adapted from the method used by Wei et al.[76] In a round bottom flask, compound **2g** (0.247 g, 0.743 mmol), 4-methylumbelliferone (0.262 g, 1.485 mmol) were suspended in 3 mL of anhydrous $CHCl_3$, under an argon atmosphere. The solids began to dissolve with the addition of 0.1 mL of triethylamine (0.743 mmol, 1 eq). The reaction was stirred at room temperature for 3 hours. 0.46 mL of boron trifluoride diethyl etherate (3.713 mmol, 5 eq) was diluted with minimal dry DCM and added to the reaction flask slowly, dropwise. The reaction was stirred at room temperature overnight. Once disappearance of starting material is confirmed with TLC, 4 mL of saturated $NaHCO_3$ solution was added to the reaction flask, and stirred until frothing was ceased. The organic layer was separated, and aqueous layer was extracted with DCM (3 × 10 mL). The combined organic layer was dried over Na_2SO_4. A pure product **3g** (0.072 g, 22%) was isolated by using silica gel flash column chromatography. Rf: 0.37 (EtOAc:Hex, 1:1 V:V)

^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.53 (t, *J* = 8.3 Hz, 2H, H-5'), 7.07 (d, *J* = 2.4 Hz, 1H, H-8'), 7.04 (dd, *J* = 8.7, 2.5 Hz, 1H, H-6'), 6.19 (q, *J* = 1.4 Hz, 1H, H-3'), 5.75 (d, *J* = 7.7 Hz, 1H, H-1), 5.65 (dd, *J* = 10.3, 9.4 Hz, 1H, H-3), 5.03 (dd, *J* = 10.3, 3.6 Hz, 1H, H-2), 4.92 (dt, *J* = 19.5, 9.6 Hz, 2H, H-4), 3.96 (dq, *J* = 10.0, 6.3 Hz, 1H, H-5), 2.41 (t, *J* = 1.1 Hz, 5H, H-Me), 2.10 – 2.02 (m, 9H, 3 × H-OAc), 1.17 (d, *J* = 6.3 Hz, 3H, H-6a, H-6b, H-6c).

^{13}C NMR (126 MHz, Chloroform-*d*) δ 170.26 (C-OAc), 169.76 (C-OAc), 169.34 (C-OAc), 162.81 (C-2'), 160.11 (C-7'), 156.61, 155.72, 125.78 (C-5'), 115.80 (C-4'), 113.44 (C-6'), 113.12 (C-3'), 104.45 (C-8'), 94.43 (C-1), 73.47 (C-4), 70.60 (C-2), 69.88 (C-3), 66.49 (C-5), 20.74 (C-OAc), 20.69 (C-OAc), 20.65 (C-OAc), 18.67 (C-Me), 17.26 (C-6).

HRMS (ESI^{+}) *m/z* [M+H]$^{+}$: Calc'd for $C_{22}H_{25}O_{10}$ 449.1448, found 449.1306.

4-Methylumbeliferyl 6-deoxy-β-D-glucopyranoside (4g)

Compound **4g** was prepared according to the general procedure for deacetylation reaction reported by the synthesis of **5a**. Compound **3g** (0.067 g, 0.15 mmol) was reacted with catalytic K_2CO_3 (0.002 g, 0.015 mmol, 0.1 eq) in MeOH (1.7 mL). A crude product was purified with silica gel flash column chromatography to give a pure compound **4g** (0.023 g, 47%) as a white powder. Rf: 0.37 (DCM:MeOH, 9:1 V:V)

^{1}H NMR (500 MHz, Methanol-d_4) δ 7.74 (d, *J* = 8.8 Hz, 1H, H-5'), 7.11 (dd, *J* = 8.8, 2.3 Hz, 1H, H-6'), 7.04 (d, *J* = 2.3 Hz, 1H, H-8'), 6.23 (s, 1H, H-3'), 5.05 (d, *J* = 7.3 Hz, 1H, H-1), 3.59 (m, 1H, H-5), 3.50 (m, 1H, H-2), 3.46 (t, *J* = 8.9 Hz, 1H, H-3), 3.12 (t, *J* = 9.0 Hz, 1H, H-4), 2.48 (s, 3H, H-Me), 1.35 (d, *J* = 6.2 Hz, 3H, H-6a, H-6b, H-6c).

^{13}C NMR (126 MHz, MeOD) δ 163.33 (C-2'), 162.00, 155.73 (C-7'), 155.49, 127.32 (C-5'), 115.92 (C-4'), 114.88 (C-6'), 112.87 (C-3'), 104.81 (C-8'), 101.55 (C-1), 77.56 (C-3), 76.71 (C-4), 75.01 (C-2), 73.62 (C-5), 18.63 (C-Me), 18.09 (C-6).

HRMS (ESI$^+$) *m/z* [M+H]$^+$: Calc'd for $C_{16}H_{19}O_7$ 323.1131, found 323.1095.

2.5.9. Synthesis of 4-methylumbelliferyl 2-*O*-methyl-β-D-glucopyranoside (8) and 4-methylumbelliferyl 3-*O*-methyl-β-D-glucopyranoside (9)

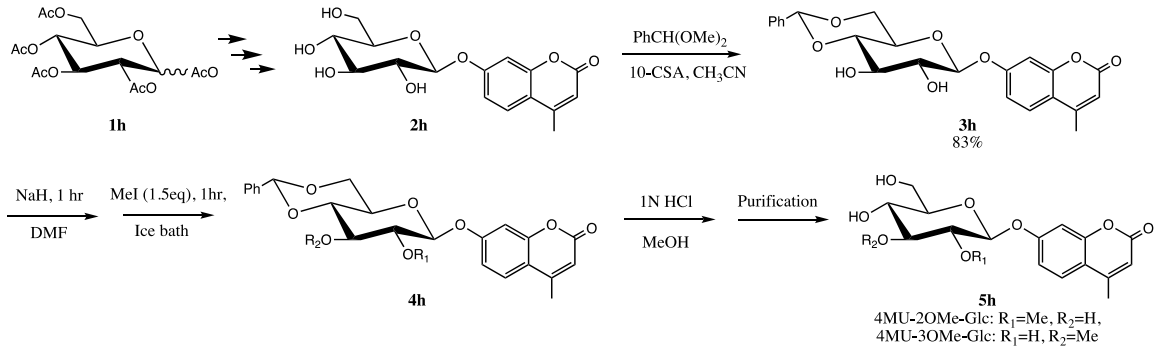


Figure 2.13. Synthetic scheme of 4MU-2OMe-Glc and 4MU-3OMe-Glc

4-Methylumbelliferyl-β-D-glucopyranoside (2h)

Compound **2h** was prepared following the general procedure for bromination, glycosylation, and deacetylation reactions reported by the synthesis of compound **5a**. A pure compound **2h** (0.927 g) was obtained.

4-Methylumbelliferyl 4,6-*O*-benzylidene-β-D-glucopyranoside[93] (3h)

The benzylidene acetal protection conditions were adopted from the method used by Driguez et al.[82] Compound **2h** (0.76 g, 2.25 mmol) was dissolved in 43 mL of anhydrous acetonitrile under an atmosphere of argon. 10-Camphorsulfonic acid (0.052 g, 0.225 mmol, 0.1 eq), dissolved in minimal solvent, and benzaldehyde dimethyl acetal (1.2 mL, 7.865 mmol, 3.5 eq) were added to the reaction flask. The reaction mixture was refluxed and stirred for 1.5 hour, at which all of the reagents dissolved and the solution became clear. The flask was cooled to room temperature, then the reaction was quenched with 0.62 mL of triethylamine (2 eq). The solvent was concentrated *in vacuo*. A pure compound **3h** (0.810 g, 85%) was isolated using silica gel flash column chromatography (Eluted at DCM:MeOH 9:1). Rf: 0.67 (DCM:MeOH, 9:1 V:V)

^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.55 – 7.35 (m, 6H, H-Ph, H-5'), 7.03 – 6.97 (m, 2H, H-6', H-8'), 6.18 (d, *J* = 1.4 Hz, 1H, H-3'), 5.55 (s, 1H, 3°H), 5.11 (d, *J* = 7.6 Hz, 1H, H-1), 4.39 (dd, *J* = 10.6, 4.1 Hz, 1H, H-6a), 3.87 (t, *J* = 9.0 Hz, 1H, H-4), 3.83 – 3.73 (m, 2H, H-6b, H-5), 3.67 – 3.59 (m, 2H, H-2, H-3), 2.40 (d, *J* = 1.2 Hz, 3H, H-Me).

HRMS (ESI$^+$) *m/z* [M+H]$^+$: Calc'd for $C_{23}H_{23}O_8$ 427.1393, found 427.1246.

4-Methylumbelliferyl 4,6-*O*-benzylidene-2-*O*-methyl-β-D-glucopyranoside (4h)

Compound **3h** was co-evaporated with toluene and MeOH, then placed under a high vacuum overnight to ensure maximum dryness of the compound. Compound **3h** (0.810 g, 1.900 mmol) was dissolved in 20 mL of anhydrous DMF under an atmosphere of argon. NaH solid (60% NaH in mineral oil) (0.137 g, 5.701 mmol, 3 eq) was added to the reaction flask. The reaction was stirred for 1 hour at room temperature. The flask was cooled to 0 °C, then 0.18 mL of MeI (2.850 mmol, 1.5 eq) was added dropwise to the reaction. The reaction mixture was stirred at 0 °C for 1 hour. 10 mL of saturated aqueous ammonium chloride solution was added to the reaction to quench the excess NaH. The reaction was extracted with DCM (3 × 10 mL), then washed with sat. NH_4Cl (3 × 10 mL),

water (3 × 10 mL), brine (3 × 10 mL) in the following order. The organic layer was dried over Na_2SO_4, then concentrated *in vacuo*. The crude product contained some unreacted starting material (Rf: 0.50, DCM:MeOH, 9:1 V:V), two mono-methylated products: 4MU-2OMe-Glc (Rf: 0.63, DCM:MeOH, 9:1 V:V) and 4MU-3OMe-Glc (Rf: 0.59, DCM:MeOH, 9:1 V:V), and a di-methylated product (Rf: 0.82, DCM:MeOH, 9:1 V:V). The starting material **3h** (0.195 g, 25%) was recovered using silica gel flash column chromatography (eluted at DCM:MeOH 97:3).

Isolation was the mono- and di- methylated compounds were difficult due to low polarity of the compounds, therefore the compounds were unable to be isolated at this step.

4-Methylumbelliferyl 2-*O*-methyl-β-D-glucopyranoside and 4-methylumbelliferyl 3-*O*-methyl-β-D-glucopyranoside (5h)

Crude product **4h** (0.386 g) containing two mono-methylated at 2O and 3O, and a di-methylated compound at those positions were suspended in 8.7 mL of MeOH and 0.5 mL water. 1 M HCl (0.9 mL) was added to the reaction. The reaction was stirred at 55 °C for 3 hours. Upon checking for disappearance of starting materials, saturated $NaHCO_3$ solution was added to the reaction dropwise until the pH reached 7. The solution was azeotrope-dried with toluene. The two mono-methylated products, **4MU-2OMe-Glc** (0.034 g, 11%) and **4MU-3OMe-Glc** (0.045 g, 15%), and the di-methylated product (0.022 g, 7%) were isolated using silica gel flash column chromatography. **4MU-2OMe-Glc** Rf: 0.3, **4MU-3OMe-Glc** Rf: 0.4, di-methylated product Rf: 0.6 (DCM:MeOH, 9:1 V:V)

4MU-2OMe-Glc

^{1}H NMR (600 MHz, Chloroform-*d*) δ 7.54 – 7.49 (m, 1H, H-5'), 6.99 – 6.92 (m, 2H, H-6', H-8'), 6.16 (t, *J* = 1.2 Hz, 1H, H-3'), 5.01 (d, *J* = 7.7 Hz, 1H, H-1), 3.93 – 3.78 (m, 2H, H-6a, H-6b), 3.66 (d, *J* = 0.9 Hz, 3H, H-OMe), 3.60 – 3.54 (m, 2H, H-3, H-4), 3.44 (dt, *J* = 9.3, 3.7 Hz, 1H, H-5), 3.26 (dd, *J* = 8.9, 7.3 Hz, 1H, H-2), 2.39 (d, *J* = 1.3 Hz, 3H, H-Me).

^{13}C NMR (151 MHz, $CDCl_3$) δ 161.54 (C-2'), 159.73, 154.88 (C-7'), 152.83, 125.89 (C-5'), 115.22 (C-4'), 113.53 (C-8'), 112.82 (C-3'), 104.11 (C-6'), 100.79 (C-1), 82.67 (C-2), 76.00 (C-5), 75.95 (C-3), 69.75 (C-4), 61.70 (C-6), 61.01 (C-OMe), 18.78 (C-Me).

HRMS (ESI$^+$) *m/z* [M+H]$^+$: Calc'd for $C_{17}H_{21}O_8$ 353.1236, found 353.1177.

4MU-3OMe-Glc

^{1}H NMR (600 MHz, Chloroform-*d*) δ 7.48 (d, *J* = 8.5 Hz, 1H, H-5'), 6.94 (d, *J* = 8.1 Hz, 2H, H-6', H-8'), 6.11 (s, 1H, H-3'), 4.92 (d, *J* = 7.7 Hz, 1H, H-1), 3.77 (m, 2H, H-6a, H-6b), 3.63 (s, 3H, H-OMe), 3.57 (m, 2H, H-2, H-4), 3.40 (dt, *J* = 9.8, 3.9 Hz, 1H, H-5), 3.19 (t, *J* = 9.1 Hz, 1H, H-3), 2.35 (s, 3H, H-Me).

^{13}C NMR (151 MHz, $CDCl_3$) δ 161.92 (C-2'), 160.12 (C-7'), 154.67, 153.26 (C-4'), 125.80 (C-5'), 115.04, 113.66 (C-8'), 112.45 (C-3'), 104.13 (C-6'), 100.44 (C-1), 85.70 (C-3), 76.22 (C-5), 73.02 (C-2), 69.47 (C-4), 61.40 (C-6), 60.81 (C-OMe), 18.65 (C-Me).

HRMS (ESI$^+$) *m/z* [M+H]$^+$: Calc'd for $C_{17}H_{21}O_8$ 353.1236, found 353.1187.

Di-methylated product

^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.53 (m, 1H, H-5'), 6.95 (m, 2H, H-6', H-8'), 6.18 (q, *J* = 1.2 Hz, 1H, H-3'), 5.03 (d, *J* = 7.6 Hz, 1H, H-1), 3.89 (ddd, *J* = 53.0, 12.0, 4.1 Hz, 2H, H-6a, H-6b), 3.68 (s, 3H, H-3OMe), 3.66 (s, 3H, H-2OMe), 3.63 (t, *J* = 9.3 Hz, 1H, H-4), 3.51 (m, 1H, H-5), 3.34 (dd, *J* = 9.1, 7.6 Hz, 1H, H-2), 3.25 (t, *J* = 9.0 Hz, 1H, H-3), 2.41 (d, *J* = 1.3 Hz, 3H, H-Me).

^{13}C NMR (126 MHz, $CDCl_3$) δ 161.11 (C-2'), 159.78 (C-7'), 155.08, 152.39, 125.90 (C-5'), 115.36 (C-4'), 113.38 (C-8'), 113.11 (C-3'), 104.31 (C-6'), 101.08 (C-1), 85.64 (C-3), 83.43 (C-2), 75.59 (C-5), 69.95 (C-4), 62.42 (C-6), 61.30 (C-3OMe), 60.62 (C-2OMe), 18.82 (C-Me).

HRMS (ESI$^+$) *m/z* [M+H]$^+$: Calc'd for $C_{18}H_{23}O_8$ 367.1393, found 367.1178.

Chapter 3. Enzyme kinetic analysis of β-glucosidase natural substrate analogs

3.1. Introduction

This chapter describes the results obtained during analysis of the synthetic substrates 1 to 9 (**Figure 2.1**), with GBA1, GBA2, and GBA3. First, the overview of fundamental enzyme kinetics is provided, followed by a brief description of the methods used in these measurements, followed by the results and discussion of these experiments, and finally, future directions that can be pursued based on the findings.

3.1.1. Enzymatic reactions

A common single substrate enzyme-catalyzed reaction follows the general scheme: $E + S \rightleftharpoons ES \rightarrow E + P$, consisting of two steps. In the first step, the substrate binds to the enzyme and forms the ES complex, and the second step involves chemical changes of the ES complex to form the product and free enzyme. In the case with the synthetic substrates with retaining β-glycosidases, which follows the Koshland double displacement mechanism, adopts the following scheme: $E + S \rightleftharpoons [ES^{\ddagger}] \rightleftharpoons E\text{–}S \rightleftharpoons [ES^{\ddagger}] \rightarrow E + P$. This scheme involves two consecutive inversions at the anomeric carbon ($\beta \rightarrow \alpha \rightarrow \beta$) and an E–S covalent glycosyl-enzyme intermediate flanked by the two ES‡ transition states. The formation of the ES‡ goes through Michaelis-Menten (M-M) ES complex. The formation of the M-M ES complex is very rapid and reversible, and at this step, the substrate is held at the active site of the enzyme by physical and/or electrostatic forces.

3.1.2. Steady-state assumption

Enzymatic reactions are traditionally measured under steady-state condition. Initially, when an enzyme is mixed with a much larger quantities of substrate, the reaction rate steadily increases with the increased concentration of substrate-enzyme bound complex, which is called pre-steady state. After the pre-steady state, the reaction reaches a steady state where the change in concentration of substrate-enzyme bound complex is negligible, and the reaction rate is assumed to be constant. Measuring

enzymatic activity in a pre-steady state gives information on the chemical mechanisms of the enzyme. When studying enzyme kinetics, one usually performs experiments under steady-state conditions since it mimics the dynamic steady-state of catalytic activities occurring in biological systems. It is under this steady-state assumption that which the M-M equation is valid.

3.1.3. Michaelis-Menten plot, equation and kinetic expressions

Two biochemists Leonor Michaelis and Maud Menten, published a paper in 1913 describing M-M kinetics. Since then, M-M kinetics has been the most commonly used model when studying enzyme metabolism. It explains the dependence of the rate of an enzyme changes with concentrations of enzyme and substrate and evaluates the kinetic characteristics of an enzyme regarding a substrate. M-M kinetics is measured at initial velocities (V_0) of the enzymatic reactions at different substrate concentrations. Measured at an initial rate of formation of ES complex includes minimal reversible reactions and excludes inhibition of an enzyme by the products. Moreover, V_0 corresponds to a known [S] where not an appreciable amount of substrate is depleted, thus keeping the quasi-steady-state valid. V_0 values obtained at various substrate concentrations are plotted to fit the M-M curve. The reaction rates are linearly increasing at low [S], then as [S] gets higher, the reaction rate slows down until the rate becomes zeroth-order in substrate concentration. At high [S], all the enzymes are saturated with substrates, and the rate reached its maximum velocity attainable (V_{max}). The value of V_{max} is dependent on enzyme concentration.

The M-M equation represents the rates of an enzymatic reaction (E + S ⇌ ES → E + P) in relation to enzyme and substrate concentrations.

$$V = \frac{[E]_0[S]k_{cat}}{K_m+[S]} \quad (\mathbf{3.1})$$

Kinetic expressions that describe the characteristics of the particular enzyme can be obtained from the M-M plot. Substrate concentration at which the rate is half of V_{max} is K_m. K_m is an important parameter when assessing an enzyme as it describes how fast the rate increases with an increase in [S], and is a measure of substrate's affinity to the enzyme: Lower K_m indicates higher protein affinity. K_m is independent of enzyme concentration, and its value is only valid for the particular enzyme with the particular

substrate at reaction conditions it is measured in. Several factors in reaction conditions affect the K_m value, such as pH, temperature, and ionic strengths of the buffer. The rate constant of the second step of a single-substrate enzymatic reaction, k_2, is also known as k_{cat}. k_{cat} is a first-order rate constant, based on the assumption that the enzyme is saturated with a substrate such that $[E]_0 = [ES]$ and the reaction is first order overall with respect to the enzyme but zero-order with respect to the substrate. It represents the number of substrate molecules converted into the products by a molecule of enzyme per unit time. Because it has a unit of s^{-1}, it is also called the turnover number. From the M-M curve, when substrate concentration is significantly lower than K_m ($[S] << K_m$), the M-M equation (**3.1**) is reduced to

$$V = \frac{k_{cat}}{K_m}[E]_0[S] \qquad (\mathbf{3.2})$$

Equation **3.2** represents the initial curve at low [S] on the M-M graph, and k_{cat}/K_m is the slope of the initial curve. At significantly low [S], enzymes are mostly unbound, therefore total enzyme concentration is roughly equal to initial enzyme concentration ($[E] \approx [E]_0$). k_{cat}/K_m is a second-order rate constant: first-order in $[E]_0$ and first-order in [S]. It is a useful value when evaluating the catalytic efficiency of an enzyme with various substrates, as the value combines the effectiveness of substrate binding to the enzyme plus the effectiveness of the transformation of the bound product.

3.2. Kinetic evaluation of the synthetic substrates

3.2.1. General experimental outline

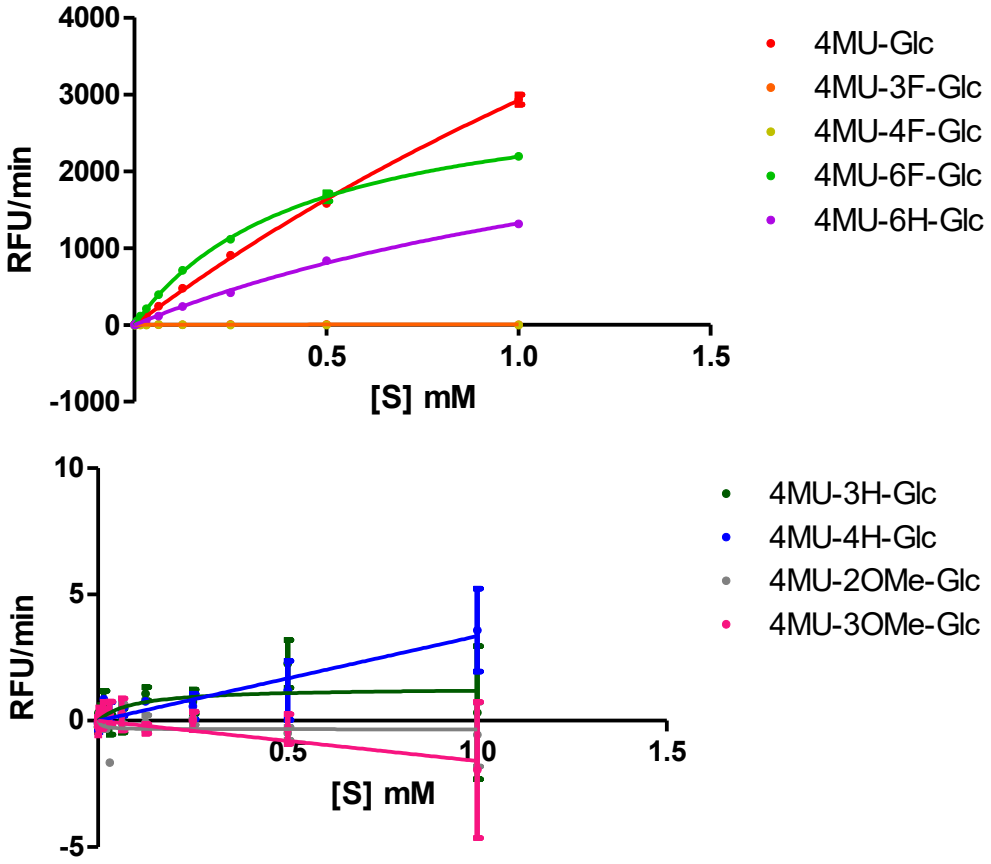
Each of the eight synthetic substrates prepared, compounds 2-9 (**Figure 2.1**) along with the parent compound, 4MU-Glc, as a reference, were tested against each glucocerebrosidase to find whether any compounds would be turned over more rapidly by non-lysosomal GBA2 as compared to GBA1 and GBA3. It would be ideal to find a substrate with significant selectivity for GBA2; however, a compound with modest selectivity would still be very useful as it provides a starting point for further development of more highly selective substrates for GBA2.

We expected that the enzyme would not turn over all substrates. Therefore, we first screened a panel of compounds to identify those each GBA enzyme can turn over

using either stopped kinetic assay or continuous assay to measure the extent of substrate hydrolysis (**Graph 3.1, 3.2**, **Table 3.1** - **3.3**). These assays provide the basis for a more detailed analysis of each substrate with each of these enzymes. Following this screen, we performed a more detailed analysis to determine kinetic values at the optimum *in vitro* buffer conditions for each of these enzymes, which differ in buffer composition, temperature, enzyme concentration, and plate reader settings. The data obtained for the candidate substrates were evaluated by plotting M-M curves and obtaining kinetic parameters for each of the three enzymes. Those substrates displaying inhibition were fitted substrate inhibition curves, and the inhibition parameters were obtained **(Graph 3.4**- **3.6**). The kinetic parameters obtained describe the catalytic efficiency of the compounds with each of the GBA enzyme as outlined below (**Table 3.4** - **3.6**). All enzymatic assay measurements were done in triplicates.

3.2.2. Initial screening to identify substrates each GBA enzyme can turn over

GBA1



Graph 3.1. Kinetic graphs of eight synthetic substrates along with 4MU-Glc tested against GBA1 shows a turnover of compounds

From the data, one can see that 4MU-6F-Glc and 4MU-6H-Glc are both turned over efficiently when using a standard concentration of GBA1 (2.5 nM), at rates comparable to the known substrate 4MU-Glc. These observations are consistent with previous studies showing that GBA1 can tolerate modifications at C6-position, whereas GBA2 and GBA3 are less tolerant of substitutions at this position.[20] The initial assays using the other substrates gave no significant rates; therefore, to determine kinetic parameters governing turnover, it required using high enzyme concentrations in the assays. Therefore, for this experiment, I performed the assay using 20-fold higher GBA1

concentration (50 nM) with higher substrate concentrations (500 μM) and performed the assay in a continuous manner to measure small rates accurately.

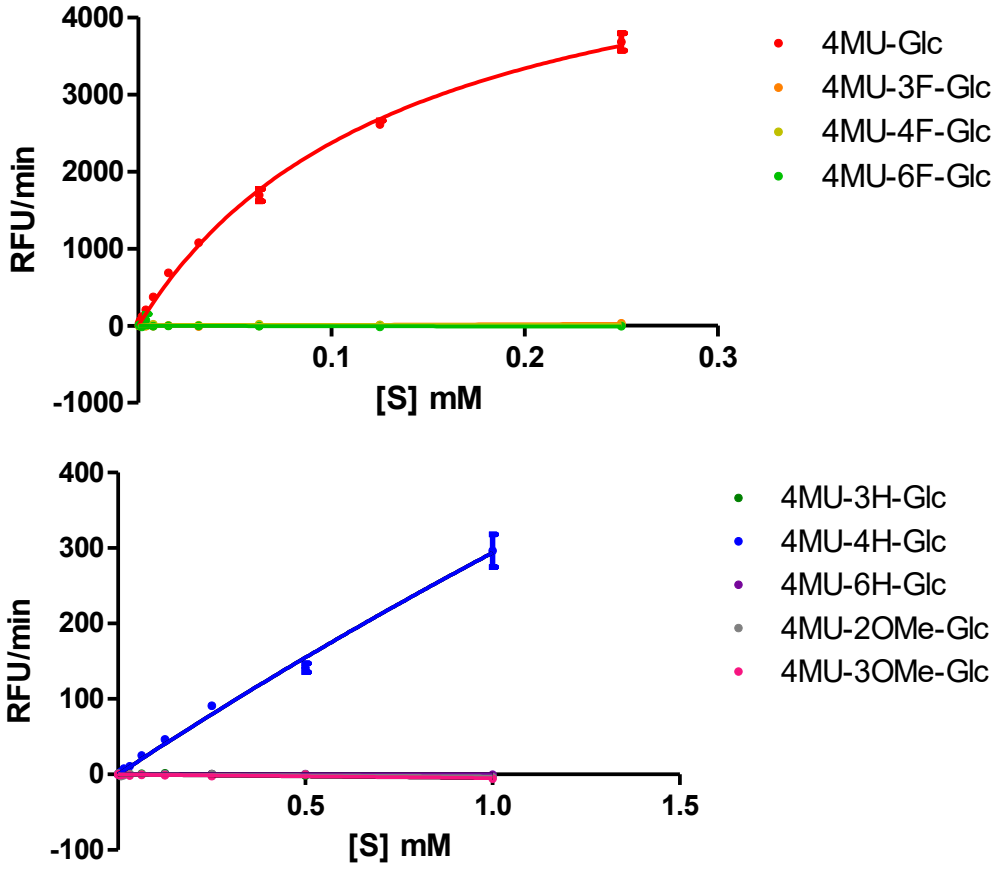
These data showed 4MU-4H-Glc was the best of these substrates, with 4MU-3F-Glc and 4MU-3H-Glc also being turned over but only more slowly. None of the other compounds were turned over at all, and though not tested, given the data, it is doubtful that 2,3-*O*-di-methylated compound would be turned over, and was therefore, not tested. The identified synthetic substrates turned over by GBA1 are summarized in **Table 3.1**.

Table 3.1. Selection of compounds GBA1 turned over

Compound	Turnover	Compound	Turnover	Compound	Turnover
4MU-3F-Glc	+	**4MU-3H-Glc**	+	**4MU-Glc**	+++
4MU-4F-Glc	-	**4MU-4H-Glc**	++	**4MU-2OMe-Glc**	-
4MU-6F-Glc	+++	**4MU-6H-Glc**	+++	**4MU-3OMe-Glc**	-

Note: Plus signs indicate the rate of hydrolysis: +++ for high, ++ for moderate, + for low. – indicates no detectable hydrolysis.

GBA2



Graph 3.2. Kinetic graphs of eight synthetic substrates along with 4MU-Glc tested against GBA2 shows a turnover of compounds

All eight synthetic substrates, along with 4MU-Glc, were tested in a stopped assay with a standard concentration of GBA2 (2.5 nM), except for 4MU-6H-Glc, which was tested with a higher enzyme concentration (10 nM). Only 4MU-4H-Glc was turned over under these conditions out of all synthetic substrates tested. The turnover of the other substrates could not be detected using even 20-fold higher enzyme concentrations (50 nM). The identified synthetic substrates that are turned over by GBA2 are summarized below in **Table 3.2**.

Table 3.2. Selection of synthetic compounds GBA2 turned over

Compound	Turnover	Compound	Turnover	Compound	Turnover
4MU-3F-Glc	-	4MU-3H-Glc	-	4MU-Glc	+++
4MU-4F-Glc	-	4MU-4H-Glc	++	4MU-2OMe-Glc	-
4MU-6F-Glc	-	4MU-6H-Glc	-	4MU-3OMe-Glc	-

Note: Plus signs indicate the rate of hydrolysis: +++ for high, ++ for moderate, + for low. – indicates no detectable hydrolysis.

GBA3

Finally, the eight synthetic substrates were incubated with GBA3 with final concentrations of 500 μM and 20 nM of the substrates and enzymes, respectively. The fluorescence signal was observed continuously for 1.5 hour at 25 ºC, and the graph of RFU vs time was plotted. From the graph, the relative rate of hydrolysis was determined with the slope at the initial, linear region of the reaction progress curve. Compound 4MU-4H-Glc was hydrolyzed the fastest, followed by substrates 4MU-6H-Glc and 4MU-6F-Glc with hydrolysis rates roughly 2.4-fold and 4.3-fold lower than compound 4MU-4H-Glc, respectively. Compound 4MU-4F-Glc was also turned over at a significantly lower rate as it was 164-fold slower than 4MU-4H-Glc. The results are summarized in **Table 3.3**. GBA3 turned over substrates with 4OH- and 6OH-group modifications. These results agree with the known fact on broad substrate specificity of GBA3 as it is reported to hydrolyze various β-glycosides including other β-glucosides and β-galactosides.[30]

Table 3.3. Selection of synthetic compounds that GBA3 turned over

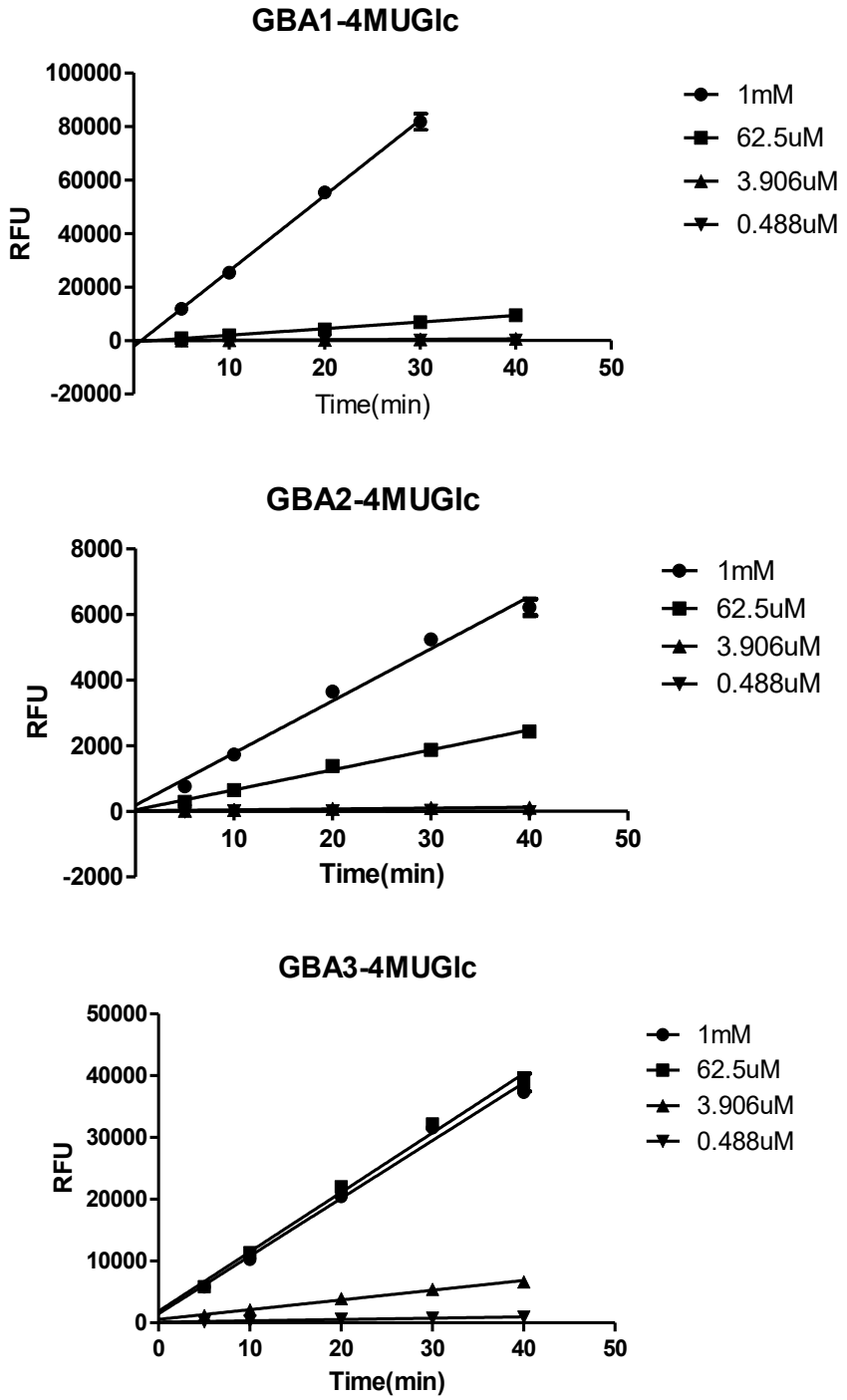
Compound	Turnover	Compound	Turnover	Compound	Turnover
4MU-3F-Glc	-	4MU-3H-Glc	-	4MU-Glc	+++
4MU-4F-Glc	+	4MU-4H-Glc	++	4MU-2OMe-Glc	-
4MU-6F-Glc	++	4MU-6H-Glc	++	4MU-3OMe-Glc	-

Note: Plus signs indicate the rate of hydrolysis: +++ for high, ++ for moderate, + for low. – indicates no detectable hydrolysis.

3.2.3. Test to find out stop time point for stopped assay

Based on these collective data, those compounds showing any turnover as summarized in **Table 3.1 - 3.3** were selected for assays with each enzyme. Next, the detailed kinetic parameters for each substrate with each enzyme were determined using

their optimal *in vitro* assay conditions. Notably, assays using substrates with 4-MU as a fluorescence reporter group usually are done in a stopped assay format, where a base is used to terminate the reaction.[67] For GBA1, GBA2, and GBA3, the buffers used were pH 5.2, pH 5.5, and pH 6.5, respectively. Because the pK_a of 4-MU is around 7.8 and the pH values of the buffers are lower than the pK_a value of 4-MU, only a fraction of 4-MU will exist as the deprotonated fluorescent phenoxide form. Though continuous assays could be performed, performing them in a stopped assay format maximized the signal intensity and simplified comparing rate data since there was no need to correct the extent of ionization of 4-MU at each pH value in the different buffers. Therefore, the M-M kinetic experiments of the selected substrates against each GBA enzyme were done as stopped assays. The time the reaction needs to be stopped should be a time point that lies within the initial part of the progress curve that is still linear. Therefore, a preliminary assay was performed using four different concentrations of 4MU-Glc (1 mM, 62.5 μM, 3.906 μM, 0.488 μM) tested against 2.5 nM of each GBA1, GBA2, and GBA3. These preliminary experiments were carried out and stopped at different time points to determine the linear range of the assay with respect to time.



Graph 3.3. Graph of increasing RFU over time of 4MU-Glc with each GBA enzyme at four different substrate concentrations. All measurements were done in triplicates.

All of the data points on **Graph 3.3** of all GBA enzymes fit linear increases in fluorescence up until the 40 min time point with R^2 values greater than 0.94. The fits

obtained were better at higher substrate concentrations and worse at lower concentrations due to larger errors associated with lower signal intensities. On the graph of GBA1 activity, at 1 mM concentration of 4MU-Glc, the 40 min time point exceeded the maximum signal detection intensity for the plate reader, so the value was excluded. However, the 20 min time point lies well within the initial linear region for all three GBA enzymes. Therefore, an assay time of 20 minutes was used for measuring the initial velocities.

It was observed in the above studies that most of the synthetic substrates were hydrolyzed more slowly than 4MU-Glc except for 4MU-6F-Glc, which only had about twice the faster initial rate with GBA1 (**Graph 3.1**). Therefore, it was reasoned that the initial period over which 4MU-Glc showed linear increases in fluorescence would also be linear for the other substrates. Therefore, all of the assays with all substrates were also terminated at 20 min after initiation of the enzyme.

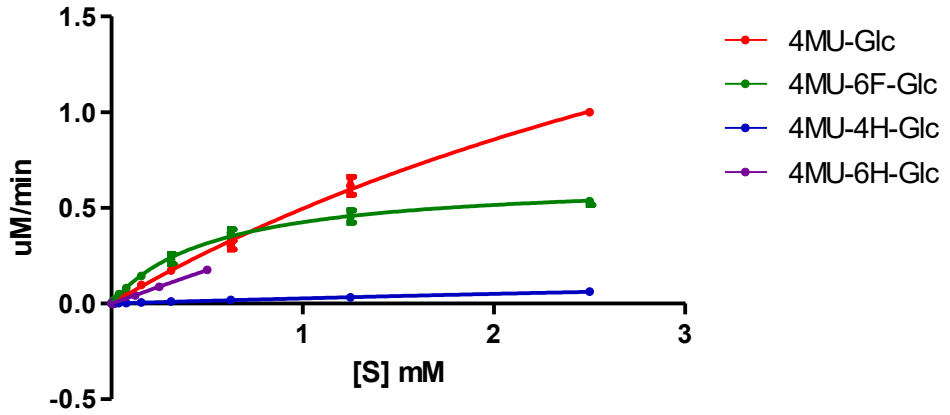
3.2.4. Michaelis-Menten plots of the candidate compounds with each GBA enzyme

The selected substrates (**Table 3.1** - **3.3**) were tested with each GBA enzyme to produce M-M plots. Using the M-M plots, kinetic parameters were obtained, which allowed the comparison of the catalytic efficiency of the substrates. Second-order rate constants (k_{cat}/K_m) were obtained for all compounds tested. For most synthetic substrates, however, we were unable to obtain parameters V_{max}, K_m, and k_{cat} due to compounds being of poor solubility at higher concentrations. Where possible, V_{max} and K_m values were obtained from the fitted curves shown on the graph, and k_{cat} was calculated by dividing V_{max} by the concentration of enzyme used in the assay. As mentioned above, GBA2 was observed in studies to process β-galactosides and GBA3 is also known to turn over β-galactosides as well as other β-glycosides.[28,94,11] To confirm these reports and to further explore their substrate tolerances, 4-methylumbelliferyl β-D-galactoside (4MU-Gal) was tested with all GBA enzymes.

GBA1

The M-M curve with 4MU-6F-Glc, 4MU-4H-Glc, 4MU-6H-Glc, and the parent compound 4MU-Glc against GBA1, were plotted (**Graph 3.4**). No hydrolysis of 4MU-Gal

was seen with GBA1. 4MU-3F-Glc and 4MU-3H-Glc, which were hydrolyzed more slowly than the other substrates, were subjected to detailed analysis to determine M-M kinetics. Because these were not hydrolyzed by other GBA enzymes, including GBA2, this was judged as not being required.



Graph 3.4. Michaelis-Menten graphs of 4MU-Glc, 4MU-6F-Glc, 4MU-4H-Glc, and 4MU-6H-Glc against GBA1

Table 3.4. Reported kinetic parameters of the compounds with GBA1

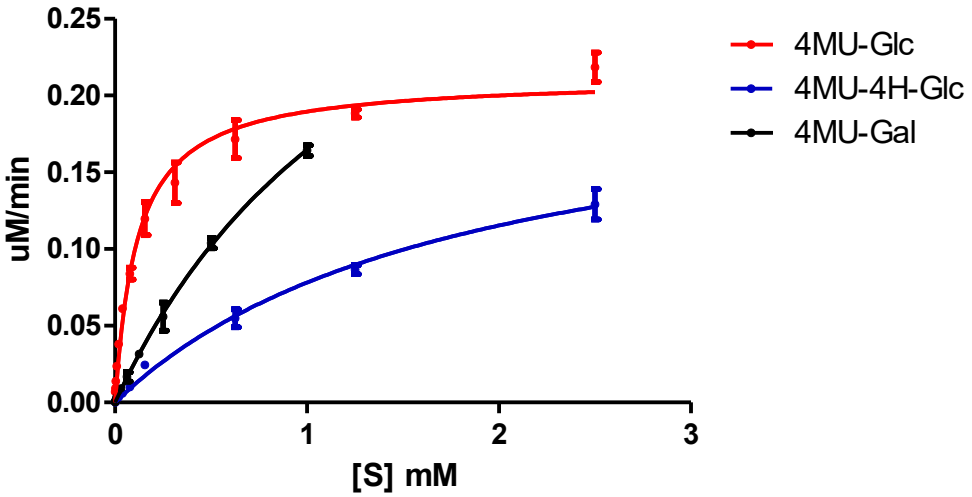
GBA1	K_m (mM)	k_{cat} (min^{-1})	k_{cat}/K_m (mM^{-1} min^{-1})
4MU-Glc	n.d.	n.d.	0.61 ± 0.01
4MU-6F-Glc	0.535	261	1.22 ± 0.04
4MU-4H-Glc	n.d.	n.d.	0.0245 ± 0.0006
4MU-6H-Glc	n.d.	n.d.	0.350 ± 0.006

Concentrations of substrate 4MU-6F-Glc reached saturation, enabling the fitting to the M-M curve (**Graph 3.4**) The V_{max}, K_m, and k_{cat} values of 4MU-Glc, 4MU-4H-Glc and 4MU-6H-Glc were unable to be obtained due to saturation of the compounds at higher concentrations. However, k_{cat}/K_m values for all compounds tested above were determined, allowing direct comparisons of the catalytic efficiencies of these substrates.

GBA1 is known to tolerate modification of the 6-hydroxyl group.[42] Interestingly, the second-order rate constant for 4MU-6F-Glc was the highest, with its value two-fold higher than the parent compound, 4MU-Glc, whereas the second-order rate for 4MU-6H-Glc was two-fold slower than 4MU-Glc. This result emphasizes the importance of the

hydrogen bonding at the 6-hydroxyl of the substrate in catalysis. Although 4MU-6F-Glc is processed faster than the parent substrate, it can be predicted from inspection of the M-M plots (**Graph 3.4**) that the k_{cat} value is significantly lower than that of 4MU-Glc, which indicates GBA1 can more efficiently processes 4MU-Glc once bound to the enzyme. The K_m value for 4MU-6F-Glc is expected to be lower than that of 4MU-Glc by the trend of the M-M curve, which suggests that 4MU-6F-Glc has a higher affinity for the enzyme than 4MU-Glc. The second-order rate constant for turnover of 4MU-4H-Glc is 25-times slower than the value determined for the parent substrate, 4MU-Glc.

GBA2



Graph 3.5. Michaelis-Menten graphs of 4MU-Glc, 4MU-4H-Glc, and 4MU-Gal against GBA2

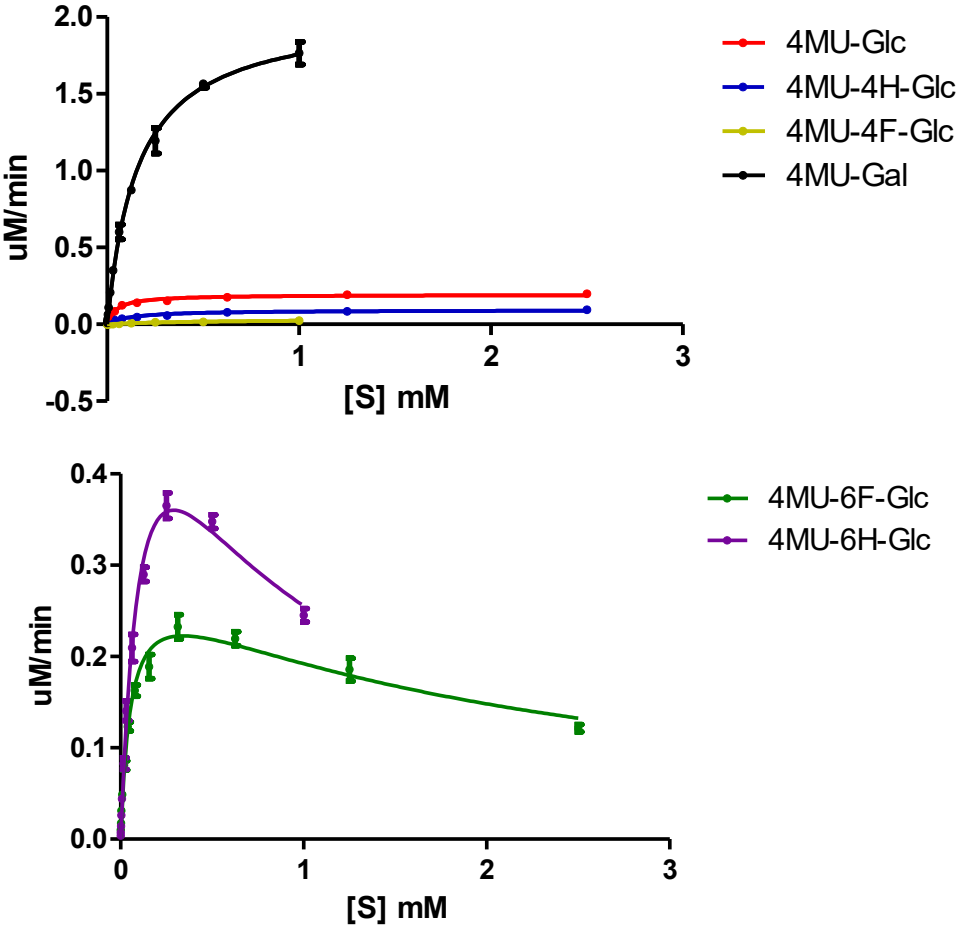
Table 3.5. Reported kinetic parameters of the compounds with GBA2

GBA2	K_m (mM)	k_{cat} (min^{-1})	k_{cat}/K_m (mM^{-1} min^{-1})
4MU-Glc	0.117	84.7	1.44 ± 0.05
4MU-4H-Glc	n.d.	n.d.	0.145 ± 0.007
4MU-Gal	n.d.	n.d.	0.31 ± 0.01

Out of all the synthetic substrates, only 4MU-4H-Glc was hydrolyzed by GBA2. Although the H-bond interaction diagrams (**Figure 1.4, 1.5**) show GBA1 and GBA2 having a similar degree of involvement of H-bond interaction at the active site, above

data reports that GBA2 is less tolerating than GBA1 on substrate OH-group modifications. Due to the inability to observe saturation kinetics from using high concentrations of 4MU-4H-Glc, K_m and k_{cat} values for this substrate were could not be obtained. The second-order rate constant for 4MU-4H-Glc was around 10-fold slower than the value measured for 4MU-Glc. As in some studies that suggest that GBA2 also can cleave β-galactosides, 4MU-Gal tested against GBA2 shows its second-order rate constant about 4.5-times slower than 4MU-Glc, and 2-times greater than 4MU-4H-Glc.[28,94] By inspection of the M-M curve for 4MU-Gal, k_{cat} and K_m values of 4MU-Gal are predicted to be larger than 4MU-Glc, indicating lower affinity of 4MU-Gal with GBA2. However, significant hydrolysis of 4MU-Gal, along with studies observing β-galactoside hydrolysis by GBA2, suggests that GBA2 might also be involved in the degradation of β-galactosides within cells.

GBA3



Graph 3.6. TOP: Michaelis-Menten graphs of 4MU-Glc, 4MU-4H-Glc, 4MU-4F-Glc, 4MU-6F-Glc, 4MU-6H-Glc, and 4MU-Gal with GBA3. BOTTOM: 4MU-6F-Glc and 4MU-6H-Glc show substrate inhibition curves with K_i values of 2.06 mM and 0.649 mM, respectively.

Table 3.6. Reported kinetic parameters of the compounds with GBA3

GBA3	K_m (mM)	k_{cat} (min^{-1})	k_{cat}/K_m (mM^{-1} min^{-1})
4MU-Glc	0.0485	77.2	4.4 ± 0.1
4MU-4F-Glc	n.d.	n.d.	0.052 ± 0.003
4MU-6F-Glc	0.0562	29.6	5.7 ± 0.3
4MU-4H-Glc	0.122	37.1	1.09 ± 0.07
4MU-6H-Glc	0.130	137	6.5 ± 0.2
4MU-Gal	0.160	408	15.5 ± 0.9

The initial screen showed that GBA3 hydrolyzed only the C4 and C6 modified compounds among the synthetic substrates. Surprisingly, C6-modified substrates displayed higher k_{cat} and K_m values, and a higher k_{cat}/K_m value than the parent substrate 4MU-Glc. Moreover, the M-M curves obtained for the substrates 4MU-6F-Glc and 4MU-6H-Glc revealed substrate inhibition.

Substrate inhibition of the kind seen here is observed when higher substrate concentrations lead to decrease in velocity of the enzymatic reaction rather than asymptotically reaching saturation. The equation for substrate inhibition model is shown in equation (**3.3**). At low substrate concentrations, the term $[S]/K_i$ reaches zero, and the early parts of the substrate inhibition curve are determined by K_m. At high substrate concentrations, the $K_m/[S]$ value approaches zero, and the later parts of the inhibition curve are determined by K_i.

$$V = \frac{k_{cat}\,[E]_0}{\frac{K_m}{[S]} + 1 + \frac{[S]}{K_i}} \qquad (\mathbf{3.3})$$

Substrate inhibition is a phenomenon usually observed when high concentrations of a substrate are used in enzymatic assays and in some cases with natural substrates involved in regulating metabolism in living organisms.[95] In other cases, substrate inhibition can arise when the observed rate is an aggregate of hydrolysis and transglycosylation to another saccharide that binds adjacent to the active site. When the substrate is modified to remove a hydroxyl group, the transglycosylation can no longer occur, leading to apparent substrate inhibition.[96] However, the cause for the substrate inhibition cannot be distinguished between these alternatives at this time. Compound 4MU-4H-Glc displays 2-fold lower k_{cat} and 2.5-fold higher K_m as compared to 4MU-Glc, and its second-order rate was only 4-fold slower than the parent substrate 4MU-Glc. 4MU-4F-Glc has an 85-fold lower k_{cat}/K_m value relative to 4MU-Glc.

Although most of the studies suggest that GBA3 may be acting as β-glucosidase in cells, many studies show that GBA3 has broad substrate specificity including β-galactosidase activity and reports on conflicting data of whether GBA3 possess higher β-galactosidase or β-glucosidase activity.[30,11,97] When 4MU-Gal was tested against GBA3,

surprisingly, my result showed that 4MU-Gal was processed more efficiently than 4MU-Glc, with k_{cat}/K_m 3.5-times higher, k_{cat} value 5-times higher, and K_m 3-times higher than 4MU-Glc. These results agree with a precedented study by Daniels et al., where they observed higher β-galactosidase activity than β-glucosidase activity of GBA3.[97] These results suggest that the primary role of GBA3 in biological systems may not be that of a β-glucosidase but rather a β-galactosidase. Clarifying this point would require further investigation to determine its primary function in living cells.

Table 3.7. Summary of rate constants of selected compounds with GBA1, GBA2, and GBA3

Substrate	GBA1 k_{cat}/K_m (mM^{-1} min^{-1})	GBA2 k_{cat}/K_m (mM^{-1} min^{-1})	GBA3 k_{cat}/K_m (mM^{-1} min^{-1})
4MU-Glc	0.61 ± 0.01	1.44 ± 0.05	4.4 ± 0.1
4MU-4F-Glc	N/A	N/A	0.052 ± 0.003
4MU-6F-Glc	1.22 ± 0.04	N/A	5.7 ± 0.3
4MU-4H-Glc	0.0245 ± 0.0006	0.145 ± 0.007	1.09 ± 0.07
4MU-6H-Glc	0.350 ± 0.006	N/A	6.5 ± 0.2
4MU-Gal	N/A	0.31 ± 0.01	15.5 ± 0.9

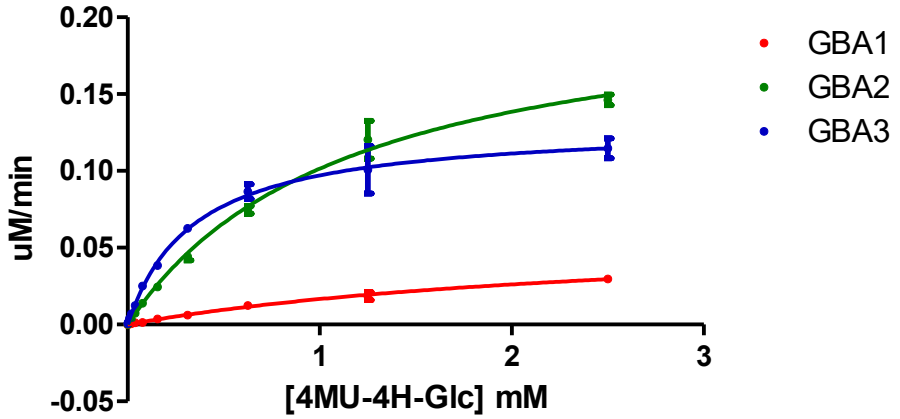
Note: N/A denotes not applicable

The second-order rate describes the catalytic efficiency of the enzymatic reaction on going from the free enzyme and substrate to the transition state of the first irreversible step of the overall reaction. The second-order rates of the analyzed substrates with three GBA enzymes are summarized in **Table 3.7**. C6 modified substrates were tolerated by both GBA1 and GBA3 but not by GBA2, and it is supported by H-bond interactions of compounds with the active site residues of each GBA enzyme. GBA1 shows C6 hydroxyl group having only one H-bond interaction with Asn 396, and that of GBA3 has no H-bonding interaction, whereas C6 of GBA2 has two H-bond with Arg 786 and one H-bond with Glu 777 (**Figure 1.4**, **1.5**).[14] As mentioned earlier, out of the modified synthetic substrates, only 4MU-4H-Glc is turned over by GBA2. The catalytic efficiency for turnover of 4MU-4H-Glc with GBA2 is 7.5-times lower than with GBA3 and around 6-times higher than with GBA1. Notably, both GBA1 and GBA2 are abundantly expressed within brain tissue, whereas GBA3 levels are negligible.[97] Based on these considerations, it seems that 4MU-4H-Glc could have the potential to serve as

a selective substrate for GBA2 for use with brain tissues where GBA3 is not abundant. Therefore, the relative processing of 4MU-4H-Glc was further considered as a standard and commonly used substrate used to measure GBA2 activity.

3.2.5. Test of 4MU-4H-Glc with each GBA enzyme in GBA2 buffer

A GBA2 cell lysate study using 4MU-4H-Glc would be done using optimal buffer for GBA2 activity. The assays described above were done using buffer systems that are optimal for each enzyme. Therefore, 4MU-4H-Glc was tested as a substrate for GBA1, GBA2, and GBA3 using the optimal buffer for GBA2 to see how the selectivity and kinetic parameters of compound 4MU-4H-Glc would be affected.



Graph 3.7. Michaelis-Menten graph of 4MU-4H-Glc against each GBA enzyme

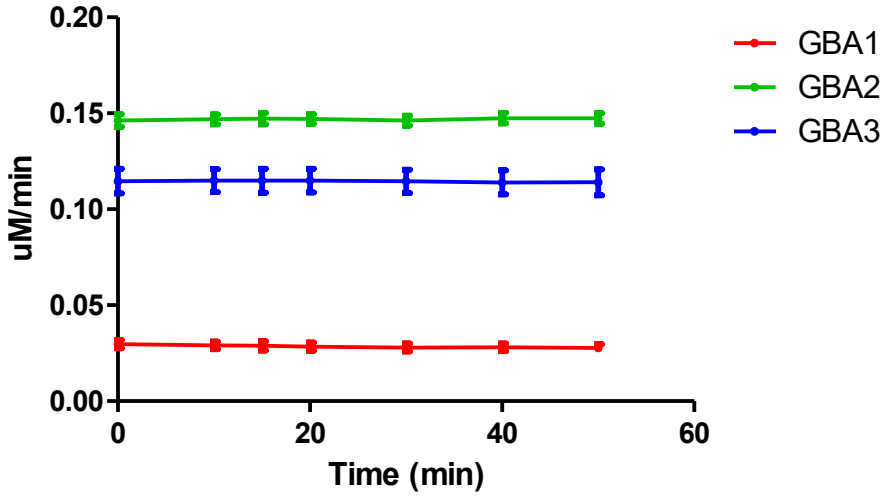
Table 3.8. Second-order rates of 4MU-4H-Glc with each GBA enzyme

Substrate	GBA1 k_{cat}/K_m (mM^{-1} min^{-1})	GBA2 k_{cat}/K_m (mM^{-1} min^{-1})	GBA3 k_{cat}/K_m (mM^{-1} min^{-1})
4MU-4H-Glc	0.020 ± 0.001	0.159 ± 0.004	0.32 ± 0.01

Table 3.8 shows the second-order rate constants for 4MU-4H-Glc tested against each GBA enzyme in the GBA2 buffer. These values were compared with the second-order rates of 4MU-4H-Glc with each GBA enzyme in their corresponding optimal buffers (**Table 3.7**) The k_{cat}/K_m values for 4MU-4H-Glc with GBA1 in GBA1 buffer and GBA2 buffer are quite close to each other with only 0.005 mM^{-1} min^{-1} difference since pH of GBA1 buffer (pH 5.2) and GBA2 buffer (pH 5.5) are pretty close. The k_{cat}/K_m value

determined for turnover of 4MU-4H-Glc by GBA3 in GBA2 buffer is 3.4-fold lower than in its optimal buffer (pH 6.5).

4MU-4H-Glc tested with each GBA enzyme using GBA2 buffer (pH 5.5) showed the highest second-order rate constant with GBA3 with the rate 2-fold higher than by GBA2. The catalytic efficiency of 4MU-4H-Glc by GBA1 was 8-times lower than by GBA2. Although 4MU-4H-Glc exhibits significant turnover by GBA3, 4MU-4H-Glc was not discouraged from its potential as a GBA2 selective substrate in brain tissues since there is a negligible GBA3 present. 4MU-4H-Glc shows 8-times higher selectivity towards GBA2 over GBA1 is a good indicator for 4MU-4H-Glc as a potential GBA2 selective substrate.

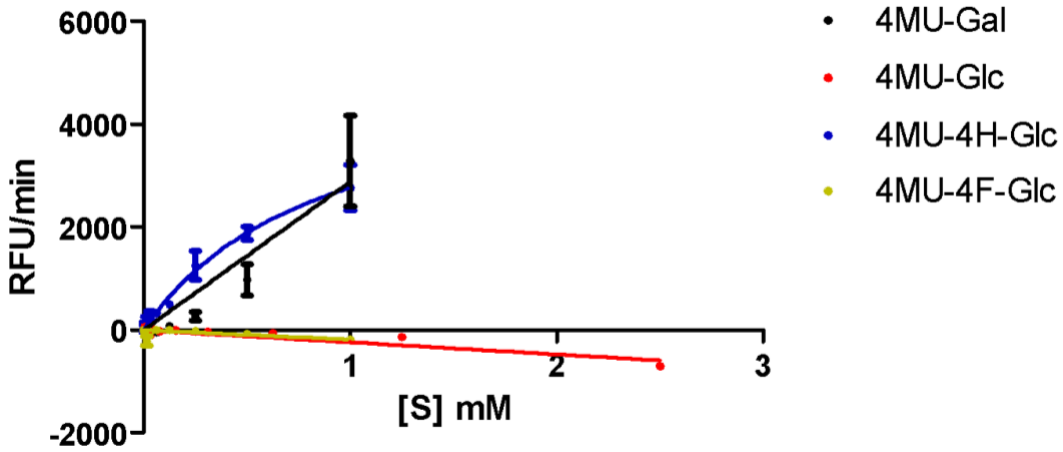


Graph 3.8. No change in fluorescence was observed at 10, 15, 20, 30, 40, and 50 minutes after the stop solution was added (0 min) to the reaction. 2.5 mM of 4MU-4H-Glc was used with each GBA enzyme.

The use of 1.0 M Tris buffer (pH 9.0) to terminate the reaction was based on a study by Deen et al..[20] Nevertheless, to confirm the suitability of this stop buffer, further experiment was carried out to ensure that the fluorescent signal was stable in the stopped reactions. To do this, enzyme reactions using 4MU-4H-Glc with GBA1, GBA2, and GBA3 in GBA2 buffer was carried out and then terminated with a stop buffer. The resulting stopped reactions were then monitored on a plate reader at 10, 15, 20, 30, 40, and 50 minutes after the reaction was stopped (0 minute). **Graph 3.8** showed the fluorescence signal did not increase after the stop solution was added, indicating that the enzyme activities for each of the three enzymes under these conditions were negligible.

3.3. Prospects of 4MU-4H-Glc as a potential GBA2 selective probe for brain cell lysate study

The *in vitro* studies to assess the selectivity of 4MU-4H-Glc with the three GBA enzymes reveal a preference for GBA2 in brain tissues where GBA3 is in negligible amount. However, if 4MU-4H-Glc is to be developed further as a cell lysate probe, other enzymes present in cell lysate need to be considered. One complicating factor is that substrate 4MU-4H-Glc is not only a 4-deoxy analog of glucose but also an analog for galactose as well. GalC is a retaining β-galactosylceramidase located in lysosomes and largely expressed in brain tissue and many other types of tissues in the human body. If GalC can turn over 4MU-4H-Glc efficiently, this could present an obstacle in the targeted and selective measurement of GBA2 activity in cell lysates. Though GalC is a lysosomal enzyme and is, therefore more active in acidic conditions, it was judged that this issue should be addressed. To test this possibility, I tested 4MU-4H-Glc as a substrate for human GalC enzyme *in vitro* using the GBA2 optimal buffer conditions (pH 5.5). The assay was done as a continuous assay with 4MU-Gal and 4MU-Glc for comparisons. 4MU-4F-Glc was also tested to see if GalC would tolerate a substrate with a fluorine substitution at C4-position with a different stereochemistry.



Graph 3.9. The kinetic graph of substrates tested against GalC shows a turnover of 4MU-4H-Glc.

As expected, compounds 4MU-Glc and 4MU-4F-Glc were not hydrolyzed by GalC. Unfortunately, as it is seen on **Graph 3.9**, 4MU-4H-Glc is processed by GalC with a rate roughly comparable to 4MU-Gal. Therefore, for when considering using 4MU-4H-

Glc as a GBA2 selective substrate for brain cell lysate assay, an inhibitor of GalC needs to be added to the assay to remove GalC activity from being measured.

GBA2 has higher optimal pH (pH 5.5 -6.0)[24] than that of GalC (pH 4.0 – 4.3) [98] and GBA1 (pH 4.0 – 5.5).[11] An assay of 4MU-4H-Glc in brain cell lysate could be done in GBA2 buffer at pH 6.0, which could decrease unwanted enzyme activities and increase selectivity towards GBA2.

3.4. Methods

3.4.1. Materials

Recombinant human enzymes GBA1, GBA2 and GBA3, were purchased from R&D systems. Buffers were prepared for each GBA assay as listed in the table below.

Table 3.9. Buffers used for GBA1, GBA2, and GBA3 enzymatic assays

Enzyme	Buffers used
GBA1	McIlvaine buffer pH 5.2, with 0.1%TritonX-100, 0.25% w/w sodium taurodeoxycholate hydrate
GBA2	McIlvaine buffer pH 5.5, with 1% C10E6
GBA3	0.1 M MES buffer, pH 6.5

For termination of the reaction, 1.0 M Tris solution (pH 9.0) was used as a stop solution. The fluorescence of all assays was measured using BioTek Neo 2 Plate reader with excitation and emission wavelengths of 355 nm and 450 nm, respectively. The data were analyzed using GraphPad Prism 5 software.

3.4.2. Experimental procedures

Screening of compounds that GBA enzymes can turn over using stopped assay

To a 96-well mixing plate containing 50 μL solution of substrates (Compounds 1-4, 6, 7) for GBA1 and the substrates (Compounds 1-9) for GBA2, 50 μL of GBA1 (5 nM) and GBA2 (5 nM) were added to corresponding experimental wells containing substrates, while control wells received corresponding buffers without the enzymes. Final concentrations of substrates were in 2-fold dilutions from 1 mM to 0.488 μM with a

final enzyme concentration of 2.5 nM. The reaction plates were left at room temperature, and after 20 mins, 100 μL of stop solution was added to every well to terminate the reaction. The solutions were aliquoted onto a black 384-well Nunc flat bottom plate in 45 μL triplicates, and fluorescence was measured using BioTek Neo 2 Plate reader. Kinetic graphs were plotted using GraphPad Prism 5 software.

Screening of compounds that GBA enzymes can turn over using continuous assay

In a 96-well mixing plate, 75 μL aliquots of 1 mM substrates (Compounds 2, 3, 5, 6, 8, 9) for GBA1 and 1 mM of substrates (Compounds 2-5, 8, 9) for GBA2 were added to rows 1, 2 (control), and rows 3, 4 (control), respectively. Then, to the control wells, 75 μL aliquots of GBA1 buffer and GBA2 buffer were added to row 2 and row 4, respectively. For experimental wells, 75 μL of GBA1 (100 nM) and GBA2 (100 nM) were added to row 1 and row 3, respectively. The final substrate and enzyme concentrations were 500 μM and 50 nM, respectively, with 0.5% DMSO. The reaction solutions were aliquoted in 45 μL duplicates onto a black 384-well Nunc flat bottom plate, and fluorescence was measured using BioTek Neo 2 Plate reader continuously for 1.5 hour at 25 °C. The background hydrolysis signals of the controls were subtracted from the raw experimental signals. Graphs of RFU vs time were obtained, and the change in RFU corresponding to each compound determined the substrate hydrolysis.

Compound screening for GBA3 was done similarly to above. 30 μL of 1 mM substrate solutions (Compounds 2-9) were added in two aliquots into row 1 and row 2 (control). 30 μL of GBA3 enzyme solutions (40 nM) were added to all experimental wells in row 1, and 30 μL of GBA3 buffers were added to all control wells in row 2. Final concentrations of substrate and enzyme were 500 μM and 20 nM, respectively, with 0.5% DMSO.

Determination of the reaction termination time for stopped assays

In four centrifuge tubes, each containing 300 μL of different concentration of 4MU-Glc, 300 μL of GBA1 solutions were added to each tube, giving final substrate concentrations of 1 mM, 62.5 μM, 3.906 μM, 0.488 μM, and enzyme concentration of 2.5 nM with 1% DMSO. After 5, 10, 20, 30, and 40 minutes of reaction, 100 μL of the

reaction mixture in each tube were transferred into a new centrifuge tube containing 100 μL of stop solution. Immediately after the reaction was terminated, the solution was transferred onto a black 384-well Nunc flat bottom plate in 45 μL triplicates, and the fluorescence signal was measured using BioTek Neo 2 Plate reader. The control samples were performed in the same manner except for GBA1 buffer without the enzyme to obtain background hydrolysis signals. The raw experimental signals were background corrected by subtracting background hydrolysis signals. The fluorescence signals were divided by the time course of the corresponding reactions to get RFU per min, then the graph of RFU vs time was plotted using GraphPad Prism 5 software. 4MU-Glc was assayed with GBA2 and GBA3 and graphed in the same fashion.

Michaelis-Menten kinetics of the selected compounds

To a 96-well mixing plate containing 50 μL of compounds 4MU-Glc, 4MU-6F-Glc, and 4MU-4H-Glc distributed in twelve 2-fold dilutions, 50 μL of GBA1 solution was added to all wells containing substrates. This yielded the final highest substrate concentrations of 2.5 mM and final enzyme concentration of 2.5 nM with 1% DMSO. To the control wells, GBA1 buffer without the enzyme was added. The reaction mixtures were incubated at 27 °C for 20 mins, then aliquoted in 45 μL triplicates onto a black 384-well Nunc flat bottom plate. The fluorescence signal was measured using BioTek Neo 2 Plate reader. The standard curve was produced using 4-methylumbelliferone, with twelve 3-fold dilutions with from its highest concentration of 25 μM. The raw fluorescence signals were background corrected by subtracting the background hydrolysis signals measured from the controls. The fluorescence signals were divided by 20, then by slope of the standard curve (RFU/μM) to plot the Michaelis-Menten graph of μM/min vs [S]. 4MU-6H-Glc and 4MU-Gal were tested with GBA1 in the same manner, using different concentrations; 4MU-6H-Glc used the highest substrate concentration of 0.5 mM with 2.5 nM GBA1, and 4MU-Gal used the highest substrate concentration of 1 mM with 10 nM GBA1.

4MU-Glc and 4MU-6H-Glc with substrate concentrations of 2.5 mM were tested with 2.5 nM GBA2, and 4MU-6H-Glc and 4MU-Gal with substrate dilutions from 1 mM were tested with 10 nM GBA2. All substrates with GBA2 were assayed and graphed in the same manner as above.

4MU-Glc and 4MU-4H-Glc, with substrate dilutions from 2.5 mM were tested with 2.5 nM GBA3, and 4MU-6F-Glc with substrate dilutions from 2.5 mM was tested with 10 nM GBA3. 4MU-4F-Glc with substrate dilutions from 1 mM was tested with 20 nM GBA3. 4MU-6H-Glc and 4MU-Gal with substrate dilutions from 1 mM were tested with 5 nM GBA3. All substrates with GBA3 were assayed and graphed in the same manner as above.

Michaelis-Menten kinetics of 4MU-4H-Glc with GBA enzymes in GBA2 buffer

In a 96-well mixing plate, 50 μL of 4MU-4H-Glc was mixed with 50 μL of each GBA1, GBA2, and GBA3. The substrate 4MU-4H-Glc were distributed in twelve 2-fold dilutions with the highest concentration of 2.5 mM. Enzyme concentrations used were 50 nM for GBA1, and 2.5 nM for GBA2 and GBA3 with 1% DMSO. The reaction mixtures were incubated at 27 °C for 20 mins, then transferred onto a black 384-well Nunc flat bottom plate in 45 μL triplicates. Fluorescence was measured using BioTek Neo 2 Plate reader. The fluorescence of the samples was measured 10, 15, 20, 30, 40, and 50 mins after the stop solution was added. A standard curve was produced with 4-methylumbelliferone. The fluorescence of the raw experimental signals was background corrected by subtracting the fluorescence of controls. The fluorescence values were divided by 20, then by the slope of the standard curve to produce a M-M plot of μM/min vs [S].

References

1. Schnaar, R. L. Glycosphingolipids in cell surface recognition. *Glycobiology* **1**, 477–485 (1991).

2. Bansode, S. *et al.* Glycation changes molecular organization and charge distribution in type I collagen fibrils. *Sci. Rep.* **10**, 3397 (2020).

3. Prescher, J. A. & Bertozzi, C. R. Chemical technologies for probing glycans. *Cell* **126**, 851–854 (2006).

4. Zhu, Y., Zajicek, J. & Serianni, A. S. Acyclic forms of [1-^{13}C]aldohexoses in aqueous solution: quantitation by ^{13}C NMR and deuterium isotope effects on tautomeric equilibria. *J. Org. Chem.* **66**, 6244–6251 (2001).

5. HERS, H. G. α-Glucosidase deficiency in generalized glycogen-storage disease (Pompe's disease). *Biochem. J.* **86**, 11–16 (1963).

6. Huang, W. J., Zhang, X. & Chen, W. W. Gaucher disease: A lysosomal neurodegenerative disorder. *Eur. Rev. Med. Pharmacol. Sci.* **19**, 1219–1226 (2015).

7. Sardi, S. P. *et al.* Augmenting CNS glucocerebrosidase activity as a therapeutic strategy for parkinsonism and other Gaucher-related synucleinopathies. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 3537–3542 (2013).

8. D'Angelo, G., Capasso, S., Sticco, L. & Russo, D. Glycosphingolipids: Synthesis and functions. *FEBS J.* **280**, 6338–6353 (2013).

9. Bektas, M. & Spiegel, S. Glycosphingolipids and cell death. *Glycoconj. J.* **20**, 39–47 (2003).

10. Körschen, H. G. *et al.* The non-lysosomal β-glucosidase GBA2 is a non-integral membrane-associated protein at the endoplasmic reticulum (ER) and Golgi. *J. Biol. Chem.* **288**, 3381–3393 (2013).

11. Hayashi, Y. & Ito, M. Klotho-related protein KLrP: Structure and functions. *Vitam. Horm.* **101**, 1–16 (2016).

12. Fredj Ben, B., Artola, M., Overkleeft, H. S., Ubbink, M. & Johannes, M. F. G. A. Distinguishing the differences in β-glycosylceramidase folds, dynamics, and actions informs therapeutic uses. *J. Lipid Res.* **59**, 2262–2276 (2018).

13. Do, J., McKinney, C., Sharma, P. & Sidransky, E. Glucocerebrosidase and its relevance to Parkinson disease. *Mol. Neurodegener.* **14**, 1–16 (2019).

14. Hayashi, Y. *et al.* Klotho-related protein is a novel cytosolic neutral β-glycosylceramidase. *J. Biol. Chem.* **282**, 30889–30900 (2007).

15. Dopeso-Reyes, I. G. *et al.* Glucocerebrosidase expression patterns in the non-human primate brain. *Brain Struct. Funct.* **223**, 343–355 (2018).

16. Wilkening, G., Linke, T. & Sandhoff, K. Lysosomal degradation on vesicular membrane surfaces: Enhanced glucosylceramide degradation by lysosomal anionic lipids and activators. *J. Biol. Chem.* **273**, 30271–30278 (1998).

17. Romero, R. *et al.* Mechanism of glucocerebrosidase activation and dysfunction in Gaucher disease unraveled by molecular dynamics and deep learning. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 5086–5095 (2019).

18. Dvir, H. *et al.* X-ray structure of human acid-β-glucosidase, the defective enzyme in Gaucher disease. *EMBO Rep.* **4**, 704–709 (2003).

19. Orwig, S. D. *et al.* Binding of 3,4,5,6-tetrahydroxyazepanes to the acid-β-glucosidase active site: Implications for pharmacological chaperone design for Gaucher disease. *Biochemistry* **50**, 10647–10657 (2011).

20. Deen, M. C. *et al.* Selective fluorogenic β-glucocerebrosidase substrates for convenient analysis of enzyme activity in cell and tissue homogenates. *ACS Chem. Biol.* **15**, 824–829 (2020).

21. Witte, M. D. *et al.* Ultrasensitive in situ visualization of active glucocerebrosidase molecules. *Nat. Chem. Biol.* **6**, 907–913 (2010).

22. Matern, H., Gartzen, R. & Matern, S. β-Glucosidase activity towards a bile acid glucoside in human liver. *FEBS Lett.* **314**, 183–186 (1992).

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

24. Ridley, C. M. *et al.* β-Glucosidase 2 (GBA2) activity and imino sugar pharmacology. *J. Biol. Chem.* **288**, 26052–26066 (2013).

25. Yildiz, Y. *et al.* Mutation of β-glucosidase 2 causes glycolipid storage disease and impaired male fertility. *J. Clin. Invest.* **116**, 2985–2994 (2006).

26. Boot, R. G. *et al.* Identification of the non-lysosomal glucosylceramidase as β-glucosidase 2. *J. Biol. Chem.* **282**, 1305–1312 (2007).

27. Matern, H., Boermans, H., Lottspeich, F. & Matern, S. Molecular cloning and expression of human bile acid β-Glucosidase. *J. Biol. Chem.* **276**, 37929–37933 (2001).

28. Charoenwattanasatien, R. *et al.* Bacterial β-glucosidase reveals the structural and functional basis of genetic defects in human glucocerebrosidase 2 (GBA2). *ACS Chem. Biol.* **11**, 1891–1900 (2016).

29. Yahata, K. *et al.* Molecular cloning and expression of a novel klotho-related protein. *J. Mol. Med.* **78**, 389–394 (2000).

30. De Graaf, M. *et al.* Cloning and characterization of human liver cytosolic β-glycosidase. *Biochem. J.* **356**, 907–910 (2001).

31. Koshland Jr., D. E. Stereochemistry and the mechanism of enzymatic reactions. *Biol. Rev.* **28**, 416–436 (1953).

32. Beavan, M. S. & Schapira, A. H. V. Glucocerebrosidase mutations and the pathogenesis of Parkinson disease. *Ann. Med.* **45**, 511–521 (2013).

33. Foundation, P. Statistics @ www.parkinson.org. (2019).

34. Sidransky, E. *et al.* Multicenter analysis of glucocerebrosidase mutations in Parkinson's disease. *N. Engl. J. Med.* **361**, 1651–1661 (2009).

35. Goker-Alpan, O. *et al.* Glucocerebrosidase mutations are an important risk factor for

Lewy body disorders. *Neurology* **67**, 908–910 (2006).

36. Walden, C. M. *et al.* Accumulation of glucosylceramide in murine testis, caused by inhibition of β-glucosidase 2: Implications for spermatogenesis. *J. Biol. Chem.* **282**, 32655–32664 (2007).

37. Massimo, A. *et al.* Current and novel aspects on the non-lysosomal β-glucosylceramidase GBA2. *Neurochem. Res.* **41**, 210–220 (2016).

38. Marques, A. R. A. *et al.* Reducing GBA2 activity ameliorates neuropathology in Niemann-pick type C mice. *PLoS One* **10**, 1–18 (2015).

39. Burke, D. G. *et al.* Increased glucocerebrosidase (GBA) 2 activity in GBA1 deficient mice brains and in Gaucher leucocytes. *J. Inherit. Metab. Dis.* **36**, 869–872 (2013).

40. Mistry, P. K. *et al.* Glucocerebrosidase 2 gene deletion rescues type 1 Gaucher disease. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 4934–4939 (2014).

41. Schonauer, S. *et al.* Identification of a feedback loop involving β-glucosidase 2 and its product sphingosine sheds light on the molecular mechanisms in Gaucher disease. *J. Biol. Chem.* **292**, 6177–6189 (2017).

42. Yadav, A. K. *et al.* Fluorescence-quenched substrates for live cell imaging of human glucocerebrosidase activity. *J. Am. Chem. Soc.* **137**, 1181–1189 (2015).

43. Shayman, J. A. & Larsen, S. D. The development and use of small molecule inhibitors of glycosphingolipid metabolism for lysosomal storage diseases. *J. Lipid Res.* **55**, 1215–1225 (2014).

44. Overkleeft, H. S. *et al.* Generation of specific deoxynojirimycin-type inhibitors of the non-lysosomal glucosylceramidase. *J. Biol. Chem.* **273**, 26522–26527 (1998).

45. Lahav, D. *et al.* A fluorescence polarization activity-based protein profiling assay in the discovery of potent, selective inhibitors for human nonlysosomal glucosylceramidase. *J. Am. Chem. Soc.* **139**, 14192–14197 (2017).

46. Gu, X. *et al.* Structure–activity studies of *N*-butyl-1-deoxynojirimycin (NB-DNJ) analogues: Discovery of potent and selective aminocyclopentitol inhibitors of GBA1 and

GBA2. *ChemMedChem* **12**, 1977–1984 (2017).

47. Scott, M. E. & Viola, R. E. The use of fluoro- and deoxy-substrate analogs to examine binding specificity and catalysis in the enzymes of the sorbitol pathway. *Carbohydr. Res.* **313**, 247–253 (1998).

48. Street, I. P., Armstrong, C. R. & Withers, S. G. Hydrogen bonding and specificity. Fluorodeoxy sugars as probes of hydrogen bonding in the glycogen phosphorylase-glucose complex. *Biochemistry* **25**, 6021–6027 (1986).

49. Zhou, Y. *et al.* Next generation of fluorine-containing pharmaceuticals, compounds currently in phase II-III clinical trials of major pharmaceutical companies: New structural trends and therapeutic areas. *Chem. Rev.* **116**, 422–518 (2016).

50. Wang, J. *et al.* Fluorine in pharmaceutical industry: Fluorine-containing drugs introduced to the market in the last decade (2001-2011). *Chem. Rev.* **114**, 2432–2506 (2014).

51. Street, I. P., Rupitz, K. & Withers, S. G. Fluorinated and deoxygenated substrates as probes of transition-state structure in glycogen phosphorylase. *Biochemistry* **24**, 1581–1587 (1989).

52. Namchuk, M. N., McCarter, J. D., Becalski, A., Andrews, T. & Withers, S. G. The role of sugar substituents in glycoside hydrolysis. *J. Am. Chem. Soc.* **122**, 1270–1277 (2000).

53. Hudson, K. L. *et al.* Carbohydrate-aromatic interactions in proteins. *J. Am. Chem. Soc.* **137**, 15152–15160 (2015).

54. Withers, S. G., Percival, M. D. & Street, I. P. The synthesis and hydrolysis of a series of deoxyfluoro-α-D-glucopyranosyl phosphates. *Carbohydr. Res.* **154**, 127–144 (1986).

55. Dunitz, J. D. & Taylor, R. Organic fluorine hardly ever accepts hydrogen bonds. *Chem. - A Eur. J.* **3**, 89–98 (1997).

56. Jiménez-Barbero, J. *et al.* Fluorinated carbohydrates as chemical probes for molecular recognition studies. Current status and perspectives. *Chem. Soc. Rev.* **49**, 3863–3888 (2020).

57. O'hagan, D. Understanding organofluorine chemistry. An introduction to the C–F bond.

Chem. Soc. Rev. **37**, 308–319 (2008).

58. Gunbas, G. *et al.* Extreme oxatriquinanes and a record C-O bond length. *Nat. Chem.* **4**, 1018–1023 (2012).

59. Withers, S. G., Street, J. P. & Rettig, S. J. The preferred conformation of 2-fluoro-2-deoxy β-D-mannopyranosyl fluoride. An X-ray crystallographic and 2-dimensional proton nuclear magnetic resonance study. *Can. J. Chem.* **64**, 232–236 (1986).

60. Arnott, J. A. & Planey, S. L. The influence of lipophilicity in drug discovery and design. *Expert Opin. Drug Discov.* **7**, 863–875 (2012).

61. Vermersch, P. S., Tesmer, J. J. G. & Quiocho, F. A. Protein-ligand energetics assessed using deoxy and fluorodeoxy sugars in equilibrium binding and high resolution crystallographic studies. *J. Mol. Biol.* **226**, 923–929 (1992).

62. Vickman, A. E. & Pohl, N. L. B. Probing deoxysugar conformational preference: A comprehensive computational study investigating the effects of deoxygenation. *Carbohydr. Res.* **475**, 17–26 (2019).

63. Snavely, B. B., Peterson, O. G. & Reithel, R. F. Blue laser emission from a flashlamp-excited organic dye solution. *Appl. Phys. Lett.* **11**, 275–276 (1967).

64. Mead, J. A. R., Smith, J. N. & Williams, R. T. The biosynthesis of the glucuronides of umbelliferone and 4-methylumbelliferone and their use in fluorimetric determination of β-glucuronidase. *Biochem. J.* **61**, 569–574 (1955).

65. Fink, D. W. & Koehler, W. R. pH effects on fluorescence of umbelliferone. *Anal. Chem.* **42**, 990–993 (1970).

66. Perry, J. D. *et al.* Evaluation of novel fluorogenic substrates for the detection of glycosidases in Escherichia coli and enterococci. *J. Appl. Microbiol.* **101**, 977–985 (2006).

67. Pritsch, K. *et al.* A rapid and highly sensitive method for measuring enzyme activities in single mycorrhizal tips using 4-methylumbelliferone-labelled fluorogenic substrates in a microplate system. *J. Microbiol. Methods* **58**, 233–241 (2004).

68. Manafi, M., Kneifel, W. & Bascomb, S. Fluorogenic and chromogenic substrates used in

bacterial diagnostics. *Microbiol. Rev.* **55**, 335–348 (1991).

69. Sardi, S. P. *et al.* CNS expression of glucocerebrosidase corrects α-synuclein pathology and memory in a mouse model of Gaucher-related synucleinopathy. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 12101–12106 (2011).

70. Withers, S. G., Rupitz, K. & Street, I. P. 2-Deoxy-2-fluoro-D-glycosyl fluorides. A new class of specific mechanism-based glysosidase inhibitors. *J. Biol. Chem.* **263**, 7929–7932 (1988).

71. Keenan, T. *et al.* Profiling substrate promiscuity of wild-type sugar kinases for multi-fluorinated monosaccharides. *Cell Chem. Biol.* **27**, 1199-1206.e5 (2020).

72. Slámová, K. *et al.* 4-Deoxy-substrates for β-*N*-acetylhexosaminidases: How to make use of their loose specificity. *Glycobiology* **20**, 1002–1009 (2010).

73. Stick, R. V. & Williams, S. J. Synthesis and Protecting Groups. *Carbohydrates: The essential molecules of life* (2009). doi:10.1016/b978-0-240-52118-3.00002-8

74. Schmidt, R. R. & Michel, J. Direct *O*-glycosyl trichloroacetimidate formation. Nucleophilicity of the anomeric oxygen atom. *Tetrahedron Lett.* **25**, 821–824 (1984).

75. Vo, Q. V. *et al.* Synthesis of aromatic and indole alpha-glucosinolates. *Carbohydr. Res.* **455**, 45–53 (2018).

76. Wei, X. *et al.* An improved helferich method for the α/β-stereoselective synthesis of 4-methylumbelliferyl glycosides for the detection of microorganisms. *Molecules* **20**, 21681–21699 (2015).

77. Hunsen, M., Long, D. A., D'Ardenne, C. R. & Smith, A. L. Mild one-pot preparation of glycosyl bromides. *Carbohydr. Res.* **340**, 2670–2674 (2005).

78. Wu, Z., Fu, X. ling, Yang, N. & Wang, Q. Synthesis and fluorescence properties of coumarin glycosides and triazoylglycosides. *Chem. Res. Chinese Univ.* **29**, 460–465 (2013).

79. Singh, R. P. & Shreeve, J. M. Recent advances in nucleophilic fluorination reactions of organic compounds using Deoxofluor and DAST. *Synthesis (Stuttg).* 2561–2578 (2002).

doi:10.1055/s-2002-35626

80. Maki, Y. *et al.* Acceleration and deceleration factors on the hydrolysis reaction of 4,6-*O*-benzylidene acetal group. *J. Org. Chem.* **85**, 15849–15856 (2020).

81. Ferro, V., Mocerino, M., Stick, R. V. & Tilbrook, D. M. G. An improvement in the preparation of some carbohydrate benzylidene acetals. *Aust. J. Chem.* **41**, 813–5 (1988).

82. Driguez, H., McAuliffe, J. C., Stick, R. V., Tilbrook, M. G. & Williams, S. J. A new approach to some 1,6-dideoxy 1,6-epithio sugars. *Aust. J. Chem.* **49**, 343–348 (1996).

83. Mukobata, T., Ochi, Y., Ito, Y., Wada, S. & Urata, H. Facile and efficient approach for the synthesis of *N*-dimethylaminomethylene-2′-*O*-methylguanosine. *Bioorganic Med. Chem. Lett.* **20**, 129–131 (2010).

84. Magnusson, G., Noori, G., Dahmén, J., Frejd, T. & Lave, T. BF_3-Etherate induced formation of 2,2,2-trichloroethyl glycopyranosides. Selective visualization of carbohydrate derivatives on TLC plates. *Acta Chemica Scandinavica* **35**, 213–216 (1981).

85. Sarkar, S., Dutta, S. & Sen, A. K. Expedient and facile one-pot syntheses of triazole-linked glycoconjugates under microwave irradiation. *Synthesis (Stuttg).* **44**, 1079–1089 (2012).

86. Goto, K. *et al.* Synthesis of 1,5-Anhydro-D-fructose derivatives and evaluation of their inflammasome inhibitors. *Bioorganic Med. Chem.* **26**, 3763–3772 (2018).

87. Egli, M. *et al.* Synthesis, improved antisense activity and structural rationale for the divergent RNA affinities of 3′-fluoro hexitol nucleic acid (FHNA and Ara-FHNA) modified oligonucleotides. *J. Am. Chem. Soc.* **133**, 16642–16649 (2011).

88. Liu, Z. *et al.* 3-Fluoro-azetidine carboxylic acids and trans,trans- 3,4-difluoroproline as peptide scaffolds: Inhibition of pancreatic cancer cell growth by a fluoroazetidine iminosugar. *J. Org. Chem.* **80**, 4244–4258 (2015).

89. Xu, G. *et al.* Design, synthesis and biological evaluation of (2S,3R,4R,5S,6R)-5-fluoro-6-(hydroxymethyl)-2-aryltetrahydro-2H-pyran-3,4-diols as potent and orally active SGLT dual inhibitors. *Bioorganic Med. Chem. Lett.* **28**, 3446–3453 (2018).

90. Roth, A. G., Redmer, S. & Arenz, C. Development of carbohydrate-derived inhibitors of acid sphingomyelinase. *Bioorganic Med. Chem.* **2**, 939–944 (2010).

91. Selrnan, S., Karoui, S. & Singer, E. 4-Deoxy-l-thio-β-D-glucose and its (1:1)-complex. *Swiss Chem. Soc.* **6**, 2079–2082 (1973).

92. Song, Z., Meng, L., Xiao, Y., Zhao, X., Fang, J., Zeng, J. Calcium hypophoshite mediated deiodination in water: Mechanistic insight and applicatins in large scale synthesi of D-quinovose and D-rhamnose. *R. Soc. Chem.* **5**, 1122–1127 (2019).

93. Malet, C. *et al.* Synthesis of 4-methylumbelliferyl-β-D-glucan oligosaccharides as specific chromophoric substrates of (1 → 3),(1 → 4)-β-D-glucan 4-glucanohydrolases. *Carbohydr. Res.* **274**, 285–301 (1995).

94. Akiyama, H. *et al.* Glucocerebrosidases catalyze a transgalactosylation reaction that yields a newly-identified brain sterol metabolite, galactosylated cholesterol. *J. Biol. Chem.* **295**, 5257–5277 (2020).

95. Penner, M. H. & Liaw, E.-T. Kinetic consequences of high ratios of substrate to enzyme saccharification systems based on trichoderma cellulase. *American chemical society* 363–371 (1994). doi:10.1021/bk-1994-0566.ch018

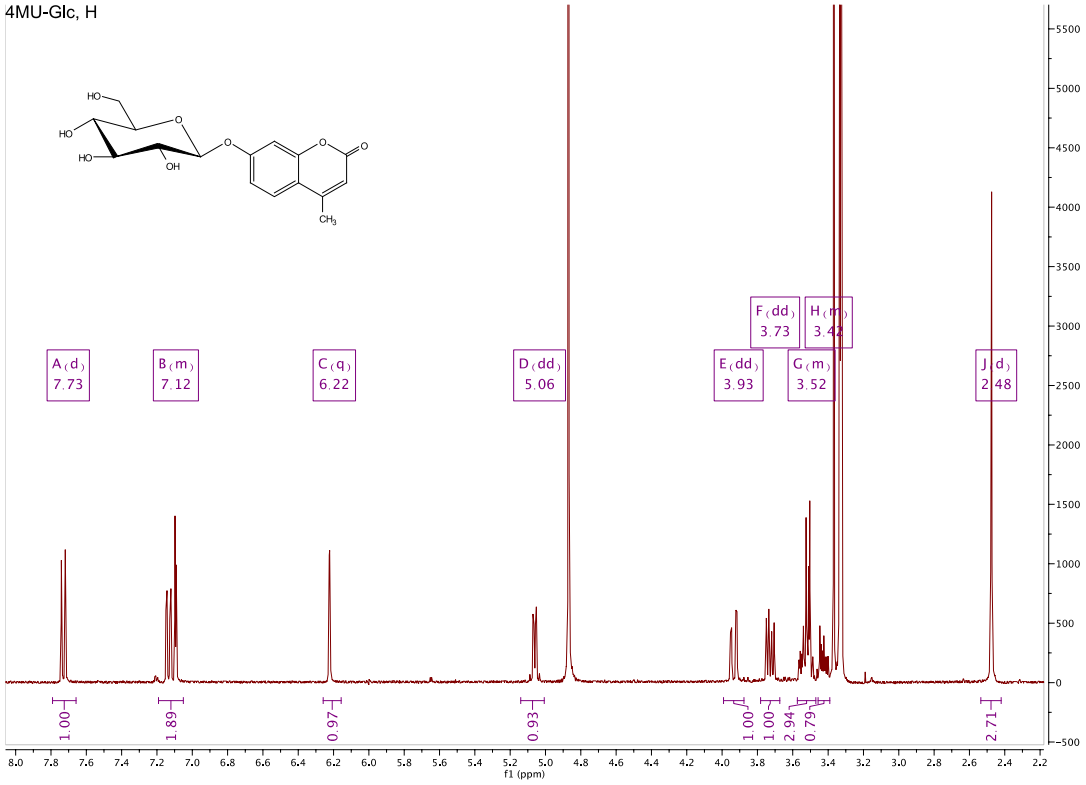
96. Kawai, R., Igarashi, K., Kitaoka, M., Ishii, T. & Samejima, M. Kinetics of substrate transglycosylation by glycoside hydrolase family 3 glucan (1→3)-β-glucosidase from the white-rot fungus Phanerochaete chrysosporium. *Carbohydr. Res.* **339**, 2851–2857 (2004).

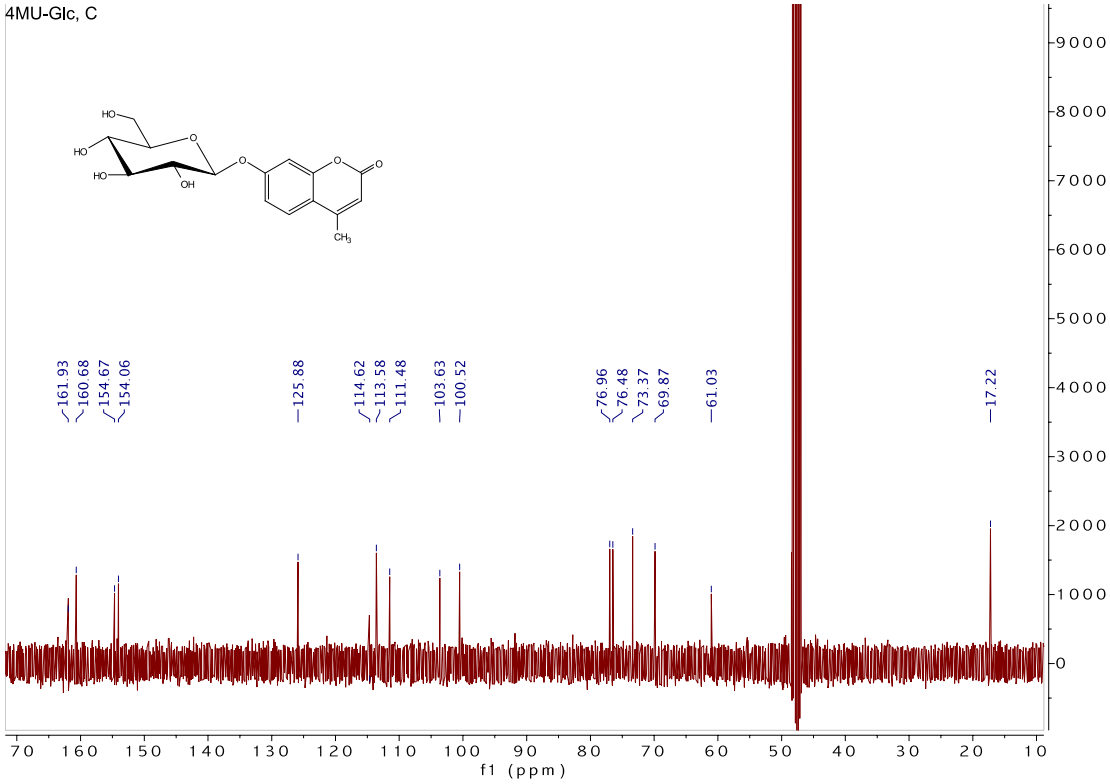
97. Daniels, L. B., Coyle, P. J., Chiao, Y. B., Glew, R. H. & Labow, R. S. Purification and characterization of a cytosolic broad specificity β-glucosidase from human liver. *J. Biol. Chem.* **256**, 13004–13013 (1981).

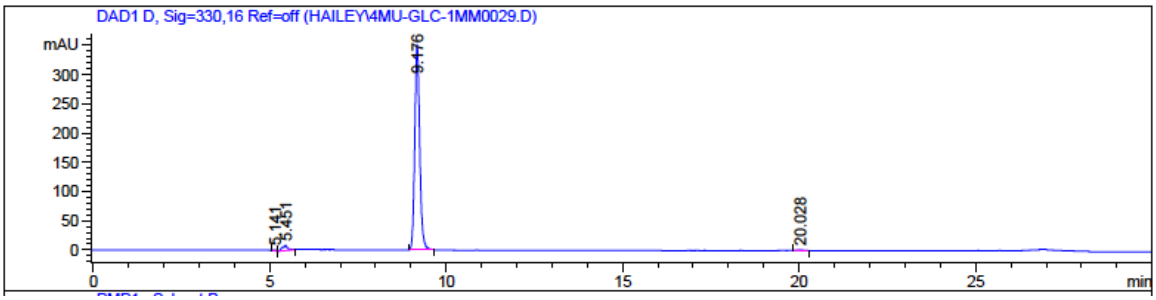
98. Ullal, A. J. *et al.* Fluorimetric assay with a novel substrate for quantification of galactocerebrosidase activity in dried blood spot specimens. *Pract. Lab. Med.* **18**, e00141 (2020).

Appendix

NMR spectra and HPLC traces of final compounds

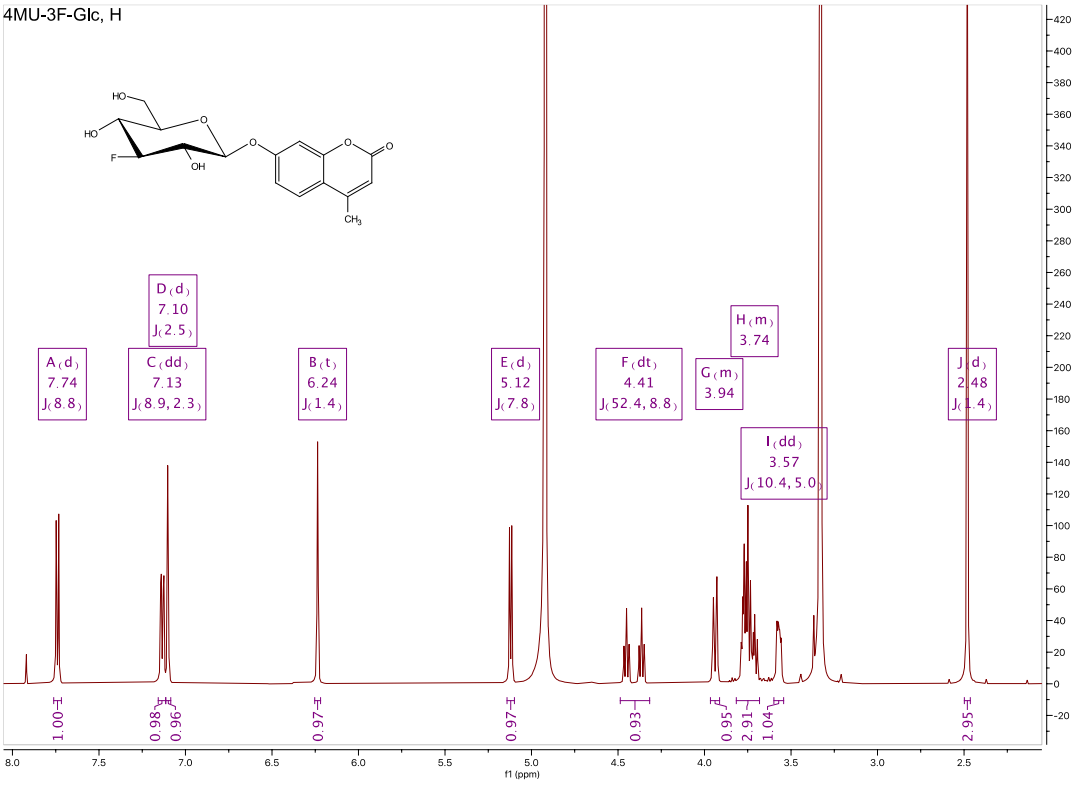

4MU-Glc, H
A (d) 7.73
B (m) 7.12
C (q) 6.22
D (dd) 5.06
E (dd) 3.93
F (dd) 3.73
G (m) 3.52
H (m) 3.42
J (d) 2.48
1.00
1.89
0.97
0.93
1.00
1.00
2.94
0.79
2.71
f1 (ppm)

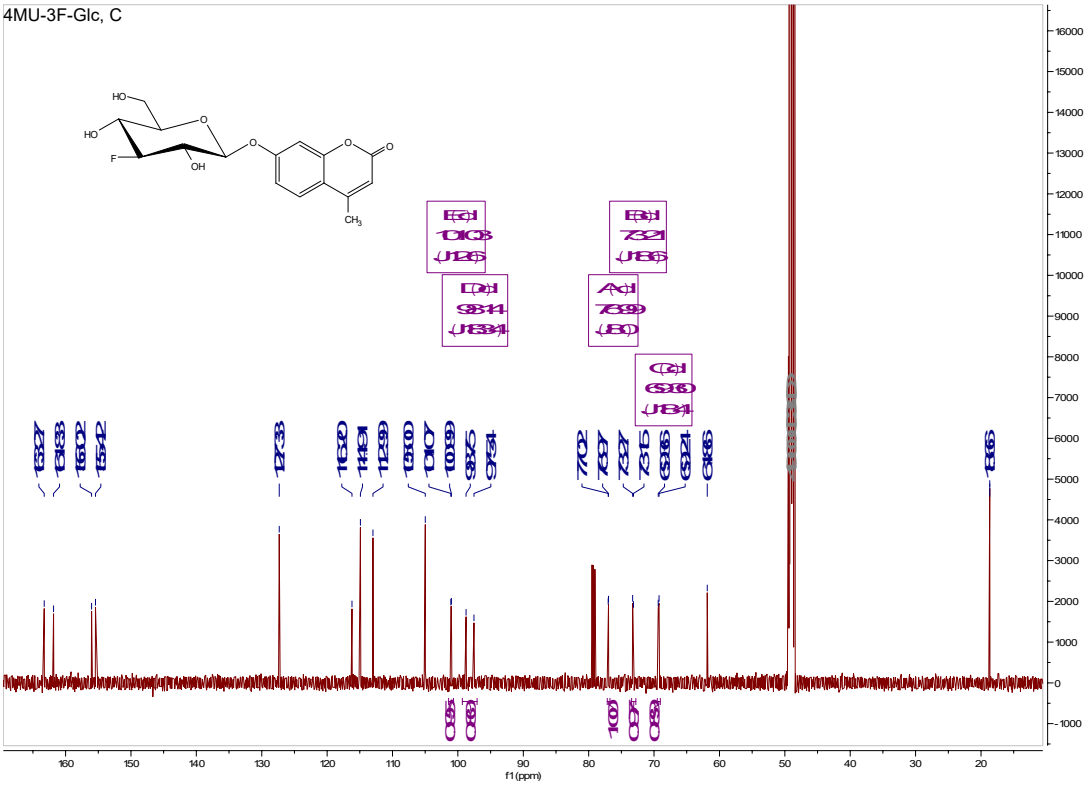

4MU-Glc, C
161.93
160.68
154.67
154.06
125.88
114.62
113.58
111.48
103.63
100.52
76.96
76.48
73.37
69.87
61.03
17.22
f1 (ppm)

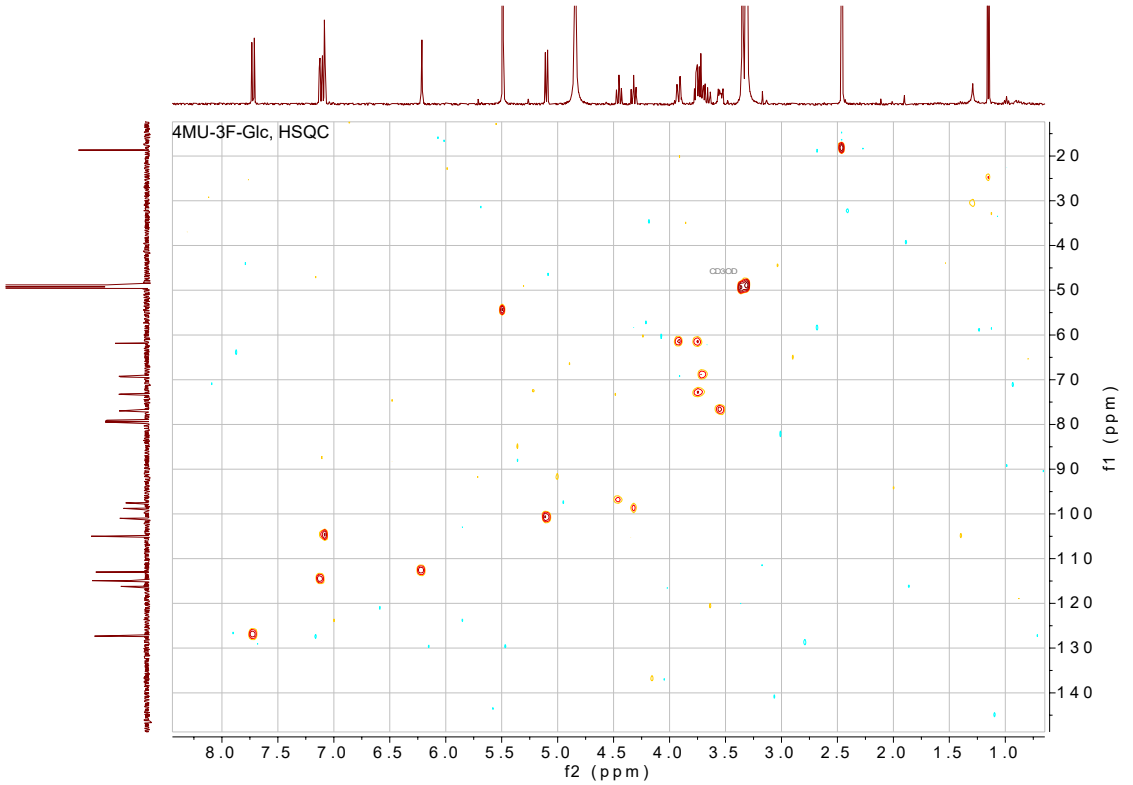


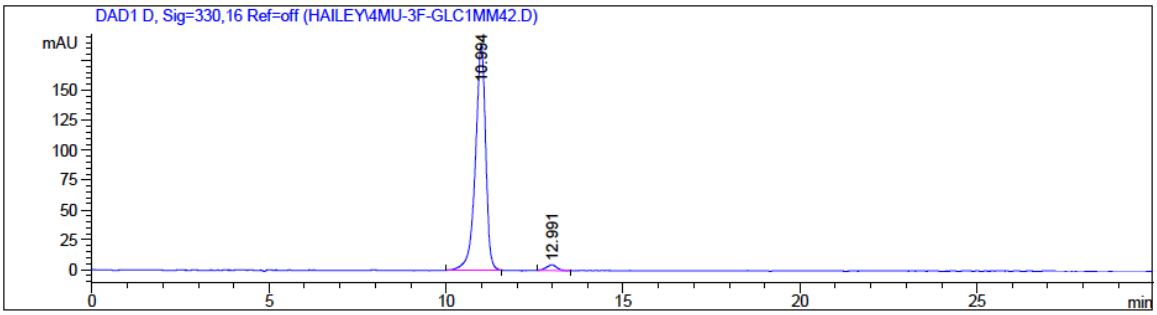
Retention time (min)	Area (mAU ∗ s)	Height (mAU)	Area %
5.141	11.736	2.193	0.3266
5.451	98.540	9.029	2.7426
9.176	3468.569	351.396	96.5392
20.028	14.067	1.141	0.3915

Figure S. 1. ^{1}H, ^{13}C NMR spectra and HPLC trace of 4MU-Glc (1)


4MU-3F-Glc, H
A (d) 7.74 J (8.8)
C (dd) 7.13 J (8.9, 2.3)
D (d) 7.10 J (2.5)
B (t) 6.24 J (1.4)
E (d) 5.12 J (7.8)
F (dt) 4.41 J (52.4, 8.8)
G (m) 3.94
H (m) 3.74
I (dd) 3.57 J (10.4, 5.0)
J (d) 2.48 J (1.4)
1.00
0.98
0.96
0.97
0.97
0.93
0.95
2.91
1.04
2.95
f1 (ppm)

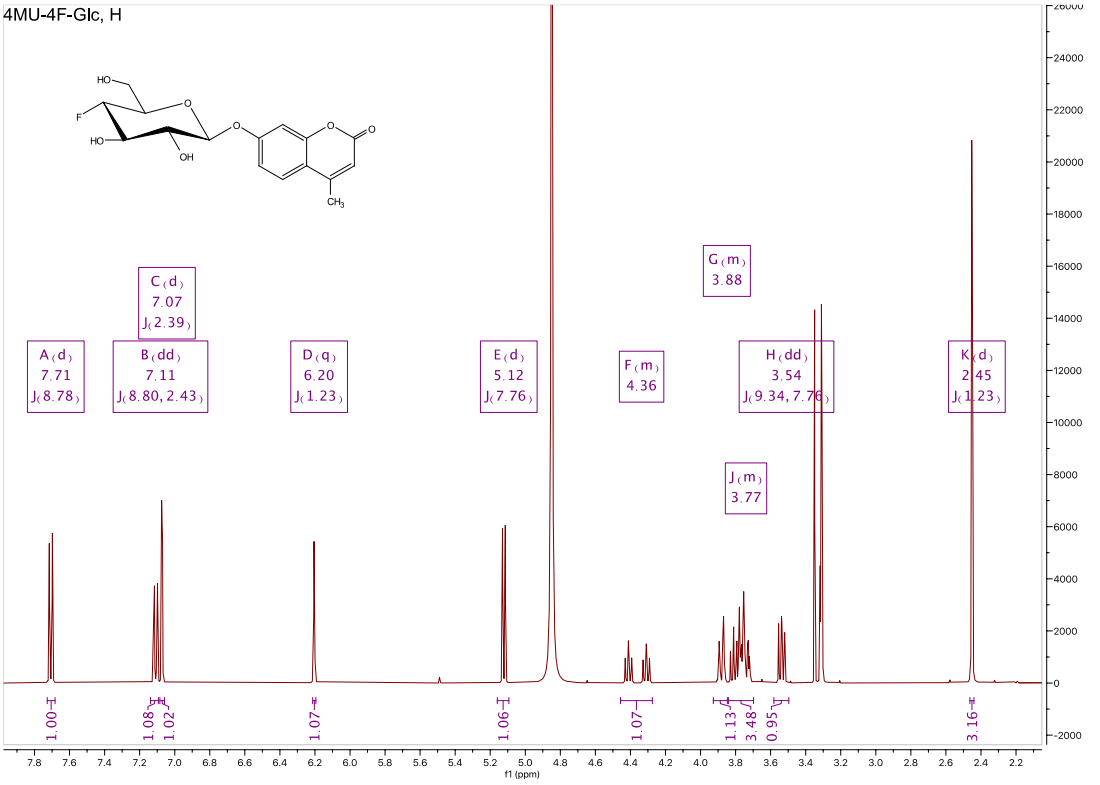

4MU-3F-Glc, C
f1 (ppm)

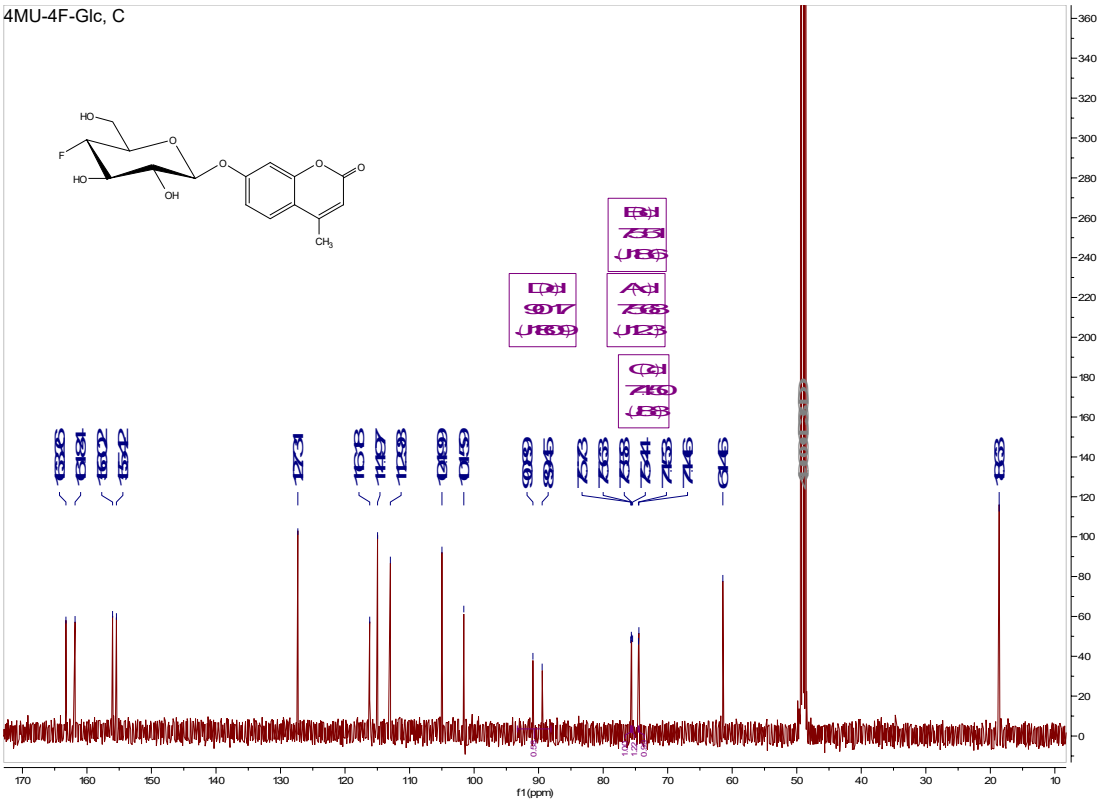


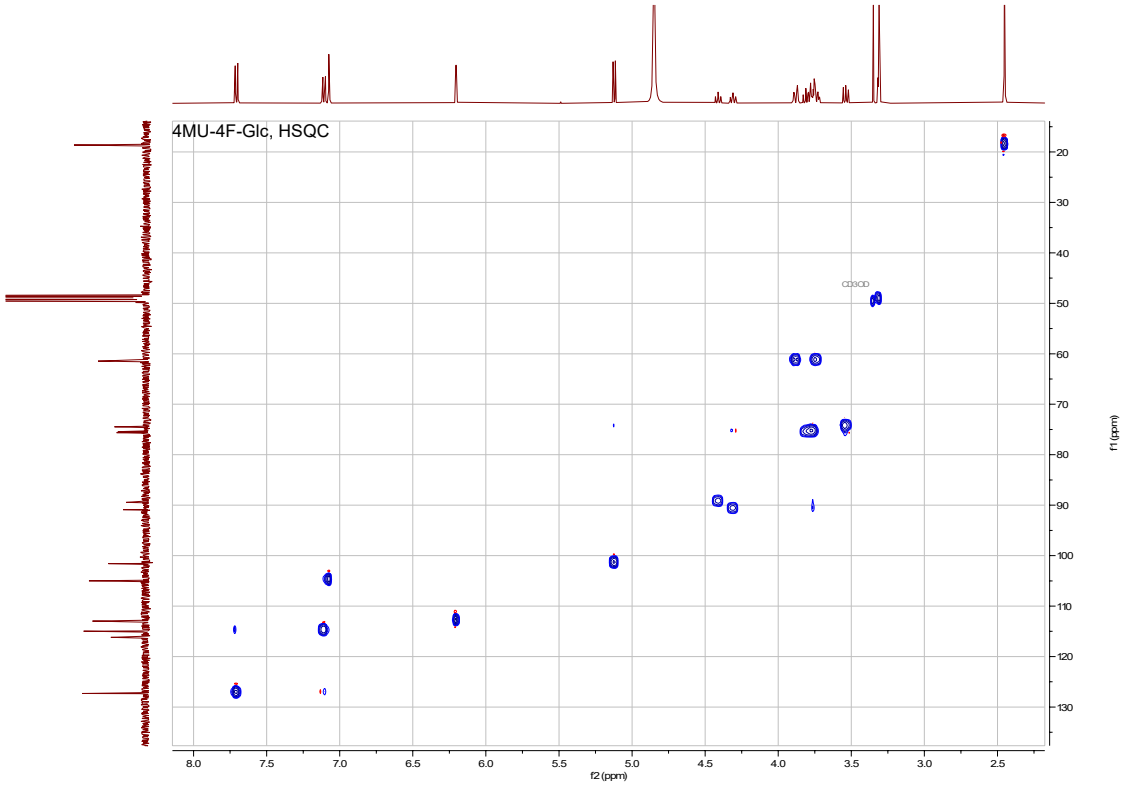


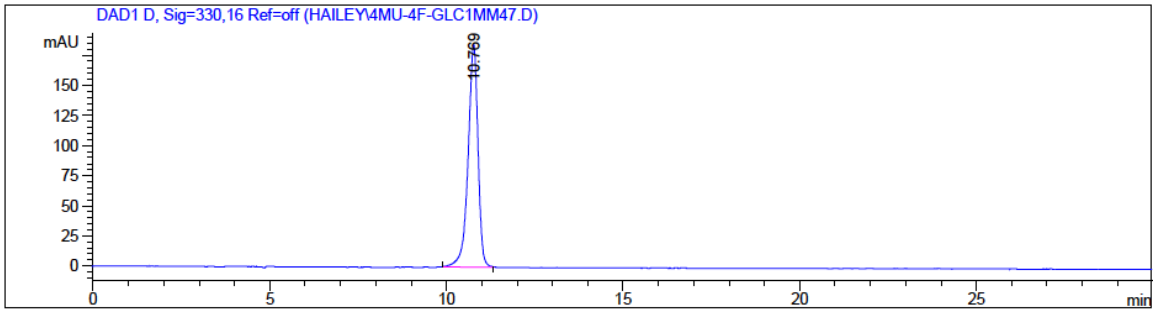
Retention time (min)	Area (mAU ∗ s)	Height (mAU)	Area %
10.994	3817.739	188.562	97.4542
12.991	99.730	4.706	2.5458

Figure S. 2. ^{1}H NMR, ^{13}C NMR, HSQC spectra and HPLC trace of 4MU-3F-Glc (2)


4MU-4F-Glc, H
A (d) 7.71 J (8.78)
B (dd) 7.11 J (8.80, 2.43)
C (d) 7.07 J (2.39)
D (q) 6.20 J (1.23)
E (d) 5.12 J (7.76)
F (m) 4.36
G (m) 3.88
H (dd) 3.54 J (9.34, 7.76)
J (m) 3.77
K (d) 2.45 J (1.23)
f1 (ppm)

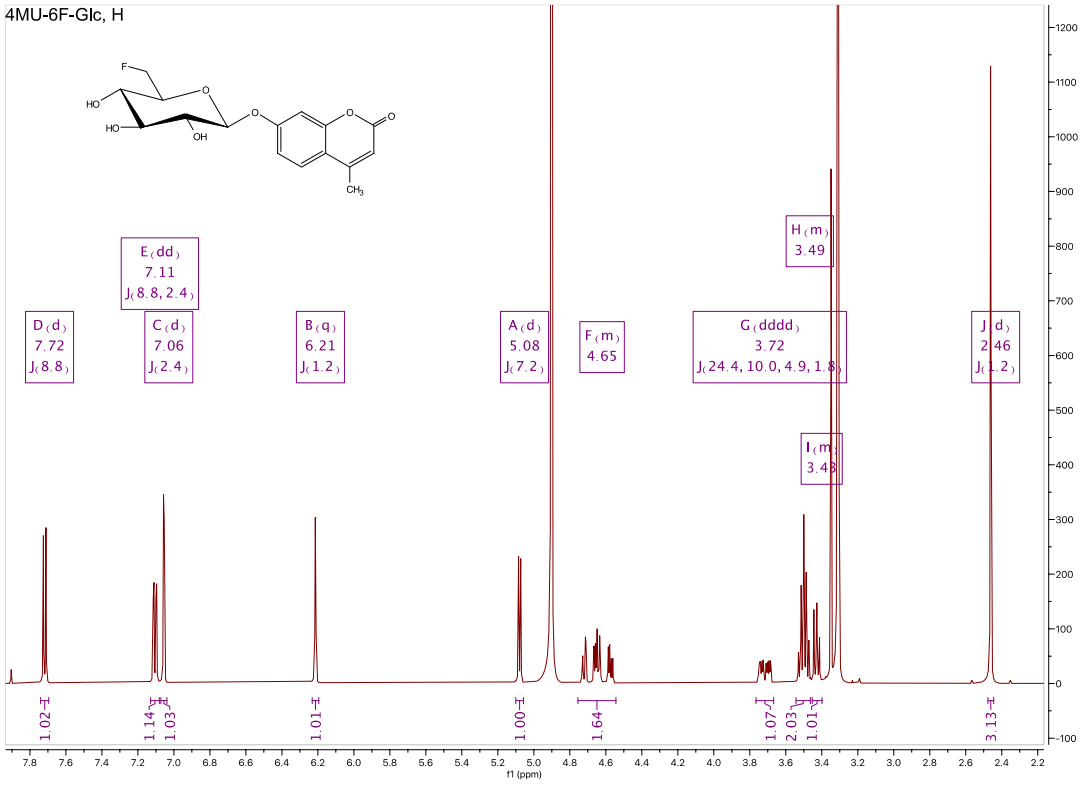

4MU-4F-Glc, C
f1 (ppm)

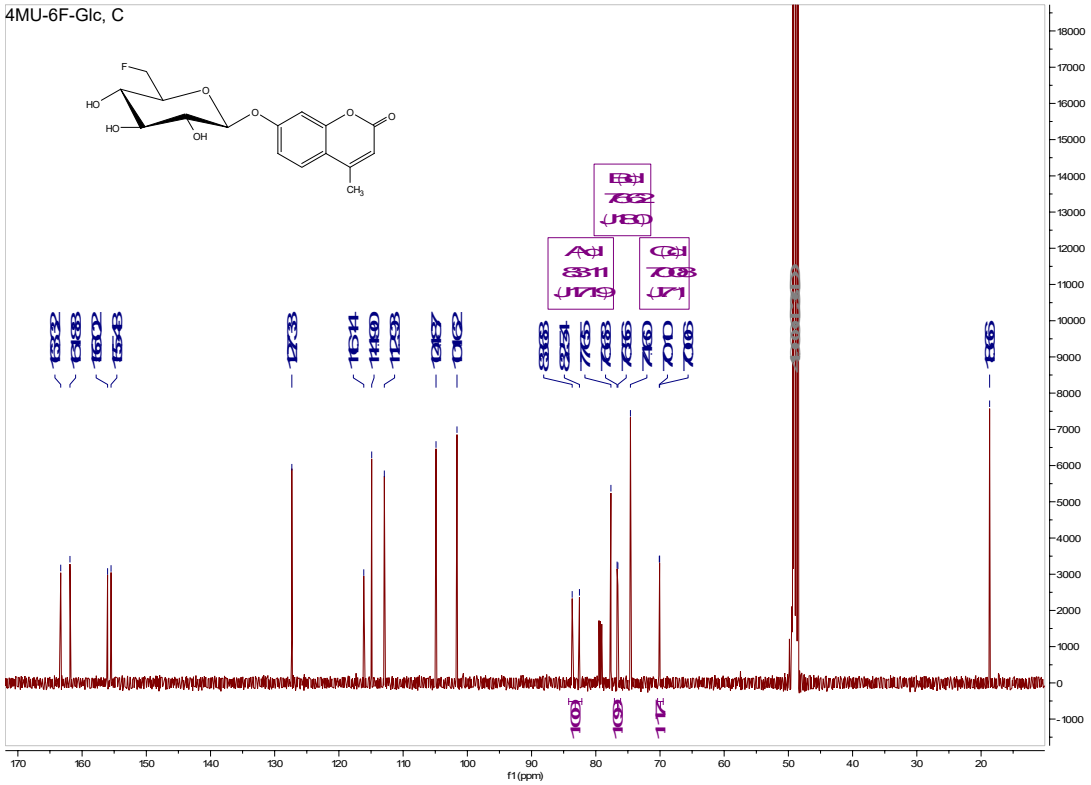


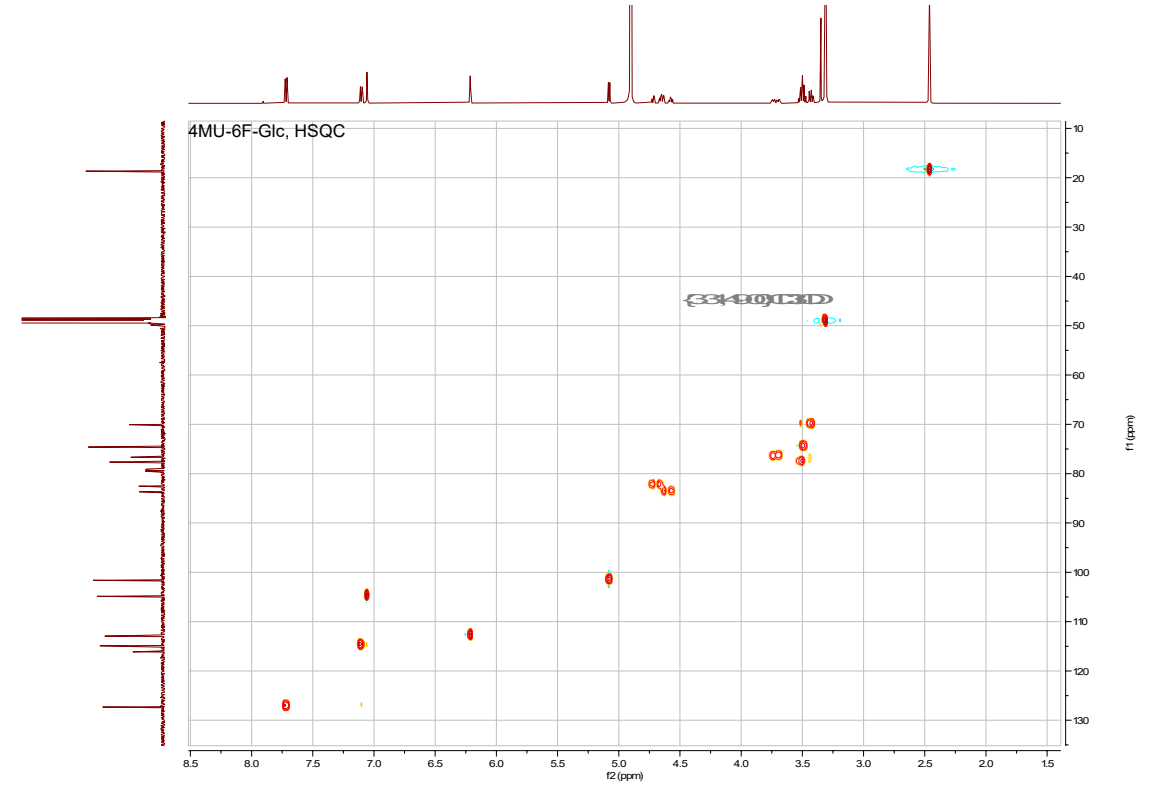


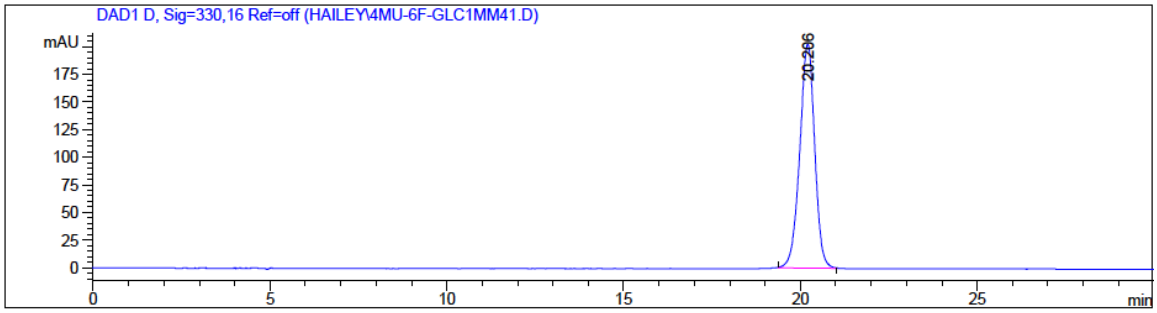
Retention time (min)	Area (mAU ∗ s)	Height (mAU)	Area %
10.769	3621.027	185.413	100.0000

Figure S. 3. ^{1}H NMR, ^{13}C NMR, HSQC spectra and HPLC trace of 4MU-4F-Glc (3)


4MU-6F-Glc, H
D (d) 7.72 J(8.8)
E (dd) 7.11 J(8.8, 2.4)
C (d) 7.06 J(2.4)
B (q) 6.21 J(1.2)
A (d) 5.08 J(7.2)
F (m) 4.65
G (dddd) 3.72 J(24.4, 10.0, 4.9, 1.8)
H (m) 3.49
I (m) 3.43
J (d) 2.46 J(1.2)
1.02
1.14
1.03
1.01
1.00
1.64
1.07
2.03
1.01
3.13
f1 (ppm)

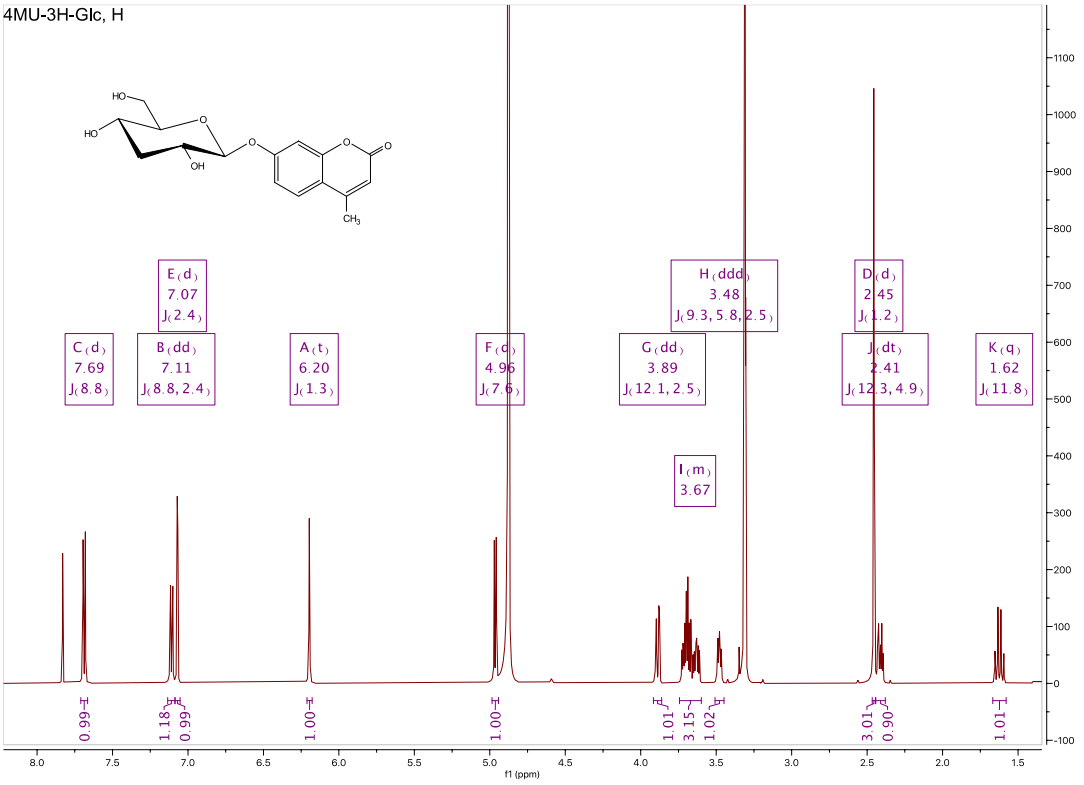

4MU-6F-Glc, C
f1 (ppm)

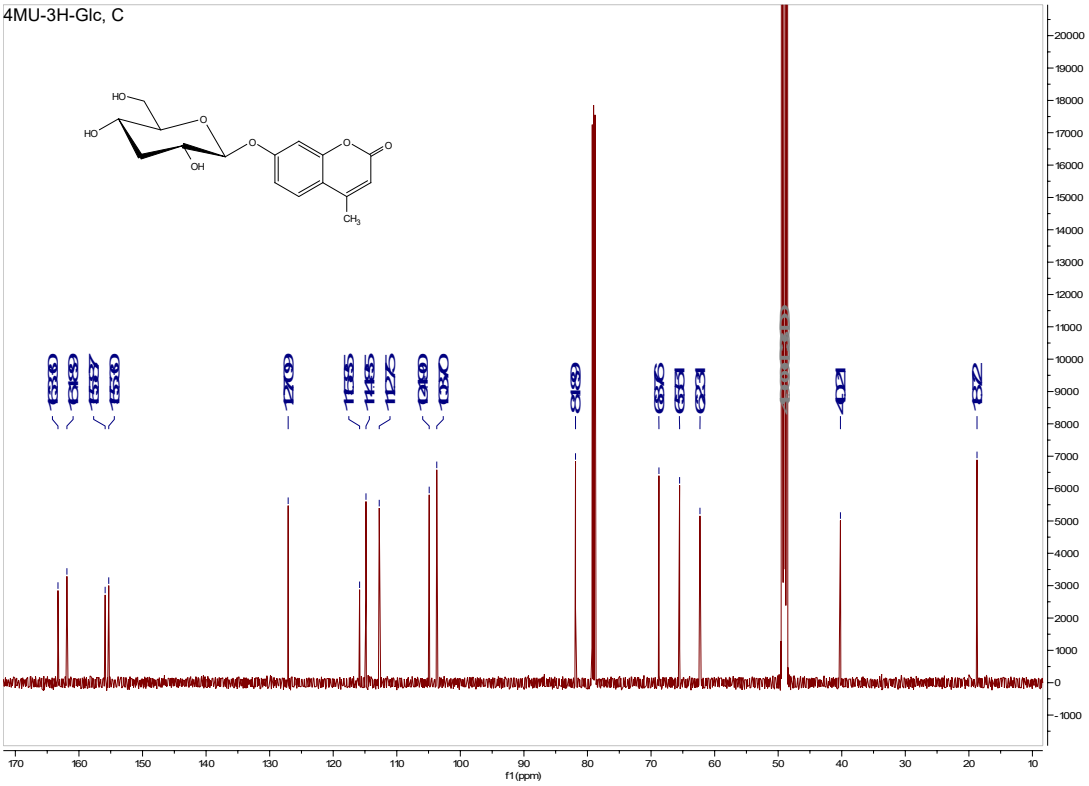


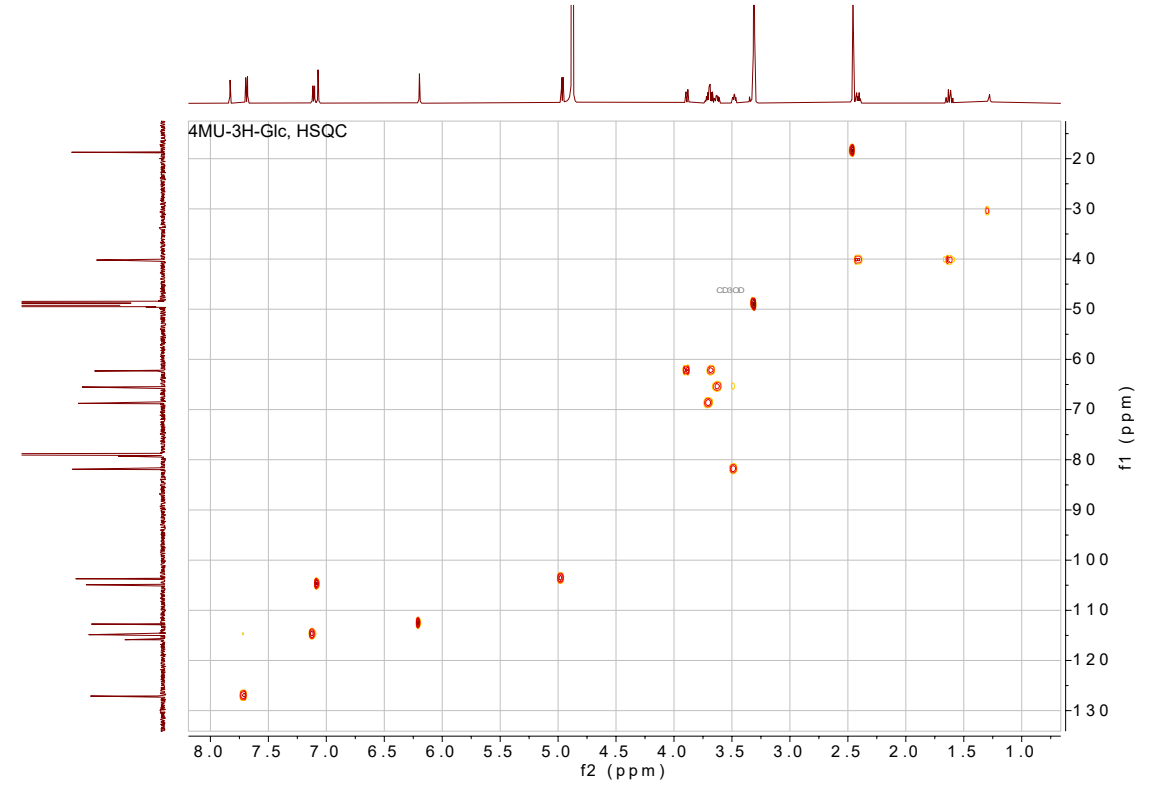


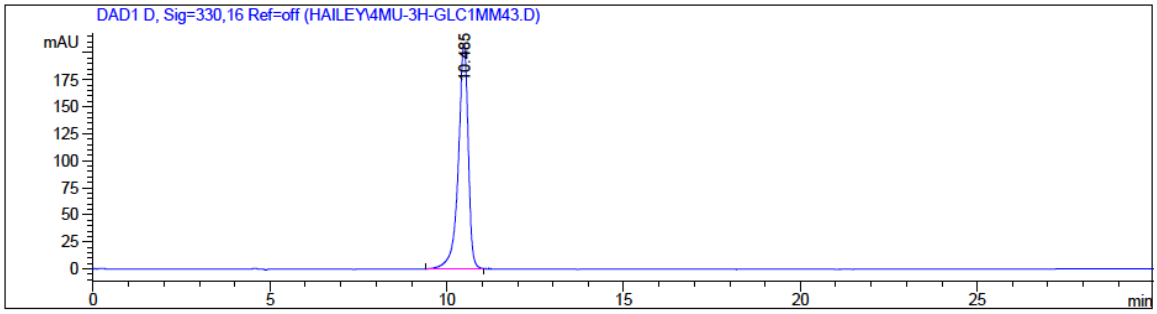
Retention time (min)	Area (mAU ∗ s)	Height (mAU)	Area %
20.206	5956.047	202.609	100.0000

Figure S. 4. ^{1}H NMR, ^{13}C NMR, HSQC spectra and HPLC trace of 4MU-6F-Glc (4)


4MU-3H-Glc, H
E (d) 7.07 J (2.4)
C (d) 7.69 J (8.8)
B (dd) 7.11 J (8.8, 2.4)
A (t) 6.20 J (1.3)
F (d) 4.96 J (7.6)
H (ddd) 3.48 J (9.3, 5.8, 2.5)
G (dd) 3.89 J (12.1, 2.5)
I (m) 3.67
D (d) 2.45 J (12.2)
J (dt) 2.41 J (12.3, 4.9)
K (q) 1.62 J (11.8)
0.99
1.18
0.99
1.00
1.00
1.01
3.15
1.02
3.01
0.90
1.01
f1 (ppm)

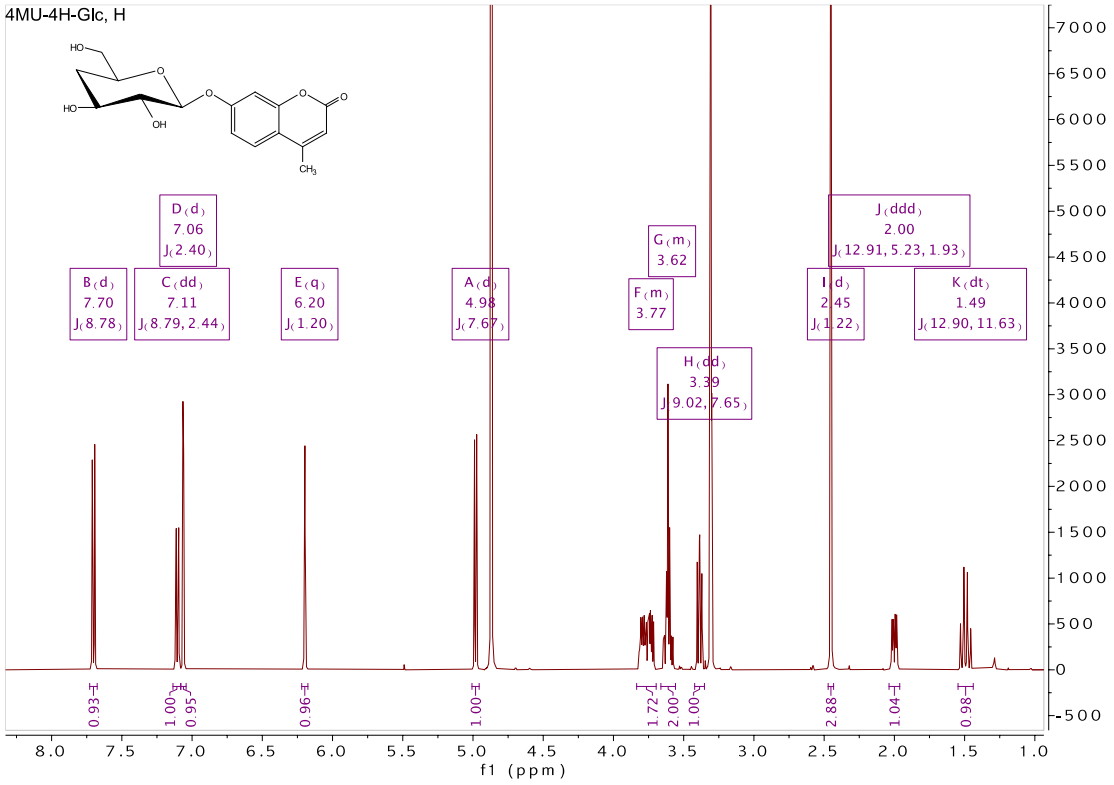

4MU-3H-Glc, C
f1 (ppm)

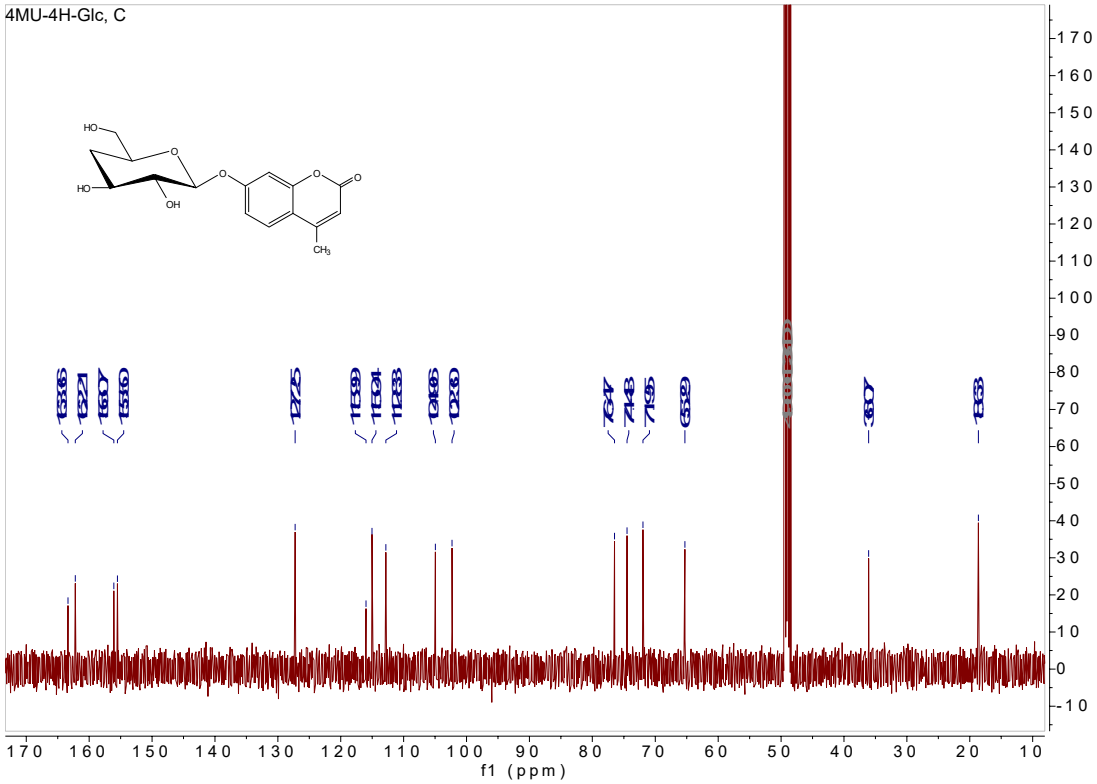


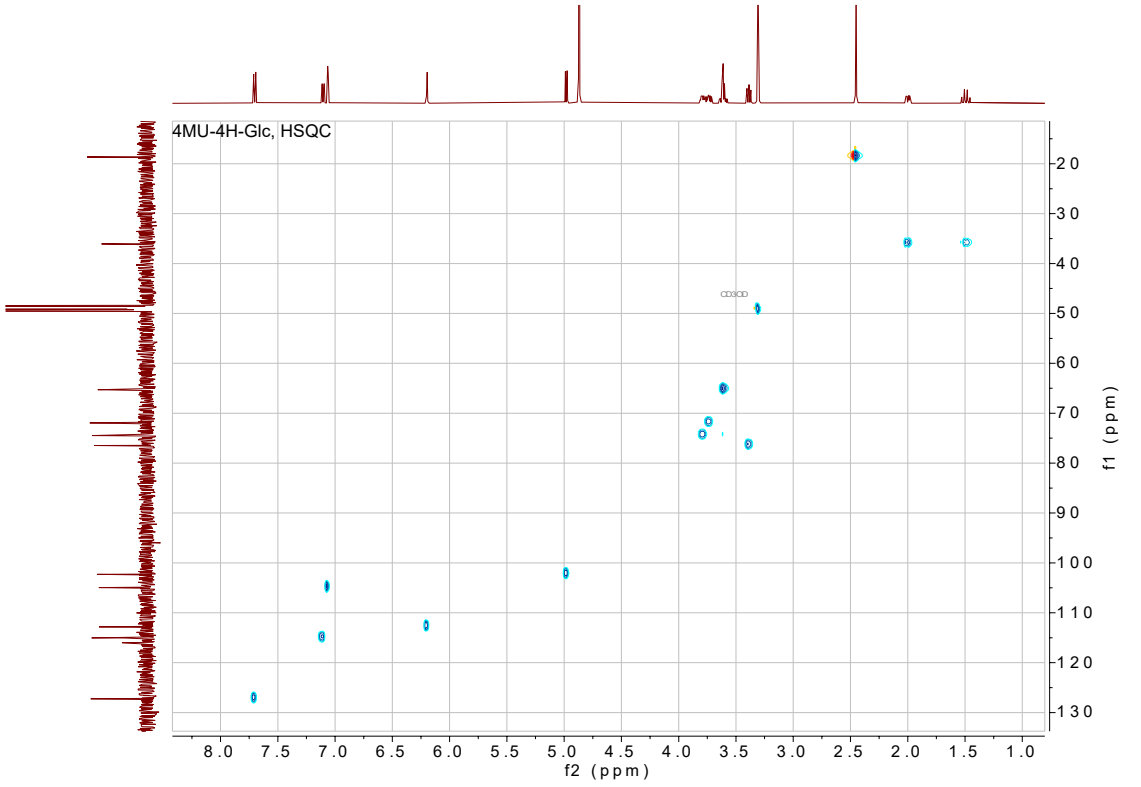


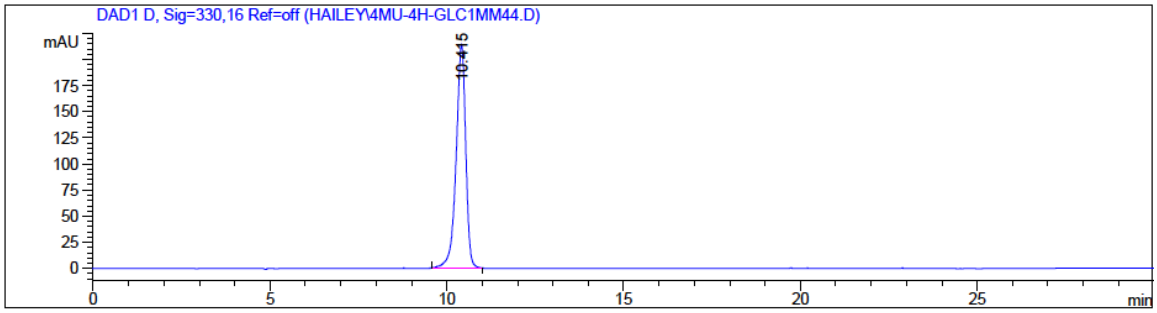
Retention time (min)	Area (mAU ∗ s)	Height (mAU)	Area %
10.485	4057.851	207.915	100.0000

Figure S. 5. ^{1}H NMR, ^{13}C NMR, HSQC spectra and HPLC trace of 4MU-3H-Glc (5)


4MU-4H-Glc, H
B (d) 7.70 J (8.78)
D (d) 7.06 J (2.40)
C (dd) 7.11 J (8.79, 2.44)
E (q) 6.20 J (1.20)
A (d) 4.98 J (7.67)
G (m) 3.62
F (m) 3.77
H (dd) 3.39 J (9.02, 7.65)
J (ddd) 2.00 J (12.91, 5.23, 1.93)
I (d) 2.45 J (1.22)
K (dt) 1.49 J (12.90, 11.63)
0.93
1.00
0.95
0.96
1.00
1.72
2.00
1.00
2.88
1.04
0.98
f1 (ppm)

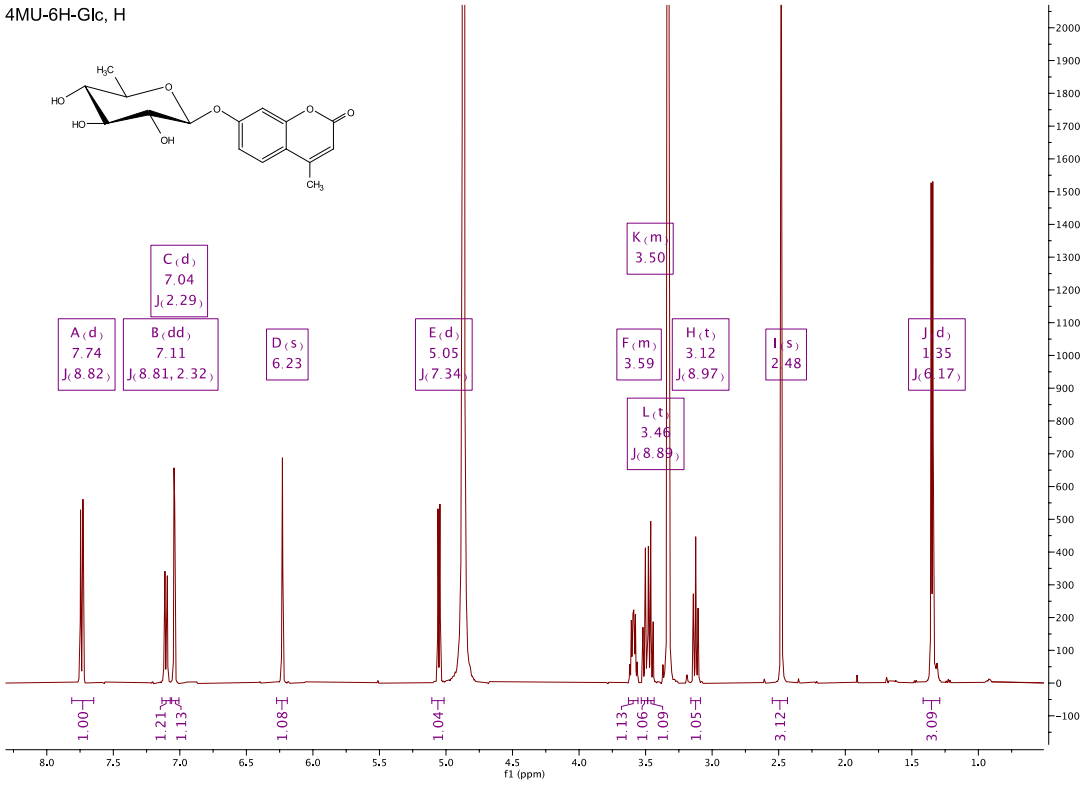

4MU-4H-Glc, C
f1 (ppm)

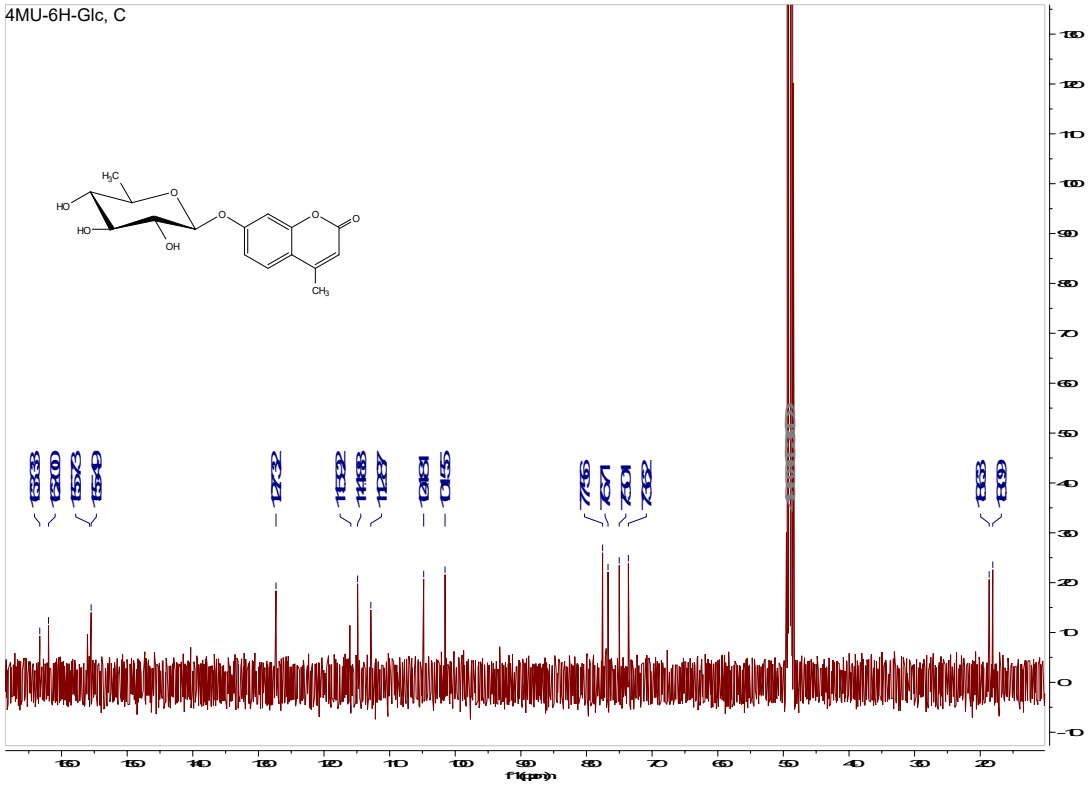


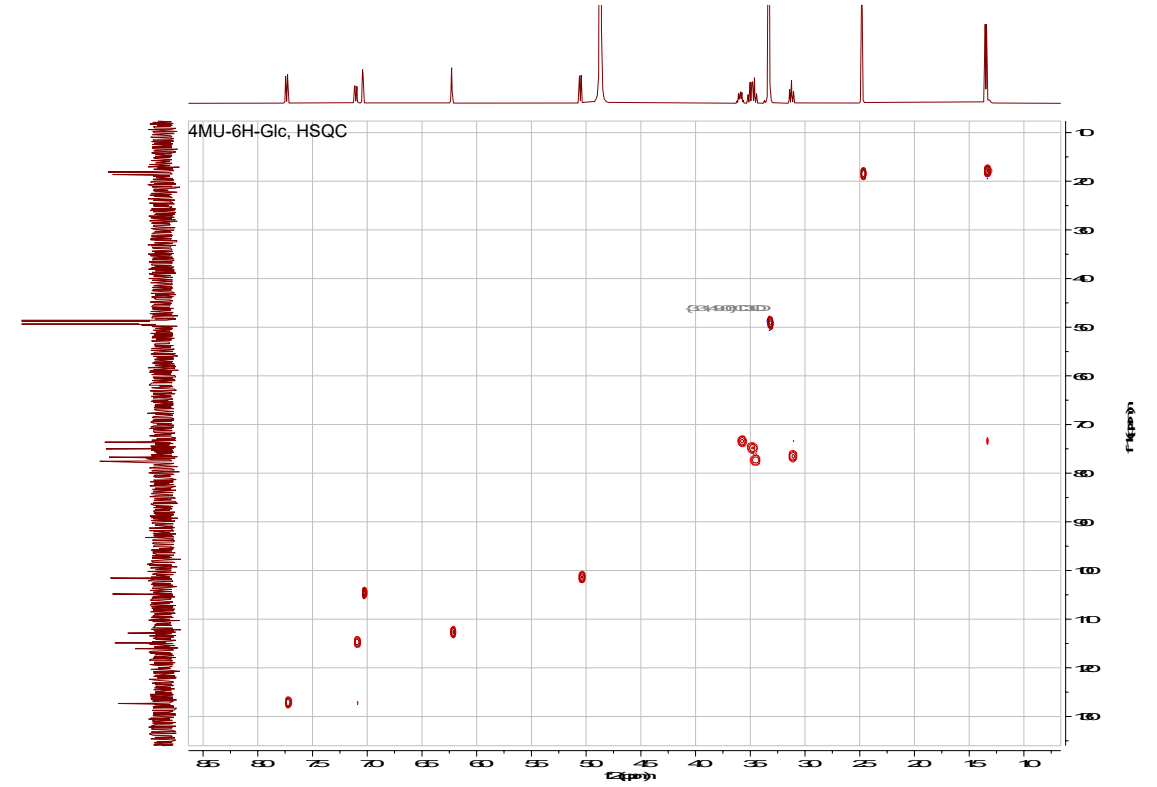


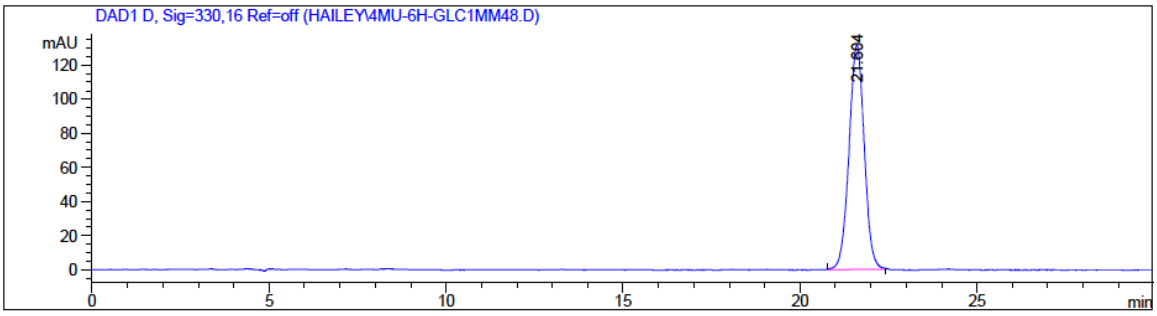
Retention time (min)	Area (mAU ∗ s)	Height (mAU)	Area %
10.415	4053.964	214.617	100.0000

Figure S. 6. ^{1}H NMR, ^{13}C NMR, HSQC spectra and HPLC trace of 4MU-4H-Glc (6)


4MU-6H-Glc, H
A (d) 7.74 J(8.82)
B (dd) 7.11 J(8.81, 2.32)
C (d) 7.04 J(2.29)
D (s) 6.23
E (d) 5.05 J(7.34)
F (m) 3.59
H (t) 3.12 J(8.97)
I (s) 2.48
J (d) 1.35 J(6.17)
K (m) 3.50
L (t) 3.46 J(8.89)
1.00
1.21
1.13
1.08
1.04
1.13
1.06
1.09
1.05
3.12
3.09
f1 (ppm)

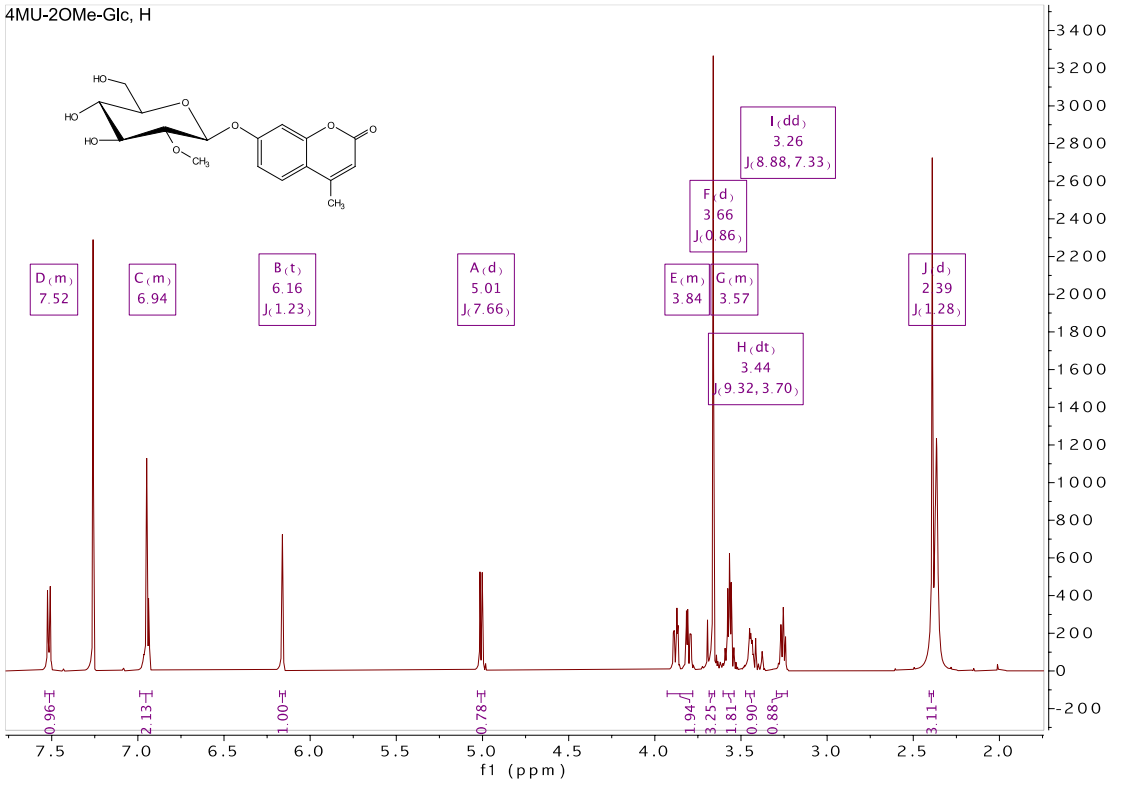

4MU-6H-Glc, C

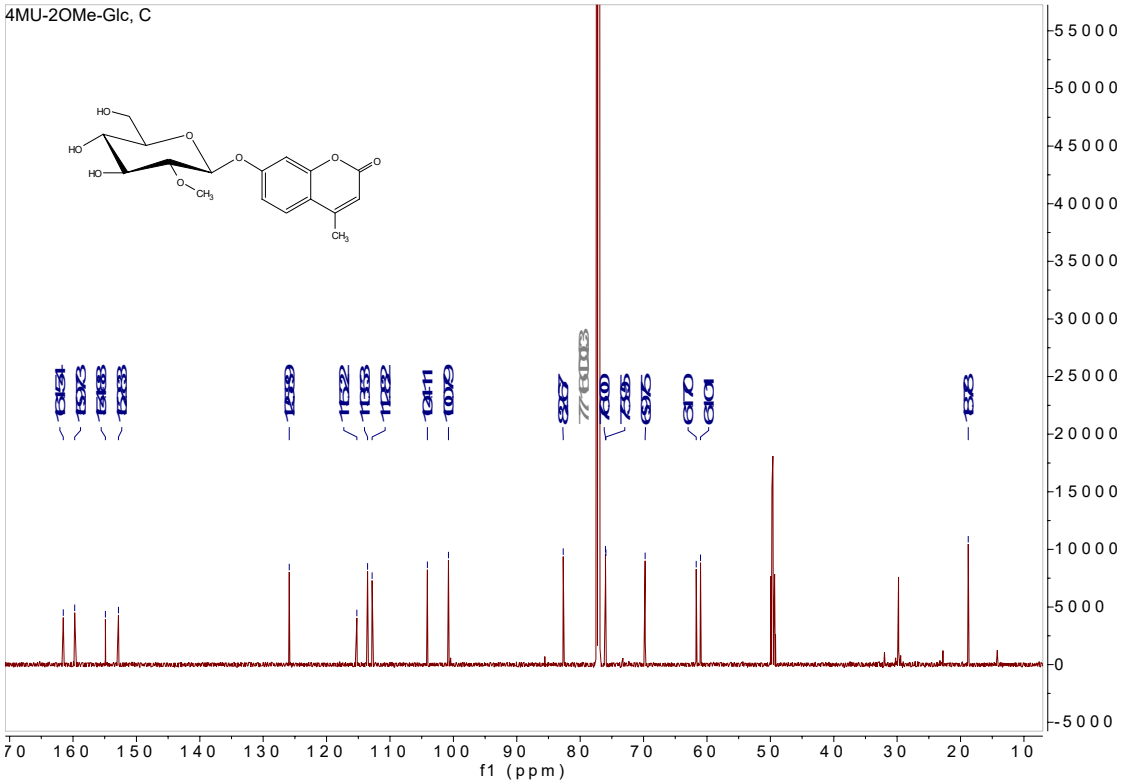


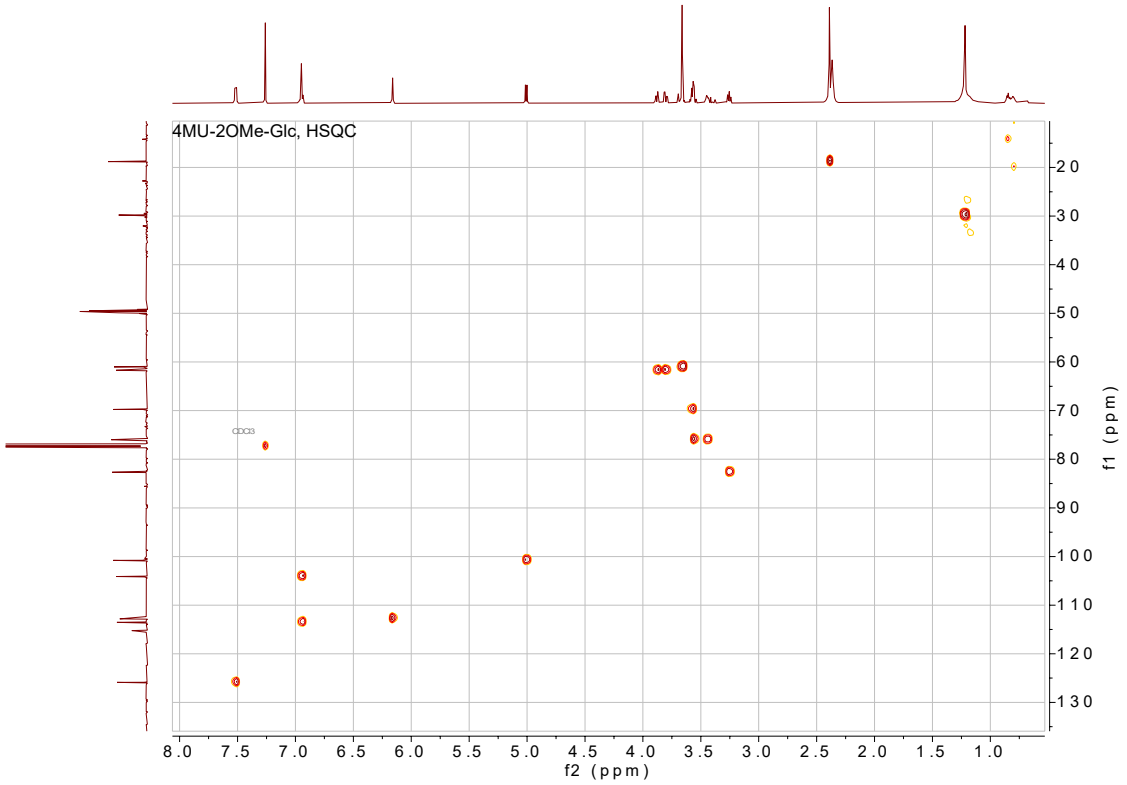


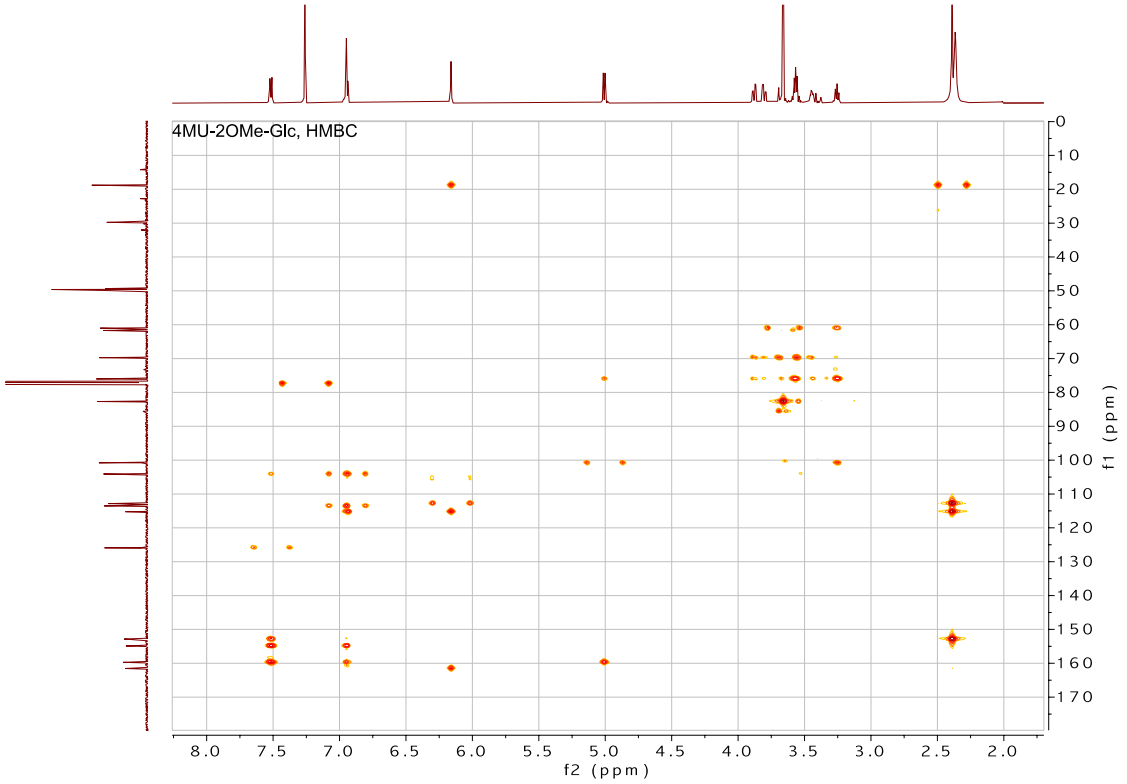
Retention time (min)	Area (mAU ∗ s)	Height (mAU)	Area %
21.604	3936.242	131.593	100.0000

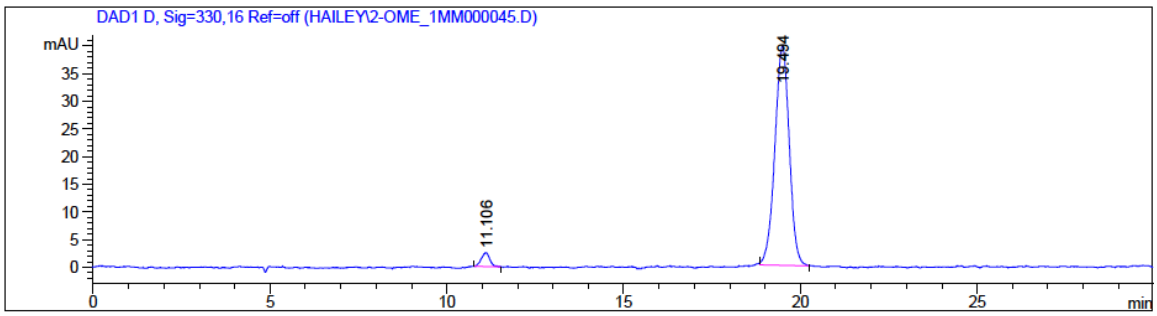
Figure S. 7. ^{1}H NMR, ^{13}C NMR, HSQC spectra and HPLC trace of 4MU-6H-Glc (7)


4MU-2OMe-Glc, H
D (m) 7.52
C (m) 6.94
B (t) 6.16 J(1.23)
A (d) 5.01 J(7.66)
E (m) 3.84
G (m) 3.57
F (d) 3.66 J(0.86)
I (dd) 3.26 J(8.88, 7.33)
H (dt) 3.44 J(9.32, 3.70)
J (d) 2.39 J(1.28)
0.96
2.13
1.00
0.78
1.94
3.25
1.81
0.90
0.88
3.11
f1 (ppm)


4MU-2OMe-Glc, C
f1 (ppm)

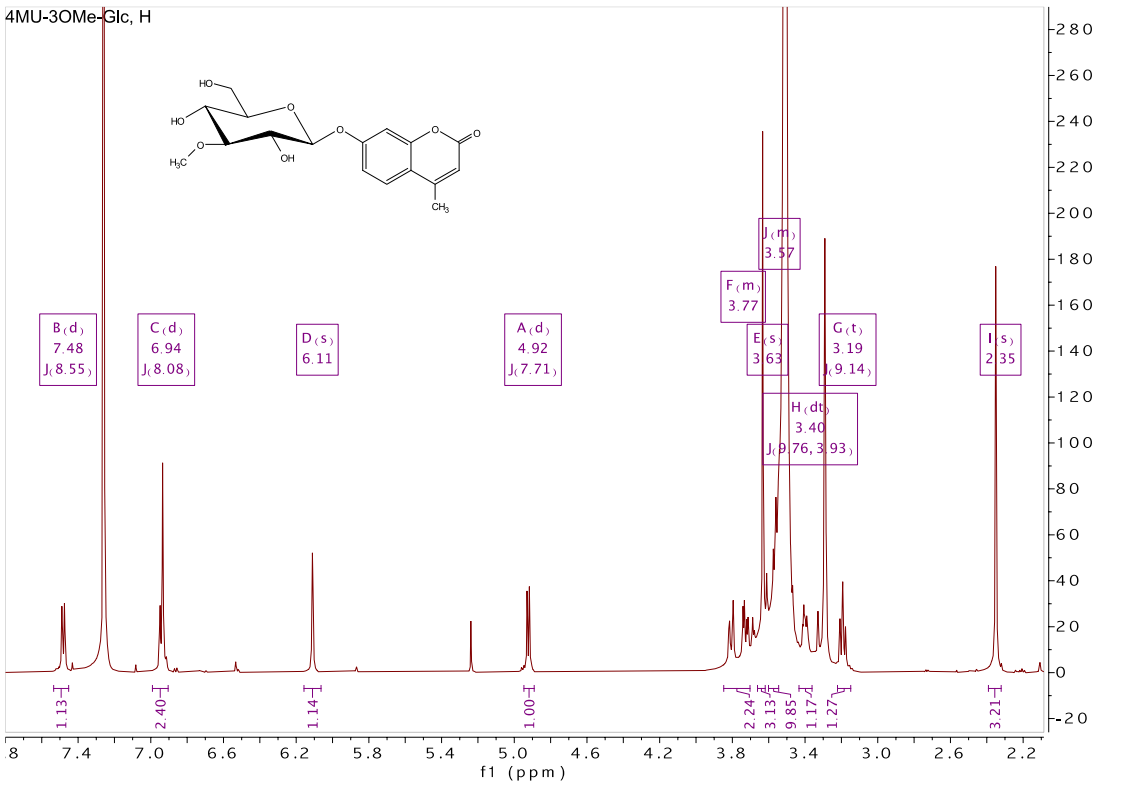

4MU-2OMe-Glc, HSQC
f2 (ppm)
f1 (ppm)

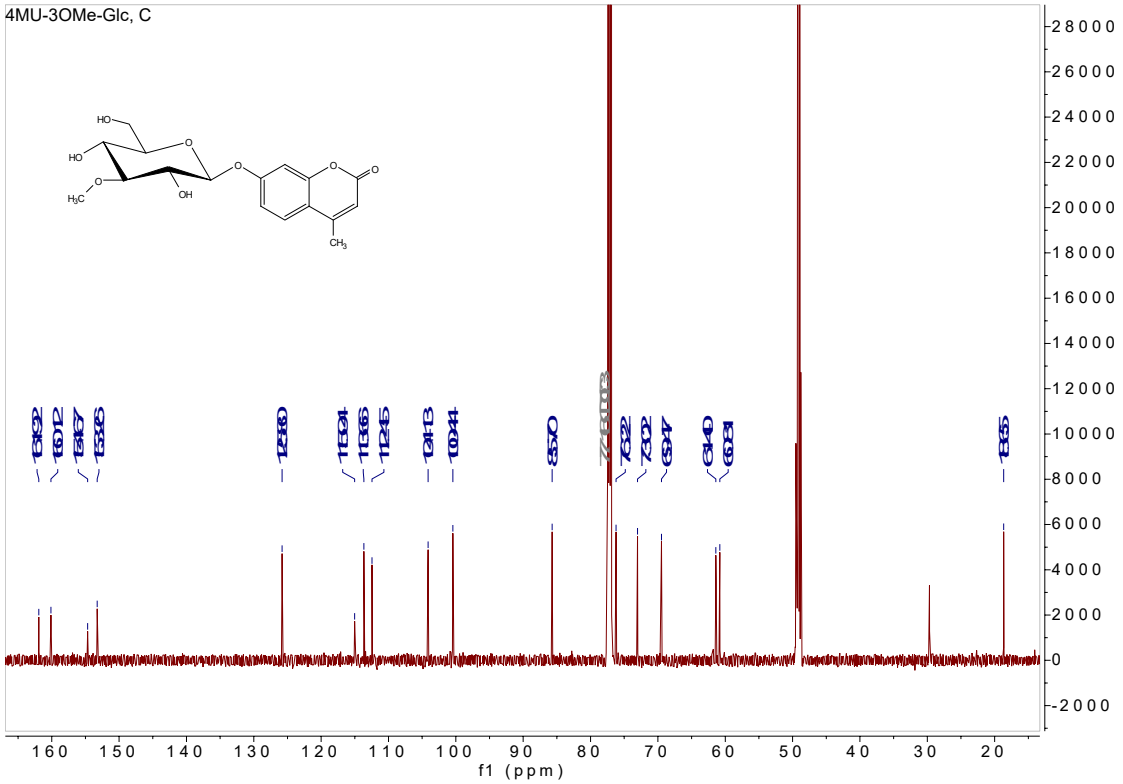

4MU-2OMe-Glc, HMBC
f2 (ppm)
f1 (ppm)

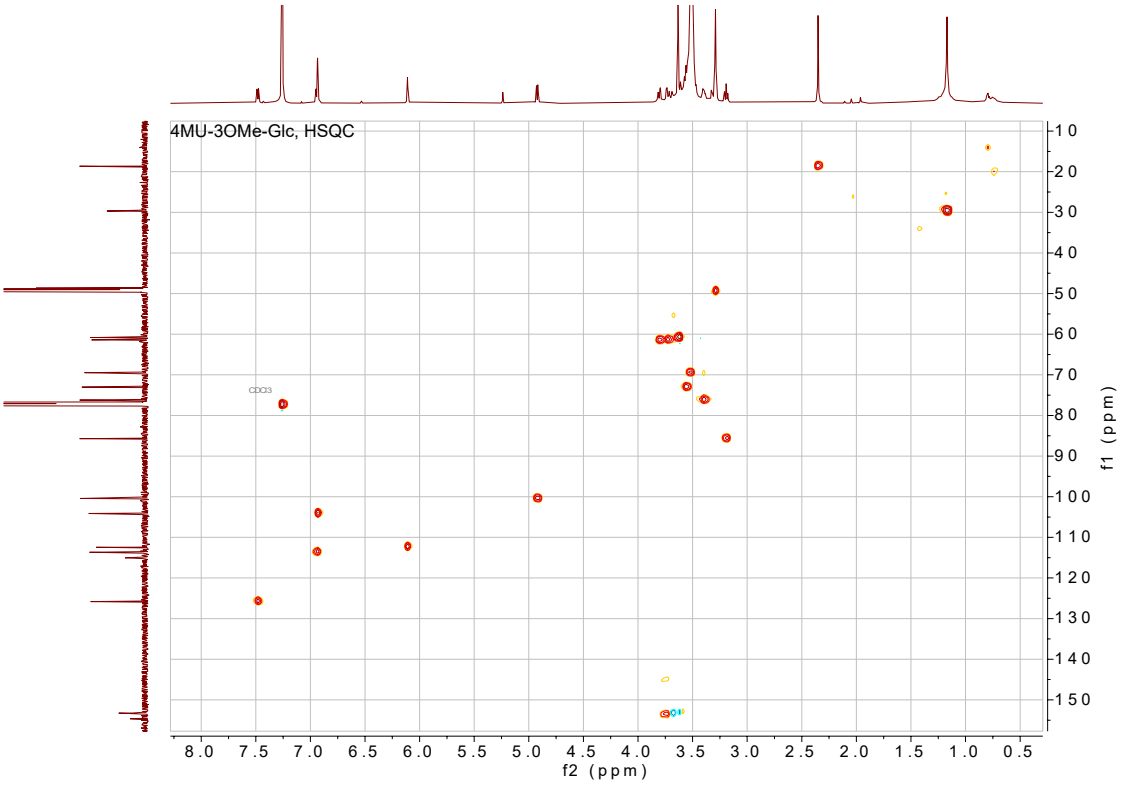


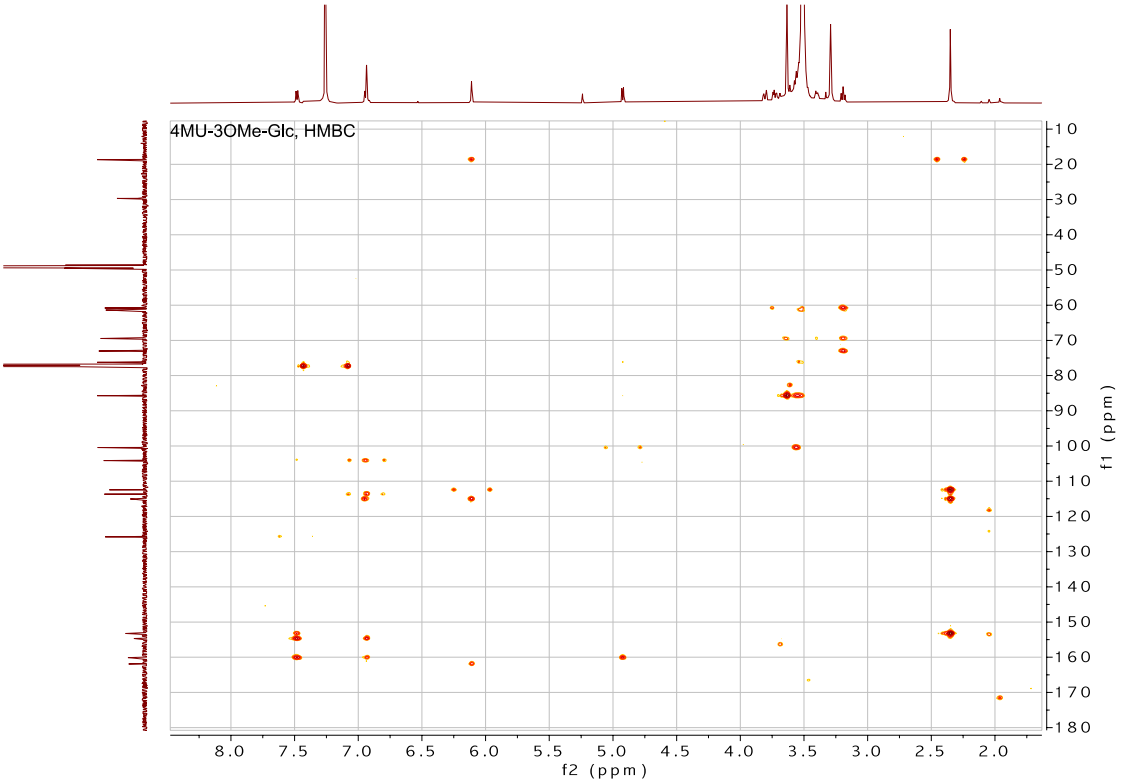
Retention time (min)	Area (mAU ∗ s)	Height (mAU)	Area %
11.106	42.289	2.538	3.6287
19.494	0.4310	1123.121	96.3713

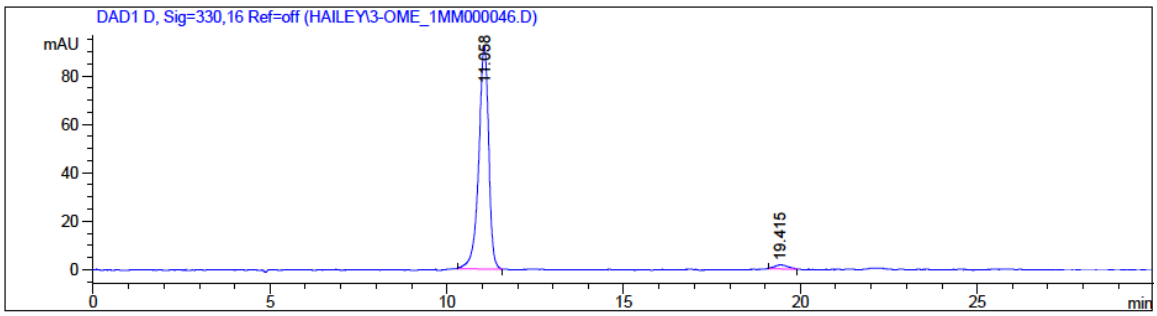
Figure S. 8. ^{1}H NMR, ^{13}C NMR, HSQC, HMBC spectra and HPLC trace of 4MU-2OMe-Glc (8)


4MU-3OMe-Glc, H
B (d) 7.48 J(8.55)
C (d) 6.94 J(8.08)
D (s) 6.11
A (d) 4.92 J(7.71)
J (m) 3.57
F (m) 3.77
E (s) 3.63
G (t) 3.19 J(9.14)
H (dt) 3.40 J(9.76, 3.93)
I (s) 2.35
1.13
2.40
1.14
1.00
2.24
3.13
9.85
1.17
1.27
3.21
f1 (ppm)


4MU-3OMe-Glc, C
f1 (ppm)


4MU-3OMe-Glc, HSQC
CDCl3
f2 (ppm)
f1 (ppm)


4MU-3OMe-Glc, HMBC
f2 (ppm)
f1 (ppm)



Retention time (min)	Area (mAU ∗ s)	Height (mAU)	Area %
11.058	1808.712	92.057	97.6214
19.415	44.069	1.662	2.3786

Figure S. 9. ^{1}H, ^{13}C, HMBC NMR spectra and HPLC trace of 4MU-3OMe-Glc (9)