Development of Chemical Tools for Studying Human *O*-GlcNAcase Activity

by

Nevena Cekic

B.Sc., University of Ottawa, 2012

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 2015

Approval

Name: **Nevena Cekic**

Degree: **Master of Science (Chemistry)**

Title: ***Development of Chemical Tools for Studying Human O-GlcNAcase Activity***

Chair: Krzysztof Starosta
Associate Professor

Examining Committee:

David J. Vocadlo
Senior Supervisor
Professor

Andrew J. Bennet
Supervisor
Professor

Tim Storr
Supervisor
Associate Professor

George R. Agnes
Internal Examiner
Professor

Date Defended/Approved: November 18, 2015

Abstract

In recent years, the post-translational modification of nuclear and cytoplasmic proteins with *O*-linked *N*-acetylglucosamine (*O*-GlcNAc) has emerged as playing diverse roles in health and disease. Interestingly, this modification is regulated by only two enzymes, *O*-GlcNAc transferase (OGT) and *O*-GlcNAcase (OGA). A method to study the effects of elevated levels of *O*-GlcNAc is to selectively target OGA. Herein, we describe the structure-activity relationships of a family of 2'-aminothiazoline-based inhibitors, one of which shows to be among the most potent inhibitors of human OGA (hOGA) known to date. We present the selectivity ratios of these compounds for hOGA over the structurally-related lysosomal β-hexosaminidases, define them as transition state analogues and rationalize their potencies by using linear free energy analyses. We also identify two fluorescence quenched substrates for hOGA bearing thioamide quenchers having different fluorogenic leaving groups, which reveal design features for substrates to monitor hOGA activity in live cells.

Keywords: Glycoside hydrolases; *O*-GlcNAcase; *O*-GlcNAc modification; transition state analogy; quenched substrates

.

I dedicate this thesis to my parents, Dragan and Slavojka, and my little brother, Marko.

Acknowledgements

I would like to thank my supervisor, Dr. David Vocadlo, for his mentorship and words of wisdom throughout the past few years. Working in the Vocadlo research group has truly been amazing and I am grateful for the skills which I have acquired along the way. I thank David for his optimistic outlook on research - and life in general - and for his continuous patience and encouragement. I would also like to thank my supervisory committee members, Dr. Andrew Bennet and Dr. Tim Storr for taking the time to oversee my research and provide feedback throughout, as well as my internal examiner, Dr. George R. Agnes.

I thank Dr. Andrew Lewis and Colin Zhang for their assistance with NMR spectroscopy, and the wonderful graduate secretaries during my time at SFU, Nathalie Fournier and Lynn Wood.

I sincerely thank Dr. Renee Mosi for being an incredible role model and taking me under her wing so many times - I am forever grateful for her guidance with enzyme kinetics, her constant encouragement, generosity and invaluable friendship.

I express enormous gratitude to the members of the Vocadlo lab, particularly Dr. Samy Cecioni for his patience, guidance in chemistry and for being a great friend. I thank Dr. Razieh Eskandari for her input in chemistry, for her friendship and for sharing an office with me this entire time! Thank you to David Shen for his help with enzymatic assays and for always sharing his delicious food with me - you were a blast to work with. Thank you to all of the Vocadlo lab members for their collaborative and positive attitudes, hilarious times together, wine-and-cheese, poker nights, pub nights, squash mornings and many other memorable times - I could not have asked for more. I thank you all (current and former members): Isaac, Evan, Nancy, Hong-Yee, Dave, Lianne, Xiaoyang, Samy, Matt A., Laleh, Steve, Matt D., Mike, Ta-Wei, Yan, Wes, Zarina, Scott, Kelsey, Tom and Hayden. I also must acknowledge all of the wonderful members of the Bennet lab who made great allies with us in the carbohydrate wing. Finally, thank you to my dear friend Viviana Cerda for her constant support and for always being good company. Your friendship means the world to me!

Table of Contents

List of Tables

List of Figures

List of Schemes

Glossary

∆E	Change in electric potential
4MU	4-methylumbelliferone
4MU-GlcNAc	4-methylumbelliferyl *N*-acetyl-β-D-glucosaminide
4-MUGlu	4-methylumbelliferyl β-D-glucoside
6-Ac-Cas	6-acetamido-6-deoxy-castanospermine
ABP	Activity-based probe
ABPP	Activity-based protein profiling
Ac	Acetyl group
Acetyl CoA	Acetyl coenzyme A
AcOH	Acetic acid
AD	Alzheimer Disease
AGM	*N*-acetylglucosamine phospho-glucomutase
AMI	Acute myocardial infarction
a-PET	Acceptor-excited photoinduced electron transfer
ATP	Adenosine triphosphate
BACE1	β-secretase 1
BHQ	Black hole quencher
BODIPY	Boron dipyrromethene
CAZy database	Carbohydrate-Active Enzymes database
CBE	Conduritol-based epoxide
$CDCl_3$	Deuterated chloroform
DAB	Dimethylaminophenylazophenyl
DABNAc	2-acetamido-1,4-imino-1,2,4-trideoxy-D-arabinitol
DAJNAc	2-acetamido-1,2-dideoxy-D-allonojirimycin
DCM	Dichloromethane
DDIM	2,5-dideoxy-2,5-imino-D-mannitol
DGJNAc	2-acetamido-1,2-dideoxy-D-*galacto*-nojirimycin
DIPEA	*N,N*-Diisopropylethylamine
DMAP	*N,N*-Dimethylamino pyridine
DMF	Dimethylformamide
DNJ	Deoxynojirimycin

DNJNAc	2-acetamido-1,2-dideoxy-D-gluconojirimycin
d-PET	Donor-excited photoinduced electron transfer
EC_{50}	Half maximal effective concentration
ECM	Extracellular matrix
EDANS	5-(2-aminoethyl) amino-1-napthalene-sulfonic acid
ERT	Enzyme replacement therapy
ESI/MS/MS	Electrospray ionization tandem spectrometry
EtOAc	Ethyl acetate
F6P	Fructose-6-phosphate
FDGlcNAc	fluorescein di(*N*-acetyl-β-D-glucosaminide)
Fmoc	*N*-fluorenylmethyloxycarbonyl
FRET	Förster Resonance Energy Transfer
G	Guanine
G6P	Glucose-6-phosphate
GBA1, GBA2 and GBA3	Glucocerebrosidase 1, 2 and 3
GFAT	Glutamine fructose-6-phosphate amidotransferase
GH	Glycoside hydrolase
GlcN6P	Glucosamine-6-phosphate
GlcNAc	*N*-acetylglucosamine
GlcNAc1P	*N*-acetylglucosamine-1-phosphate
GlcNAc6P	*N*-acetylglucosamine-6-phosphate
GNAT	Glucosamine-6-phosphate acetyltransferase
HaCaT cells	Spontaneously immortalized human keratinocyte cells
HBSP	Hexosamine biosynthetic pathway
HBTU	2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HEK 293 cells	Human embryonic kidney 293 cells
hHexB	Human lysosomal β-hexosaminidase
HK-1	Hexokinase-1
hOGA	Human OGA
HOMO	Highest occupied molecular orbital
IC_{50}	50% Maximal inhibitory concentration
IPC	Ischemic preconditioning

IR	Infrared
K_2CO_3	Potassium carbonate
k_{cat}	Catalytic rate constant (turnover number)
k_{cat}/K_m	Catalytic efficiency; second-order rate constant
K_i	Inhibition constant
k_{inact}	Rate of inactivation for mechanism-based inhibitors
K_m	Michaelis constant
K_m^{app}	Apparent Michaelis constant of a substrate in the presence of an inhibitor
$KMnO_4$	Potassium permanganate
LABNAc	2-acetamido-1,4-imino-1,2,4-trideoxy-L-arabinitol
LacZ	*Escherichia coli* β-galactosidase
LE	Ligand efficiency
LFER	Linear free energy relationship
LPH	Lactase/phlorizin hydrolase
LR	Lawesson's reagent
LSD	Lysosomal storage disorder
LUMO	Lowest unoccupied molecular orbital
MBI	Mechanism-based inhibitor
MEB-4 cells	Mouse melanoma cells
$MgSO_4$	Magnesium sulfate
mRNA	Messenger ribonucleic acid
NaOMe	Sodium methoxide
NAT2	*N*-acetyl transferase 2
NButGT	1,2-dideoxy-2'-propyl-α-D-glucopyranoso-[2,1-d]-Δ2'-thiazoline
N-butyl-DNJ	*N*-butyl-Deoxynojirimycin
NLGCase	Glucosylceramidase
NMR	Nuclear magnetic resonance
OGA	β-*N*-acetylglucosaminidase
O-GlcNAc	*O*-linked *N*-acetylglucosamine
OGT	UDP-*N*-acetyl-D-glucosamine:polypeptide-*N*-acetylglucosaminyl transferase
ϕ	Quantum yield
PEG	Poly(ethylene glycol)

PET imaging	Positron emission tomography imaging
PGI	Phosphoglucose isomerase
pK_a	-Log of acid dissociation constant, K_a
PUGNAc	*O*-(2-acetamido-2-deoxy-D-glycopyranosilidene)amino-*N*-phenylcarbamate
RAW cells	Mouse leukaemic monocyte macrophage cells
R_o	Förster radius
S_1	Excited singlet state
Ser	Serine
$SnCl_4$	Tin (IV) chloride
S_o	Ground state
TBAHS	Tetrabutylammonium hydrogen sulfate
TFA	Trifluoroacetic acid
ThiamEt-G	1,2-dideoxy-2'-ethylamino-α-D-glucopyranoso-[2,1-d]-Δ2'-thiazoline
Thr	Threonine
TLC	Thin layer chromatography
Trp	Tryptophan
TS	Transition state
UDP	Uridine diphosphate
UDP-[^{3}H] galactose	Radiolabelled UDP-galactose at H-3
UDP-GlcNAc	UDP-β-*N*-acetylglucosamine
UDP-GlcNAcPP	UDP-*N*-acetylglucosamine pyrophosphorylase
UTP	Uridine triphosphate
UV	Ultraviolet
V_{max}	Maximal velocity of an enzyme; first-order rate constant
β-MAP	β-secretase membrane-anchored probe
λ_{max}	Absorption maximum

Chapter 1.

An overview of glycoside hydrolases and their inhibitors, with a focus on the *O*-GlcNAc modification and *O*-GlcNAc hydrolase

1.1. *O*-GlcNAc is a dynamic post-translational modification

"Carbohydrate", a word which originated more than 100 years ago, describes any sugar-based substance. The word stems from the original term "hydrate of carbon" and represents naturally occurring compounds with the formula $C_x(H_2O)_n$ which also contain a ketone or an aldehyde.[1] Commonly, when carbohydrates are considered in everyday situations, what comes to mind for most people is their diet and often the avoidance of carbohydrates for health or weight-control based reasons. However, carbohydrates play crucial roles beyond serving as an energy source for organism survival. Sugars can exist as monosaccharides (*e.g*. glucose), disaccharides, or polymers referred to as oligo- or polysaccharides. They may also be conjugated to macromolecules such as proteins or lipids through α- or β-linked glycosidic linkages (Figure 1.1 A, B, respectively). Carbohydrates linked together are often referred to as glycans and when a monosaccharide or glycan is linked to another biomolecule, the resulting compound is termed a glycoconjugate. Many different glycoconjugates are densely concentrated on the cell surface and also make up a large part of the extracellular matrix (ECM). These species play vital roles in cell communication, cell viability, immune response and as receptors essential for the binding of pathogens.[2]

2-Acetamido-2-deoxy-D-glucopyranose, which is also known by the non-systematic name *N*-acetylglucosamine (GlcNAc), is a monosaccharide that is structurally similar to glucose but performs drastically different functions (Figure 1.1, C). GlcNAc plays a critical role as a post-translational modification of both intracellular and

extracellular proteins.[3] One of the most abundant forms of glycosylation is found within the nucleocytoplasm and mitochondria of cells in all multicellular eukaryotes studied to date, ranging from humans through to plants.[1, 4] This protein modification, known as *O*-GlcNAc or *O*-linked GlcNAc, involves the attachment of GlcNAc to the hydroxyl group of serine and threonine residues of proteins (Figure 1.2) through a β-glycosidic linkage. It contrasts with *N*-glycans which are formed by the addition of glycans containing GlcNAc at their core to specific asparagine residues found within particular amino acid sequences of proteins expressed within the secretory pathway of cells.[5, 6]

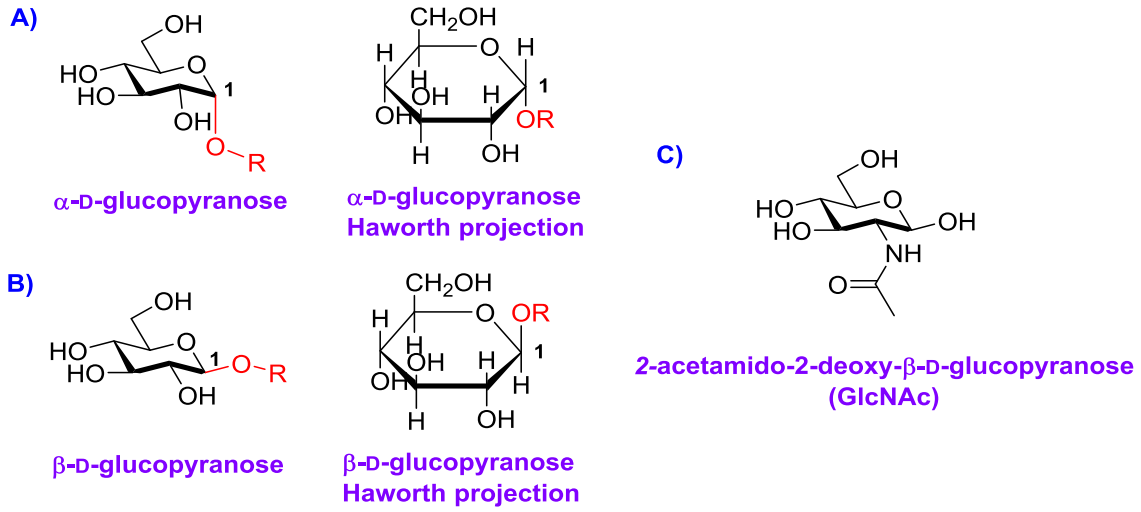


Figure 1.1. Types of glycosidic linkages for D-sugars and the structure of GlcNAc.
A) An α-glycosidic linkage has the hydroxyl group on the **C**-1 anomeric centre projecting downwards in a Haworth projection. B) A β-glycosidic linkage has the hydroxyl group on the **C**-1 anomeric centre projecting upwards in a Haworth projection. C) The structure of *N*-acetylglucosamine (GlcNAc).

The importance of understanding the *O*-GlcNAc protein modification stems from the many important roles played by *O*-GlcNAc within cells, including modification of other proteins which are responsible for various cellular functions such as cellular signalling[5, 7-9], DNA transcription[2, 10], regulation of cytoskeletal structure[11] and stress response[12-15]. For example, cytoskeletal proteins[16-20], kinases[21-23] and nucleoporins[24-26] themselves are all modified with *O*-GlcNAc, which emphasizes the direct impact of *O*-GlcNAc on proteins crucial for cellular structure and function. The *O*-GlcNAc modification was discovered by Torres and Hart in 1984 when terminal GlcNAc

residues were described on B-, T- and thymic lymphocyte populations, by using a β-1-4 galactosyltransferase radiolabeled donor substrate (UDP-[^{3}H] galactose) with bovine milk galactosyltransferase to probe terminal GlcNAc residues on the lymphocyte surface.[27]

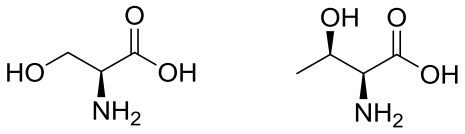


Figure 1.2. Amino acids modified with GlcNAc.
L-serine (left) and L-threonine (right) on various proteins are linked to GlcNAC via an *O*- β-glycosidic linkage.

Interestingly, the *O*-GlcNAc modification is regulated only by two enzymes, UDP-*N*-acetyl-D-glucosamine:polypeptide-*N*-acetylglucosaminyl transferase (OGT)[28, 29] and β-*N*-acetylglucosaminidase (OGA)[30, 31]. This is in contrast to serine and threonine phosphorylation which is regulated by hundreds of kinases and phosphatases that together comprise genes encoding more than 2% of the genome.[32]

Both *O*-GlcNAc processing enzymes are found within nucleocytoplasmic and mitochondrial cellular compartments[33, 34], with a higher abundance of OGT in the nucleus and OGA within the cytosol, but curiously, a higher abundance of OGA in the nucleolus, from which OGT is excluded[35, 36]. The source of GlcNAc used by OGT is the nucleotide sugar UDP-β-*N*-acetylglucosamine (UDP-GlcNAc), which is derived from glucose and is synthesized via the hexosamine biosynthetic pathway (HBSP) (Figure 1.3). Once glucose is actively transported across the cell membrane, it is either committed to glycolysis to form pyruvate, is stored as glycogen for later use, or is transformed by the HBSP to form UDP-GlcNAc. This last pathway consumes ~2-5% of glucose taken up by cells,[37] and has been proposed to be a nutrient-sensing pathway since the amount of UDP-GlcNAc produced has been experimentally shown to depend on the amount of dietary glucose[38-40]. Increased glucose flux through the HBSP has been found to contribute to glucose-induced insulin resistance in animals, cells and tissues.[38-42]

Figure 1.3 shows the detailed HBSP through which UDP-GlcNAc is synthesized. After active transport into the cell via a Glut transporter, glucose is first transformed into

glucose-6-phosphate (G6P) by hexokinase 1 (HK-1) and is then converted into fructose-6-phosphate (F6P) by phosphoglucose isomerase (PGI). F6P continues into glycolysis to form pyruvate or enters the HBSP to be transformed into glucosamine-6-phosphate (GlcN6P) by glucosamine-fructose 6-phosphate amino transferase (GFAT), during the rate-limiting step in the HBSP.[43] This last transformation involves a transamination of ammonia from a glutamine donor to the carbonyl of F6P, followed by an isomerization from the imine-containing open chain furanose to the pyranose GlcN6P. Acetyl Coenzyme A (Acetyl CoA) resulting from fatty acid metabolism is then used as an acetyl group donor that is transferred onto GlcN6P by glucosamine-6-phosphate acetyltransferase (GNAT) to yield *N*-acetylglucosamine-6-phosphate (GlcNAc6P). This is followed by an isomerization catalysed by *N*-acetylglucosamine phospho-glucomutase (AGM) which transfers the phosphate group from the 6-position of the sugar to the α-anomeric position, producing *N*-acetylglucosamine-1-phosphate (GlcNAc1P). Lastly, transfer of uridine diphosphate (UDP) from uridine triphosphate (UTP), which is derived from nucleic acid metabolism, onto GlcNAc1P in a process catalyzed by UDP-*N*-acetylglucosamine pyrophosphorylase (UDP-GlcNAcPP) yields UDP-GlcNAc. This nucleotide donor sugar can then be used as a substrate by OGT to transfer GlcNAc onto proteins. Acting in opposition, OGA catalyses the hydrolysis of GlcNAc from proteins.

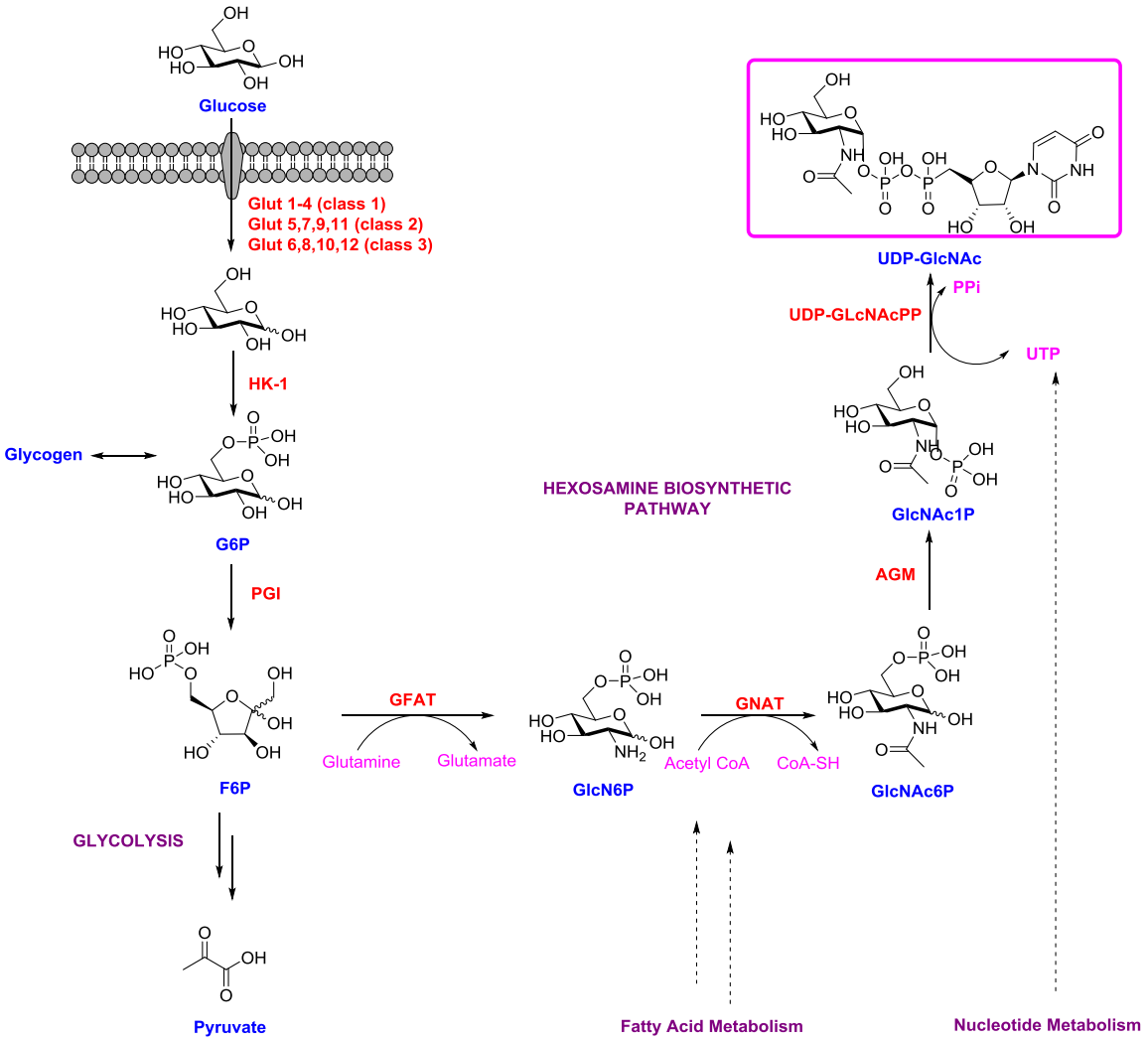


Figure 1.3. **The Hexosamine Biosynthetic Pathway.**
After glucose is modified by hexokinase 1 (HK-1) to make glucose-6-phosphate (G6P) and phosphoglucose isomerase (PGI), to make fructose-6-phosphate (F6P), it can enter the hexosamine biosynthetic pathway (HBSP), a pathway which consumes 2-5% of glucose. Enzymes in the HBSP: glutamine-fructose-6-phosphate-transaminase (GFAT); gluosamine-6-phosphate-acetyl transferase (GNAT); *N*-acetylglucosamine phospho-glucomutase (AGM); UDP-*N*-acetylglucosamine pyrophosphorylase (UDP-GlcNAcPP); Products of the HBSP: glucosamine-6-phosphate (GlcN6P); *N*-acetylglucosamine-6-phosphate (GlcNAc6P); *N*-acetylglucosamine-1-phosphate (GlcNAc1P); UDP-*N*-acetylglucosamine (UDP-GlcNAc).

1.1.1. The *O*-GlcNAc modification is a regulatory mechanism coupled with phosphorylation

Interestingly, *O*-GlcNAc has been shown to be able to modify proteins at the same or proximal sites as protein phosphorylation (Figure 1.4).[30, 41, 44, 45] It has been known since the early 1990's that there are some *O*-GlcNAc modified proteins that exist in a dynamic equilibrium with phosphorylation [46, 47], which has been described as a fundamental regulatory mechanism that may influence a spectrum of nucleocytoplasmic proteins[48-50]. This potential competition between *O*-GlcNAc or phosphate may occur on the same serine/threonine sites in a reciprocal and dynamic manner, as is displayed in Figure 1.4. Alternatively, modification can be competitive, and can occur on proximal sites, or simultaneously, where the protein can be both phosphorylated and GlcNAcylated at once on separate sites.[35] For example, the nuclear phosphoprotein and transcription factor, c-Myc, plays the important role of regulation of gene expression and affects processes such as apoptosis and cell proliferation. c-Myc is an example of a protein which is glycosylated and phosphorylated in a reciprocal fashion on its Thr-58 site, a known site which is frequently mutated in lymphomas.[51, 52]

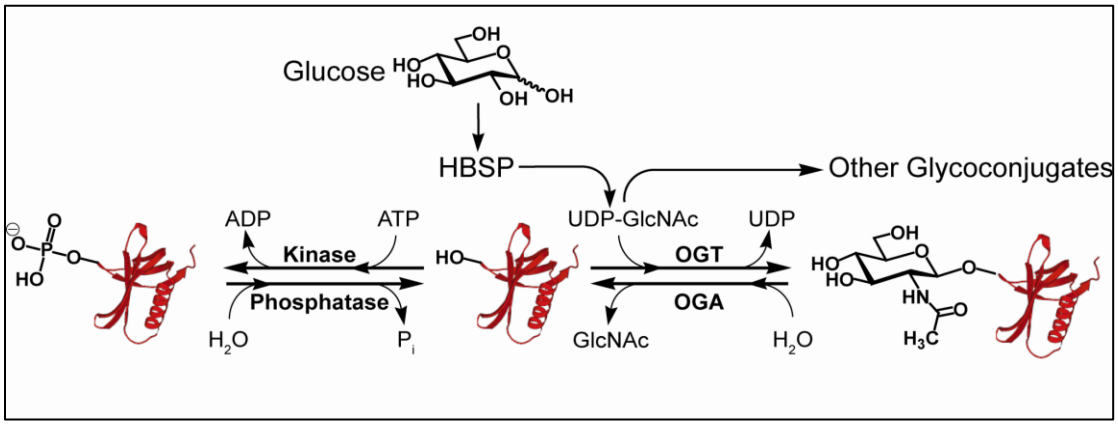


Figure 1.4. The reciprocal relationship between phosphorylation and *O*-GlcNAc modification of some intracellular proteins.
The protein being modified, and UDP-GlcNAc emerging from the HBSP pathway, serve as substrates for OGT. The transfer of GlcNAc onto the protein serine/threonine residue results in the formation of a β-glycosidic linkage between the sugar and protein. OGA hydrolyzes this glycosidic linkage and the unmodified protein is either re-modified with GlcNAc or phosphorylated by a kinase where ATP (Adenosine triphosphate) is a substrate. Phosphatases can hydrolyze this linkage, liberating the free phosphate and the unmodified protein which is later modified again many times throughout the protein lifespan.

This cross talk between phosphorylation and *O*-GlcNAcylation also plays an important regulatory role in chronic diseases such as neurodegenerative diseases[53-55], diabetes[56-58] and different types of cancers[59-61]. Also, the implications of *O*-GlcNAc on cardio-protection and cardiovascular dysfunction has been of growing interest in the past decade. Ischemic preconditioning (IPC) is a cardio-protective mechanism which occurs when short ischemic episodes protect the myocardium in the heart against subsequent ischemias by periodically restoring blood flow to the organ. After acute myocardial infarction (AMI) occurs, this may result in myocardial ischemia-reperfusion-injury, which is the tissue damage caused from the temporary lack of blood supply or oxygen.[62] It has been shown on numerous occasions that elevation of *O*-GlcNAc prior and subsequent to ischemia on perfused hearts reduces the amount of cell death both *in vitro* and *in vivo* during cardiomyocyte survival experiments[63-65] and in animal models[66]. These studies demonstrate that IPC could be dependent on *O*-GlcNAc and this idea has been formally suggested.[62] The cardio-protective effect of *O*-GlcNAc has been suggested to stem from the glycosylation of mitochondrial proteins

such as voltage-dependent anion channel[67], which constitutes a part of the mitochondrial permeability transition pore that has been shown to play a key role in damage caused by ischemia[68]. Accordingly, *O*-GlcNAc is a valuable area to explore in biological systems, and obtaining a better understanding of *O*-GlcNAc regulation in various acute and chronic diseases could have significant implications in medical research.

1.2. A closer look at glycoside hydrolases: the enzymes responsible for cleavage of sugar modifications from glycans

As previously mentioned, the enzymes responsible for the *O*-GlcNAc modification of proteins are OGT and OGA, which fall within different enzyme superfamilies. OGT falls within the glycosyltransferase superfamily, which are enzymes responsible for catalysing the transfer of sugars onto biomolecular substrates, such as proteins or lipids. OGA is part of the superfamily of enzymes known as the glycoside hydrolases, which cleave the glycosidic linkage between sugars or a sugar and biomolecule. Since OGA is the enzyme of central focus throughout this thesis, a general overview of glycoside hydrolases is described in this section in order to provide background information useful for the reader.

1.2.1. Glycoside Hydrolases

Glycoside hydrolases (GHs) are extremely proficient catalysts which are able to cleave glycosidic linkages between sugars and proteins or lipids with rate constants of up to 1000 s^{-1}.[69] In 1998, it was established just how impressive these catalytic efficiencies are, after Wolfenden concluded that the half life for the spontaneous hydrolysis of some glycosidic bonds between sugars is about 5 million years.[70] During the late 1990's and early 2000's, major mechanistic[71] and structural advances[72] were made for this class of enzymes.

GHs can be classified as *endo*- or *exo*-acting, depending on what part of the substrate they turn over. *Endo*-acting GHs cleave non-terminal glycosidic linkages on

conjugated or free oligo- or polysaccharides,[73] whereas *exo*-acting GHs cleave the terminal glycosidic linkage from free or conjugated oligo- or polysaccharides at the reducing or non-reducing end[74]. The structures of these categories of enzymes are differentiated by their topologies, since *endo*-acting GHs contain a shallow groove on the enzymatic surface, while the *exo*-acting GHs often contain a deeper binding pocket to be able to accommodate the terminal end of the substrate.[2]

In order to better understand the structure and function relationships for GHs, the sequence-based family classification tool, The Carbohydrate-Active Enzymes database (CAZy; http://www.cazy.org) has been widely utilized by researchers since it was first established in 1999.[75, 76] The purpose of the CAZy database is to correlate the 3D structure, sequence, and molecular mechanism of CAZymes in order to predict substrate specificity. Currently, there are 133 families of GHs,[75] which is a substantial increase since year 2000 when only 80 families were reported[77]. While the range of GH families has expanded over the last 5 years, the number of 3D structures has increased even more drastically, which should undoubtedly allow easier prediction of substrate specificity of these enzymes in the future.

1.2.2. Stereochemical outcomes of glycosidic bond hydrolysis: inversion or retention

An additional classification of GHs can be used which predicts the stereochemical outcome of the product after hydrolysis. The stereochemical outcome of a hydrolysis reaction of a glycosidic linkage can be either an inversion or retention of the configuration of the substrate at the anomeric centre.[78-80] Within enzymes from a specific GH family the same stereochemistry is almost always operative. For an enzyme catalyzed inversion, a single displacement mechanism is followed, where the nucleophilic attack of the water onto the anomeric centre is facilitated in a concerted fashion (Figure 1.5.). The scissile bonds within the transition state (TS) are polarized by two carboxyl residues within the active site, which catalyze departure of the leaving group. In the enzyme's resting state, one of the residues is de-protonated and acts in its carboxylate form as a general base, while the other carboxyl is protonated and acts in its carboxylic acid form as a general acid. In the transition state, while the glycosidic linkage

is being cleaved, partial positive charge develops at the anomeric centre which results in the development of an oxocarbenium ion-like features on the pyranose ring where the positive charge build-up is stabilized by the sugar oxygen lone pairs.[32] The net result of this reaction is a sugar hemiacetal having the opposite stereochemistry at the anomeric position as compared to the substrate, hence the use of the term "inverting mechanism".

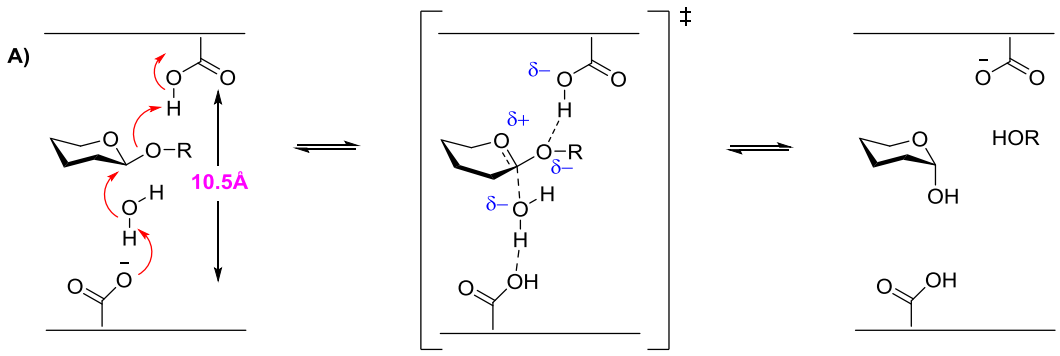


Figure 1.5. The inverting mechanism of GHs.

Retaining GHs generally act through a two step double displacement mechanism where the departure of the aglycone is typically coupled to attack of an enzymatic nucleophile at the anomeric centre, leading to glycosylation of the enzyme. This step is followed by the nucleophilic attack of water at the anomeric centre to break down the glycosyl-enzyme intermediate and liberate the sugar hemiacetal with overall anomeric retention of configuration, hence the term "retaining mechanism" (Figure 1.6). Both steps are facilitated by a second enzymatic residue - most often a carboxylic acid - which acts as a general acid in the first step to aid departure of the leaving group. In the second step, this same residue acts as a general base to aid the attack of water at the anomeric centre. The general mechanism used by retaining GHs was initially proposed by Koshland in 1953[80] and clarified over time[81].

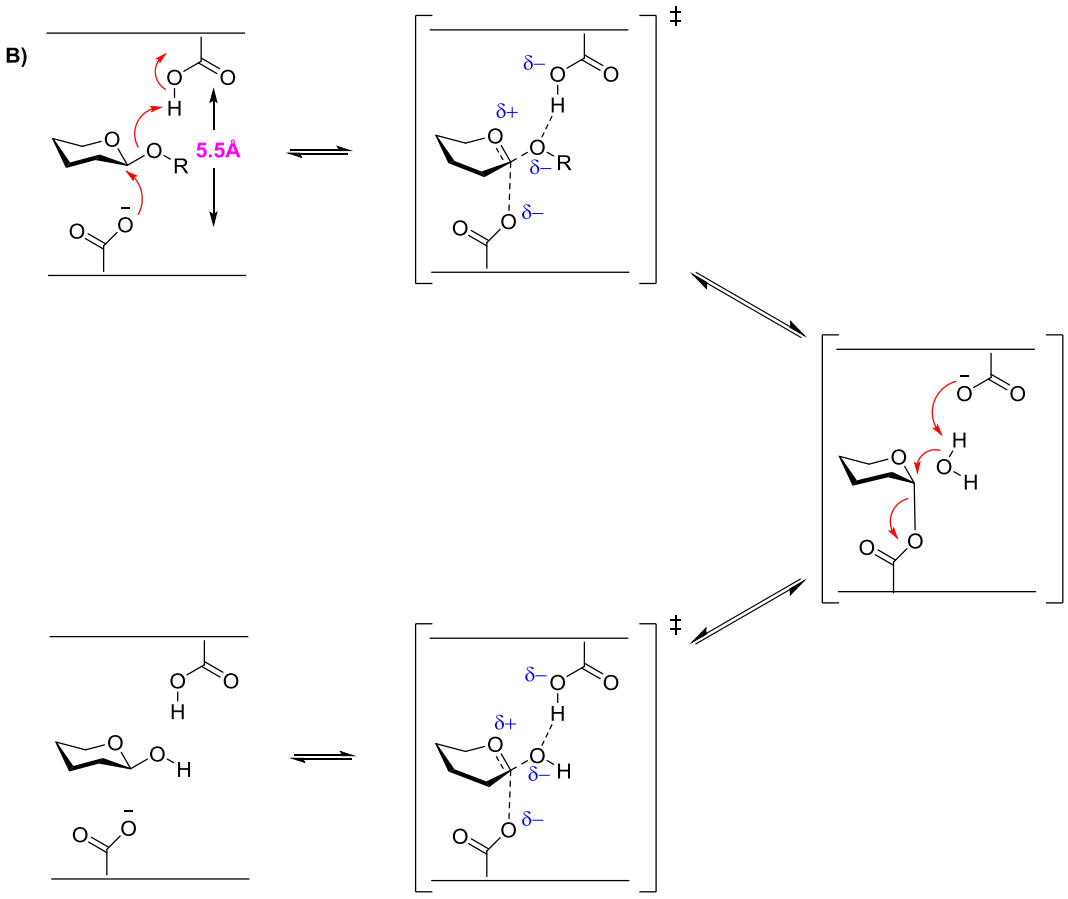


Figure 1.6. The retaining mechanism of GHs.

The distance between these carboxylates in the active site of GHs varies, depending on the type of mechanism which is used by the enzyme. The carboxylates in retaining enzymes are typically around 5.5 Å apart, and in inverting enzymes, they are approximately 10.5 Å apart.[82] This general guideline does not, however, always hold. Nevertheless, this general guideline regarding the distances between the carboxylates is not entirely surprising, since inverting enzymes require extra room to be able to accommodate a molecule of water in the active site between the general base and the anomeric centre, which contrasts with the case for retaining enzymes.

There is an alternative mechanism which accounts for the retention of stereochemistry at the anomeric position that involves neighbouring group participation from an acetamido group at the 2- position of the substrate pyranose ring (Figure 1.7). Contrary to the retaining mechanism described in Figure 1.6, this situation does not

involve the formation of a covalent glycosyl-enzyme intermediate. Instead, a carboxylate within the active site both orients and polarizes the 2-acetamido group to increase its nucleophilicity. So activated, the amide carbonyl oxygen attacks the anomeric carbon at the same time as the glycosidic bond is being cleaved through a process involving general acid catalysis. The net result is formation of an oxazoline intermediate. In the second step of the catalytic cycle, this bicyclic oxazoline intermediate undergoes nucleophilic attack by a molecule of water at the anomeric position. The general acid in the first step now serves as a general base, aiding the nucleophilic attack of water to open the oxazoline ring, leading to formation of the sugar hemiacetal with overall retention of stereochemistry at the anomeric centre.[83]

In 2005, Macauley *et al* reported that OGA uses substrate-assisted catalysis by applying Taft-like linear free energy analysis using a series of 4-Methylumbelliferyl 2-deoxy-2-*N*-fluoroacetamido-β-D-glucopyranoside substrates and mechanism-inspired inhibitors to assess whether decreasing the basicity of the 2-acetamido group on the substrate resulted in a slower turnover.[84] Ultimately, this mechanistic study concluded that OGA uses anchimeric assistance, consistent with the mechanism observed for human lysosomal *N*-acetylhexosaminidases. These latter enzymes are often referred to as β-hexosaminidases or simply hexosaminidases and are a functionally related family of enzymes which cleave terminal GlcNAc off of glycoconjugates in the lysosome. A more detailed discussion of these GHs is included in Chapter 2 of this thesis.

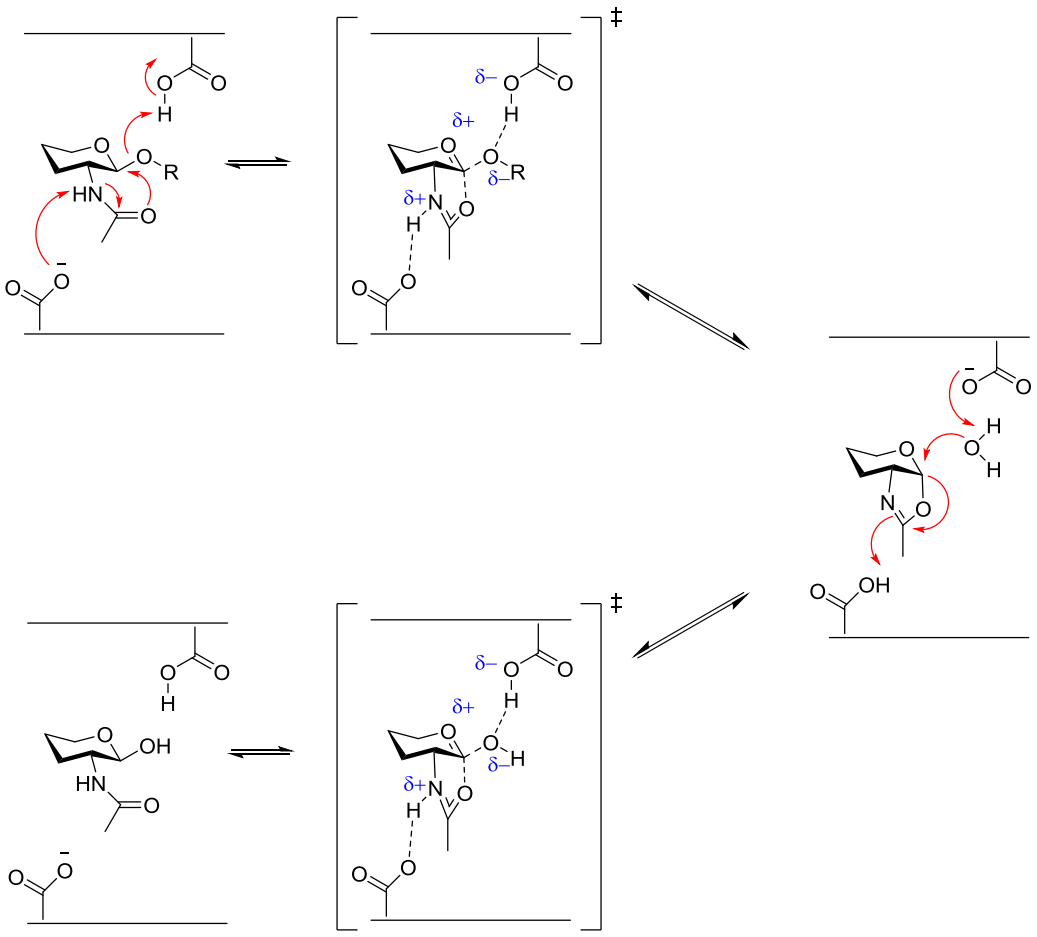


Figure 1.7. The retaining mechanism of GHs involving anchimeric assistance (substrate-assisted catalysis).
OGA and functionally related hexosaminidases use this method of catalysis.

1.3. Inhibitors of glycoside hydrolases

There are both chemical and genetic strategies that have been commonly used to manipulate the levels of GHs to study the resulting physiological effects. A classic method which has advanced the field is to manipulate expression levels of these enzymes using genetic methods to either delete genes, reduce target messenger ribonucleic acid (mRNA) levels, or increase protein levels by generating transgenic animals or inducing stable or transient over-expression. An obvious limitation of the knockout approach is that all enzyme activity is abolished, as is the enzyme itself. Critical protein-protein interactions may be lost as a result and, further, the complete loss

of activity may not be tolerated. These effects could lead to the organism no longer being viable. Furthermore, developmental effects may complicate interpretation of the results obtained using genetic methods and obscure the effect of losing some enzyme activity within the adult organism. All of these challenges may make it somewhat difficult to explore the effects of altered glycan processing using genetic methods.

The alternative - and medicinally relevant - approach to manipulating enzymes is to use inhibitors. The clear advantage to this strategy is that enzyme activity can in principle be selectively down-regulated without completely abolishing its activity as compared to gene knock out. Ideally, a high quality inhibitor will be potent towards the target enzyme with a low K_i value *in vitro* when using buffer conditions mimicking the *in vivo* environment. Even more importantly, the inhibitor will be selective towards its target and would show minimal or no activity towards enzymes that are functionally related. Stability in aqueous solutions is also crucial because this property is important for use of inhibitors *in vivo*. Additionally, a modest cLogP value (logarithm of the partition coefficient of the compound between *n*-octanol and water) of between 1 and 4 benefits absorption and biodistribution of the inhibitor *in vivo*.[85] Analytical studies should be used to reveal changes in appropriate biomarkers such as specific types of glycosylation. Since a large part of this thesis is focused on the development of OGA inhibitors, I will provide a brief overview of rational GH inhibitor design, as well as relevant inhibitors, both covalent and non-covalent, which target GHs.

1.3.1. GH rational inhibitor design

In order to rationally design inhibitors for GHs one point of consideration is the mechanism of action of the specific enzyme (i.e. retaining or inverting) and whether the substrate being cleaved contains an α- or β- linkage at the anomeric centre. Additionally, it is important to examine whether the GH is *endo*-[86, 87] acting or *exo* acting[88, 89] as well as *syn*- or *anti*-protonating with respect to the leaving group[90]. These factors are relevant during the design of inhibitor structure because they collectively provide an idea of the type of structure that the enzyme active site is able to accommodate.

More importantly, the structures of many effective GH inhibitors are often influenced by considering the proposed TS structure which is stabilized by the targeted enzyme. In 1946, Linus Pauling introduced his hypothesis that enzymes are "molecules that are complementary in structure to the activated complexes of the reactions that they catalyze".[91] He proposed that the binding interactions between the enzyme to the fleeting TS decreases the activation energy, leading to a dramatic rate enhancement that is observed with enzymes including GHs. Richard Wolfenden subsequently expanded on this topic by relating Pauling's suggestion to the theory of absolute reaction rates.[92] He compared the reaction rates of comparable non-enzymatic and enzymatic reaction and he then described the latter as having a reaction rate which is higher by a factor of F. The non-enzymatic reaction proceeds through a TS that he termed TX, while the enzymatic reaction proceeds through a TS that he termed ETX (Figure 1.8). He concluded that an ideal inhibitor, regardless of the binding forces involved, would mimic the interaction occurring between the activated enzyme and activated substrate in the ETX complex. Furthermore, an ideal inhibitor would not mimic the activated substrate complex TX in the un-catalyzed reaction, since the rate of this non-enzymatic reaction is much slower in comparison.

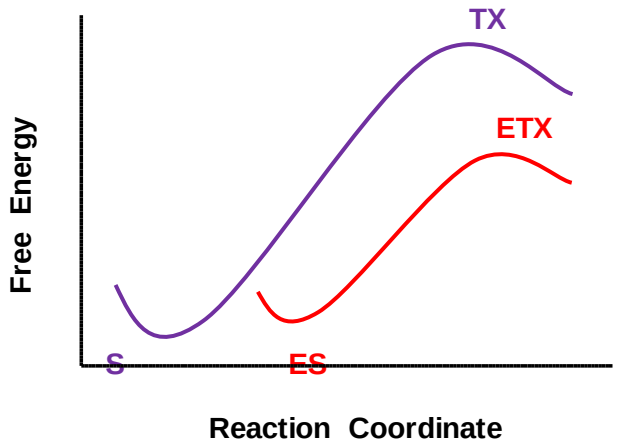


Figure 1.8. Wolfenden's comparison of the transition states of a non-enzymatic reaction and catalyzed enzymatic reaction.
Purple: Non-enzymatic reaction; Red: Catalyzed enzymatic reaction; S = free substrate, TX = transition state for the non-enzymatic reaction, ES = Michaelis complex, ETX = transition state for the enzymatic reaction. The force of attraction between E and TX was described by Wolfenden to increase by a factor F, compared to the binding constant of the substrate in the Michaelis complex.[92]

Wolfenden applied this concept through the analysis of known inhibitors of triose phosphate isomerase, an enzyme involved in glycolysis, where he pointed out that 2-phosphoglycolate (K_i = 6.8 x 10^{-6} M in chicken muscle; K_i = 6.0 x 10^{-6} M in rabbit muscle) is the only known inhibitor which is tighter binding than glycerol-3-phosphate, a molecule with similar tight binding properties as the enzyme's natural substrate (K_i =1.1 x 10^{-4} M in chicken muscle; K_i= 2.3 x 10^{-4} M in rabbit muscle). He reasoned that this tighter binding of ~100-fold is a result of the resemblance of 2-phosphoglycolate to the proposed ene-diolate intermediate which was at the time speculated to form during catalysis with triose phosphate isomerase (Figure 1.9), and later confirmed to do so[93]. Overall, in this line of research Wolfenden outlined a method to approaching tight inhibitor design by considering the structural changes that occur to the substrate as it approaches the TS where it is bound most tightly by the enzyme.

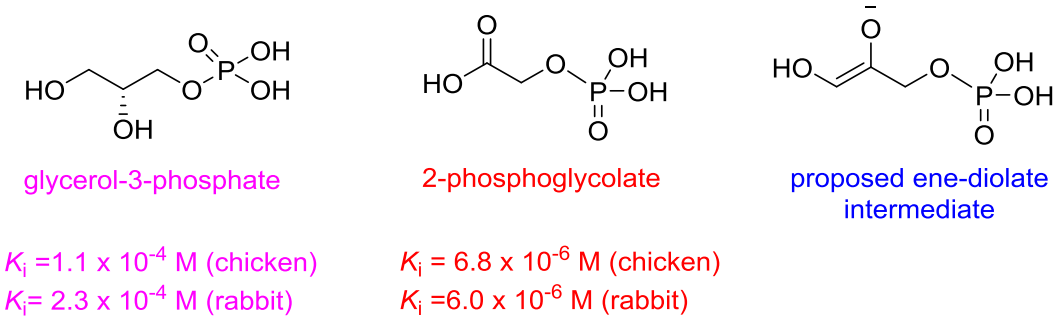


Figure 1.9. Triose-phosphate isomerase plays an important role in glycolysis to reversibly interconvert the triose phosphate isomers glycerol-3-phosphate and D-glyceraldehyde 3-phosphate.
2-phosphoglycolate (middle) binds even tighter to triose-phosphate isomerase than glycerol-3-phosphate since it presumably imitates the proposed transition state ene-diolate intermediate (right).

Taking these factors into consideration, it is possible to design both non-covalent and covalent inhibitors of GHs, and the decision to pursue either depends on the needs of the research studies as outlined briefly below.

1.3.2. Non-covalent Inhibitors of GHs

The interactions of an enzyme with non-covalent inhibitors are reversible. As a result, the types of forces involved during binding to the enzymatic pocket are Van der Waals forces, hydrogen bonding and ionic interactions. Non-covalent inhibitors are

generally more suitable for use as therapeutics and in drug design, since the enzyme being targeted is not irreversibly inhibited.

Throughout the last few decades a profusion of non-covalent GH inhibitors have surfaced with tremendous therapeutic potential for a spectrum of diseases including lysosomal storage disorders (LSDs), diabetes, and viral infections. Reviews by Gloster *et al* [85] and Lahiri *et al* [94] elegantly highlight the recent advances which have been made in designing selective inhibitors for GHs for therapeutic and research purposes. An additional review by Gloster *et al* gives a thorough overview of small-molecule inhibitors which serve as chemical tools for GHs.[2]

As noted above, the design of non-covalent inhibitors is often based on TS analogy via mimicry of the TS geometry or charge. Most GHs go through oxocarbenium-like transition states, regardless whether they are retaining or inverting enzymes. The O-5 oxygen lone pairs donate electron density into the empty *p*-orbital of C1, resulting in partial double-bond and sp^2-like character at the endocyclic oxygen and anomeric carbon.[95] Consequently, there is distortion of the pyranose ring of the substrate when approaching the TS, such that it adopts a planar conformation which resembles a half boat or chair.[95, 96] As a result, GH inhibitors generally mimic this canonical proposed TS in two categorical ways: either by imitating the partial positive charge at C1 and O5 which occurs during the TS by incorporating a basic atom within the core inhibitor ring structure or by incorporating an sp^2 hybridized centre at C1 or O5 to replicate the TS geometry around the C5-O5-C1-C2 bonds.[97]

Natural product iminosugars, where a nitrogen atom occupying the O-5 or C-1 position mimics the partial positive charge in the oxocarbenium-like GH TS, are inhibitor scaffolds that are commonly used in GH inhibitor design. Nojirimycin**,** first isolated in the 1960s from a cultured *Streptomyces* bacterial broth[98], was the first piperidine-based iminosugar bearing a nitrogen atom at the O5 position which was found to inhibit GHs including both α- and β-glucosidases.[99] Since then, the number of natural and synthetic iminosugar GH inhibitors with therapeutic potential has grown significantly. The natural product and close analogue of nojirimycin, deoxynojirimycin (DNJ),[100] has also been commonly used to target both α- and β-glucosidases. The synthetic GlcNAc

analogue of DNJ, 2-acetamido-1,2-dideoxy-D-gluconojirimycin (DNJNAc), and its derivatives have been effective at targeting *β*-hexosaminidases[101, 102], including OGA[103] with K_i values in the sub-micromolar range and high micromolar range, respectively[104]. The other commonly used synthetic derivative of DNJ is 2-acetamido-1,2-dideoxy-D-galactonojirimycin (DGJNAc), the galactose epimer, which is a potent and selective inhibitor targeting mainly α-galactosaminidases, with K_i values as low as 0.081 μM with chicken liver α-galactosaminidase.[105] Recently, 2-acetamido-1,2-dideoxy-D-allonojirimycin (DAJNAc) was synthesized and shown to be a moderate selective inhibitor of β-hexosaminidases, with K_i values ranging from 2.6-5.7 μM with human placenta, bovine kidney, and jack bean β-hexosaminidases, comparable in activity to DNJNAc and DGJNAc, for which the K_i values range from 1.8-8.3 μM.[106] Overall, the broad family of iminosugars has many potential therapeutic applications in various human diseases. Notable examples are *N*-butyl-deoxynojirimycin (Zavesca™), (K_i = 7μM)[107] an inhibitor targeting glycosylceramide synthase used for substrate reduction therapy in the LSD known as Gaucher's disease[108, 109] and *N*-hydroxyethyldeoxynojirimycin (Glyset™), an inhibitor of α-glucosidases (for example, K_i = 0.21 μM with maltase glucoamylase) used for treating type II diabetes[110], both of which are already drugs on the market. The structures of some iminosugar inhibitors are presented in Figure 1.10 A.

Other rationally designed TS analogue piperidine-based iminosugar inhibitors are isofagomine[111, 112] and its analogue, azafagomine[113] (Figure 1.10, B), which has a hydrazine functionality incorporated into the ring structure. It was found that by moving the nitrogen in DNJ to the anomeric position to make isofagomine, an even more potent β-glucosidase inhibitor was created. Isofagomine inhibits β-glucosidase more strongly compared to α-glucosidase and other GHs (K_i = 0.11 μM in almond),[114] and has been shown to act as an inhibitory pharmacological chaperone of β-glucosidase in Gaucher cell fibroblasts, increasing enzyme activity by 3-fold in N370S cells which mimic the type 1 mutation in Gaucher disease.[115] 1-Azafagomine inhibits almond β-glucosidase in almonds between pH 5.0 and 6.8 with K_i values ranging from 0.65 to 1 μM. Notably, it also displayed inhibition against α-glucosidase with a reasonable potency of K_i = 3.9 μM[113], unlike isofagomine, which targets α-glucosidase less efficiently (K_i = 86 μM in yeast)[116]. 1-azafagomine derivatives have since been developed which have been

rationally modified to improve potency for β-glucosidase[117] or increase selectivity for α-glucosidase.[118]

Pyrrolidine iminosugar natural products have also been exploited as TS analogues of GHs, and their structures are often used as a starting point for new synthetic analogues which have become increasingly more selective and potent for various GHs. 2,5-dideoxy-2,5-imino-D-mannitol[119] (DDIM) is a potent inhibitor of both α- and β-glucosidases with an IC_{50} of 0.2 μM for both types of enzymes at a pH above 6.5, predominantly inhibiting in its de-protonated state.[120] A combinatorial approach which involved modification of the amine of DDIM subsequently led to the discovery of a potent low-nanomolar inhibitor of *N*-acetyl β-hexosaminidase with a K_i of 2.6 nM where the functionality on the N-1 of the pyrrolidine was a heptamine.[121] Other synthetic pyrrolidines which have shown to be hexosaminidase inhibitors are enantiomers 2-acetamido-1,4-imino-1,2,4-trideoxy-L-arabinitol (LABNAc) and 2-acetamido-1,4-imino-1,2,4-trideoxy-D-arabinitol (DABNAc) and their respective *N*-benzyl analogues. LABNAc and *N*-benzyl LABNAc both target human placenta β-*N*-acetyl-D-hexosaminidase with low micromolar K_i's (15 and 3.7 μM respectively), while *N*-benzyl DABNAc is less effective, having a K_i of 180 μM.[122] The structures of the pyrrolidine inhibitor family for GHs are summarized in Figure 1.10,C.

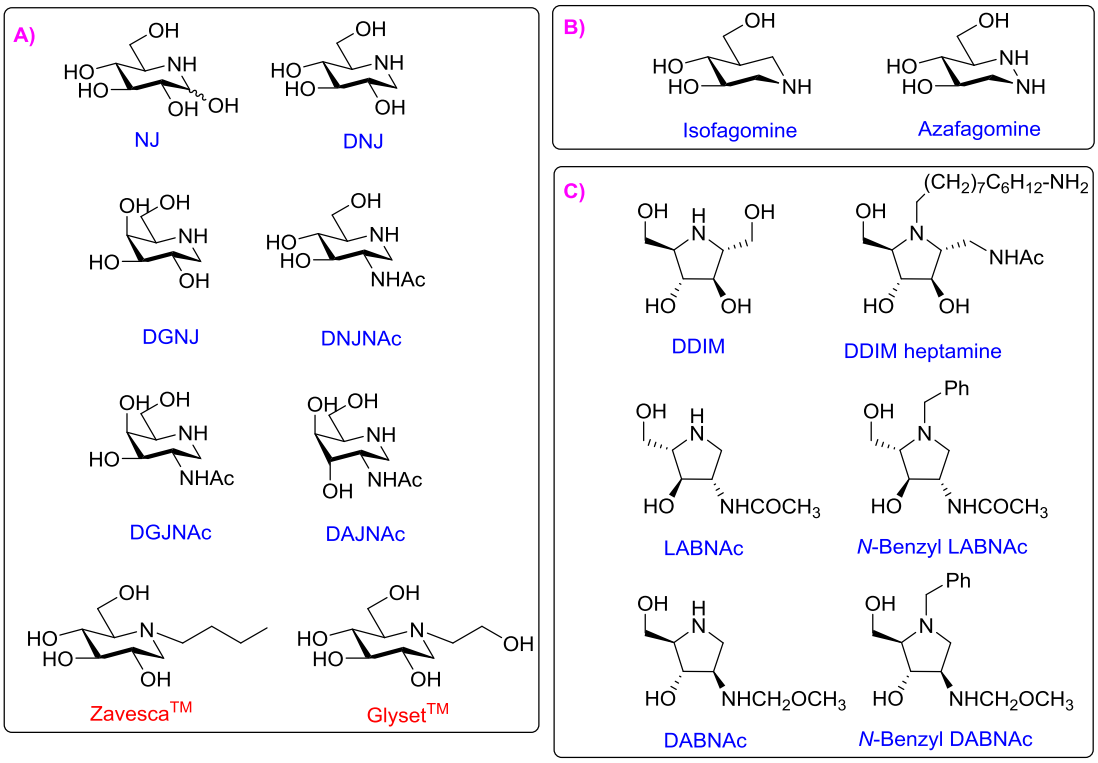


Figure 1.10. Examples of non-covalent iminosugar inihibitors of GHs. **A) Iminosugar GH inhibitors stemming from the structure of nojirimycin (NJ) which targets both α- and β-glucosidases.** DNJ: 1-deoxynojirimycin (α- and β-glucosidases); DGNJ: 1-deoxygalactonojirimycin (α- and β-galactosidases); DNJNAc: 2-acetamido-1,2-dideoxy-D-nojirimycin (β-hexosaminidases); DGJNAc: 2-acetamido-1,2-dideoxy-D-galactonojirimycin (α-galactosaminidases and β-hexosaminidases); DAJNAc: 2-acetamido-1,2-dideoxy-D-allonojirimycin (β-hexosaminidases). Zavesca™ and Glyset™ are current therapeutics for Gaucher's disease and type II diabetes, respectively. Zavesca™ is an inhibitor of glycosyl ceramide synthase, and Glyset™ of α-glucosidase. This class of compounds mimics the oxocarbenium ion in the GH TS. **B) Isofagomine and azafagomine, rationally designed TS mimic iminosugar inhibitors of β-glucosidases.** In contrast to nojirimycin and analogues, isofagomine and azafagomine resemble the anomeric carbon resonance form of the TS oxocarbenium. **C) Pyrrolidine-based iminosugar GH inhibitors.** DDIM: 2,5-dideoxy-2,5-imino-D-mannitol (α and β-glucosidases); DDIM heptamine is a potent and specific inhibitor of *N*-acetyl-β-hexosaminidase; LABNAc: 2-acetamido-1,4-imino-1,2,4-trideoxy-L-arabinitol; DABNAc: 2-acetamido-1,4-imino-1,2,4-trideoxy-D-arabinitol; enantiomers and their *N*-benzyl analogues target *N*-acetyl-β-hexosaminidases.

Castanospermine, another natural iminosugar with an indolizine alkaloid backbone isolated from the tree *Castanospermum australe,* is used as an effective

competitive inhibitor of many α- and β-glycosidases.[123, 124] More recently, an analogue of castanospermine, 6-acetamido-6-deoxy-castanospermine (6-Ac-Cas) was shown to be a hexosaminidase inhibitor (K_i = 300 nM towards OGA and 250 nM for HexB). 6-Ac-Cas was applied to demonstrate that insulin resistance is not a direct result of increase in *O*-GlcNAc levels. 3T3-L1 adipocytes were observed to have increased *O*-GlcNAc levels upon treatment with 6-Ac-Cas, without observed insulin resistance.[125] In a different line of research, the indolizine alkaloid natural product swainsonine has been shown to target Golgi α-mannosidase II, an enzyme involved in *N*-linked oligosaccharide processing.[126] Swainsonine has gained much attention for its potential use as an anti-cancer agent, showing a reduction in tumour metastasis and enhancement of the immune response towards cancer cells.[127, 128] More recently, pochonicine, a polyhydroxylated pyrrolizidine alkaloid natural product which comes from *Pochonia suchlasporia* var. *suchlasporia TAMA 87* fungus was reported[129] to be a selective potent inhibitor of β-*N*-acetyl-hexosaminidases in a range of organisms. K_i values were in the low nanomolar range spanning from 0.1 nM to 141 nM, depending on the enzyme in question. Since then, synthetic analogues and stereoisomers have been further explored for inhibition optimization[130], identifying potent pochonicine stereoisomers which also target β-*N*-acetylgalactosaminidases in the low micromolar or mid-nanomolar range. The bicyclic natural products which serve as non-covalent TS analogue GH inhibitors are summarized in Figure 1.11.

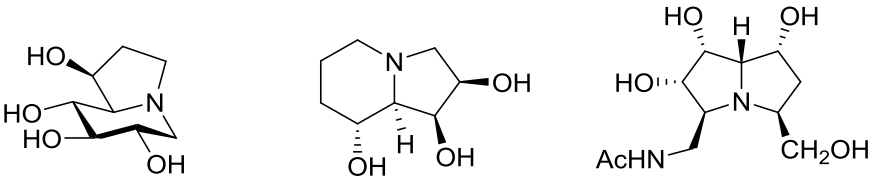


Figure 1.11. Natural product iminosugar GH inhibitors which have a bicyclic ring system backbone.
Castanospermine and swainsonine are indolizine alkaloids which target α- and β-glycosidases and α-mannosidase II, respectively. Pochonicine is a potent inhibitor of β-*N*-acetyl hexosaminidases.

Instead of mimicking the charge build-up proposed to occur in the GH TS, there are effective GH inhibitors which mimic the planar geometry of the TS by incorporation of an sp^2 hybridized centre at what would be the anomeric centre of the substrate. A

notable example is the 1,5-hydroximolactone derivative, *O*-(2-acetamido-2-deoxy-D-glycopyranosilidene)amino-*N*-phenylcarbamate (PUGNAc), which targets β-*N*-acetylhexosaminidases, including human OGA. PUGNAc is a nanomolar inhibitor of OGA (K_i = 46 nM)[131] which has commonly been used for evaluation of increased *O*-GlcNAc levels in cells[132], despite this compound lacking selectivity over other functionally related enzymes. Imidazole-based sugar derivative nagstatin, a natural product initially isolated from *Streptomyces amakusaensi,*[133] is another strong β-*N*-acetylhexosaminidase inhibitor which mimics the TS geometry. The fused imidazole ring locks the 6-membered ring into a planar conformation at the pseudo-anomeric centre, while the imidazole is sufficiently basic to enable its protonation within the enzyme active site.[134] GlcNAcstatin, a glucoimidazole analogue, is a hexosaminidase inhibitor which is also not selective for OGA over other mammalian hexosaminidases[135] and shows a K_i of 4.6 pM against a bacterial homologue of OGA (NagJ from *Clostridium perfringens*). Derivatives of gluconagstatin, such as GlcNAcstatin-C[136], have been shown to have selectivity for OGA and to be able to raise *O*-GlcNAc levels in human embryonic kidney 293 (HEK 293) cells more effectively than PUGNAc, making it a useful chemical tool for studying human OGA in cell culture. The TS analogues designed to incorporate aspects of GH transition states that have been mentioned are summarized in Figure 1.12.

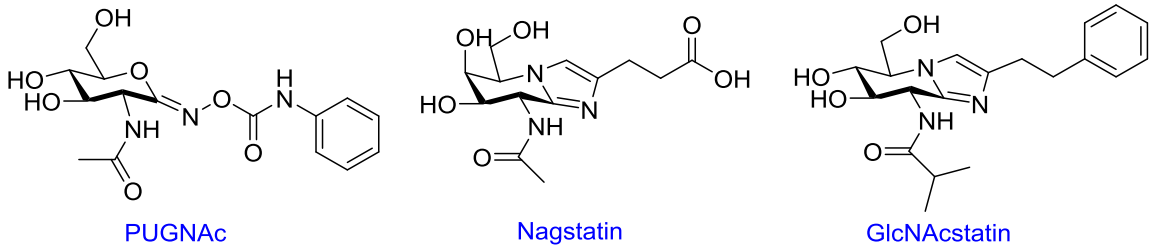


Figure 1.12. Inhibitors which are TS analogues of GHs that mimic the sp^2 hybridized anomeric centre.
O-(2-acetamido-2-deoxy-D-glycopyranosilidene)amino-*N*-phenylcarbamate (PUGNAc) and nagstatin are both β-*N*-acetylhexosaminidase inhibitors. GlcNAcstatin is a potent and selective inhibitor of OGA.

An alternative GH inhibitor family targeting *N*-acetyl hexosaminidases, including OGA[84] which mimics the oxazoline intermediate formed during substrate-assisted catalysis[137] is the thiazoline family, stemming from the structure of NAG-thiazoline, which was first prepared by Knapp *et al*[138]. This inhibitor has been shown to be a TS

analogue[131] with a potency of 70 nM for family 20 β-hexosaminidases and family 84 β-*N*-acetylglucosaminidases, which includes OGA. 1,2-dideoxy-2'-propyl-α-D-glucopyranoso-[2,1-d]-Δ2'-thiazoline (NButGT) and 1,2-dideoxy-2'-ethylamino-α-D-glucopyranoso-[2,1-d]-Δ2'-thiazoline (ThiamEt-G), two increasingly more selective competitive inhibitors of human OGA have been developed. NButGT bears a bulkier alkyl chain in place of the 2'-methyl group observed for NAG-thiazoline and has a K_i value of 70 nM for human OGA[84]. NButGT has been used in cell culture[139] as as well as *in vivo* to show that increased O-GlcNAc levels do not induce insulin resistance[140]. ThiamEt-G was shown to be 10 times more potent (K_i = 21 nM), which was accomplished by installation of an exocyclic amine at the 2' position of the thiazoline ring. ThiamEt-G has also been successfully used to study the relationship between the *O*-GlcNAcylation and hyperphosphorylation of the tau protein in an Alzheimer disease *in vivo*, overall showing that an increase in *O*-GlcNAc levels decreases the toxic hyperphosphorylation of tau[54, 141, 142] and protects against amyloid pathologies[143]. This family of inhibitors will be more extensively discussed in Chapter 2, as the inhibitors which were synthesized for the purposes of this research stem from the structure of ThiamEt-G. The structures of NAG-Thiazoline, NButGT and ThiamEt-G are summarized in Figure 1.13.

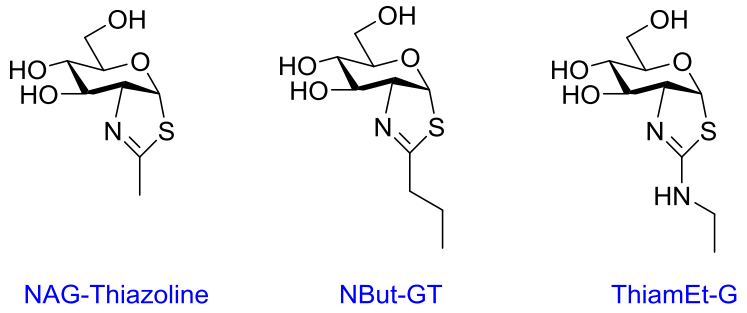


Figure 1.13. Structurally-related inhibitors of *N*-acetyl hexosaminidases and OGA, NAG-thiazoline, NButGT and ThiamEt-G.
NAG-thiazoline is a potent, non-selective inhibitor for both human lysosomal hexosaminidases and OGA (K_i = 70 nM); NButGT is selective for human OGA (K_i = 250 nM)[131] while ThiamEt-G, still selective for human OGA is even more potent (K_i = 21 nM)[54].

1.3.3. Covalent inhibitors of GHs

In contrast to non-covalent inhibitors, there are also a variety of GH inhibitors that bind to the active site and then form a covalent bond with the enzyme, resulting in long lived or even irreversible inactivation of the target enzyme. Consequently, these covalent types of inhibitors can be used as probe molecules in different ways as compared to non-covalent inhibitors. Mechanism-based covalent inhibitors (MBIs) are commonly used for identification of site residues[144, 145] and studying the catalytic mechanisms of GHs[146]. A review by Rempel and Withers elaborates on the many uses of covalent inhibitors with a focus on their biological significance when used with GHs[32], and a more recently published book chapter by Kallemejin, *et al* reviews this topic as well[147] These families of covalent inhibitors can be further subdivided into affinity-based or MBIs as alluded to just above.

MBIs are chemically inert molecules until they are bound within the enzyme active site and activated by the enzyme machinery. For GHs this activation of the probe typically occurs during the glycosylation step when the nucleophile within the active site attacks the anomeric centre leading to cleavage of the glycosidic linkage. At this stage, the MBI forms a covalent bond with a chemically reactive amino acid residue within the active site. Often, this residue is the active-site nucleophile and hence, incubation with certain MBIs can be used to identify this catalytic residue. Some examples of the mechanism base of MBIs are carbasugar epoxides[148-151], fluorosugar glycosides[152], conduritol aziridines[153, 154] and newer cyclopropylcarbasugars[155] most of which lead to labeling of the catalytic nucleophile within the active sites of retaining GHs.

Conduritol-based epoxide (CBEs) (Figure 1.14) is an example of an MBI which has long been used for studying retaining GHs. Interestingly, the structural symmetry of CBE allows it to act as an electrophile for both retaining α- and β-glucosidases, although it is particularly selective, and most commonly used for studying human lysosomal β-glucocerebrosidase 1 (GBA1) to gain insight into Gaucher's disease.[156] GBA1 is the enzyme which is deficient in the genetic disease known as Gaucher's, which stems from a loss of function of GBA1 and as a result, an accumulation in the lysosome of its substrate, glucocerebroside. The mechanism of enzyme inactivation involves a

nucleophilic attack by the catalytic nucleophile on the pseudo-anomeric position of CBE, which results in the opening of the epoxide with simultaneous protonation of the oxygen atom by the general acid/base catalytic residue and consequent formation of a stable covalent bond (Figure 1.14). An example of a significant application of CBE was the exploitation of its specificity for GBA1, and more specifically its lack of reactivity with a non-lysosomal integral membrane protein glucosylceramidase (NLGCase), being used to uncover that the GBA2 gene encodes for NLGCase.[157] 4-Methylumbelliferyl β-D-glucoside (4-MUGlu), a substrate for both GBA1 and GBA2, activity in GBA2 transfected mouse leukaemic monocyte macrophage (RAW) cells showed increased β-glucosidase activity that could not be ablated on treatment of the cells with CBE. Conversely, treating cultures with *N*-butyl-DNJ (Zavesca^{TM}), a strong inhibitor of GBA2 resulted in the same level of β-glucosidase activity regardless of whether transfection was performed. Additionally, the subcellular localization of GBA2 was evaluated in RAW cells and mouse melanoma (MEB-4) cells. The sub-cellular localization of GBA2 was done in parallel using a glucosylceramide fluorogenic substrate, which upon enzymatic cleavage could be localized using fluorescence microscopy. Applications of fluorogenic substrates such as this one will be further described in Section 1.4 of the thesis.

Since CBE is able to inactivate both α- and β-glucosidases, there is clearly a limitation in terms of its ability to selectively target β-glucosidases. Cyclophellitol, however, is also an epoxide MBI which is structurally similar to CBE except that it contains a hydroxymethyl moiety at the C5 position instead of a hydroxyl group (Figure 1.14), a substitution that breaks the symmetry of the molecule. This natural product was isolated from the culture filtrate of a mushroom, *Phellinus sp*., and was established to be a specific inhibitor of β-glucosidases with inactivation constants of K_i = 0.34 mM and k_{inact} = 2.38 min^{-1} for almond β-glucosidase and K_i = 0.055 mM and k_{inact} = 1.26 min^{-1} for *Agrobacter. sp*. β-glucosidase.[150] In 2007, the crystal structure of covalently bound cyclophellitol in the active site of the β-glucosidase from *Thermatoga maritima (Tm*GH1) was determined, confirming that a covalent bond is formed with the catalytic nucleophile within the glycosidase active site.[151] The demonstrated biological applications of cyclophellitol are more limited than CBE, but this MBI has extensively been used to selectively inhibit human lysosomal GBA1 and induce an enzyme deficient-like state. The motivation to induce this physiological state is to explore a "Gaucher-like" phenotype

in an animal model, which is the disease arising from GCase deficiency as noted above. The effects of GCase inhibition using either CBE[158] or cyclophellitol[159] have been explored in animal models in this manner to take a closer look at the physiological effects of loss of function of GBA1[160].

In addition to CBE and cyclophellitol, the aziridine analogue of CBE (Figure 1.13) has also been established as an inactivator of α (K_i = 3 mM, k_{inact} = 0.077 min^{-1}) and β-glucosidases (K_i = 9.5 mM, k_{inact} = 0.39 min^{-1}).[153] The design of aziridine-based inactivators stems from the notion that the nitrogen in the ring when protonated, should have a higher affinity for the negatively charged GH active site since it would be positively charged, although this has not been borne out by studies.[32] Aziridines, however, have evolved to emerge as effective activity-based probes for human GBA1[161], which is a topic further elaborated on in Section 1.3.4, as well as other enzymes[162, 163].

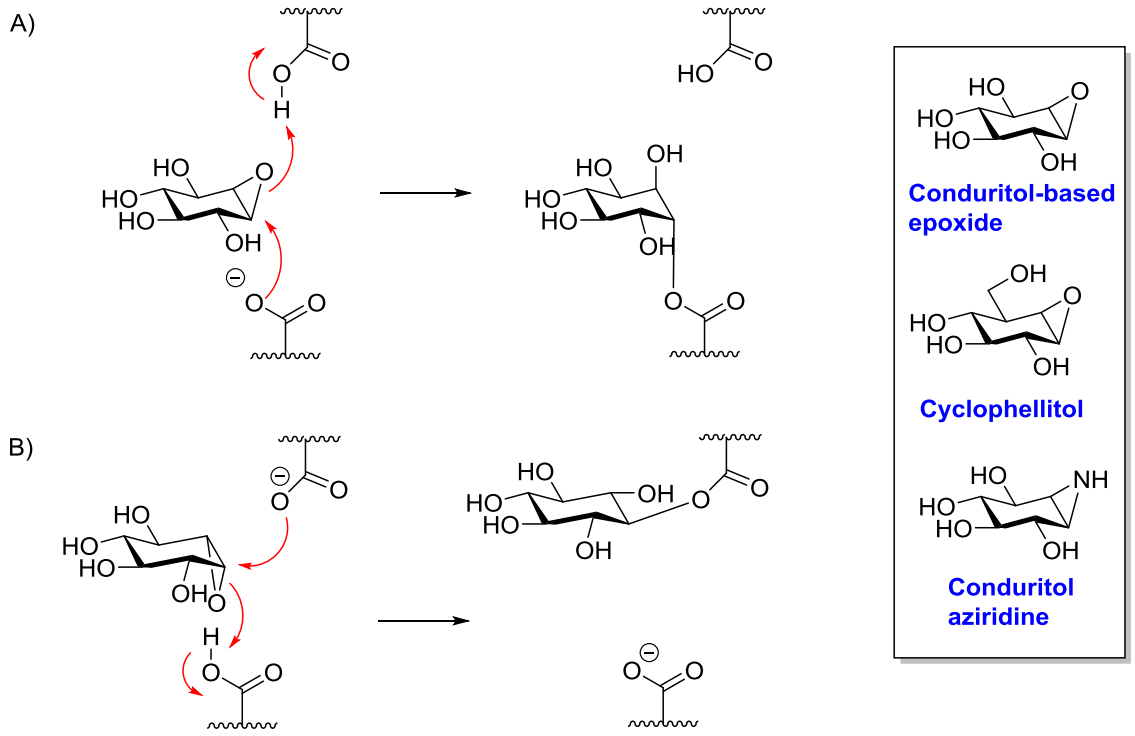


Figure 1.14. Mechanism of inhibition of A) retaining β-glucosidases and B) retaining α-glucosidases by conduritol-based epoxide (CBE). Cyclophellitol and CBE are among the most common MBIs of glucosidases used to study enzymatic mechanisms and identify catalytic nucleophiles within glycosidase active sites. Both inhibitors have been applied in various biological studies which target GBA1 and the lysosomal storage disease, Gaucher's disease. Conduritol aziridine is the aza-analogue of CBE which is an MBI of both β- and α-glucosidase, with a higher affinity for α-glucosidase.

Recently, in 2014, Chaklandar *et al* presented two bicyclo[4.1.0]heptyl analogues of galactose as carbocyclic MBIs for retaining galactosidases from GH families 27 and 36.[155] Activation of the D-*galacto* MBI within the active site of GH36 family α-galactosidase *Thermotoga maritima* (Figure 1.15, A) leads to inactivation through the formation of a bicyclobutenium ion intermediate (Figure 1.15, B) that then collapses with the catalytic nucleophile residue in the active site, as determined by electrospray ionization tandem spectrometry (ESI/MS/MS). This is so far the sole example of a GH forming a carbocationic intermediate that does not involve the stabilization by the endocyclic O-5 atom of positive charge build up on the anomeric centre in the TS.

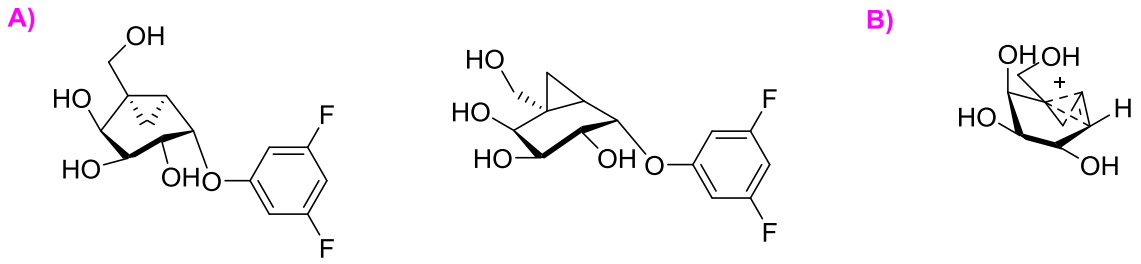


Figure 1.15. Mechanism-based inhibitors involving formation of a non-classical carbocation.
Bicyclo[4.1.0]heptyl-based D-*galacto* (A, left) and L-*altro* (A, right) analogues which were the MBIs synthesized form a bicyclobutenium ion in the transition state (B), leading to enzyme alkylation and inactivation. These MBIs were used to identify the catalytic aspartic acid residue for two different retaining α-galactosidases from GH families 27 and 36.[155]

A widely explored family of MBIs which have been extensively reported on in the literature are the 2-deoxy-2-fluoroglycosides, whose mechanism of action is depicted in Figure 1.16.[152] The addition of the electronegative fluorine at the 2-position of the pyranose ring increases the free energy of both the glycosylation and de-glycosylation transition states, which have significant oxocarbenium-like character. As a result, both catalytic steps are slowed. However, through the installation of a good aglycone leaving group, the glycosylation step of the reaction can be sped up such that it becomes faster than the de-glycosylation step ($k_1 > k_2$) and the resulting covalent intermediate stably accumulates. This concept has been useful for obtaining crystal structures of glycosyl enzyme intermediates and identification of catalytic nucleophiles in retaining glycosidases having an equatorial leaving group, as seen for example with β-glucosidases.[32, 164, 165]

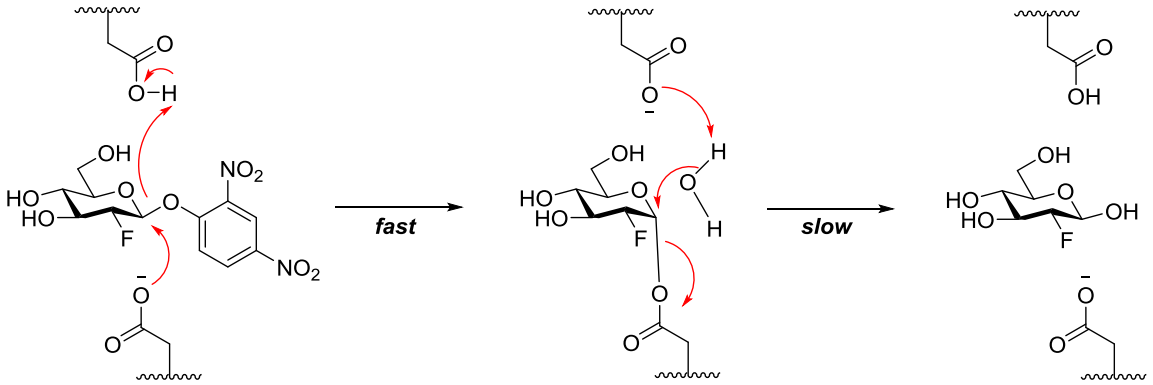


Figure 1.16. The inhibitory mechanism of 2-deoxy-2-fluoroglucosides with retaining glycosidases.

1.3.4. Covalent GH inhibitors as activity-based probes (ABPs)

The concept of GH inhibition has been described above primarily in the context of enzymatic *in vitro* assays where inhibitors have been used for mechanistic studies with GHs. However, several MBIs have been converted into chemical probes that are able to covalently label GHs in cell lysates, intact cells, or even in tissues and whole animals. The labeled GH can then be detected or identified using techniques such as mass spectrometry and western blotting. This strategy has been particularly useful in the field of activity-based protein profiling (ABPP) where the covalent enzyme inhibitor acts as a probe enabling enzymes having an activity of interest to be detected or identified. One example is the application of a fluorescently-labelled cyclophellitol and a cyclophellitol-aziridine, which have been used successfully as activity based probes (ABPs) for visualization of active GBA1 *in vitro* and *in vivo* by labelling the active enzyme. Witte, M.D. *et al* described an elegant approach for *in situ* labeling of GBA1 with cyclophellitol-based ABPs labeled with a boron dipyrromethene (BODIPY) fluorophore (Figure 1.17, **1** and **2**) to monitor GBA1 activity in fibroblasts from Gaucher patients and in a mouse animal model.[166] These ABPs covalently label the catalytic nucleophile of GBA1 (residue E340), and appear to be extremely tight binding ligands for GBA1, with IC_{50} values of 7 nM and 8 nM for **1** and **2**, respectively. A few years later, Kallemeijn *et al* described the ABPs **3** and **4** (Figure 1.17) based on the structure of cyclophellitol-aziridine MBIs which in contrast to **1** and **2** are able to target other β-exoglucosidases besides GBA1.[167] These probes are also slightly more potent than **1** and **2**, inhibiting

GBA1, GBA2 and GBA3 in the low nanomolar and picomolar range (IC_{50} = 0.98, 0.130 and 3.96 nM, respectively). **3** was shown to be effective for labelling GBA1 in living HepG2 cells and in COS-7 cells over-expressing GBA2, GBA3, and lactase/phlorizin hydrolase (LPH), in studies where the cell lysates were incubated with the probe and then analyzed by western blot. Additionally, western blot analysis showed that **3** was able to label retaining β-glucosidases in live mice when tissue homogenates from various organs were taken and analyzed 2 hours after injection with the probe. Both sets of probes are useful methods for monitoring GBA1 levels in cell lysates and for imaging functional GBA1 within fixed cells.

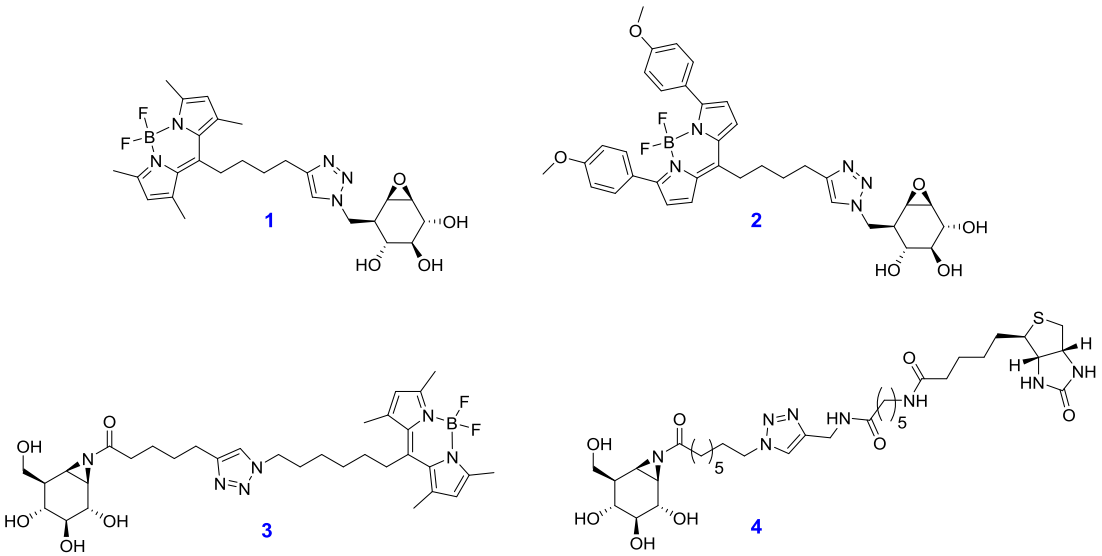


Figure 1.17. Activity-based probes (ABPs) based on the structures of epoxide or aziridine MBI families.
1 and **2** are based on the MBI cyclophellitol structure, connected with a linker to boron dipyrromethene (BODIPY) fluorophores and have been used to allow visualization of active GBA1 enzyme in cells and living mice.[166] ABPs **3** and **4** are based on the aziridine MBI structure linked to a biotin molecule as the reporter and target a wider range of retaining β-glucosidases which include GBA1, GBA2 and GBA3.[167]

Another example related to the MBIs described previously is the 2-deoxy-2-fluoroglycoside concept. One of the earliest examples of a MBI adapted for use as an ABP was published by Vocadlo and Bertozzi in 2004 and used a 2-deoxy-2-fluorogalactoside derivative.[168] This study describes the labelling of an endogenous β-galactosidase in cell lysates using an azide-functionalized 2-fluorosugar ABP. Subsequent labelling of the *Escherichia coli* β-galactosidase (LacZ) using the Staudinger

ligation permitted downstream detection of the enzyme by western blot analysis (Figure 1.18). This technique was a pioneering approach for observation and isolation of glycosidases in complex cellular mixtures.

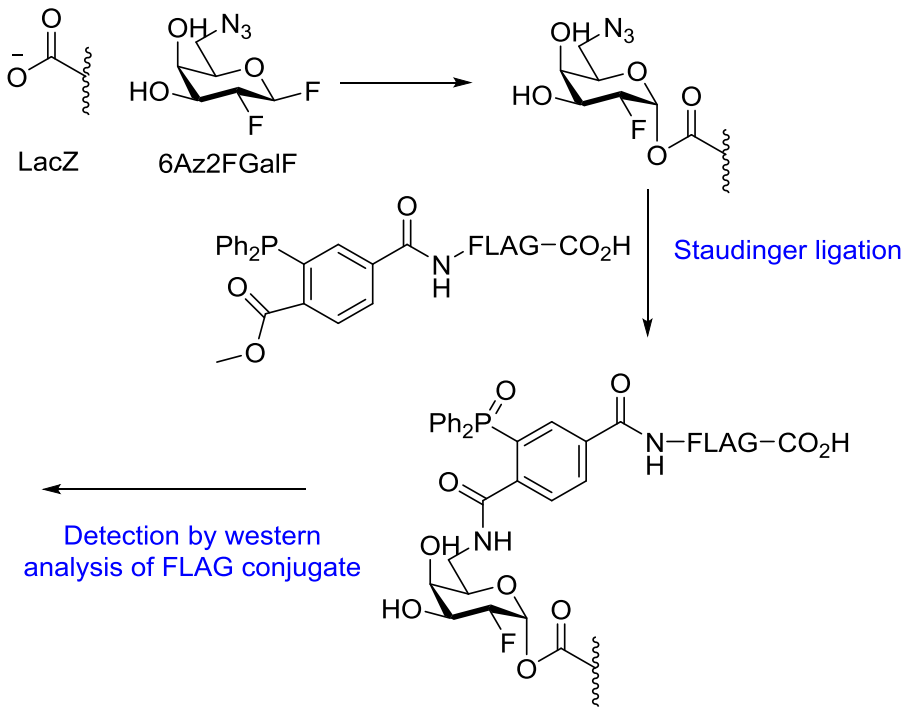


Figure 1.18. An azidosugar ABPP used in combination with the Staudinger ligation to label glycosidases.
Incubation of cell lysates with 6-azido-2,6-dideoxy-2-fluoro-β-D-galactopyranosyl fluoride (6Az2FGalF) covalently labeled LacZ. Subsequent Staudinger ligation allowed tagging with phosphine-FLAG (DYKDDDDK peptide sequence), which was detected using western analysis.

An extension of the 2-fluorosugar approach has been described in conjunction with radioisotopic labelling for *in vivo* studies of enzyme biodistribution in a mouse animal model. GBA1 was labeled *in vitro* using an ^{18}F-labelled fluorosugar and then injected into mice. Positron Emission Tomography (PET) imaging was used to follow distribution of the enzyme, allowing the direct *in vivo* assessment of GBA which could be used to better optimize the current standard of care for Gaucher disease, which is Enzyme Replacement Therapy (ERT) with recombinant enzyme (Figure 1.19).[169]

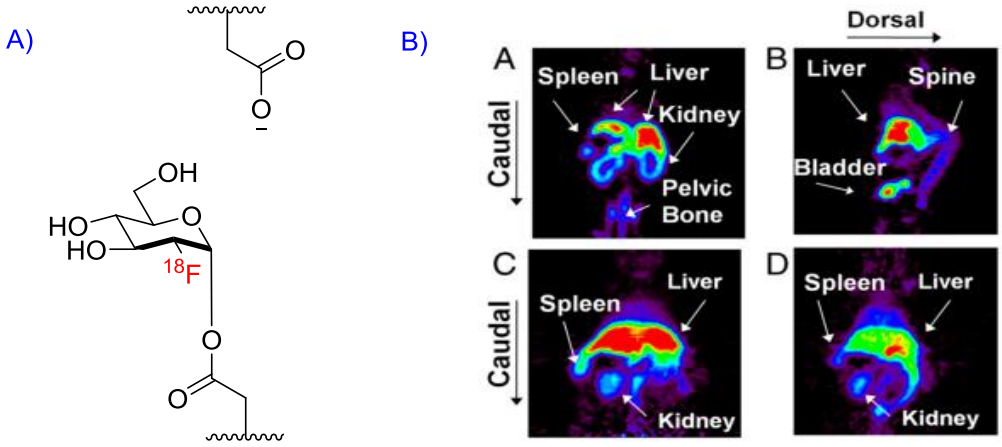


Figure 1.19. The 2-fluorosugar approach applied during radioisotopic labeling *in vivo*.
A) The structure of the radioisotopically labelled GBA (^{18}F-Glc-GCase) catalytic site. **B)** The PET images obtained upon injection of the radioactive inhibitor into mice. This image was reproduced from PNAS v.107 n.24, p.10848. Copyright Christopher P. Phenix *et al*. Used with permission.[169]

The types of ABPP of GHs has been consistently expanding over the last decade, with many examples elegantly highlighted in a recent review by Willems *et al*.[154] As valuable of a technique as using MBIs as ABPs is, throughout the course of the research described in this thesis, we have instead focused on probe substrates that could enable monitoring OGA activity in live cells. This is a largely un-explored field for GHs in general and permits monitoring enzyme activity rather than simply enzyme levels. For the purposes of appreciating the objectives of this thesis, it is necessary to further expand upon the use of GH substrates as chemical tools for evaluation of GH function.

1.4. Use of fluorescent substrates as probes to monitor GH activity

Fluorescent substrates are frequently used for *in vitro* enzymatic assays to monitor enzyme activity in buffers that can mimic the natural environment of the enzyme. For GHs there are many examples of fluorogenic substrates in which a fluorophore has been conjugated to the recognition element of the substrate, such that upon cleavage of the glycosidic bond a dramatic increase in a fluorescent signal results. One of the most

common fluorophores currently employed for this purpose is 4-methylumbelliferone (4-MU) (Figure 1.20), which has been conjugated to natural substrates to probe a wide range of GH activities. 4-MU has been used to assay enzymes in this way as far back as 65 years ago[170, 171] and is currently used to assay enzymes such as GBA1, GBA2 and GBA3[157, 166], lysosomal β-hexosaminidases[172, 173] and human OGA[84]. One of the most problematic features of using 4-methylumbelliferone as a fluorogenic leaving group is its high signal to noise ratio, making background fluorescence a concern. Additionally, since the phenolate anion of free 4-methylumbelliferone is the fluorescent species and the pK_a of 4-methylumbelliferone is 7.8, [174, 175], the pH of the *in vitro* assay may need to be raised by quenching the reaction with a basic buffer in order to generate measurable signal, depending on the enzyme being studied. There are many other fluorophores that have been applied to probing GH activity *in vitro*, such as resorufin (Figure 1.20), which fluoresces at longer wavelengths. The main advantages of resorufin are its low pK_a (5.8), large extinction coefficient (ε 62000 M^{-1} cm^{-1}) and high fluorescence quantum yield (ϕ = # of photons emitted/ # photons absorbed; ϕ=0.74)[176] of the phenolate anion.[177] Examples of *in vitro* assays using resorufin glycosides include the use of resorufin β-D-cellobioside to monitor cellulose activity[178] and probing the activity of GBA1[179] with resorufin β-D-glucopyranoside. A basic outline of how these fluorophores are used in assays is summarized in Figure 1.20.

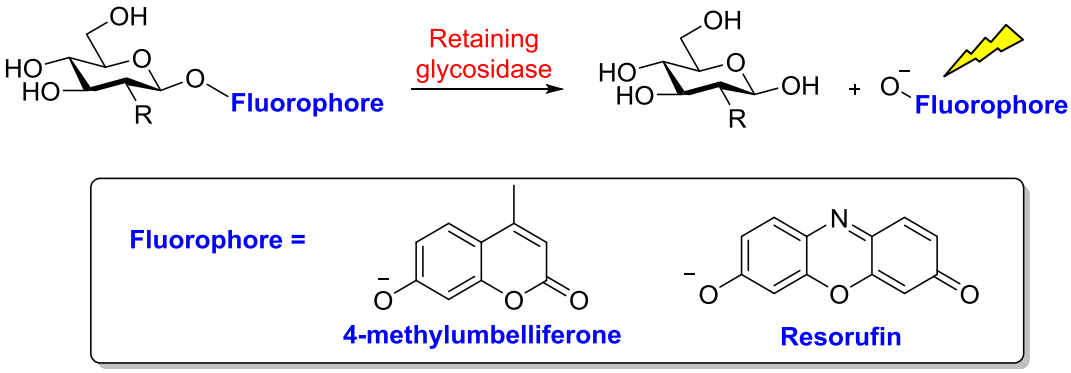


Figure 1.20. The basic detection method of GH hydrolysis during *in vitro* kinetic studies commonly performed with 4-methylumbelliferone and resorufin.
Upon hydrolysis with the enzyme, the fluorescence of the liberated anionic fluorophore can be detected.

Although conjugating fluorophores onto substrates has been a valuable approach to probing GH activity *in vitro*, most of these have not proven useful for use *in vivo*. This issue arises in large part because of their background fluorescence and in some cases the low wavelengths of their excitation and emission bands, which leads to damage to the cell or is confounded by cellular auto-fluorescence, respectively. An alternative method for achieving *in vivo* or *in cellulo* monitoring of activity while minimizing background signal is to use substrates that show Förster Resonance Energy Transfer (FRET) when intact between two different fluorophores or FRET quenching in the case of using a fluorophore and quencher pair. The FRET is then relieved upon hydrolysis by the GH of interest. There are limited examples of successful application of this approach in the literature as applied to GHs.[177, 180, 181] A major challenge with this approach is that the active site architecture of GHs, in particular GHs that cleave terminal sugars from glycoconjugates, is pocket shaped and does not accommodate the simultaneous installation of two fluorophores. One example where the FRET-based quenched substrate approach has been applied to detect the activity of *endo*-chitinases and -chitobiosidases is where a bifunctionalized chitooligosaccharide substrate was used in which the fluorophore and quencher were installed at either end of the oligomer.[182] Such a design is tolerated by *endo*-acting GHs that cleave the middle of oligosaccharide chains because they have a groove shaped active site. This chitooligosaccharide

substrate was used to detect *endo*-chitinase activity within a complex mixture since *exo*-chitinases are not able to hydrolytically cleave the substrate. Using this approach, it was shown that the chitobiosidase in question actually show *endo* hydrolytic activity, which was significant since that class of enzymes was previously considered as exclusively *exo*-acting (Figure 1.21).

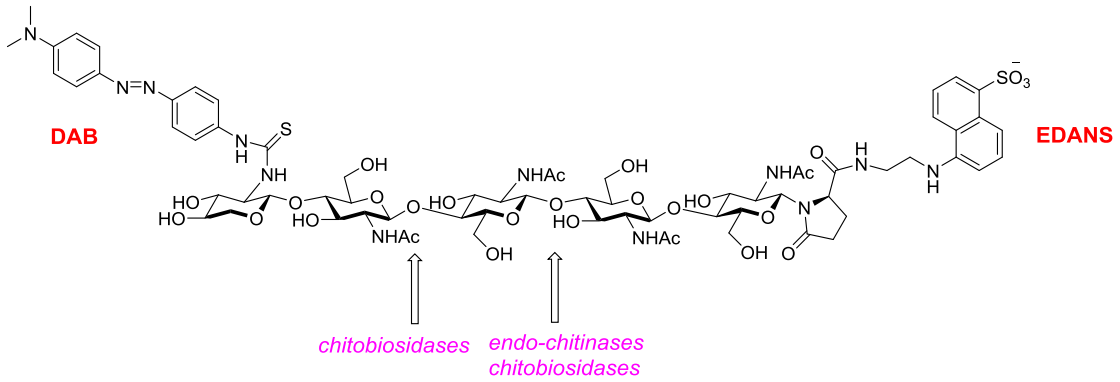


Figure 1.21. A chitooligosaccharide substrate bi-functionalized with an EDANS/DAB fluorophore/quencher pair.
EDANS: 5-(2-aminoethyl) amino-1-napthalene-sulfonic acid; DAB: dimethylaminophenylazophenyl.

In contrast to this *in vitro* quenched substrate approach, a recent and pioneering example of monitoring GH activity *in cellulo* has been reported by Yadav *et al*. This approach permits the time-dependent monitoring of GBA1 activity in live cells and assessment of its inhibition or activation in live cells (Figure 1.22).[181] From a more general perspective, direct visualization of GBA1 activity within cells should enable the optimization of current therapeutics for Gaucher's[183] and Parkinson's diseases[184, 185], which both result from inadequate levels of active GBA1 in the lysosome, as well as identification of new therapeutic strategies or active molecules.

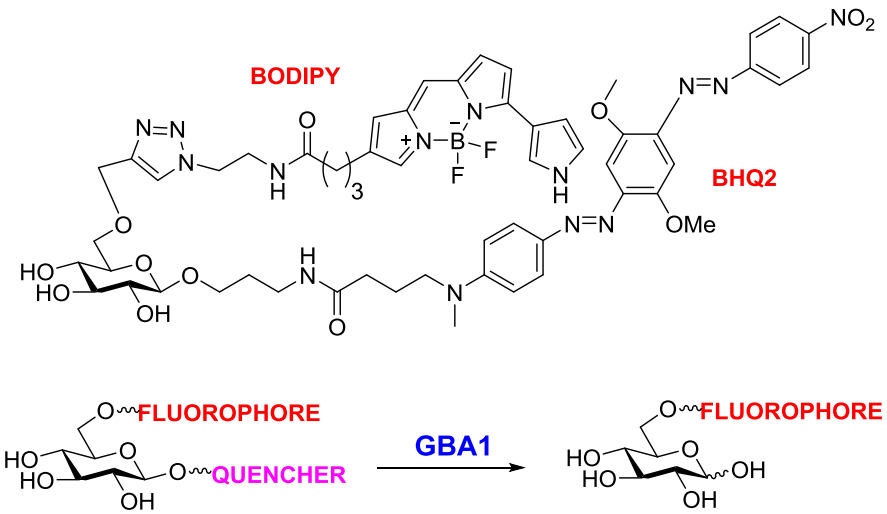


Figure 1.22. The fluorescence quenched substrate used for monitoring the activity of endogenous GBA1 in live cells.
BODIPY (λ_{ex}=576 nm/ λ_{em}=589 nm) is a fluorophore compatible with use of red-shifted light for the excitation, which minimizes background cell autofluorescence and cell damage from irradiation. The absorption wavelength of Black Hole Quencher 2 (BHQ2) overlaps with the emission of the BODIPY fluorophore, enabling efficient FRET.

1.5. Goals and aims of this thesis

As discussed in Section 1.1, the *O*-GlcNAc modification is a biologically important protein modification that is implicated in several disease states. The overall motivation for the research in this thesis was to develop tools that could be used to probe the functional roles of the *O*-GlcNAc modification by manipulating human OGA in cells and *in vivo* and thereby help uncover the biological functions of *O*-GlcNAc. The specific aims of this thesis were to:

1) Synthesize and evaluate the structure-activity relationships of selective non-covalent OGA inhibitors based on the 2′-aminothiazoline inhibitor family described in Section 1.3.2. Inhibitor characterization involved measurement of K_i values, determining selectivity for human OGA over the lysosomal hexosaminidases, pK_a measurements using nuclear magnetic resonance (NMR) spectroscopy and analysis of TS analogy using linear free energy relationships (LFERs). This research is described in detail in Chapter 2.

2) Synthesize and evaluate two new designs for quenched fluorogenic substrates containing different fluorophores and assess whether they would be compatible for *in vivo* monitoring of OGA. The two sugar substrates synthesized contain either 4-methylumbelliferone or resorufin as the fluorophore positioned at the anomeric position and incorporated thioamide functionalities as the quencher positioned at the 2-position of the pyranose ring. The kinetic properties governing hydrolysis of these compounds catalyzed by OGA have been characterized. The synthesis and kinetic studies are described in Chapter 3.

It is expected that both sets of these chemical tools could be used for various purposes to study a range of biological processes which could help expand our knowledge of the *O*-GlcNAc modification.

Chapter 2.

Structure-activity relationships of aminothiazoline inhibitors of human *O*-GlcNAcase reveals they are sub-nanomolar transition state analogues

The final manuscript draft is currently in preparation for this project. The designated numbers given to synthetic compounds in this Chapter relate only to Chapter 2.

2.1. Contributions

Julia Heinonen and Ernest J. McEachern established the synthetic route for the 2'-aminothiazoline hOGA inhibitors. Julia Heinonen performed the preliminary studies for the K_i measurements with hOGA. Dr. Keith Stubbs synthesized the 4-methylumbelliferone substrates used in the TS analogy study. I re-synthesized the inhibitors, established the K_i values of each inhibitor with hOGA and the human lysosomal hexosaminidase B, performed the TS analogy assays and determined pK_a values using NMR spectroscopy.

2.2. Abstract

In recent years the modification of nuclear and cytoplasmic proteins with *O*-linked *N*-acetylglucosamine (*O*-GlcNAc) has emerged as playing diverse roles in health and disease. Inhibitors of *O*-GlcNAc hydrolase (OGA), the enzyme that removes O-GlcNAc from proteins, lead to increased *O*-GlcNAc levels in cells and *in vivo* and brain permeable OGA inhibitors have been found to modify progression of Alzheimer Disease (AD) pathologies in animal models of AD. Here we uncover the structure-activity

relationships for a series of tight binding 2′-aminothiazoline inhibitors of human OGA (hOGA). The potency and selectivity trends observed include the greater than 1 000 000-fold selectivity observed for hOGA over functionally related human lysosomal β-hexosaminidase for one of these compounds. Using linear free energy analyses, we show a strong linear correlation between binding of these 2′-aminothiazoline inhibitors and the pK_a of the aminothiazoline system. We further show, using an analogous set of synthetic substrates, that 2′-aminothiazoline inhibitors are transition state analogues of hOGA with remarkable ligand efficiencies observed of up to 0.88 kcal/mol/heavy atom for one of these inhibitors. The resulting insights regarding the basis for the high potency and selectivity of these tool compounds should enable their general use in the field and enable the development of compounds with still more desirable properties for use *in vivo* and within clinical applications.

2.3. Introduction

The modification of serine and threonine residues of nuclear and cytoplasmic proteins with terminal *O*-linked β-N-acetylglucosamine (*O*-GlcNAc) has been found on hundreds of proteins.[186-188] Occurring both post-[189] and co-translationally,[190] *O*-GlcNAc has been found in all multi-cellular eukaryotes studied, ranging from humans through to plants[4]. Multiple cycles of addition and removal of the *O*-GlcNAc modification can occur over the lifetime of a protein[17] and global *O*-GlcNAc levels have been shown to vary in response to cellular nutrient availability and stress[191]. Notably, *O*-GlcNAc has been found in some cases to influence protein phosphorylation.[26, 45, 191] These observations have stimulated interest in the physiological roles played by *O*-GlcNAc and research has implicated this modification in various cellular processes, including for example, proteosomal degradation[10, 192, 193] and transcriptional regulation[194]. Additionally, a growing body of literature has implicated *O*-GlcNAcylation as influencing the progress of chronic diseases including neurodegeneration[55, 195-198] and cancer[59, 60, 127, 199-201]. Given the mounting potential in targeting protein *O*-GlcNAcylation for therapeutic benefit, there has been a rising interest in the identification and characterization of small molecule modulators of this pathway in mammals.

Only two enzymes primarily regulate levels of *O*-GlcNAc within cells. The glycosyltransferase uridine diphospho-*N*-acetylglucosamine:peptide β-*N*-acetylglucosaminyl transferase (OGT) installs *O*-GlcNAc residues using uridine diphosphate *N*-acetylglucosamine (UDP-GlcNAc) as the sugar substrate donor.[28, 29] The enzyme responsible for removing *O*-GlcNAc from proteins is *O*-GlcNAcase (OGA),[202, 203] which is a member of glycoside hydrolase family 84 (GH84) of the CAZy classification system[75, 76, 204]. Consistent with the reversible nature of protein *O*-GlcNAcylation, inhibitors of these enzymes have been shown to induce time-dependent changes in cellular *O*-GlcNAc levels.[54, 84, 132, 135] Small molecule inhibitors of OGA, in particular, have emerged as useful research tools that can be used for evaluating the phenotypic effects of increased *O*-GlcNAc levels in cultured cells,[58, 125] as well as *in vivo*[54, 140].

Among the first reported inhibitors of OGA is *O*-(2-acetamido-2-deoxy-D-gluco-pyranosylidene) amino-*N*-phenylcarbamate (PUGNAc, Figure 2.1 C, **1**)[205, 206] (hOGA K_i = 46 nM). This inhibitor, however, has well-described off target effects[125] including the inhibition of the functionally related lysosomal β-hexosaminidases HEXA and HEXB[207] from family GH20. These two lysosomal enzymes cleave β-linked terminal *N*-acetylhexosamine residues from various glycoconjugates including gangliosides. Genetic deficiency of these hexosaminidases result in Tay-Sachs and Sandhoff diseases.[208] More recently identified inhibitors such as 6-acetamido castanospermine (6-Ac-Cas)[125] (hOGA K_i = 300 nM) (Figure 2.1 C, **2**), and 1,2-dideoxy-2'-methyl-α-D-glucopyranoso-[2,1-*d*]-Δ2'-thiazoline (NAG-thiazoline)[138, 209] (hOGA K_i = 70 nM) (Figure 2.1 C, **3**) while reasonably potent, are also non-selective. Given that gangliosides play varied roles in cellular processes ranging from cell membrane structure to cell signaling[210], and some gangliosides play a role in insulin resistance[211, 212], interest has emerged for the development of selective inhibitors of OGA as useful probe molecules[84, 135]. The importance of on-target selectivity and associated potency of OGA inhibitors is reflected by the variation in results obtained when using different inhibitors to interrogate the connection between increased *O*-GlcNAcylation and insulin resistance. The non-selective inhibitor PUGNAc induces insulin resistance[58] whereas other inhibitors, including the non-selective 6-Ac-Cas as well as the selective 1,2-dideoxy-2'-propyl-α-D-glucopyranoso-[2,1-*d*]-Δ2'-thiazoline (NButGT)[209] (Figure 2.1, **4**)

have no effect on glucohomeostasis either in cells or *in vivo*[125, 139, 140] and NButGT has been shown to be well tolerated over a period of months[84, 139, 213].

One approach that has been pursued to create selective OGA inhibitors is to rationally design OGA inhibitors, based on knowledge of the catalytic mechanism and structure of OGA.[209] Detailed mechanistic studies[84, 209] coupled with structural studies of bacterial homologues of hOGA[137, 214] have provided clear support for a catalytic mechanism involving substrate-assisted catalysis in which the 2-acetamido group of the substrate serves as a catalytic nucleophile to generate a transient enzyme-bound oxazoline intermediate (Figure 2.1, A). Notably, in this catalytic mechanism, two aspartate residues (Asp^{174} and Asp^{175}) play key roles as general acid/base catalytic residues.[215] Asp^{174} serves to orient and polarize the 2-acetamido group to aid its attack at the anomeric centre while Asp^{175} facilitates cleavage of the glycosidic bond (Figure 2.1, A,B). Given the clear resemblance of NAG-thiazoline[138] to the oxazoline intermediate, analogues of this molecule in which the 2′-position of the thiazoline ring was modified were generated and showed fair selectivities coupled with moderate nanomolar potencies as exemplified by NButGT (hOGA K_i = 230 nM, K_i(HEX)/K_i(hOGA) = 1,500)[84].

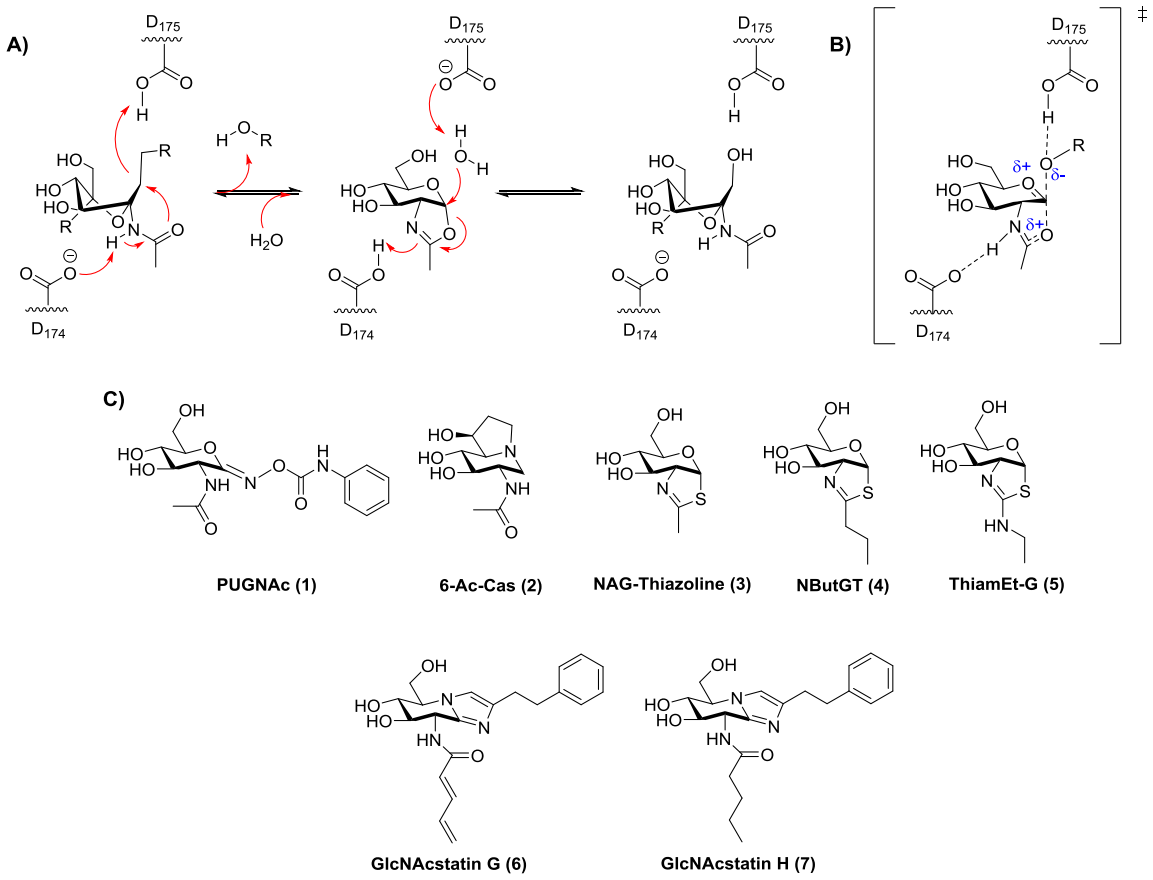


Figure 2.1. The catalytic mechanism of hOGA and its common inhibitors. **A)** hOGA uses a substrate-assisted catalytic mechanism through formation of an oxazoline intermediate.[84] Polarization of the 2-acetamido group by D^{174} facilitates the cyclization of the 2-acetamido carbonyl oxygen onto the anomeric centre, where the leaving group is displaced and its departure is facilitated by the D^{175} general acid, forming an oxazoline ring intermediate. The oxazoline is then broken down by the attack of water at the anomeric position, which is facilitated through general acid catalysis of D^{175} and through the polarization of the acetamido group by general base D^{174}. **B)** The TS for oxazoline intermediate formation. **C)** Inhibitors which have been used to study elevated levels of *O*-GlcNAc in *vitro* and in cellular assays.

More selective hOGA inhibitors have since been generated including a bioisostere of NButGT, 1,2-dideoxy-2'-ethylamino-α-D-glucopyranoso-[2,1-d]-Δ2'-thiazoline (ThiamEt-G) (hOGA K_i = 21 nM, K_i(HEX)/K_i(hOGA) = 37,000)[54] (Figure 2.1 C, **5**). This conveniently accessible inhibitor shows clear effects on brain *O*-GlcNAc levels in mammals and has been used by several groups to show chronic OGA inhibition and increased *O*-GlcNAcylation over several months has no deleterious effects and also

protects against both tau[141, 142, 216] and amyloid pathologies[55, 143] as well as neurodegeneration and cognitive defects in several different transgenic rodent models of AD.

Given the utility of ThiamEt-G and its potential as a disease modifying therapy for AD, we set out to synthesize a panel of 2'-aminothiazoline hOGA inhibitors and to use these to gain an understanding of the basis for potent hOGA inhibition by ThiamEt-G, as well as define structure-activity relationships for such aminothiazoline OGA inhibitors. Herein, we report on the synthesis and characterization of aminothiazoline inhibitors with hOGA and human HexB (hHexB), demonstrate the strongly pK_a-dependent inhibition of hOGA by such inhibitors, and reveal these inhibitors are genuine TS analogues – which explains the picomolar binding and exceptional ligand efficiencies (LE) of the best representatives from this inhibitor family.

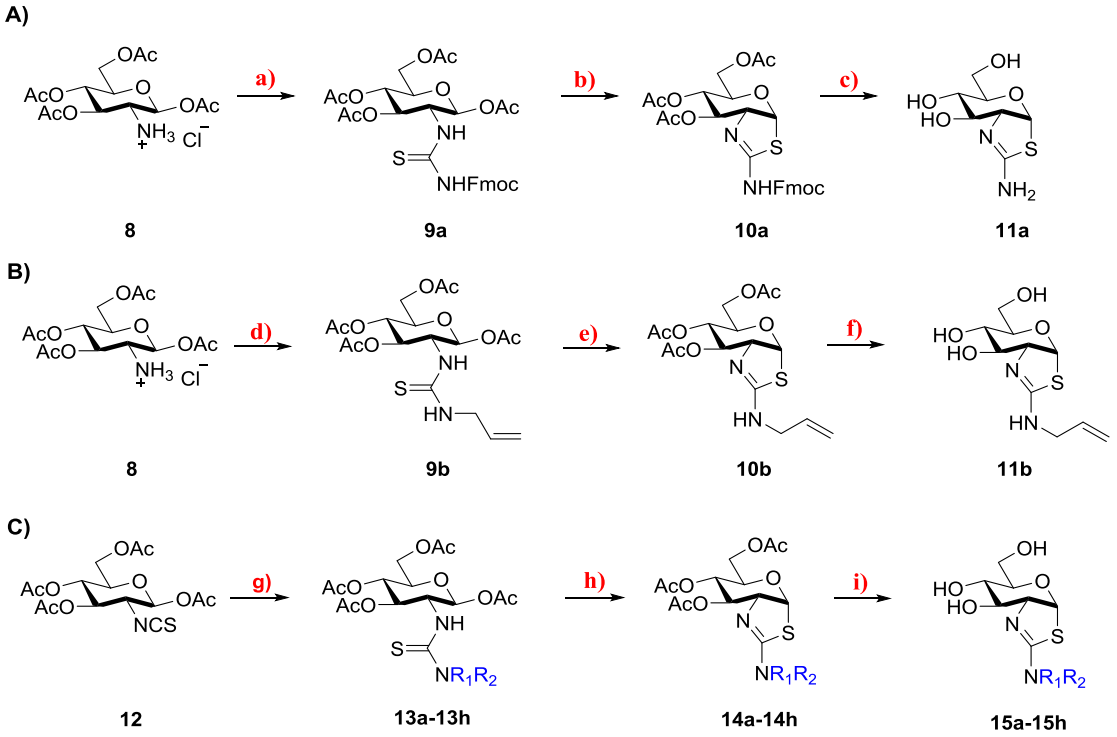
2.4. Results and Discussion

2.4.1. Preparation and evaluation of 2'-alkylaminothiazoline hOGA inhibitors

Given the increasing interest in OGA inhibitors as potential therapeutic agents for a range of pathological processes, we were interested in understanding the basis for the potency of the orally bioavailable ThiamEt-G (Figure 2.1 C, **5**), particularly since it shows significantly improved inhibition of OGA over the isosteric NButGT (Figure 2.1 C, **4**). Early studies showed that varying the 2′-alkyl substituent of NAG-thiazoline resulted in increased selectivity for OGA over the lysosomal hexosaminidases at the slight expense of potency.[84] This trend is observed in a series of glucoimidazole inhibitors[135, 136, 217] whereas it does not hold for PUGNAc analogues which show only modest selectivity[218, 219]. Exploiting a similar approach of increasing steric bulk of the acetamido group, more recent selective hOGA inhibitors include the modified cell-penetrant *gluco*-imidazole inhibitors GlcNAcstatin G (hOGA K_i = 4.1 nM, K_i(HEX)/K_i(hOGA) > 900,000) (Figure 2.1 C, **6**) and GlcNAcstatin H (hOGA = 2.6 nM, K_i(HEX)/K_i(hOGA) = 35,000) (Figure 2.1 C, **7**), the most selective compounds for hOGA from the GlcNAcstatin series A-H[217].

Nevertheless, structural studies have suggested that selective carbohydrate based OGA inhibitors derive their selectivity from structural differences between the active sites of these enzymes in the region that serves to position the 2-acetamido group of the *N*-acetylglucosamine residue of the substrate. We therefore set out to gain an understanding of the detailed structure-activity relationships between the size of substituents at the 2′ position of a series of 2′-aminothiazoline inhibitors and the influence on binding of altered electronic properties of the 2′-aminothiazoline system.

We synthesized a series of analogues using three different approaches (Schemes 2.1 and 2.2). Using the common intermediate hydrochloride salt of 1,3,4,6-*tetra*-*O*-acetyl-2-amino-2-deoxy-β-D-glucopyranose **8**[220] (Schemes 2.1 A,B), which was accessed in three steps following the procedure outlined by Gonzalez *et al*[220], we prepared target compounds **11a** and **11b** by reacting either *N*-fluorenylmethyloxycarbonyl (Fmoc)-protected isothiocyanate or *N*-allyl isothiocyanate with **8** in the presence of triethylamine to generate the respective thiourea intermediates **9a** and **9b**. Subsequent cyclization with excess TFA or $SnCl_4$ followed by a two step de-protection using catalytic NaOMe in anhydrous methanol and piperidine catalyzed removal of the Fmoc group afforded analogue **11a**. Compound **11b** was obtained from protected aminothiazoline **10b** in a one step de-protection using catalytic K_2CO_3 in anhydrous methanol.



Scheme 2.1. Synthesis of 2'-alkylaminothiazoline OGA inhibitors.
A) Synthesis of OGA inhibitor **11a** from intermediate **8.** a) i) NEt_3, DCM; ii) Fmoc-NCS, pyridine, NEt_3; b) $SnCl_4$, pyridine, NEt_3; c) i) NaOMe, MeOH; ii) Piperidine, DMF; **B)** Synthesis of OGA inhibitor **11b** from intermediate **8**. d) allyl-NCS (2 eq), NEt_3 (2 eq), CH_3CN; e) TFA (7.5 eq), DCM; f) K_2CO_3 (5% weight), MeOH; **C)** Synthesis of OGA inhibitors **15a-h** from isothiocyanate intermediate **12**. g) NHR_1R_2.HCl (1.2 eq), NEt_3 (1.2 eq), CH_3CN; h) TFA (7.5 eq), DCM i) K_2CO_3 (5% weight), MeOH; **13a-15a**: R_1 = H, R_2 = CH_3; **13b-15b**: R_1 = CH_3, R_2 = CH_3; **13c-15c**: R_1 = H, R_2 = CH_2CH_3; **13d-15d**: R_1 = H, R_2 = $(CH_2)_2CH_3$; **13e-15e**: R_1 = H, R_2 = $(CH_2)_3CH_3$; **13f-15f**: R_1 = H, R_2 = $(CH_2)_2F$; **13g-15g**: R_1 = H, R_2 = CH_2CHF_2; **13h-15h**: R_1 = H, R_2 = CH_2CF_3.

Inhibitors **15a-h** were synthesized using an alternate route from the common isothiocyanate intermediate **12** (Scheme 2.1, C), which was prepared from **8** via a biphasic reaction with thiophosgene. Reaction of isothiocyanate **12** with a series of alkylamine and dialkylamine hydrochloride salts yielded thioureas **13a**-**h**. Cyclization of these thioureas using excess TFA provided protected aminothiazolines **14a**-**h**, which after de-protection with K_2CO_3 in dry methanol provided **15a**-**h** in overall yields over three steps ranging from 22 to 51%.

2.4.2. K_i determination of 2'-alkylaminothiazoline inhibitors with hOGA and human lysosomal β-hexosaminidase

We then set out to determine the inhibition constants (K_i) for inhibition of hOGA by the series of compounds **11a-b** and **15a**-**h**. Initial experiments to establish these K_i values involved use of traditional Michaelis-Menten inhibition kinetics.[221] However, an assumption of the Michaelis-Menten method is that the change in the concentration of free inhibitor is negligible upon the addition of enzyme. Accordingly, it is generally accepted that the enzyme concentration in the assay should be $<1/7^{th}$ the inhibitor concentration in order to avoid systematic errors. Given the low activity of hOGA coupled with the modest sensitivity of the assay being used, the use of sub-nanomolar hOGA concentrations resulted in low and unreliable initial rates. Nevertheless, this approach proved amenable to study the less potent inhibitors **15e-h,** and provided K_i values and double-reciprocal plots that were consistent with a competitive mode of inhibition, as expected (Appendix A, Figure A1).

Given this limitation of the Michaelis-Menten method for assessing tight-binding inhibitors for which the inhibitor concentration approaches the enzyme concentration, we judged that the K_i values obtained in this way were increasingly unreliable as the potency of the inhibitor increased. We therefore turned to using the non-linear fitting method described by Morrison[222, 223], which can be used to determine K_i values for inhibitors that are present in assays at concentrations similar to the enzyme being studied. The initial mathematical analysis of tight binding inhibitors[222] has been elaborated[222, 224] and further refined[225] to enable fitting of data to a quadratic equation (Equation 2.1), which enables defining the K_i value for an inhibitor through the use of relative rates. This approach depends only on knowing the initial free enzyme and inhibitor concentrations and does not depend on the free inhibitor concentration, $[I]_f$, being comparable to total inhibitor concentration, $[I]_t$. Use of the Morrison approach allowed us to use hOGA concentrations in the 1-20 nM range, ensuring accurate rate measurements. Kuzmic *et al*[226] addressed the determination of K_i values and precise enzyme concentration from dose-response curves using the Morrison approach. Of practical relevance to establishing structure-activity relationships is that in cases where the K_i value is $<0.1\times[E]_o$ or $>1.0\times[E]_0$ ($[E]_0$ = the active enzyme concentration) the precise enzyme concentration cannot be accurately established using this model,

however, the K_i value can nevertheless be accurately determined. *In silico* studies[227] have revealed the ideal experimental conditions for the most accurate determination of K_i, in cases where the mode of inhibition is known. We followed the guidance provided by Murphy to select inhibitor concentrations for our K_i value determinations with the reasonable assumption that all inhibitors would be competitive in nature.

To verify the accuracy of this method we determined K_i values for inhibitor **11b** using both methods and found the results were in reasonable concordance (Figure 2.2).

$$\frac{v_i}{v_o} = 1 - \frac{([\mathrm{E}]_\mathrm{T} + [\mathrm{I}]_\mathrm{T} + \mathrm{K}_\mathrm{i}^{\mathrm{app}}) - \sqrt{([\mathrm{E}]_\mathrm{T} + [\mathrm{I}]_\mathrm{T} + \mathrm{K}_\mathrm{i}^{\mathrm{app}})^2 - 4[\mathrm{E}]_\mathrm{T}[\mathrm{I}]_\mathrm{T}}}{2[\mathrm{E}]_\mathrm{T}}$$

Equation 2.1. The Morrison quadratic equation for determination of K_i for tight-binding competitive inhibitors.

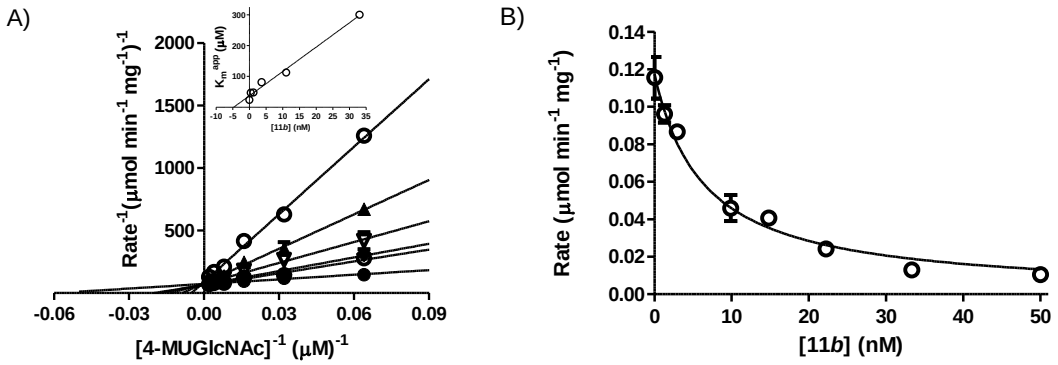


Figure 2.2. Comparison of K_i values determined for compound 11b with hOGA using the Michaelis-Menten and Morrison approaches.
A) Michaelis-Menten analysis of the inhibition of hOGA-catalyzed hydrolysis of 4-methylumbelliferyl 2-acetamido-2-deoxy-β-D-glucopyranoside (4MU-GlcNAc) reveals a pattern of competitive inhibition (K_i = 2.3 nM ± 0.3). The concentrations of **11b** used in the assays were 33.0, 11.0, 3.7, 1.2, 0.4 and 0.0 nM, and that of hOGA was 0.8 nM. **Inset:** graphical determination of the K_i value obtained by plotting K_m apparent (K_m^{app}) values against the concentration of **11b** affords a K_i value of 4.7 nM. **B)** Determination of the K_i value using the Morrison approach (K_i = 3.2 nM ± 0.4). The concentration of hOGA used in the assay was 10 nM.

Having established that the method was sound, we found the K_i values for compounds **11a-b** and **15a-h** toward hOGA ranged from the sub-nanomolar to low nanomolar range (Table 2.1). Notably, by using the Morrison strategy we find that the K_i value for ThiamEt-G was 10-fold lower (K_i = 2.1 nM) than that previously determined[54] using the Michaelis-Menten method (K_i = 21 nM) and is therefore over 100-fold more potent than the isosteric NButGT (K_i = 230 nM). The most tight-binding compound with a K_i of 0.51 nM is **15a** (ThiamMe-G) and is among the most potent glycoside hydrolase inhibitors known as well as the most potent hOGA inhibitor reported to date. Interestingly, we note that there is only a slight decrease in potency for hOGA upon increasing the volume of the 2′-aminoalkyl substituent to the point where the alkyl group is propyl (**15d**, K_i = 2 nM). A butyl chain, however, leads to a greater than 100-fold loss of potency (**15e**, K_i = 350 nM). In concordance, structures of bacterial OGA homologues, in which the active site residues are completely conserved with hOGA[214, 228] show a discretely sized pocket having a volume that nicely accommodates the propyl substituent of **15d.**

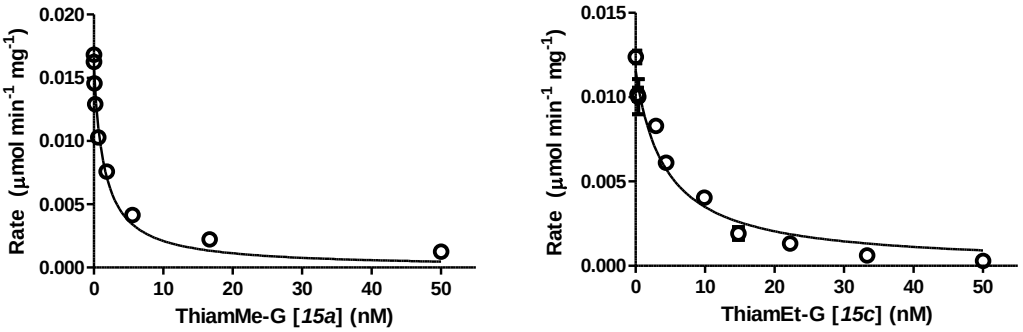


Figure 2.3. Morrison data and fitted data used to obtain K_i values for the tight binding hOGA inhibitors ThiamMe-G (15a) and ThiamEt-G (15c).
15a: K_i = 0.51 ± 0.05 nM; The substrate used in this assay was 4MU-GlcNAc (200 μM) and concentration of hOGA used was 10 nM. Inhibitor concentrations used are successive 3-fold dilutions from 50 to 0.02 nM; **15c**: K_i = 2.1 ± 0.3 nM; The hOGA concentration used was 15 nM and the concentration of 4MU-GlcNAc in the assay was 170 μM. Inhibitor concentrations used are successive 1.5-fold dilutions from 50 to 0.17 nM.

We next set out to assess the selectivity of this series of inhibitors for hOGA over the lysosomal β-hexosaminidases, which are comprised of combinations of α and β subunits that are products of the highly homologous HEXA and HEXB genes. Using purified human hexosaminidase B (hHexB, a kind gift from Michael Tropak, Sick Children's Hospital) we determined the range-finder K_i values for compounds **11a** to **15h** using Dixon plot analysis[229] (Appendix A, Figure A4) and found remarkably high inhibitor selectivities ranging from 1100 to 1, 850, 000-fold preference for hOGA (Table 1). To confirm these Dixon plot analyses we used the Michaelis-Menten method to measure K_i values for hHexB for compounds **11a** and **15a**, since these bear the smallest 2′-aminoalkyl substituents (Appendix A, Figure A3). Remarkably, we observed that inhibitors **11a** and **15a** still retain 1100 and 3400-fold selectivity for hOGA, respectively, despite their similarity in size to NAG-thiazoline, which itself demonstrated no selectivity. Accordingly, the presence of the 2′-amino substituent, on its own, confers at least 1000-fold selectivity for hOGA over hHexB. We also noted that the selectivity ratio for hOGA increases as the 2′-aminoalkyl chain length increases to the three-carbon propyl derivative (**15d**), but this trend reverses once the chain length increases beyond this point, as seen for the 2′-aminobutyl analogue (**15e**). Since the active site pocket for hHexB is more constrained in the vicinity of the acetamido group, these observations are

consistent with structural observations of both bacterial hOGA homologues[214] and hHexB[230] and suggest that once the 2′ substituent passes the volume that can be accommodated in the active site of hOGA, increases in bulk are even more deleterious for hOGA as compared to hHexB.

One additional observation of interest is that ThiamEt-G binds over 100-fold more tightly than NButGT. The basis for this increase in potency and selectivity was previously speculated[54] to stem from an increase in the basicity of ThiamEt-G as compared to NButGT. Detailed mechanistic studies in combination with pH-rate profiles of wild-type and mutant hOGA revealed the key catalytic residue Asp^{174} in the OGA catalytic site (Asp^{242} in the bacterial OGA, *Bt*GH84) acts as a general base to assist the attack of the substrate 2-acetamido group onto the anomeric centre. The kinetic pK_a of this residue was determined to be 5.2 so that at physiological pH this residue is expected to be in its carboxylate form[215] and therefore suitably ionized in the resting enzyme to facilitate catalysis. Accordingly, this inhibitor is unlikely to engage in productive electrostatic ionic interactions with the carboxylate residue Asp^{174}. Given that aminothiazolines are known to be more basic than thiazolines,[231] it was speculated that installation of the 2′-alkylamino group would increase the basicity of ThiamEt-G as compared to NButGT and thereby contribute to its enhanced potency relative to inhibitors bearing 2′-alkyl groups at physiological pH. Structural data of ThiamEt-G bound within the active site of *Bt*GH84 was consistent with this proposal, revealing that both the endo and exocyclic amines interacted with Asp242[54].

We set out to evaluate this proposal and assess the importance of the inhibitor pK_a on potency by preparing and studying a series of 2′-aminothiazoline inhibitors in which increasing numbers of fluorine substituents were progressively substituted for hydrogen at the terminal methyl group (**15f-h**). These inhibitors were synthesized as described above (Scheme 2.2) in good overall yields over three steps ranging from 40-51% (starting from isothiocyanate intermediate **12**). Evaluation of the potency of these compounds revealed a progressive increase in K_i values upon increasing substitution with fluorine (K_i = 2.1 (CH_3, **15c**) 15 (CH_2F, **15f**), 60 (CHF_2, **15g**) and 1000 (CF_3, **15h**) nM), which is consistent with the electron withdrawing fluorine substituents leading to a decrease in the pK_a value of the protonated 2′-aminothiazolium inhibitors and supports

the proposal that an ionic interaction occurs between Asp^{174} and the 2′-aminothiazoline inhibitor.

Table 2.1. K_i selectivity ratios of inhibitors 11a to 15h for hOGA over hHexB.

Inhibitor	hOGA K_i (nM)[a]	hHexB K_i (μM)[b]	(hHexB/hOGA)[c]
11a: R_1=R_2=H	4.7 ± 0.3	5.0 ± 0.6[d]	1100
15a: R_1=H, R_2=CH_3	0.51 ± 0.05	1.7 ± 0.2[d]	3300
15c: R_1=H, R_2=CH_2CH_3	2.1 ± 0.3	740 ± 60[54]	350000
15b: R_1=R_2=CH_3	2.4 ± 0.2	13.0 ± 3.8	5400
11b: R_1=H, R_2=CH_2CHCH_2	3.2 ± 0.4	2850 ± 570	950000
15d: R_1=H, R_2= $(CH_2)_2CH_3$	2.0 ± 0.2	3700 ± 670	1850000
15e: R_1=H, R_2= $(CH_2)_3CH_3$	350 ± 90[d]	4800 ± 763	13700
15f: R_1=H, R_2=$(CH_2)_2F$	15 ± 5[d]	180 ± 44	12000
15g: R_1=H, R_2=CH_2CHF_2	60 ± 10[d]	150 ± 50	2500
15h: R_1=H, R_2=CH_2CF_3	1000 ± 200[d]	4200 ± 1525	4200

[a]Determined through a continuous in vitro enzymatic assay at pH 7.4 using 4MU-GlcNAc as the substrate. K_i values were established with the Morrison K_i fit if the values are in the low-nanomolar range.
[b]Determined from an *in vitro* assay at pH 4.25 using 4MU-GlcNAc as a substrate and fitted using Dixon plot analysis.
[c]Selectivity ratios representing the favored selectivity for hOGA over hHexB.
[d]Determined from an *in vitro* assay at pH 4.25 (hHexB) or pH 7.4 (hOGA) using 4MU-GlcNAc as a substrate and fitted using Michealis-Menten K_i competitive inhibition analysis.

2.4.3. Determination of pK_a values for ThiamEt-G, NButGT and fluorinated 2'-alkylaminothiazoline inhibitors using ^{13}C NMR spectroscopy

To understand better the relationship between inhibitor basicity and potency we used ^{13}C NMR titration to determine the pK_a values of these compounds. Traditional methods of measuring pK_a include performing potentiometric,[232] conductivity-based[233] and spectrophotometric titrations[234, 235]. The use of NMR spectroscopy is a convenient technique that is operationally facile and enables parallel determination of the pK_a values of several compounds at one time, enabling highly accurate measurements. The chemical shifts of ^{1}H and ^{13}C resonances in molecules, as well as other nuclei, have long been known to be affected by changes in pH.[236] This pH dependence of chemical shift permits establishing the pK_a of a compound by plotting a titration curve consisting of chemical shift as a function of pH and fitting it to the

Henderson-Hasselbalch equation. The chemical shift is related to the mol fraction of protonated and de-protonated species at a specific pH as described by Equation 2.2.

$$\delta = x_p\delta_p + x_d\delta_p$$

Equation 2.2. The relationship between chemical shift (δ) and the mol fraction of protonated (x_p) and de-protonated (x_d) species at a specific pH.

Additionally, NMR titration can be used to determine the ratio of pK_a values between two species including, for example, a compound having an unknown pK_a and a reference compound having a well established pK_a value[237, 238] or between two structurally similar compounds[237, 239]. This approach enables measuring the relative pK_a (Δ pK_a) between one or more compounds as first described by Robinson and Ellison.[240] The main advantage of this relative measurement approach is that it does not require repeated pH determination throughout the titration, which makes it convenient to measure very small pK_a differences between compounds since the differences in the frequencies of resonances can be determined in a highly accurate manner.[240, 241] A non-linear plot of the difference between the frequencies of resonances of interest of the two compounds (Δ ppm) as a function of the fractional protonation of the reference compound (n) can then be used to determine the ratio (R) of the acid dissociation constants between the two compounds. It is from this R value which is obtained by fitting the mentioned set of data to Equation 2.3, that the Δ pK_a can be determined.[242] The important requirement for accurate fitting of the data to Equation 2.3 is to be able to precisely define Δ_d (the difference in the chemical shifts for the fully de-protonated reference compound *r* and fully de-protonated compound of interest, *x*), Δ^x (the difference in the chemical shifts for the fully deprotonated and protonated compound of interest *x*) and Δ^r (the difference in the chemical shifts for the fully deprotonated and protonated reference compound *r*). Technically, this approach benefits from not having to measure the pH throughout the titration, however, there are additional advantages in using ^{13}C NMR over ^{1}H NMR to measure Δ pK_a.: Deuterated solvents and acids or bases are not required, and mixtures of compounds with more than one protonation state can be studied due to the high selectivity of the ^{13}C nucleus to structural parameters.[243]

$$\Delta = \Delta_d + \frac{\boldsymbol{Rn}(\Delta^{\boldsymbol{x}})}{\boldsymbol{Rn} - \boldsymbol{n} + \mathbf{1}} - \boldsymbol{n}(\Delta^{\boldsymbol{r}})$$

Equation 2.3. The relationship between the difference in chemical shifts for a pair of resonances between a compound of interest *x* and a reference compound *r* (Δ) and the fractional protonation (*n*) of the reference compound *r.*
$\Delta_d = \delta_d^r - \delta_d^x$; $\Delta^x = \delta_d^x - \delta_p^x$; $\Delta^r = \delta_d^r - \delta_p^r$; $R = K_a^x/K_a^r$ and *n* = fractional concentration of the protonated form of r, where $n = (\delta-\delta_d)/(\delta_p-\delta_d)$.

Using ^{13}C NMR this method has been used to provide a useful fit for determining comparative pK_a (Δ pK_a) values of organic acids.[243] We adopted this method to determine Δ pK_a and absolute pK_a values for compounds **15c** and **15f-h,** by titrating these inhibitors in the presence of a convenient reference compound, 3-nitrophenol (pK_a = 8.42)[244], which has a single site of protonation in the range of interest and has ^{13}C resonances far downfield from the majority of the resonances for the inhibitors. We observed the 2′-carbon of the aminothiazoline system and the **C-1** carbon of 3-nitrophenol since these are most proximal to the ionizable centre. Fitting of the resulting data to Equation 2.3 (Figure 2.4) enabled us to obtain values of R for each compound and, by comparison to the reference compound, absolute pK_a values (Figure 2.4, Table 2.2). The pK_a value of NButGT was also determined in this way (Appendix 1, Figure A6, Table A2). The observed pK_a values range from 7.68 to 5.44 and fall in the expected order with the conjugate acid of the trifluoromethyl-containing analogue (**15h**) being the most acidic between compounds **15c** and **15f-h**.

Examination of the relationship between the pK_a value of the inhibitor and its corresponding K_i against hOGA reveals a clear trend. To quantify the extent to which potency depends on the pK_a value of the inhibitor, we determined the extent of correlation between the pK_a value of inhibitors (**15c, 15f-h** and NButGT) with the corresponding log K_i. The resulting linear free energy relationship (LFER) shows a strong linear correlation (R^2 = 0.9040), (Figure 2.5) which is consistent with the pK_a value of inhibitors dominating the effect of binding as compared to steric effects associated with increasing fluorine substitution. Notably, we also find that NButGT, which is isosteric to ThiamEt-G (**15c**), falls reasonably well in this correlation of 2′-aminothiazoline inhibitors. These data collectively suggest that the pK_a value of the inhibitor, either by

favouring the protonated inhibitor form or by optimizing hydrogen bonding, influences favorable interactions within the active site.

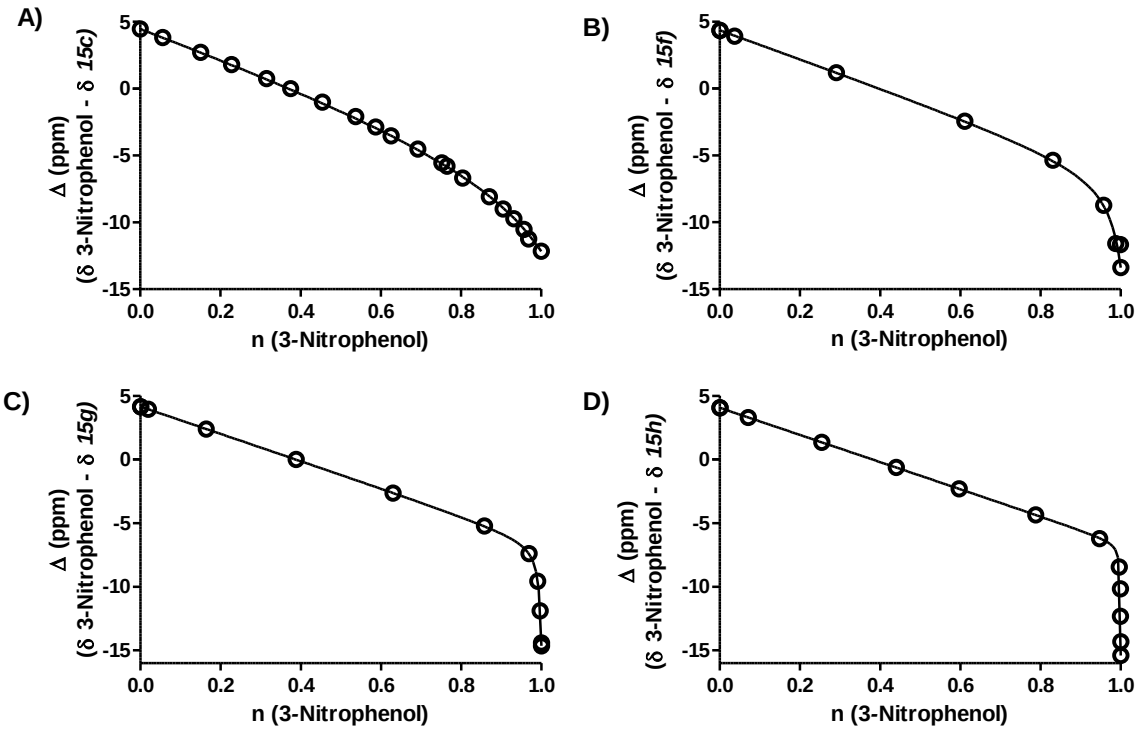


Figure 2.4. Chemical shifts of the 2′-C of the thiazoline ring resonances (Δ ppm) for compounds 15c, 15f-h as a function of the fractional protonation (n) of the reference compound, 3-nitrophenol (pK_a of 8.42).
The changes in chemical shift (Δ ppm) of **C**-1 of 3-nitrophenol were used to obtain the fractional protonation of 3-nitrophenol [n = $(\delta-\delta_d)/(\delta_p-\delta_d)$]. These data were fitted to Equation 2.3 and the fitted curves are shown in each panel with the R values summarized in Table 2.2. for each compound; A) R (K_a^{15c}/K_a^r) = 18.2 ± 0.4 x 10^{-2}; R^2 = 0.9996; B) R (K_a^{15f}/K_a^r) 3.2 ± 0.6 x 10^{-2}; R^2 = 0.9955; C) R (K_a^{15g}/K_a^r) = 5.8 ± 0.2 x 10^{-3}; R^2 = 0.999; D) R (K_a^{15h}/K_a^r) = 1.2 ± 0.2 x 10^{-3}; R^2 = 0.9942.

Table 2.2. **R, Δ pK_a, absolute pK_a, and fitting parameters determined from each titration for compounds 15c, 15f, 15g, 15h and NButGT using 3-nitrophenol as a reference compound.**

Compound	[1]Δ_d (ppm)	[2]Δ^r (ppm)	[3]Δ^x (ppm)	R (K_a^x/K_a^r)	[4]Δ pK_a	[5] pK_a
15c	4.47	10.47	-6.16	$18.2 \pm 0.4 \times 10^{-2}$	-0.74	7.68
15f	4.36	10.63	-7.12	$3.2 \pm 0.6 \times 10^{-2}$	-1.50	6.92
15g	4.15	10.65	-8.15	$5.8 \pm 0.2 \times 10^{-3}$	-2.24	6.18
15h	4.06	10.65	-8.80	$1.2 \pm 0.2 \times 10^{-3}$	-2.98	5.44
NButGT	-8.74	10.67	-10.56	$6.0 \pm 1.0 \times 10^{-4}$	-3.25	5.17

[1]The difference between the ^{13}C chemical shift of the de-protonated inhibitor (2'-aminothiazoline carbon) and 3-nitrophenol (**C**-O carbon) ($\Delta_d = \delta_d^r - \delta_d^x$); Determined for each point during the course of the titration.
[2]The chemical shift difference between the de-protonated and protonated 3-nitrophenol **C**-O ($\Delta^r = \delta_d^r - \delta_p^r$).
[3]The chemical shift difference between the de-protonated and protonated thiazoline carbon ($\Delta^x = \delta_d^x - \delta_p^x$).
[4]Δ pK_a = log R.
[5] pK_a = 8.42 - Δ pK_a for each inhibitor.

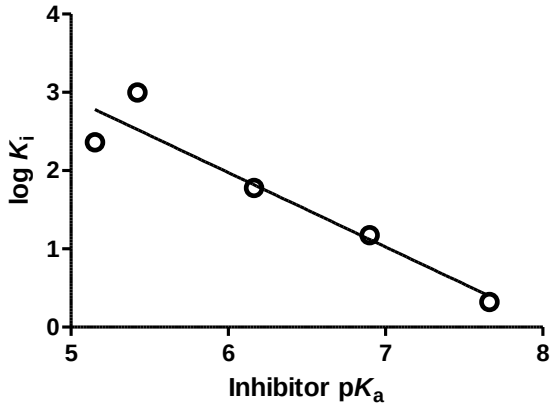


Figure 2.5. **Linear free energy relationship (LFER) analysis between the pK_a and log K_i of compounds 15c, 15f-h and NButGT.**
The slope observed is m = -0.95 ± 0.18.

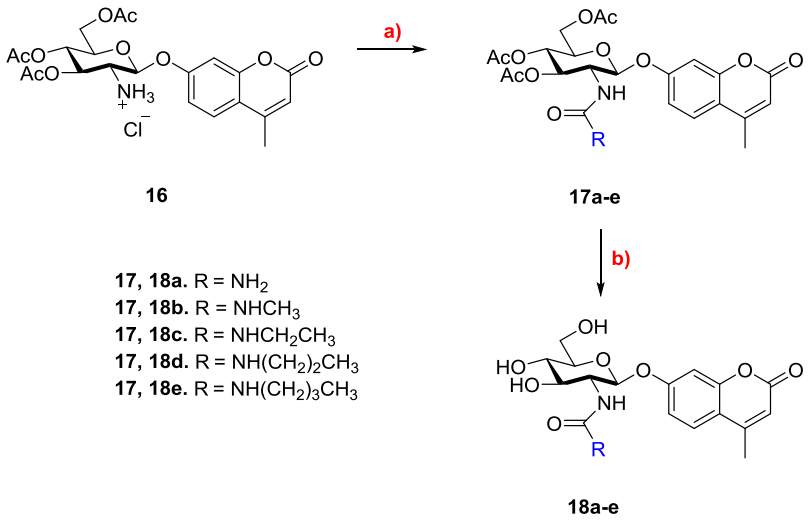
2.4.4. Assessment of 2'-aminothiazoline inhibitors as transition state analogs

The high potency of these 2′-aminothiazoline inhibitors prompted us to consider their potency in the context of their size. One widely used parameter to understand the efficiency of binding as a function of molecular weight is to consider the Ligand Efficiency (LE) of a ligand.[245, 246] This measure provides the binding affinity of the compound as a measure of the number of heavy atoms. We find remarkably high LEs of up to 0.88 kcal/mol/heavy atom for compound **15a.** Such a LE is comparable to some of the highest LEs observed for compounds of this size range (10-50 heavy atoms)[247] and suggested to us that these compounds could well be TS analogues, as had been observed for the related thiazoline inhibitors[131].

Tight-binding inhibitors that bear resemblance to enzyme substrates or intermediates have often been considered to be TS analogues simply by virtue of their potency. This tendency stems from early proposals by Pauling[91] and Wolfenden[92] who suggested that enzymes catalyze their reactions by tightly binding, and thereby stabilizing, their TS structures. By extension, stable molecules that resemble the TS in terms of their geometry and charge distribution should be tightly bound by an enzyme.[248, 249] Because it is catalytically most productive for active site interactions to stabilize the TS, rather than the ground state, TS analogues should show a correlation between the free energy change on going from the free enzyme and substrate to the enzyme stabilized TS ($\Delta G^{\neq}$) and the free energy of binding of the inhibitor ($\Delta G_{binding}$). The corollary is that no correlation should be observed between the free energy of binding of the inhibitor ($\Delta G_{I\text{-}binding}$) and the free energy of binding for the substrate ($\Delta G_{S\text{-}binding}$). Bartlett formalized this expectation and provided compelling experimental evidence in its support using LFERs.[249, 250] Using this method, genuine TS analogues yield plots of log K_m/k_{cat} values, for a series of substrates having defined structural differences, versus log K_i values, for a series of inhibitors having the analogous structural changes, that show linear correlations having a slope of unity. Furthermore, log K_m versus log k_{cat} are not correlated for TS analogues but do correlate for ground state analogues.[250] An earlier study using this LFER approach showed that NAG-thiazoline analogues are actually TS analogues despite their obvious resemblance to the oxazoline intermediate formed along the reaction pathway.[131] Given the greater than 100-fold increase in

potency we observe for the 2′-aminothiazoline inhibitors and their similarity to the oxazoline intermediate found along the reaction coordinate of hOGA as well as NAG-thiazoline, we wanted to assess whether incorporation of the 2′amino group benefitted binding through serendipitous interactions, or whether the presence of the charge included in this class of inhibitors also made them TS analogues. We therefore turned to using this Bartlett LFER approach to rigorously evaluate these 2′-aminothiazoline inhibitors as TS analogues.[225, 250]

With the series of inhibitors in hand we synthesized for use in the LFER studies a series of fluorogenic 4-methylumbelliferyl 2-deoxy-2-urea-β-D-glucopyranoside substrates (**18a-e**, Scheme 2.2) bearing *N*-alkyl substituents on the terminal urea nitrogen that correspond to those alkyl groups present on the series of 2′-aminothiazoline inhibitors (**11a, 15a, 15c-e**). We started from the known 4-methylumbelliferyl 2-amino-2-deoxy-β-D-glucopyranoside hydrochloride as a common intermediate (**16**), which was prepared as described by Roeser and Legler.[251] Per-*O*-acetylated urea substrates **17a**-**e** were prepared by reacting with the appropriate alkyl isocyanate in the presence of triethylamine. Subsequent deprotection was carried out using Zemplen conditions.[252]



Scheme 2.2. Synthesis of 4-methylumbelliferyl fluorogenic substrates 18a-e. a) R-NCO, Et_3N, CH_3CN; b) i) NaOMe, MeOH; ii) Dowex 50-H^+.

With this series of substrates (**18a-e**) in hand we determined the k_{cat}/K_m values governing their hOGA catalyzed hydrolysis (Table 2.3) and compared them to the more natural *N*-acyl substrates which were previously synthesized by Whitworth *et al* (**19a-e**)[131]. We observed ranging values which dropped off over 250-fold on going from the *N*-butyl-2-urea substrate (**18e**) to MUGlcNAc (**19a**) and 0.5-fold on going from the *N*-methyl-2-urea substrate (**18b**) to MUGlcNAc (**19a**).

Table 2.3. Summary of parameters used to determine the correlation between K_m/k_{cat} and K_i for the urea and *N*-acyl substrates and inhibitors with parallel structural alterations.

Compound	K_m (μM)	k_{cat}/K_m ($\mu M^{-1} min^{-1}$)	log K_m/k_{cat}	K_i (μM)	log K_i
18a: R=NH_2	[a]1265 ± 259	[b]1.75 x 10^{-3}	2.76	4.7 ± 3.0 x 10^{-3}	-2.32
18b: R=$NHCH_3$	[a]1825 ± 542	[b]1.36 x 10^{-2}	1.86	5.1 ± 0.5 x 10^{-4}	-3.30
18c: R=$NHCH_2CH_3$	85 ± 20	[b]3.49 x 10^{-3}	2.46	2.1 ± 0.3 x 10^{-3}	-2.68
18d: R=$NH(CH_2)_2CH_3$	120 ± 11	[b]2.55 x 10^{-3}	2.59	2.0 ± 0.2 x 10^{-3}	-2.69
18e: R=$NH(CH_2)_3CH_3$	30 ± 3	[b]3.01 x 10^{-5}	4.52	3.5 ± 0.9 x 10^{-1}	-0.46
19a: R=CH_3	---	[c]7.69 x 10^{-3}	2.11	---	---
19b: R=CH_2CH_3	---	[c]6.67 x 10^{-3}	2.18	---	---
19c: R=$(CH_2)_2CH_3$	---	[c]5.26 x 10^{-3}	2.28	---	---
19d: R=$(CH_2)_3CH_3$	---	[c]5.56 x 10^{-4}	3.25	---	---

[a]Values were estimated by non-linear regression of the Michealis-Menten data since saturation of hOGA with these substrates was not observed.

[b]Values of k_{cat}/K_m were determined from the linear regression of the second-order region of the Michealis-Menten plot.

[c]Values used were previously determined by Whitworth *et al*[131] and used for a comparison between the thiazoline and aminothiazoline series of inhibitors as transition state analogues.

Plotting these data to assess TS analogy, we observe a strong correlation (R^2 = 0.9950) with a slope of 1.08 ± 0.04 that is seen between log K_i values for the inhibitors versus the log K_m/k_{cat} values for the series of urea substrates containing the analogous substitutions (Figure 2.6, A). In contrast, we find no correlation between the log K_i values for the inhibitors and log K_m values for the series of substrates (Figure 2.6, B). These results indicate the 2′-aminothiazoline inhibitors are TS analogues for the hOGA catalyzed hydrolysis of urea substrates, with parallel changes in $\Delta\Delta G_{I\text{-binding}}$ and $\Delta\Delta G^{\neq}$ being induced by structural variations within the inhibitors and substrates. These observations are consistent with the observations of Whitworth who found that NAG-thiazoline and its analogues were transition TS analogues of the hOGA-catalyzed hydrolysis of *N*-acyl substrates.[131]

Reflecting on this data, we considered that our observations show 2′-aminothiazoline inhibitors are TS analogues for the hOGA-catalyzed hydrolysis of unnatural urea substrates (**18a-e**), however, we were curious as to whether these inhibitors would also be TS analogues for the hOGA-catalyzed hydrolysis of the more natural *N*-acyl substrates. To address this, we examined the correlation between log K_i values for the same series of 2′-aminothiazoline inhibitors and the series of *N*-acyl substrates (Figure 2.6, C, substrates **19a-e**) having the analogous structural changes for which k_{cat}/K_m values are reported (Figure 2.6, D).[131] We find a strong correlation (R^2 = 0.9768) with a slope of 2.3 ± 0.3. For this analysis we excluded the 2-aminothiazoline (**11a**) because of its unexpectedly poor inhibition of hOGA (Table 1). The steep slope observed for this series of 2-acyl substrates in correlation with the aminothiazoline inhibitors is surprising. However, slopes other than unity are precedented.[253, 254] This steep slope may be interpreted to mean that the TS for the hOGA catalyzed hydrolysis of *N*-acyl substrates bears less positive charge in the forming oxazoline ring system, perhaps because the amide proton is in flight in the TS, as compared to the *N*-urea substrates (Figure 2.6, E and F) which are expected to be more basic and therefore may lead to the formation of 2′-aminooxazolinium ion intermediates that retain their proton. The TS leading to such amino-oxazolinium ion intermediates are expected to have more positive charge than the corresponding transition state leading to the oxazoline intermediate. In keeping with this proposal, it is notable that site-directed deletion of the side chain of Asp^{174}, which is the catalytic general acid/base catalytic

residue that interacts with the acetamido group of the substrate, leads to a similar drop of between 150 to 750-fold in second order rate constant[215] as seen on going from *N*-acyl to *N*-urea substrates (250-fold).

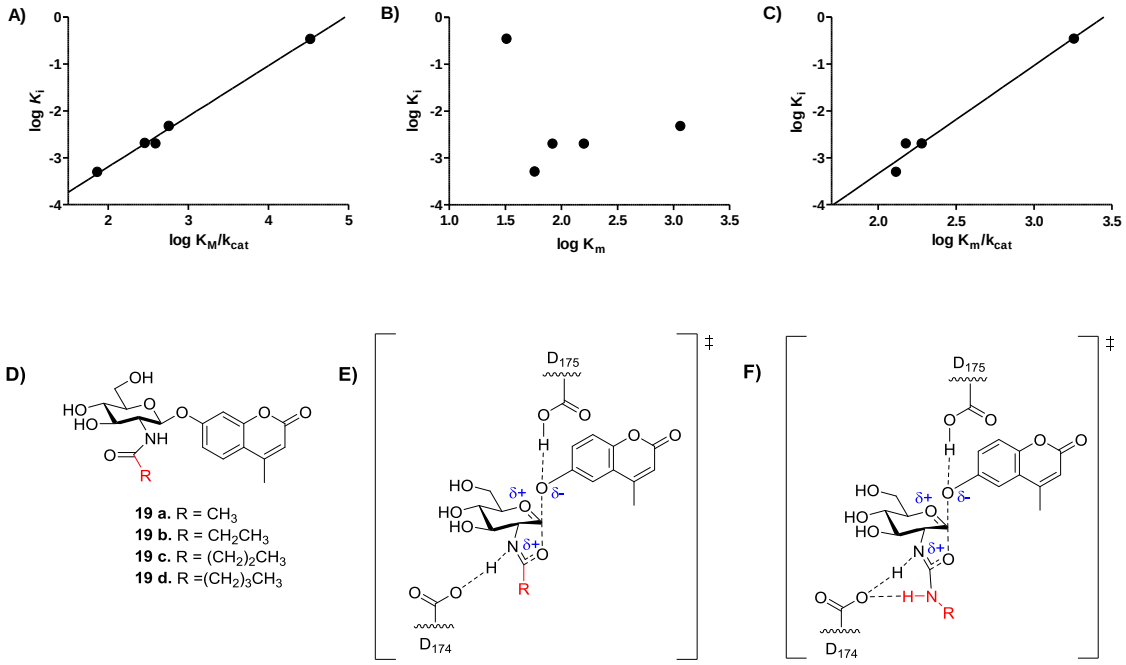


Figure 2.6. Transition state analogy diagrams for substrates 18a-e, 19a-d and inhibitors 11a, 15a, 15c-e.
A) The correlation plot between log K_m/k_{cat} of substrates **18a-e** and log K_i of inhibitors **11a**, **15a** and **15c-e** with a slope of 1.08 ± 0.04 (R^2 = 0.9950). **B)** The correlation plot between the log K_m of substrates **18a-e** and log K_i of inhibitors **11a**, **15a**, and **15c-e**. **C)** The correlation plot between the log K_m/k_{cat} of substrates **19a-d** and log K_i of inhibitors **15c-e** with a slope of 2.31 ± 0.25 (R^2 = 0.9768). **D)** Substrates **19a-d** synthesized by Whitworth *et al*[131]. **E)** Transition state in the hOGA active site of substrates **19a-d**. **F)** Transition state in the hOGA active site of substrates **18a-e.**

Accordingly, these data indicate that 2-aminothiazoline inhibitors are TS analogues for hOGA, by virtue of both their shape and general charge distribution. However, the steep slope observed for the LFER between K_m/k_{cat} values observed for *N*-acyl substrates and the log K_i values seen for the 2-aminothiazoline inhibitors suggests that this feature lends improved binding over the corresponding partial charge that likely develops for the TS found for the hOGA catalyzed processing of natural *N*-acyl-containing substrates.

2.5. Conclusions

In summary, we report the structure-activity relationships for the second generation amino-thiazoline inhibitors for human OGA, identifying a series of compounds with picomolar and low nanomolar K_i values. This potency of this inhibitor family could be attributed to their pK_a values since a clear LFER was shown between the pK_a and log K_i of 5 of the compounds with gradual structural alterations. These compounds, like the alkyl thiazoline inhibitors of the NButGT family, are also tight-binding TS analogues for hOGA. Their selectivity over the lysosomal hexosaminidases allows for efficient and selective targeting of hOGA in future enzymatic and cellular assays where elevated levels of *O*-GlcNAc are to be studied.

Overall, this family of inhibitors is a useful set of chemical tools to be able to target hOGA both *in vitro* and *in vivo* effectively and without off-target effects.[54]The high potencies of these compounds ensure that a very small amount of inhibitor would need to be used for cell or animal studies, which means that the functionally related lysosomal hexosaminidases would not be affected, particularly given the great selectivities observed. These inhibitors meet many of the criteria highlighted in Section 1.3, for the characteristics of an ideal inhibitor, as well as the criteria outlined by Frye[255] who describes that one of the qualities of an effective chemical probe is to demonstrate *in vitro* potency and selectivity data which can be translated into its cellular or *in vivo* data. Additionally, the synthetic simplicity of these inhibitors makes them readily available for others to use to answer their biological questions related to elevated *O*-GlcNAc.

2.6. Experimental Section

2.6.1. General Procedures

All dry solvents and buffer salts were used as purchased from Sigma-Aldrich and all reagents were utilized without further purification. The progress of all reactions was monitored on Merck pre-coated silica gel plates using combinations of ethyl acetate/n-hexane (EtOAc/Hexanes), ethyl acetate/dichloromethane (EtOAc/DCM), or

methanol/dichloromethane (MeOH/DCM) solvent elution systems. Flash chromatography was performed under a positive pressure with Fisher Scientific silica gel (230-400 mesh) where spots were visualized by irradiation with ultraviolet light (254 nm), staining with $KMnO_4$ and charring with 10% ammonium molybdate in 2M H_2SO_4 upon heating. Proton (^1H) and carbon (^{13}C) NMR spectra were established on either a Bruker Avance 500 (500 MHz for ^1H, 125 MHz for ^{13}C), Bruker Avance 400 (400 MHz for ^1H, 101 MHz for ^{13}C), or Bruker Avance II 600 (600 MHz for ^1H, 151 MHz for ^{13}C) using $CDCl_3$, DMSO-d_6, CD_3OD, or D_2O as the sample solvent. Chemical shifts are given in parts per million (ppm) (δ relative to a residual solvent peak for ^1H and ^{13}C). High resolution mass spectrometry (HRMS) analysis was performed using a Bruker maXis TOF LC/MS/MS instrument. hHexB was obtained from the Micheal Tropak group at the Hospital for Sick Children in Toronto, Ontario.

2.6.2. Synthesis of 2'-alkylaminothiazoline inhibitors

The per-*O*-acetylated glucosamine salt **8** was made using a known literature procedure by Cunha *et al*[256], after which its conversion to the carbohydrate isothiocyanate intermediate **12** was performed following the procedure described by Gonzalez *et al*[41]. Compound **15c** (ThiamEt-G) was previously synthesized and has been characterized and described by Yuzwa *et al*.[24]

1,3,4,6-*tetra*-*O*-acetyl-2-amino-2-deoxy-β-D-glucopyranose hydrochloride (8). ^1H-NMR (400 MHz, CD_3OD) δ (ppm) 5.87 (d, J = 8.8Hz, 1H), 5.37 (dd, J = 10.5, 9.2 Hz, 1H), 5.11 (dd, J = 10.0, 9.2 Hz, 1H), 4.33 (dd, J = 12.6, 4.6 Hz, 1H), 4.14 (dd, J = 12.6, 2.3 Hz, 1H), 4.04 (ddd, J = 10.1, 4.6, 2.3 Hz, 1H), 3.62 (dd, J = 10.5, 8.8 Hz, 1H), 2.22 (s, 3H), 2.12 (s, 3H), 2.06 (s, 3H), 2.05 (s, 3H); ^{13}C-NMR (125 MHz, H_2O+D_2O) δ (ppm) 173.53, 172.97, 172.64, 171.17, 90.41, 72.12, 70.87, 67.99, 61.42, 52.44, 20.21, 20.19, 20.08, 20.03; HRMS (m/z): [M+H]+ calculated for $C_{14}H_{22}NO_9$: 348.1295; found 348.1250; [M+Na]+ calculated for $C_{14}H_{21}NNaO_9$: 370.1114; found 370.1114; [M+K]+ calculated for $C_{14}H_{21}NKO_9$: 386.0853; found 386.0849. The characterization data matched the data described by Cunha *et al.*[256]

Preparation of isothiocyanate intermediate 12. Synthesis was performed with minor modifications to the procedure described by Gonzalez *et al.[220]* After combining amine **8** (10.0g, 26.1 mmol), and $CaCO_3$ (7.9 g, 78.3 mmol) in a round bottom flask at room temperature, DCM and H_2O (17 mL each) were added. Thiophosgene (3.0 mL, 39.2 mmol) was then added drop-wise, followed by further addition of DCM and H_2O (13 mL each). The reaction was monitored by TLC and stirred at room temperature for3 hours. The aqueous layer of the mixture was extracted twice with DCM. The organic layers were collected, dried over $MgSO_4$, filtered and concentrated to afford the crude compound **12** as a sticky dark yellow solid (8.73g, 86%). The product was used with no further purification necessary.

1,3,4,6-*tetra*-*O*-acetyl-2-deoxy-2-isothiocyanato-β-D-glucopyranose (12). ^{1}H-NMR (400 MHz, $CDCl_3$) δ (ppm) 5.69 (d, J = 8.6 Hz, 1H), 5.29 – 5.18 (t, 1H), 5.04 – 4.94 (t, 1H), 4.27 (dd, J = 12.4, 4.3 Hz, 1H), 4.13 – 4.03 (m, 1H), 3.96 (dd, J = 10.3, 8.6 Hz, 1H), 3.88 – 3.81 (m, 1H), 2.17 (s, 3H), 2.08 (s, 3H), 2.05 (s, 3H), 2.01 (s, 3H); ^{13}C-NMR (100 MHz, $CDCl_3$) δ (ppm) 170.45, 169.58, 169.48, 168.49, 142.07, 91.72, 72.93, 72.54, 67.46, 61.29, 59.40, 20.83, 20.65, 20.62, 20.50; HRMS (m/z): [M+Na]+ calculated for $C_{15}H_{19}NNaO_9S$: 412.0678; found 412.0677; [M+K]+ calculated for $C_{15}H_{19}KNO_9S$: 428.0418; found 428.0414.

Synthesis of thiourea 9a from amine 8. To a stirred solution of **8** (250 mg, 0.65 mmol) in DCM was added triethylamine (0.9 mL, 0.65 mmol). The solution was diluted with 20 mL of saturated $NaHCO_3$, and the resulting mixture extracted with DCM (3 x 10 mL). The combined organic layers were dried with Na_2SO_4 and concentrated to give 220 mg of 1,3,4,6-*tetra*-*O*-acetyl-2-amino-2-deoxy-β-D-glucopyranose which was used without purification. 1,3,4,6-*tetra*-*O*-acetyl-2-amino-2-deoxy-β-D-glucopyranose (220 mg) was then dissolved in 5 mL of pyridine and 9-fluorenylmethoxycarbonyl isothiocyanate (Fmoc) (180 mg, 0.65 mmol) and triethyl amine (0.02 mL) were added. The resulting mixture was stirred at room temperature for 16 hours. The solution was concentrated and the residue diluted with DCM (20 mL) and $NaHCO_3$ (20 mL), and subsequently extracted with DCM (3 x 10 mL). The combined organic extracts were dried with Na_2SO_4 and concentrated. Flash chromatography was performed with 75% EtOAc:hexanes to yield **9a** as a white foam (360 mg, 89% yield over two steps).

1,3,4,6-*tetra*-*O*-acetyl-2-deoxy-2-[[[(2-fluoren-9-yl methoxycarbonyl)amino]thioxomethyl]amino]-β-D-glucopyranose (9a). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.79 (d, *J* = 7.7 Hz, 2H), 7.56 (dq, *J* = 7.6, 0.8 Hz, 2H), 7.44 (td, *J* = 7.5, 1.0 Hz, 2H), 7.34 (dd, *J* = 7.5, 1.1 Hz, 2H), 5.84 (d, *J* = 8.2 Hz, 1H), 5.24 (dt, *J* = 38.6, 9.1 Hz, 2H), 5.07 (q, *J* = 9.3 Hz, 1H), 4.54 (d, *J* = 6.6 Hz, 2H), 4.31 (ddd, *J* = 12.2, 7.5, 4.6 Hz, 1H), 4.24 (t, *J* = 6.6 Hz, 1H), 4.16 (dd, *J* = 12.4, 2.6 Hz, 1H), 3.86 (ddd, *J* = 9.4, 4.7, 2.7 Hz, 1H), 2.13 (d, *J* = 0.5 Hz, 3H), 2.11 (d, *J* = 0.5 Hz, 3H), 2.06 (d, *J* = 0.5 Hz, 3H), 2.05 (d, *J* = 0.5 Hz, 3H). ^{13}C NMR (150 MHz, $CDCl_3$) δ 180.22, 170.72, 170.46, 169.34, 169.24, 152.16, 142.80, 141.37, 129.06, 128.25, 128.15, 127.83, 127.30, 125.32, 124.96, 124.84, 120.29, 120.09, 92.21, 72.87, 72.21, 68.40, 67.47, 61.65, 57.59, 46.49, 21.05, 20.77, 20.72, 20.62.

Synthesis of thiourea 9b from Amine 8. 1g (2.6 mmol) of **8** was dissolved in CH_3CN. Triethylamine (0.7 mL, 5.2 mmol) was added, followed by allyl isothiocyanate (0.5 mL, 5.2 mmol). The reaction mixture was stirred at room temperature for ~2 hours, followed by work-up with a minimal amount of saturated $NaHCO_3$ after completion. The aqueous layer formed was extracted twice with DCM and the organic layers were then combined, dried over $MgSO_4$, filtered and concentrated. The crude product was purified by silica flash chromatography (1:1 EtOAc/Hexanes), affording a solid white product (70%, 785 mg).

1,3,4,6-*tetra*-*O*-acetyl-2-deoxy-2-[[[(2-prop-1-ene)amino]thioxomethyl]amino]-β-D-glucopyranose (9b). ^{1}H NMR (500 MHz, $CDCl_3$) δ 6.34 (s, 1H), 6.21 (s, 1H), 5.91 – 5.82 (m, 1H), 5.75 (dd, J = 8.5, 1.3 Hz, 1H), 5.32 – 5.14 (m, 5H), 4.29 (ddd, J = 12.5, 4.7, 1.9 Hz, 1H), 4.19 – 4.07 (m, 2H), 3.87 (ddd, J = 9.7, 4.7, 2.3 Hz, 1H), 2.14 (s, 3H), 2.11 (s, 3H), 2.09 (s, 3H), 2.06 (s, 3H), 1.98 – 1.87 (m, 1H); ^{13}C-NMR (100 MHz, $CDCl_3$) δ (ppm) 183.51, 171.5, 171.2, 170.68, 169.27, 136.36, 117.43, 92.95, 72.93, 72.90, 67.64, 61.62, 57.77, 53.74, 21.04, 20.82, 20.74, 20.59; HRMS (m/z): [M + H]+ calculated for $C_{18}H_{27}N_2O_9S$: 447.1437 (M+H); found 447.1430; [M + Na]+ calculated for $C_{18}H_{26}N_2NaO_9S$: 469.1251; found 469.1251; [M+K]+ calculated for $C_{18}H_{26}KN_2O_9S$: 485.0996; found 485.1005.

General synthesis of thioureas 13a-b, 13f-h. Isothiocyanate **12** (1 eq) was dissolved in CH_3CN followed by addition of the desired amine hydrochloride salt (1.2-2.0 eq), then drop-wise addition of triethylamine (2.0 eq). The reaction mixture was stirred for ~4 hours by monitoring by TLC. Upon completion, it was washed with a minimal amount of saturated $NaHCO_3$. The aqueous layer was extracted twice with DCM and the organic layers were combined, dried over $MgSO_4$, filtered and concentrated. The crude product was purified by silica flash chromatography (1:1 EtOAc/hexanes or 75% EtOAc in hexanes) to afford the pure product. Products were isolated in yields ranging from 51 to 80%.

General synthesis of thioureas 13d-e. Isothiocyanate **12** (1 eq) was dissolved in DCM, followed by drop-wise addition of the desired amine (1.05 eq). The reaction mixture was stirred for ~4 hours by monitoring by TLC. Upon completion, it was washed with a minimal amount of saturated $NaHCO_3$. The aqueous layer was extracted twice with DCM and the organic layers were combined, dried over $MgSO_4$, filtered and concentrated. The crude product was purified by silica flash chromatography (1:1 EtOAc/hexanes or 75% EtOAc in hexanes) to afford the pure product. Products were isolated in yields ranging from 60 to 87%.

1,3,4,6-*tetra*-*O*-acetyl-2-deoxy-2-[(aminomethyl)thioxomethyl]amino]-β-D-glucopyranose (13a): ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.56 (s, 1H), 6.41 (s, 1H), 5.78 (d, J = 8.4 Hz, 1H), 5.27 (t, J = 9.6 Hz, 1H), 5.13 (t, J = 9.6 Hz, 1H), 4.24 (dd, J = 12.5, 4.6 Hz, 1H), 4.10 (dt, J = 14.3, 4.6 Hz, 2H), 3.85 (ddd, J = 9.9, 4.5, 2.3 Hz, 1H), 2.93 (s, 3H), 2.09 (s, 3H), 2.06 (s, 3H), 2.03 (s, 3H), 2.02 (s, 3H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 171.52, 170.75, 169.55, 169.36, 92.85, 73.14, 72.73, 67.90, 61.76, 57.63, 20.98, 20.78, 20.69, 20.55, 14.36. HRMS (m/z): [M+H]+ calculated for $C_{16}H_{25}N_2O_9S$: 421.1281 found 421.1227; [M+Na]+ calculated for $C_{16}H_{24}N_2NaO_9S$: 443.1134; found 443.1098; [M+K]+ calculated for $C_{16}H_{24}KN_2O_9S$: 459.0840; found 459.0840.

1,3,4,6-*tetra*-*O*-acetyl-2-deoxy-2[(dimethylamino)thioxomethyl]amino]-β-D-glucopyranose (13b): ^{1}H NMR (500 MHz, $CDCl_3$) δ 5.74 (t, J = 9.7 Hz, 1H), 5.70 (d, J = 9.2 Hz, 1H), 5.36 – 5.23 (m, 1H), 5.23 – 5.13 (m, 2H), 4.21 (dt, J = 15.7, 7.9 Hz, 1H), 4.12 – 4.09 (m, 1H), 3.80 (ddd, J = 9.1, 4.5, 2.2 Hz, 1H), 3.17 (s, 6H), 2.07 (s, 3H), 2.06

(s, 3H), 2.01 (d, J = 3.4 Hz, 3H), 2.00 (s, 3H). ^{13}C-NMR (125 MHz, $CDCl_3$) δ (ppm) 182.29, 171.83, 170.68, 169.94, 169.06, 93.28, 73.24, 72.99, 67.52, 61.71, 58.46, 40.63, 21.06, 20.78, 20.69, 20.53; HRMS (m/z): [M+H]+ calculated for $C_{17}H_{26}N_2O_9S$: 489.1149; found 489.1152; [M + Na]+ calculated for $C_{17}H_{26}N_2O_9S$: 511.0969; found 511.0970.

1,3,4,6-*tetra*-*O*-acetyl-2-deoxy-2-[[[(2-propyl)amino]thioxomethyl]amino]-β-D-glucopyranose (13d): ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.22 (s, 1H), 5.98 (d, J = 26.1 Hz, 1H), 5.76 (d, J = 8.5 Hz, 1H), 5.23 – 5.20 (m, 2H), 4.30 (dd, J = 12.5, 4.6 Hz, 1H), 4.27 – 4.21 (m, 1H), 4.20 – 4.10 (m, 1H), 3.87 (ddt, J = 6.8, 4.5, 2.3 Hz, 1H), 3.46 – 3.28 (m, 2H), 2.16 (s, 3H), 2.12 (s, 3H), 2.09 (s, 3H), 2.07 (s, 3H), 1.62 (qd, J = 7.3, 2.0 Hz, 2H), 0.98 (t, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 171.09, 170.95, 170.73, 169.27, 99.98, 93.09, 72.99, 67.60, 62.10, 61.68, 21.04, 20.93, 20.81, 20.75, 20.65, 20.59, 11.37, 11.30. HRMS (m/z): [M + H]+ calculated for $C_{18}H_{29}N_2O_9S$: 449.1594; found 449.1593.

1,3,4,6-*tetra*-*O*-acetyl-2-deoxy-2-[[[(2-butyl)amino]thioxomethyl]amino]-β-D-glucopyranose (13e): ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.38 (s, 1H), 5.76 (d, *J* = 8.5 Hz, 1H), 5.36 – 5.07 (m, 2H), 4.30 – 4.16 (m, 2H), 4.15 – 4.09 (m, 1H), 3.85 (ddd, *J* = 9.7, 4.6, 2.3 Hz, 1H), 3.58 – 3.13 (bs, 1H), 2.11 (s, 3H), 2.08 (s, 3H), 2.03 (s, 3H), 2.02 (s, 3H), 1.56 – 1.48 (m, 2H), 1.38 – 1.30 (m, 2H), 0.91 (t, *J* = 7.3 Hz, 3H). ^{13}C-NMR (100 MHz, $CDCl_3$) δ (ppm) 171.74 , 171.60 , 171.00 , 169.39, 169.27, 93.07, 72.97, 68.10, 67.62, 62.10, 61.69, 21.03, 20.94, 20.80, 20.75, 20.65, 20.59, 20.00, 19.97, 13.80; HRMS (m/z): [M + H] + calculated for $C_{19}H_{31}N_2O_9S$: 463.1750; found 463.1737.

1,3,4,6-*tetra*-*O*-acetyl-2-deoxy-2-[[[(2-fluoroethyl)amino]thioxomethyl]amino-β-D-glucopyranose (13f): ^{1}H-NMR (500 MHz, $CDCl_3$) δ (ppm) 6.43 (s, 1H), 5.99 (d, J = 7.6 Hz, 1H), 5.64 (d, J = 8.6 Hz, 1H), 5.20 – 5.11 (m, 2H), 4.70 – 4.50 (m, 2H), 4.31 (dd, J = 12.5, 4.5 Hz, 1H), 4.14 – 4.13 (m, 1H), 3.82 (ddd, J = 9.4, 4.5, 2.3 Hz, 2H), 2.14 (s, 3H), 2.10 (s, 3H), 2.09 (s, 3H), 2.04 (s, 3H); ^{13}C-NMR (125 MHz, $CDCl_3$) δ (ppm): 184.30, 171.94, 171.62, 169.76, 93.10, 82.77 (d, J = 166.3 Hz), 73.30, 72.91, 68.18, 61.93, 57.79, 45.44, 21.17, 21.02, 20.99, 20.84; HRMS (m/z): [M+H]+ calculated for $C_{17}H_{26}FN_2O_9S$: 453.1343; found 453.1342.

1,3,4,6-*tetra*-*O*-acetyl-2-deoxy-2-[[[(2,2-difluoroethyl)amino]thioxomethyl]amino-β-D-glucopyranose (13g): ^{1}H-NMR (400

MHz, $CDCl_3$) δ 6.55 (s, 2H), 5.99 (t, J = 56.1 Hz, 1H), 5.74 (d, J = 8.6 Hz, 1H), 5.32 – 5.21 (m, 1H), 5.11 (t, J = 9.6 Hz, 1H), 4.27 (dd, J = 12.5, 4.4 Hz, 1H), 4.14 (d, J = 6.5 Hz, 1H), 4.09 (td, J = 7.2, 1.2 Hz, 2H), 3.88 (ddd, J = 9.8, 4.2, 2.0 Hz, 3H), 2.11 (s, 3H), 2.08 (s, 3H), 2.06 (s, 3H), 2.04 (s, 3H); ^{13}C-NMR (101 MHz, $CDCl_3$) δ (ppm) 184.72, 171.57, 170.91, 169.61, 113.54 (t, J =241.5 Hz, 1C), 92.65, 72.95, 72.57, 68.05, 61.78, 57.51, 46.42 (t, J=27.2 Hz) 20.77, 20.62, 20.60, 20.48; HRMS (m/z): [M+H]+ calculated for $C_{17}H_{25}F_2N_2O_9S$: 471.1249; found 471.1242; [M+Na]+ calculated for $C_{17}H_{25}F_2NaN_2O_9S$: 493.1068; found 493.1065.

1,3,4,6-*tetra*-*O*-acetyl-2-deoxy-2-[[[(2,2,2 trifluoroethyl)amino]thioxomethyl]amino-β-D-glucopyranose (13h): ^{1}H-NMR (400 MHz, $CDCl_3$) δ (ppm) 6.62 – 6.49 (m, 2H), 5.79 (d, J = 8.5 Hz, 1H), 5.30 (t, J = 9.8 Hz, 1H), 5.16 (t, J = 9.6 Hz, 1H), 4.55 – 4.40 (m, 1H), 4.31 (dd, J = 12.4, 4.7 Hz, 2H), 4.22 – 4.11 (m, 2H), 3.94 (ddd, J = 9.8, 4.5, 2.2 Hz, 1H), 2.14 (s, 3H), 2.13 (s, 3H), 2.10 (s, 3H), 2.08 (s, 3H); ^{13}C-NMR (125 MHz, $CDCl_3$) δ (ppm) 184.83, 171.76, 171.63, 170.88, 169.62, 124.01 (q, J = 279.2 Hz), 92.44, 72.77, 72.42, 68.04, 61.66, 57.52, 45.23, 20.59, 20.52, 20.46, 20.36; HRMS (m/z): [M+H]+ calculated for $C_{17}H_{23}F_3N_2O_9S$: 489.1155; found 489.1153; [M+Na]+ calculated for $C_{17}H_{23}F_3NaN_2O_9S$: 511.0974; found 511.0970.

Synthesis of per-*O*-acetylated thiazoline 10a. Thiourea **9a** (200 mg, 0.32 mmol) was dissolved in 4 mL of DCM and $SnCl_4$ (0.5 mL, 4.0 mmol) was then added. The resulting mixture was stirred at room temperature for 16 hours. The solution was diluted with 20 mL of saturated $NaHCO_3$ and the mixture was then extracted with DCM (3 x 10 mL). The combined organic extracts were dried with Na_2SO_4 and concentrated. Silica flash chromatography was used to purify the crude mixture with 75% EtOAc/hexanes to yield **10a** as a white foam (125 mg, 65% yield).

3,4,6,-*tri*-*O*-acetyl-1,2-dideoxy-2'[(-[[[(2-fluoren-9-yl methoxycarbonyl)amino]thioxomethyl]amino]-α-D-glucopyranoso[2,1-d]-Δ2'-thiazoline (10a): ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.80 (dd, *J* = 7.6, 2.7 Hz, 2H), 7.59 (dd, *J* = 7.6, 4.3 Hz, 2H), 7.45 – 7.40 (m, 2H), 7.34 (dtd, *J* = 15.0, 7.5, 1.1 Hz, 2H), 6.01 (d, *J* = 6.9 Hz, 1H), 5.24 (t, *J* = 4.7 Hz, 1H), 4.95 (dd, *J* = 9.6, 4.2 Hz, 1H), 4.74 (s, 1H), 4.55 (dd, *J* = 10.8, 6.0 Hz, 1H), 4.29 – 4.21 (m, 2H), 4.14 – 4.12 (m, 1H), 3.81 (s, 1H), 3.70 (s,

1H), 2.12 (s, 3H), 2.04 (s, 3H), 1.89 (s, 3H). ^{13}C NMR (125 MHz, $CDCl_3$) δ 170.23, 169.63, 169.11, 161.43, 152.11, 142.86, 141.31, 130.08, 129.72, 129.61, 127.97, 127.76, 125.15, 124.04, 123.83, 120.89, 120.71, 90.23, 72.52, 71.88, 69.46, 67.92, 63.70, 57.91, 46.78, 21.35, 20.71, 20.62.

General synthesis of per-acetylated thiazolines 10b and 14a-h. Thiourea derivatives **9b, 13a-h** (1 eq) were each dissolved in DCM under argon, after which trifluoroacetic acid (7.5 eq) was added drop-wise. The reaction was stirred under an atmosphere of argon overnight (~10 hours) until judged complete by TLC analysis. The mixture was washed twice with saturated $NaHCO_3$, followed by extraction of the aqueous layers 3 times with DCM. The organic layers were combined, dried over $MgSO_4$, filtered and concentrated to obtain the crude product which was then purified using silica flash chromatography (generally 1:1 EtOAc/hexanes or 75% EtOAc/hexanes). Products were isolated in yields ranging from 55 to 84%.

3,4,6,-*tri*-*O*-acetyl-1,2-dideoxy-2'[(2-prop-1-ene)amino]-α-D-glucopyranoso[2,1-d]-Δ2'-thiazoline (10b): ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.21 (d, J = 6.5 Hz, 1H), 5.90 (ddt, J = 16.9, 10.6, 5.5 Hz, 1H), 5.38 (dd, J = 3.9, 2.8 Hz, 1H), 5.17 (ddd, J = 13.7, 11.5, 1.3 Hz, 2H), 4.94 – 4.89 (m, 1H), 4.36 – 4.31 (m, 1H), 4.13 – 4.07 (m, 3H), 3.97 – 3.89 (m, 1H), 3.83 (ddd, J = 11.4, 9.2, 5.0 Hz, 2H), 2.09 (s, 3H), 2.06 (s, 3H), 2.05 (s, 3H).; ^{13}C-NMR (125 MHz, $CDCl_3$) 170.36, 169.38, 169.18, 159.98, 133.93, 116.06, 89.30, 72.25, 71.40, 68.88, 68.12, 62.97, 46.51, 20.71, 20.58, 20.49. HRMS (m/z): [M + H]+ calculated for $C_{16}H_{23}N_2O_7S$: 387.1226 found 387.1221; [M+Na]+ calculated for $C_{16}H_{22}NaN_2O_7S$: 409.1045; found 409.1038.

3,4,6,-*tri*-*O*-acetyl-1,2-dideoxy-2'-aminomethyl-α-D-glucopyranoso[2,1-d]-Δ2'-thiazoline (14a): ^{1}H-NMR (400 MHz, $CDCl_3$) δ (ppm) 6.25 (d, J = 6.5 Hz, 1H), 5.43 (dd, J = 4.1, 3.0 Hz, 1H), 4.95 (ddd, J = 9.6, 2.8, 0.9 Hz, 1H), 4.37 (dd, J = 6.0, 4.5 Hz, 1H), 4.17-4.14 (m, 2H), 3.90-3.84 (m, 1H), 2.94 (s, 3H), 2.12 (s, 3H), 2.09 (s, 3H), 2.08 (s, 3H); ^{13}C-NMR (101 MHz, $CDCl_3$) δ (ppm) 170.65, 169.67, 169.49, 89.77, 72.40, 71.95, 69.09, 68.61, 63.15, 31.11, 20.99, 20.88, 20.77; HRMS (m/z): [M + H]+ calculated for $C_{14}H_{21}N_2O_7S$: 361.1069; found 361.1064; [M+Na]+ calculated for $C_{14}H_{20}NaN_2O_7S$:

383.0889; found 383.0882; [M+K]+ calculated for $C_{14}H_{20}KN_2O_7S$: 399.0628; found 399.0620.

3,4,6,-*tri*-*O*-acetyl-1,2-dideoxy-2'-dimethylamino-α-D-glucopyranoso[2,1-d]-Δ2'- thiazoline (14b): ^{1}H-NMR (400 MHz, $CDCl_3$) δ (ppm) 6.24 (d, J = 6.6 Hz, 1H), 5.41 (dd, J = 4.3, 3.4 Hz, 1H), 4.95 (ddd, J = 9.6, 3.3, 0.9 Hz, 1H), 4.35 (dd, J = 6.1, 4.9 Hz, 1H), 4.15, (dd, J = 8.5, 5.8 Hz, 2H), 3.89 (ddd, J = 9.1, 5.1, 3.1 Hz, 1H), 3.01 (s, 6H), 2.11 (s, 3H), 2.09 (s, 3H), 2.07 (s, 3H); ^{13}C-NMR (101 MHz, $CDCl_3$) δ (ppm) 170.63, 169.62, 169.51, 162.50, 90.17, 72.84, 72.30, 69.10, 68.53, 63.11, 39.90, 21.00, 20.86, 20.76; HRMS (m/z): [M + H]+ calculated for $C_{15}H_{23}N_2O_7S$: 375.1226; found 375.1221; [M + Na]+ calculated for $C_{15}H_{22}NaN_2O_7S$: 397.1045; found 397.1037.

3,4,6,-*tri*-*O*-acetyl-1,2-dideoxy-2'[(2-propyl)amino]-α-D-glucopyranoso[2,1-d]-Δ2'-thiazoline (14d): ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.28 (d, J = 6.5 Hz, 1H), 5.43 (dd, J = 4.3, 3.0 Hz, 1H), 4.98 (ddd, J = 9.4, 3.0, 1.0 Hz, 1H), 4.40 (ddd, J = 6.5, 4.3, 1.1 Hz, 1H), 4.21 – 4.15 (m, 2H), 3.90 (dt, J = 8.9, 4.2 Hz, 1H), 3.27 (ddt, J = 27.4, 13.1, 6.7 Hz, 2H), 2.15 (s, 3H), 2.12 (s, 3H), 2.11 (s, 3H), 1.71 – 1.61 (m, 2H), 0.99 (t, J = 7.4 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 170.53, 169.68, 169.44, 162.21, 87.67, 70.81, 69.56, 68.11, 66.18, 62.84, 29.70, 22.68, 20.90, 20.85, 20.73, 11.24. HRMS (m/z): [M + H]+ calculated for $C_{16}H_{25}N_2O_7S$: 389.1382 (M+H); found 389.1383; [M + Na]+ calculated for $C_{16}H_{24}NaN_2O_7S$: 411.1202; found 411.1199.

3,4,6,-*tri*-*O*-acetyl-1,2-dideoxy-2'[(2-butyl)amino]-α-D-glucopyranoso[2,1-d]-Δ2'-thiazoline (14e): ^{1}H NMR (500 MHz, $CDCl_3$) δ 6.29 (d, *J* = 6.5 Hz, 1H), 5.42 (dd, *J* = 4.4, 3.2 Hz, 1H), 5.01 – 4.91 (m, 1H), 4.46 – 4.33 (m, 2H), 4.19 (dd, *J* = 4.2, 2.8 Hz, 2H), 3.91 (ddd, *J* = 8.9, 5.2, 3.3 Hz, 1H), 3.41 – 3.25 (m, 2H), 2.15 (s, 3H), 2.12 (s, 3H), 2.11 (s, 3H), 1.62 (p, *J* = 7.2 Hz, 2H), 1.43 – 1.36 (m, 2H), 0.96 (t, *J* = 7.4 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 170.50, 169.85, 169.60, 162.30, 92.67, 71.75, 69.26, 68.33, 66.21, 62.08, 31.60, 20.93, 20.82, 20.66, 19.95, 16.88, 13.67. δ (ppm); HRMS (m/z): [M + H]+ calculated for $C_{17}H_{27}N_2O_7S$: 403.1539; found 403.1535.

3,4,6-*tri*-*O*-acetyl-1,2-dideoxy-2'-[(2-fluoroethyl)amino]-α-D-glucopyranoso[2,1-d]-Δ2'-thiazoline (14f): ^{1}H-NMR (400 MHz, $CDCl_3$) δ (ppm) 6.25 (d, J = 6.5 Hz, 1H), 5.40 (dd, *J* = 3.9, 2.6 Hz, 1H), 4.97 – 4.93 (m, 1H), 4.71 – 4.44 (m, 2H),

4.37 – 4.33 (m, 1H), 4.16 – 4.11 (m, 2H), 3.85 – 3.78 (m, 1H), 3.73 – 3.46 (m, 2H), 2.11 (s, 3H), 2.08 (s, 3H), 2.06 (s, 3H); ^{13}C-NMR (101 MHz, $CDCl_3$) 170.75, 169.68, 169.59, 159.80, 90.02, 82.04 (d, *J* = 166.8 Hz), 72.73, 71.78, 69.17, 68.62, 63.35, 44.59 (d, *J* = 20.2 Hz), 21.08, 20.90, 20.85; LRMS (m/z): [M + H]+ calculated for $C_{15}H_{22}FN_2O_7S$: 393.1132; found 393.1122.

3,4,6-*tri*-*O*-acetyl-1,2-dideoxy-2'-[(2,2-difluoroethyl)amino]-α-D-glucopyranoso[2,1-d]-Δ2'-thiazoline (14g): ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.25 (d, *J* = 6.5 Hz, 1H), 6.02 (tdd, *J* = 56.6, 5.0, 3.5 Hz, 1H), 5.37 (dd, *J* = 3.9, 2.5 Hz, 1H), 4.93 (ddd, *J* = 9.4, 2.4, 1.1 Hz, 1H), 4.34 (ddd, *J* = 6.5, 4.0, 1.1 Hz, 1H), 4.14 – 4.11 (m, 2H), 3.81 – 3.66 (m, 2H), 3.59 – 3.45 (m, 1H), 2.10 (s, 3H), 2.06 (s, 3H), 2.05 (s, 3H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ (ppm) 170.66, 169.56, 169.52, 159.57, 113.37 (t, *J* = 241.4 Hz), 90.31, 72.37, 71.44, 68.99, 68.55, 63.29, 46.10 (t, *J* = 27.2 Hz), 20.96, 20.73, 20.72; HRMS (m/z): [M+H]+ calculated for $C_{15}H_{21}F_2N_2O_7S$: 411.1038; found 411.1033; [M+Na]+ calculated for $C_{15}H_{20}F_2NaN_2O_7S$: 433.0857; found 433.0850.

3,4,6-*tri*-*O*-acetyl-1,2-dideoxy-2'-[(2,2,2-trifluoroethyl)amino]-α-D-glucopyranoso[2,1-d]-Δ2'-thiazoline (14h): ^{1}H NMR (500 MHz, $CDCl_3$) δ 6.22 (d, *J* = 6.6 Hz, 1H), 5.34 (dd, *J* = 3.7, 2.3 Hz, 1H), 4.91 – 4.84 (m, 1H), 4.33 – 4.28 (m, 1H), 4.09 – 3.99 (m, 3H), 3.80 (td, *J* = 9.0, 6.3 Hz, 1H), 3.76 – 3.72 (m, 1H), 2.07 (s, 3H), 2.03 (s, 3H), 2.01 (s, 3H). ^{13}C-NMR (126 MHz, $CDCl_3$) δ (ppm) 170.59, 169.70, 169.40, 159.58, 125.19 (q, *J* = 279.2 Hz), 89.61, 71.59, 71.19, 68.82, 68.19, 63.10, 44.98, 20.74, 20.54, 20.50; HRMS (m/z): [M+H]+ calculated for $C_{17}H_{20}F_3N_2O_9S$: 429.0943; found 429.0941; [M+Na]+ calculated for $C_{17}H_{19}F_3NaN_2O_9S$: 451.0763; found 451.0756.

Deprotection of amino thiazoline 10a. Per-acetylated intermediate **10a** (114 mg, 0.20 mmol) was dissolved in 2 mL of MeOH, followed by the addition of NaOMe (14 mg, 0.25 mmol). The reaction was then stirred at room temperature for 2 hours, and subsequently quenched by the addition of AcOH. Concentration gave a colorless oil which was dissolved in 3 mL of pyridine and then 0.6 mL of piperidine was added. The resulting mixture was then stirred at room temperature for 2 hours, concentrated, and any remaining piperidine was then co-evaporated with pyridine. The resulting mixture was triturated with EtOAc to yield inhibitor **11a** (38 mg, 80% yield).

1,2-dideoxy-2'-amino-α-D-glucopyranoso[2,1-d]-Δ2'-thiazoline (11a): ^{1}H NMR (500 MHz, CD_3OD) δ 6.31 (d, *J* = 6.4 Hz, 1H), 4.04 (t, *J* = 6.1 Hz, 1H), 3.91 (t, *J* = 5.3 Hz, 1H), 3.78 (dd, *J* = 11.7, 2.1 Hz, 1H), 3.67 – 3.57 (m, 2H), 3.47 (dd, *J* = 9.0, 5.0 Hz, 1H). ^{13}C NMR (101 MHz, CD_3OD) δ 173.50, 89.62, 77.75, 74.92, 69.74, 65.15. 63.26 Anal. Calcd. for $C_7H_{12}N_2O_4S$: C, 38.17; H, 5.49; N, 12.72; Found: C, 38.05; H, 5.37; N, 12.66.

General procedure for the de-protection of amino thiazolines 10b and 14a-h. A catalytic amount of K_2CO_3 (~5% w/v) was added to a solution of each per-acetylated 2'-aminothiazoline (**10b, 14a-h**) in anhydrous methanol. The reaction was stirred at room temperature under an atmosphere of argon for ~1-2 hours until judged complete by TLC analysis. The solution was filtered, and the solvent evaporated to yield the crude product which was then purified by silica flash chromatography (10-18% MeOH/DCM) to yield the final product. Products were isolated in yields ranging from 71-86%.

1,2-dideoxy-2'-[(2-prop-1-ene)amino]-α-D-glucopyranoso[2,1-d]-Δ2'-thiazoline (11b): ^{1}H NMR (500 MHz, CD_3OD) δ 6.33 (d, *J* = 6.3 Hz, 1H), 5.92 (ddt, *J* = 15.7, 10.6, 5.4 Hz, 1H), 5.27 – 5.20 (m, 1H), 5.18 – 5.10 (m, 1H), 4.08 (t, J = 6.1 Hz, 1H), 3.96 – 3.92 (m, 1H), 3.91 – 3.86 (m, 2H), 3.80 (dd, *J* = 11.7, 2.0 Hz, 1H), 3.69 – 3.60 (m, 2H), 3.49 (dd, *J* = 8.9, 5.3 Hz, 1H); ^{13}C-NMR (125 MHz, CD_3OD) δ (ppm) 162.15, 134.33, 114.87, 89.45, 75.07, 74.10, 73.83, 69.77, 61.87, 45.85; HRMS (m/z): [M+H]+ calculated for $C_{10}H_{17}N_2O_4S$: 261.0909; found 261.0908; [M+Na]+ calculated for $C_{10}H_{16}NaN_2O_4S$:283.0728; found 283.0727.

1,2-dideoxy-2'-aminomethyl-α-D-glucopyranoso [2,1-d]-Δ2'-thiazoline (15a): ^{1}H-NMR (400 MHz, D_2O) δ (ppm) 6.21 (d, *J* = 7.7 Hz, 1H), 4.11 (t, *J* = 5.9 Hz, 1H), 3.96 (t, *J* = 5.0 Hz, 1H), 3.73 (t, *J* = 7.7 Hz, 1H), 3.58 (dt, *J* = 10.0, 7.1 Hz, 2H), 3.49 (dd, *J* = 9.0, 4.5 Hz, 1H), 2.74 (s, 3H); ^{13}C-NMR (101 MHz, D_2O) δ (ppm) 163.75, 88.54, 74.25, 73.64, 73.28, 69.31, 61.42, 29.87 HRMS (m/z): [M+H]+ calculated for $C_8H_{15}N_2O_4S$: 235.0753; found 235.0750; [M+Na]+ calculated for $C_8H_{14}NaN_2O_4S$: 257.0572; found 257.0570.

1,2-dideoxy-2'-dimethylamino-α-D-glucopyranoso [2,1-d]-Δ2'-thiazoline (15b): ^{1}H-NMR (400 MHz, CD_3OD) δ (ppm) 6.39 (d, *J* = 6.5 Hz, 1H), 4.09 (t, *J* = 6.3 Hz, 1H), 3.90 (t, *J* = 5.9 Hz, 1H), 3.84-3.78 (m, 1H), 3.72-3.61 (m, 2H), 3.52-3.45 (m, 1H), 3.06 (s, 6H); ^{13}C-NMR (125 MHz, CD_3OD) δ (ppm) 164.30, 90.59, 75.02, 74.84. 74.45, 69.72, 61.81, 38.67; HRMS (m/z): [M + H]+ calculated for $C_9H_{17}N_2O_4S$: 249.0909; found 249.0908; [M+Na]+ calculated for $C_9H_{16}NaN_2O_4S$: 271.0728; found 271.0725.

1,2-dideoxy-2'-[(2-ethyl)amino]-α-D-glucopyranoso [2,1-d]-Δ2'-thiazoline (15c) ThiamEtG: Used from previous studies by Yuzwa *et al*,[54] for which the characterization data matched.

1,2-dideoxy-2'-[(2-propyl)amino]-α-D-glucopyranoso [2,1-d]-Δ2'-thiazoline (15d): ^{1}H-NMR (500 MHz, CD_3OD) δ (ppm) 6.58 (d, *J* = 6.5 Hz, 1H), 4.13 (t, *J* = 6.6 Hz, 1H), 3.91-3.79 (m, 2H), 3.76-3.63 (m, 2H), 3.50 (dd, *J* = 8.8, 6.9 Hz, 1H), 1.70-1.62 (m, 2H), 0.98 (t, *J* = 7.4 Hz, 3H); ^{13}C-NMR (125 MHz, $(CD_3)_2CO$) δ (ppm) 182.84, 93.02, 79.42, 75.18, 68.88, 64.80, 63.16, 45.96, 21.08, 10.75; HRMS (m/z): [M+H]+ calculated for $C_{10}H_{19}N_2O_4S$: 263.1066; found 263.1063; [M+Na]+ calculated for $C_{10}H_{18}NaN_2O_4S$: 285.0885; found 285.0876.

1,2-dideoxy-2'-[(2-butyl)amino]-α-D-glucopyranoso [2,1-d]-Δ2'-thiazoline (15e): ^{1}H NMR (500 MHz, CD_3OD) δ 6.34 (d, *J* = 6.4 Hz, 1H), 4.07 (t, *J* =6.2 Hz, 1H), 3.92 (t, *J* = 5.8 Hz, 1H), 3.85 – 3.77 (m, 1H), 3.71 – 3.62 (m, 2H), 3.52 – 3.46 (m, 1H), 3.31 – 3.22 (m, 2H), 1.57 (p, *J* = 7.2 Hz, 2H), 1.41 (dt, *J* = 15.0, 7.4 Hz, 2H), 0.97 (t, *J* = 7.4 Hz, 3H); ^{13}C-NMR (125 MHz, CD_3OD, δ (ppm) 163.07, 89.08, 75.17, 74.17, 72.75, 69.55, 61.73, 43.59, 31.00, 19.69, 12.72.; HRMS (m/z): [M+H]+ calculated for $C_{11}H_{20}N_2O_4S$: 277.1222 (M+H); found 277.1167; [M+Na]+ calculated for $C_{11}H_{20}NaN_2O_4S$: 299.1041; found 299.1040.

1,2-dideoxy-2'-[(2-fluoroethyl)amino]-α-D-glucopyranoso[2,1-d]-Δ2'-thiazoline (15f): ^{1}H-NMR (400 MHz, CD_3OD) δ (ppm) 6.34 (d, *J* = 6.4 Hz, 1H), 4.60-4.49 (m, 1H), 4.48-4.37 (m, 1H), 4.05 (t, *J* = 6.2 Hz, 1H), 3.88 (t, *J* = 5.8 Hz, 1H), 3.76 (dd, *J* = 11.7, 2.1 Hz, 1H), 3.66-3.43 (m, 5H); ^{13}C-NMR (100 MHz, CD_3OD) δ (ppm) 163.45, 89.20, 81.57 (d, *J* = 167.4 Hz), 75.23, 73.93, 72.20, 69.40, 61.63, 44.16 (d, *J* = 20.3 Hz); HRMS (m/z): [M+H]+ calculated for $C_9H_{16}N_2O_4SF$: 267.0815; found 267.0822.

1,2-dideoxy-2'-[(2,2-difluoroethyl)amino]-α-D-glucopyranoso[2,1-d]-Δ2'-thiazoline (15g): ^{1}H NMR (500 MHz, CD_3OD) δ 6.29 (d, *J* = 6.4 Hz, 1H), 5.96 (tt, *J* = 56.3, 4.3 Hz, 1H), 4.04 (t, *J* = 6.1 Hz, 1H), 3.89 (t, *J* = 5.4 Hz, 1H), 3.74 (dd, *J* = 11.9, 2.3 Hz, 1H), 3.64 – 3.41 (m, 5H). ^{13}C-NMR (125 MHz, D_2O) δ (ppm) 162.48, 113.81 (t, *J* = 239.6 Hz), 88.11, 73.79, 72.58, 68.78, 61.00, 48.39, 44.95 (t, *J* = 25.0 Hz); HRMS (m/z): [M+H]+ calculated for $C_9H_{15}N_2O_4SF_2$: 285.0715; found 285.0717; [M+Na]+ calculated for $C_9H_{14}NaN_2O_4SF_2$: 307.0540; found 307.0536.

1,2-dideoxy-2'-[(2,2,2-trifluoroethyl)amino]-α-D-glucopyranoso[2,1-d]-Δ2'-thiazoline (15h): ^{1}H-NMR (500 MHz, CD_3OD) δ (ppm) 6.34 (d, *J* = 6.4 Hz, 1H), 4.04-3.99 (m, 1H), 3.98-3.85 (m, 3H), 3.78 (dd, *J* = 11.9, 2.2 Hz, 1H), 3.65 (dd, *J* = 11.9, 6.1 Hz, 1H), 3.59 (ddd, *J* = 8.5, 6.2, 2.1 Hz, 1H), 3.47 (dd, *J* = 9.1, 3.5 Hz, 1H); ^{13}C NMR (151 MHz, D_2O) δ 162.61, 123.97 (q, *J* = 278.5 Hz), 88.09, 73.84, 72.63, 68.68, 60.91, 44.10; HRMS (m/z): [M+H]+ calculated for $C_9H_{14}N_2O_4SF_3$: 303.0626; found 303.0622; [M+Na]+ calculated for $C_9H_{13}NaN_2O_4SF_3$: 325.0446; found 325.0440.

2.6.3. Synthesis of 4-Methylumbelliferyl 2-acetamido-2-deoxy-β-D-glucopyranoside substrates

General procedure for the synthesis of 4-Methylumbelliferyl 2-Acetamido-2-deoxy-D-gluco-pyranosides (18a-e). 4-Methylumbelliferyl 2-amino-2-deoxy-β-D-glucopyranoside hydrochloride (16) was synthesized according to the procedure described by Roeser and Leger, and was used without further purification.[251] To a solution of the hydrochloride salt 16 (250 mg, 0.5 mmol) in CH_3CN (5 ml) was added triethylamine (2.0 eq.) and the appropriate alkyl isocyanate (2.0 eqv. [for **17a**, TMSNCO was used]) at 0° C and the solution left to stir at room temperature (3 h). The reaction was then diluted with EtOAc (25 ml) and washed with water (25 ml), $NaHCO_3$ (25 ml), brine (10 ml), dried (MgSO4), filtered and concentrated. The resultant solids were then de-protected using the procedure of Macauley *et al*[84] to give the corresponding triols **18a-e** in yields ranging from 49% - 63% over two steps.

4-Methylumbelliferyl-2-deoxy-2-[(amino)oxoamino]-β-D-glucopyranose (18a): ^{1}H NMR (600 MHz, DMSO-d_6) δ 8.60 (s, 1H), 7.71 (dd, *J* = 8.8, 3.6 Hz, 1H), 7.03 – 6.94 (m, 2H), 6.74 (s, 1H), 6.26 (d, *J* = 1.6 Hz, 1H), 6.04 (d, *J* = 8.2 Hz, 1H), 5.51 (s,

1H), 5.34 (d, *J* = 8.4 Hz, 1H), 5.21 (d, *J* = 8.2 Hz, 1H), 4.64 (s, 1H), 3.73 (d, *J* = 11.1 Hz, 1H), 3.66 (q, *J* = 9.2 Hz, 1H), 3.47 (q, *J* = 8.1, 5.8 Hz, 3H), 3.20 (q, *J* = 9.4 Hz, 1H), 2.41 (dd, *J* = 2.5, 1.3 Hz, 3H). ^{13}C NMR (151 MHz, DMSO-d_6) δ 160.16, 160.01, 158.86, 154.38, 153.25, 126.50, 114.23, 113.54, 111.77, 103.17, 98.99, 77.14, 74.69, 70.47, 60.64, 55.67, 18.11. HRMS (m/z): [M+H]+ calculated for $C_{17}H_{21}N_2O_8$: 381.1298; found 381.1293; [M+Na]+ calculated for $C_{17}H_{20}N_2NaO_8$: 403.1117; found 403.1111.

4-Methylumbelliferyl-2-deoxy-2-[(aminomethyl)oxomethyl]amino]-β-D-glucopyranose (18b): ^{1}H NMR (600 MHz, DMSO-d_6) δ 7.71 (d, *J* = 8.8 Hz, 1H), 7.02 – 6.94 (m, 2H), 6.26 (d, *J* = 1.5 Hz, 1H), 5.96 (d, *J* = 8.4 Hz, 1H), 5.81 (d, *J* = 4.9 Hz, 1H), 5.20 (d, *J* = 8.3 Hz, 1H), 5.11 (dd, *J* = 17.7, 5.3 Hz, 2H), 4.62 (t, *J* = 5.8 Hz, 1H), 3.73 (dd, *J* = 11.4, 5.4 Hz, 1H), 3.54 – 3.38 (m, 3H), 3.20 – 3.16 (m, 2H), 2.55 (d, *J* = 4.6 Hz, 3H), 2.41 (d, *J* = 1.3 Hz, 3H). ^{13}C NMR (151 MHz, DMSO-d_6) δ 160.16, 160.04, 158.72, 154.38, 153.27, 126.43, 114.12, 113.54, 111.72, 103.16, 99.03, 77.20, 74.70, 70.45, 60.70, 56.34, 26.42, 18.11. HRMS (m/z): [M+H]+ calculated for $C_{18}H_{23}N_2O_8$: 395.1454; found 395.1449; [M+Na]+ calculated for $C_{18}H_{22}N_2NaO_8$: 417.1274; found 417.1269.

4-Methylumbelliferyl-2-deoxy-2-[(aminoethyl)oxoethyl]amino]-β-D-glucopyranose (18c): ^{1}H NMR (600 MHz, DMSO-d_6) δ 7.71 (d, *J* = 8.8 Hz, 1H), 7.00 – 6.95 (m, 2H), 6.26 (d, *J* = 1.5 Hz, 1H), 5.90 – 5.86 (m, 2H), 5.21 – 5.09 (m, 3H), 4.62 (t, *J* = 5.8 Hz, 1H), 3.73 (dd, *J* = 11.7, 5.1 Hz, 1H), 3.52 – 3.45 (m, 2H), 3.41 (dd, *J* = 9.5, 5.8 Hz, 1H), 3.24 – 3.15 (m, 1H), 3.02 (dd, J = 7.5, 6.0 Hz, 2H), 2.63 (d, *J* = 2.4 Hz, 1H), 2.42 – 2.39 (m, 3H), 0.99 (td, *J* = 7.2, 1.3 Hz, 3H). ^{13}C NMR (151 MHz, DMSO-d_6) δ 160.19, 160.04, 158.07, 154.38, 153.28, 126.43, 114.28, 113.67, 111.47, 103.09, 99.09, 77.21, 74.65, 70.47, 60.71, 56.26, 34.13, 18.11, 15.61. HRMS (m/z): [M+H]+ calculated for $C_{19}H_{25}N_2O_8$: 409.1608; found 409.1611; [M+Na]+ calculated for $C_{18}H_{22}N_2NaO_8$: 431.1430; found 431.1425.

4-Methylumbelliferyl-2-deoxy-2-[(aminopropyl)oxopropyl]amino]-β-D-glucopyranose (18d): ^{1}H NMR (600 MHz, DMSO-d_6) δ 7.71 (dd, *J* = 8.9, 1.4 Hz, 1H), 7.02 – 6.94 (m, 2H), 6.26 (t, *J* = 1.3 Hz, 1H), 5.94 – 5.85 (m, 2H), 5.20 – 5.09 (m, 3H), 4.63 (t, *J* = 5.7 Hz, 1H), 3.76 – 3.70 (m, 1H), 3.55 – 3.45 (m, 2H), 3.42 (tt, *J* = 8.1, 4.6 Hz, 2H), 3.19 (s, 1H), 2.95 (ddd, *J* = 13.0, 10.4, 6.8 Hz, 2H), 2.41 (t, *J* = 1.4 Hz, 3H),

1.37 (q, *J* = 7.3 Hz, 2H), 0.82 (td, *J* = 7.4, 1.6 Hz, 3H). ^{13}C NMR (151 MHz, DMSO-d_6) δ 160.17, 160.04, 158.19, 154.38, 153.25, 126.42, 114.12, 113.51, 111.71, 103.17, 99.23, 77.21, 74.64, 70.48, 60.72, 56.27, 41.12, 23.15, 18.11, 11.31. HRMS (m/z): [M+H]+ calculated for $C_{20}H_{27}N_2O_8$:423.1767; found 423.1769; [M+Na]+ calculated for $C_{20}H_{26}N_2NaO_8$: 445.1587; found 445.1586.

4-Methylumbelliferyl-2-deoxy-2-[(aminobutyl)oxobutyl]amino]-β-D-glucopyranose (18e): ^{1}H NMR (600 MHz, DMSO-d_6) δ 7.71 (d, *J* = 8.8 Hz, 1H), 7.00 – 6.94 (m, 2H), 6.26 (d, *J* = 1.3 Hz, 1H), 5.90 – 5.84 (m, 2H), 5.20 (d, *J* = 8.3 Hz, 1H), 5.12 (dd, *J* = 21.0, 5.3 Hz, 2H), 4.63 (t, *J* = 5.7 Hz, 1H), 3.73 (dd, *J* = 11.6, 5.1 Hz, 1H), 3.54 – 3.45 (m, 2H), 3.19 (td, *J* = 9.1, 5.1 Hz, 1H), 2.99 (dp, *J* = 19.3, 6.5 Hz, 2H), 2.62 (d, *J* = 1.9 Hz, 1H), 2.41 (d, *J* = 1.3 Hz, 3H), 1.30 (dp, *J* = 47.8, 7.1 Hz, 4H), 0.85 (dd, *J* = 7.8, 6.8 Hz, 3H). ^{13}C NMR (151 MHz, DMSO-d_6) δ 160.21, 160.03, 158.16, 154.37, 153.27, 126.40, 114.20, 113.54, 111.73, 102.89, 99.24, 77.21, 74.60, 70.48, 60.72, 56.29, 38.96, 32.13, 19.48, 18.06, 13.71. HRMS (m/z): [M+H]+ calculated for $C_{21}H_{29}N_2O_8$: 437.1924; found 437.1921; [M+Na]+ calculated for $C_{21}H_{28}N_2NaO_8$: 459.1743; found 459.1738.

2.6.4. Biological evaluation

Gene expression and protein purification. hOGA was expressed and purified according to the previously described procedure by Macauley *et al*.[84] hOGA was dialyzed twice against PBS buffer after purification and its concentration determined using the Nanodrop$_{2000}$. Human β-hexosaminidase was obtained from Micheal Tropak at the Hospital for Sick Children in Toronto, without further purification.

Kinetic analysis of hOGA. All assays were carried out in 96 well plates (Thermoscientific, FloroNunc, lot # 139825) in PBS pH 7.4 buffer (0.03% BSA) in triplicate at 37°C for 20 minutes. A continuous assay procedure was performed in which the reactions were initiated by the addition of 4MU-GlcNAc substrate (50 µL or 100 µL via a multi-channel pipette). The total assay volume was 150 µL for the K_i determination or 200 µL for the K_m/V_{max} determination. The progress at the end of the reaction was determined by assessing the extent of 4-methylumbelliferone liberation as determined by fluorescence measurements using a BioTek Synergy Plate Reader. The excitation and

emission were 350 and 445 nm, respectively. The amount of fluorogenic substrate liberated was quantitatively assessed by using a standard curve for 4-methylumbelliferone under identical buffer conditions, covering concentrations 0-0.3 µM. A total of 6-8 inhibitor concentrations were tested, ranging from 1/5 to 5 times the K_i value. The hOGA concentrations in the assays ranged from 0.8 nM to 20 nM, depending on whether the K_i being determined was through classic Michaelis-Menten methods or the Morrison K_i fit. For the Morrison K_i analyses, only one substrate concentration was used (at the K_m) since the method of inhibition was already known from analysis using the Line-weaver Burke plots.

Kinetic analysis of β-hexosaminidase. All assays were carried out in 96 well plates (Thermoscientific, FloroNunc, lot # 139825) at pH 4.25 (50 mM citrate, 100 mM NaCl, 0.03% BSA) at 37°C in triplicate for 20 minutes. Enzymatic reactions were triggered by addition of substrate (20 µL or 15 µL) in a total volume of 45 or 40 µL. Reactions were quenched by the addition of a 4-fold excess (160 or 180 µL) of quenching buffer (200 mM glycine, pH 10.75) in order to detect the fluorescent signal. The progress at the end of the reaction was determined by assessing the extent of 4-methylumbelliferone liberation as determined by fluorescence measurements using a BioTek Synergy Plate Reader. The excitation and emission were 350 and 445 nm, respectively. A total of 6-8 inhibitor concentrations were tested, ranging from 1/5 to 5 times the K_i value. The hHexB concentrations in the assays ranged from 2 to 5 nM and the K_i values were assessed by linear regression of data from Dixon plot analysis.

2.6.5. pK_a determination using ^{13}C NMR

All titrations were carried out on the Bruker AVANCE II 600 MHz NMR spectrometer. The samples prepared for each titration contained 5 µL of 1,4-dioxane as an internal standard, an equal mol amount of inhibitor and 3-nitrophenol (depending on the availability of each compound) in 500 µL of H_2O and 75 µL of D_2O. Previous studies by Perrin *et al* have reported no differences in chemical shifts for varying the compound concentration in the titration assay.[237] Each sample was made basic with 25-70 µL of 2M NaOH, and prior to titration, it was ensured that there was no change in chemical shift with further addition of base in 5 µL increments, in order to accurately establish the

δ_d for each compound. The basic sample was then titrated with 10 µL of 0.5 M (**15f-15h**) or 1 M HCl (**15c,** NButGT) until there was no further change in chemical shift, indicating that the end point of the titration was reached. The point where the chemical shift stopped changing was the recorded δ_p for each compound.

Chapter 3.

Development of quenched substrates for monitoring human *O*-GlcNAcase activity

As was discussed in Chapter 1, the activities of GHs can be studied using inhibitors (both covalent and non-covalent), ABPs, and fluorogenic substrates. *In vitro* assays using inhibitors and fluorogenic substrates aid in the kinetic characterization of enzymes through determination of parameters such as the dissociation constant for the enzyme-inhibitor complex (K_i), the inhibitor concentration causing 50% inhibition of the enzyme (IC_{50}), substrate specificity of enzymes, pH and cofactor dependence, as well as evaluation of rate constants for substrate processing - kinetic concepts which were described in detail in Chapter 2. However, in general, determining enzymatic levels in cell lysates has the limitation of not providing the amount of enzyme which is active, since the cells need to be lysed prior to imaging. Also, since the cells are broken down, certain molecular effectors which augment enzyme activity may be lost. For this reason, *in vivo* assays are a more powerful approach, as they provide real time measurements of enzyme activity in the endogenous environment. Alternatively, ABPs, such as the ones noted in Chapter 1 based on cyclophellitol[166] developed by the Overkleeft group, are extremely useful chemical tools which enable monitoring of active enzymes in cellular environments. One notable example is the monitoring of active GBA1 levels in cell lysates and imaging GBA in fixed cells. This approach of using CBE-based ABPs allows for potential optimization of current therapies for Gaucher's disease since beneficial information about GBA levels in individual patients can be gained. Such ABPs are very useful tools for study of GHs, however, they report on the absolute levels of enzyme rather than enzyme activity, which can be influenced by levels of associated proteins or cofactors, as well as the sub-cellular environment itself. An elegant example of direct live cell monitoring of GBA1 activity was described by Yadav *et al* using the quenched substrate described in Chapter 1 (BODIPY/BHQ2 substrate, Figure 1.22).[181] In this

example, the use of a quenched fluorogenic substrate allows for real-time monitoring of GBA activity, and provides a platform for live cell imaging and high-throughput imaging of GBA1 activity, including in inhibitor screening. Quenched substrates have an advantage over ABPs because they are not covalently bound to the enzyme and as a result do not inhibit their activity. Accordingly, there are, as mentioned above, many advantages of applying quenched substrates to study GHs and enzymes in general, something which will be further discussed in this chapter.

3.1. Using quenched substrates is beneficial for live cell imaging

A frequently used approach to monitor enzyme activity and interactions between biomolecules employs Förster Resonance Energy Transfer (FRET) where a fluorescent donor and fluorescent acceptor, tethered to a biomolecule, with overlapping emission and absorption spectra, respectively, are studied. Alternatively, FRET quenched substrates are used for monitoring enzyme activity. These substrates consist of a fluorescent donor and a non-fluorescent acceptor covalently bound to a molecule of interest. For FRET quenching to occur, the donor and acceptor moieties must have significant spectral overlap, and must be bound in close proximity to each other. Upon cleavage of the substrate by the target enzyme, the donor and acceptor diffuse apart, preventing FRET-quenching and allowing the fluorescent donor to emit fluorescence and be detected. Using a dark non-fluorescent quencher helps to eliminate the background signal of the substrate, and therefore overcomes one of the main limitations of using simple fluorogenic substrates by allowing even small levels of substrate hydrolysis to be detected.

Optimizing fluorescence-quenched substrates to incorporate red-shifted or near-infra-red (IR) fluorophores is important for successful live cell imaging, making it easier to visualize activity using fluorescence microscopy due to background cellular auto-fluorescence that is abundant in the visible range of the electromagnetic spectrum (380 nm to 750 nm)[257]. Low auto-fluorescence combined with low background fluorescence of the substrate due to quenching is important for sensitive and accurate images of the endogenous activities of GHs. Another aspect to consider when designing probes for live

cell imaging is the need to optimize the physicochemical properties of the substrate to enable it to penetrate cell membranes and render the appropriate cellular compartment. This issue can be addressed by considering the balance between water solubility and lipophilicity and keeping the molecule as small as possible.

The chemical tools used thus far for studying increased *O*-GlcNAc levels in cells and animals are strictly inhibitors, including those mentioned in Chapters 1 and 2. Currently, there are no chemical tools available to monitor hOGA activity in live cells, which would be useful for both exploratory science and for evolution of therapeutic compounds. Thus, my research described in this chapter focuses on the goal of designing fluorescence-quenched substrates for hOGA that could be used to monitor the enzyme in live cell assays, as was achieved with GBA[181]. Fluorescent quenched substrates have not been widely explored for the live cell assessment of GH activities[181, 258] and is entirely unexplored when it comes to hOGA. Notably, there are no examples of GH quenched substrates undergoing photoinduced electron transfer (PET) quenching, as opposed to FRET quenching. In this chapter, I describe my research focused on design of hOGA substrates in which I prepare and assess fluorophore and quencher pairs by varying the fluorophore at the anomeric position of GlcNAc, and quenchers at either the 2 or 6 positions of the GlcNAc residue (Figure 3.1) in collaboration with Evan Perley-Robertson and Dr. Samy Cecioni. This thesis chapter covers the research associated with the quencher at the 2 position of the GlcNAc residue.

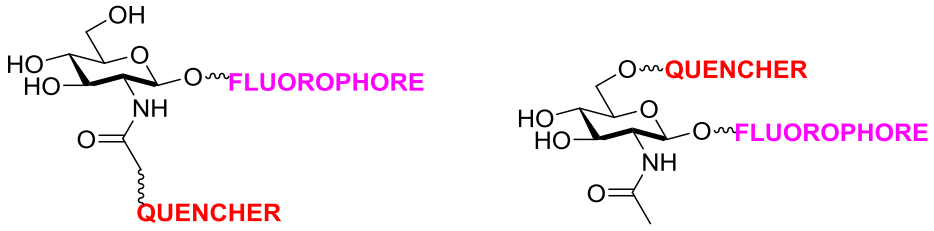


Figure 3.1. The functionalization of GlcNAc with fluorophores at the anomeric position and quenchers at the 2' (left) or 6' (right) positions.

To understand the reasoning behind the substrate design it is necessary to first elaborate on the various forms of non-radiative energy transfer responsible for

fluorescence quenching, and to assess a few relevant examples which influenced my research goals and have impact on future work.

3.2. Fluorescence quenching can occur through Förster Resonance Energy Transfer (FRET)

FRET occurs between two chromophores in a distance dependent photophysical manner, a concept described by Theodor Förster in 1948.[259] A donor chromophore in which an electron has been excited to an excited singlet state (S_1) relaxes back to the ground state (S_o) by transferring energy to a nearby acceptor chromophore through a non-radiative, long-range, dipole-dipole coupling. This coupling is enabled by the spectral overlap between the emission spectrum of the donor and absorption spectrum of the acceptor. The donor absorbs energy at a shorter wavelength, while the acceptor does so at a longer wavelength.[260] The acceptor then fluoresces, emitting a photon that can be detected and quantified (Figure 3.2). Alternatively, this energy transfer can also result in quenching, instead of fluorescence, in cases where the acceptor, in its excited singlet state (S_1) undergoes relaxation through non-radiative processes including thermal relaxation.

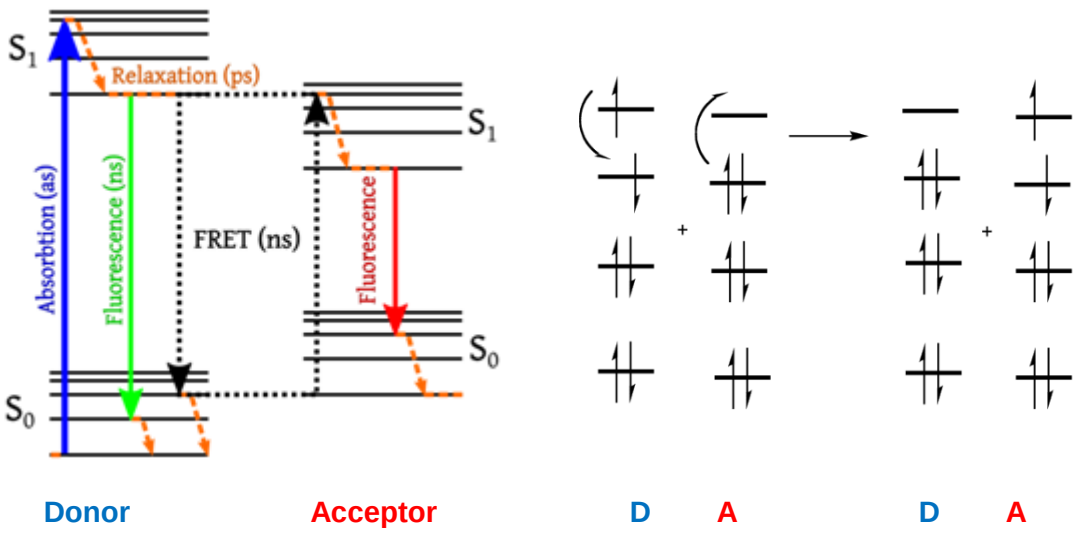


Figure 3.2. FRET occurs through non-radiative energy transfer from the excited donor to the ground state acceptor where the emission of the acceptor can be quantified.
The electronic transitions for both donor (D) and acceptor (A) are described on the right. This figure was adapted from Wikipedia.

In order to maximize FRET efficiency, the two chromophores must be within a few nanometers of one another. A distance of greater than 10 nm generally eliminates FRET. Due to the distance-dependent nature of this method, it's use has been effective in biology for the evaluation of dynamic interactions between biomolecules[261], as well as for the detection of conformational changes in biomolecules in cases where donor and acceptor chromophores are positioned on the same structure.[262] The FRET efficiency (E) depends on the donor-acceptor distance according to Equation 3.1. The Förster distance, R_0, refers to the distance at which the energy transfer efficiency between donor and acceptor is 50% and this term largely depends on the extent of overlap between the emission spectrum of the donor and absorption spectrum of the acceptor.[263-265] Noteworthy is that E is inversely proportional to the 6th power of *r,* the distance between donor and acceptor, highlighting the strong dependence of FRET efficiency on the distance between the donor/acceptor pair.[266]

$$E = \frac{1}{1 + \left(\frac{r}{R_0}\right)^6}$$

Equation 3.1. The efficiency of FRET (*E*).
R_0 represents the Förster distance with a 50% transfer efficiency, and *r* is the distance between donor and acceptor.

3.2.1. FRET-based probes have been developed for monitoring protease activity

The concepts behind fluorescent quenched substrates have been discussed from as early as the 1970's[267], and this strategy has been applied to many enzyme families including most notably for the proteases in the context of enzymatic assays, live cell imaging and imaging in organisms.[268-272] Although there have been many examples of ABPs being used for probing enzyme activity, there are limitations with this approach, the largest limitation being the covalent irreversible enzyme inhibition that results when using them. Each probe molecule labels an enzyme active site, meaning that the fluorescent signal that results cannot be amplified, potentially limiting sensitivity and leading to a measurement of active enzyme level rather than level of enzymatic activity in the cell. In the case of quenched substrates, these limitations are circumvented since there is no inhibition of the enzyme being studied, and the fluorescent signal can be amplified since one enzyme can cleave numerous substrate molecules, leading to signal amplification.[273] Consequently, the approach of using quenched substrates instead of covalent labeling strategies may have certain benefits when it comes to drug discovery because direct effects of competitive inhibition of the target enzyme can be quantified during live cell imaging studies[274] and the effect of cellular factors influencing enzyme activity rather than enzyme level can be interrogated.

An example where the advantage of using FRET quenched substrates instead of ABPs has been demonstrated in the monitoring of lysosomal cysteine cathepsin activity in live cells.[275] This family of cathepsin proteases contains different members having varying substrate specificities. For example, a pharmaceutical target of high interest is cathepsin *K*, which is involved in the progression of osteoporosis.[276] Watzke *et al* reported on the development of cell-permeable and selective quenched substrates

specifically targeting cathepsins K and S, this second enzyme being important for immune function, by replacing the electrophilic domain of known selective cathepsin inhibitors with a cleavable peptide bond. Consequently, instead of inactivating these enzymes and forming a covalent bond, the resulting quenched substrate could be turned over. Spontaneously immortalized human keratinocyte (HaCaT) cells, which possess elevated expression of cathepsins S and L in the lysosome were used for monitoring cathepsin S and L activity in live cell imaging assays by incubating cells with quenched substrates **2**, **4** and **5** (Figure 3.3), whose structures were modified from well-established cathepsin inhibitors **1** and **3.** These substrates showed great promise for future *in vivo* studies due to their fast turnover (*e.g.* k_{cat}/K_m = 15100 $M^{-1}s^{-1}$ for cathepsin S with probe **4**) and low K_m values (*e.g.* K_m = 1.3 x 10^{-6} M for cathepsin S with probe **4**).[275]

Substrate **5** was designed for cellular imaging since the photophysical properties of the fluorophore/quencher (BODIPY TMR-X/QSY-7) pair were optimized for this purpose. A clear fluorescence increase in the lyososomes of HaCaT cells was observed upon incubation with 8 µM of **5.** When 15 µM of broad-spectrum cathepsin inhibitor or 15 µM of cathepsin S and L inhibitor were pre-incubated with the HaCaT cells, there was almost complete suppression of substrate-associated fluorescence, indicating that the increased signal in the lysosome upon incubation of substrate **5** resulted from processing by cathepsins S and L. This study demonstrates how optimization of the fluorophore/quencher pair can be done to enable cellular imaging and that modifying the structure based on the well-established structure-activity relationships of known inhibitors which have already been shown to penetrate cells can be a valuable approach to rational quenched substrate design and a strategic way to achieve selectivity. This is particularly useful for proteases, since quenched reporter substrate structures are usually based on peptide substrates attached to a reporter group, and their identification is typically performed through positional scanning combinatorial libraries[277] which can be time-consuming.

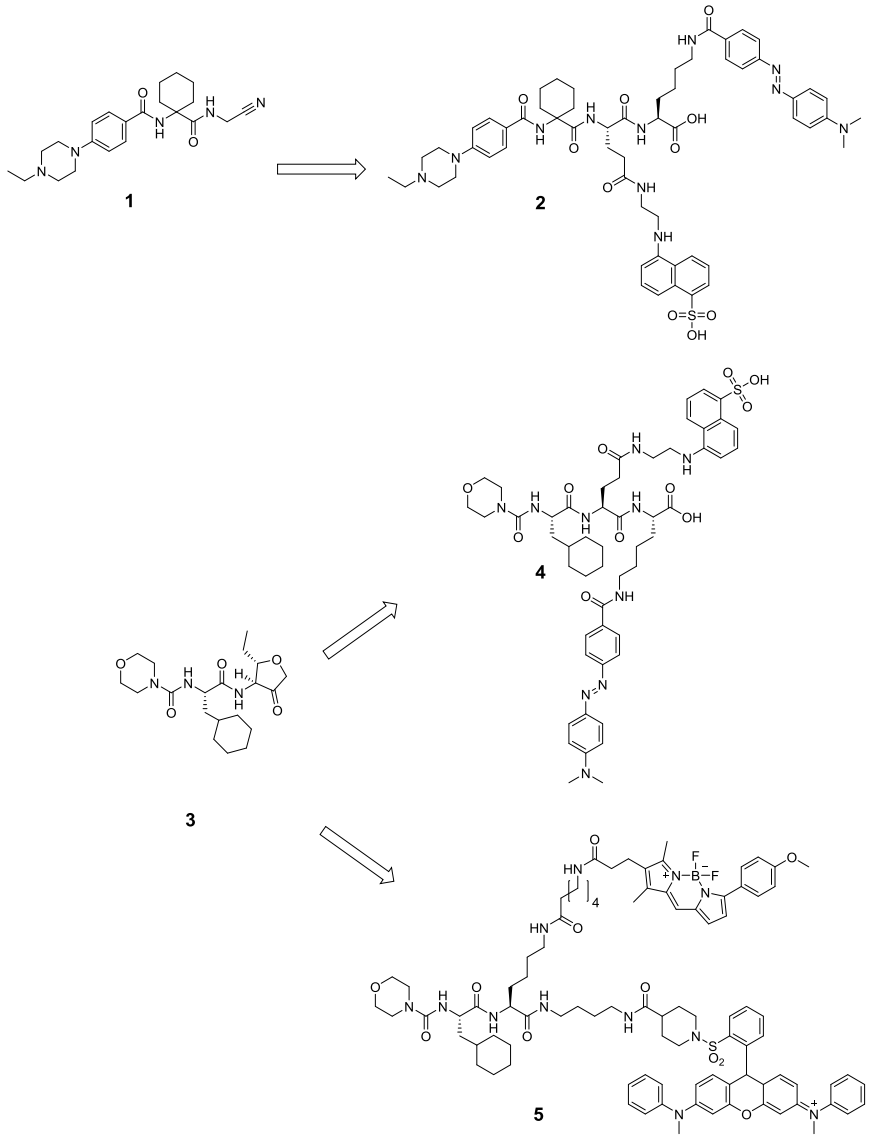


Figure 3.3. ABP substrates 2, 4 and 5 which target cathepsins S, K, B, L (probes 2 and 4) and cathepsins S and L (probe 5), whose structures were modified from known, cell-permeable, cathepsin inhibitors 1 and 3. The fluorophore/quencher pair for probes **1** and **2** are 5-((2-aminoethyl)amino)napthalene-1-sulfonic acid (EDANS) and 4-{[4-(dimethylamino)-phenyl]-benzoic acid (DABCYL), respectively. The fluorophore/quencher pair for probe **5** is BODIPY TMR-X/QSY-7, which better accommodates *in vivo* experiments by enhancing cell permeability and including a red-shifted fluorophore for cell imaging.

Another significant example of FRET quenched substrates being used is the development of such a substrate for monitoring β-secretase-1 (BACE1) activity in live cells. BACE1, one of the three proteases responsible for cleaving the amyloid precursor

protein (APP), is a membrane anchored aspartic protease which cleaves APP in the extracellular domain. Development of BACE1 inhibitors has been a target for many researchers since there is strong evidence pointing to the BACE1-catalyzed cleavage of APP being a facilitator of Alzheimer disease.[278, 279] Although it is possible to screen inhibitors of BACE1 *in vitro* using FRET substrates, the analysis of BACE1 activity in cells is much more complex due to BACE1 having to be incorporated into endosomes, where the pH is between 4 and 5, in order be activated.[280] Many inhibitors and FRET substrates are unable to be incorporated into endosomes and therefore the activity of BACE1 is difficult to analyze in this environment. To overcome these obstacles, Folk *et al* developed a β-secretase membrane-anchored probe (β-MAP) used for monitoring BACE1 activity in live cells.[281] Similar to the strategy mentioned for targeting cathepsins, known BACE1 structural features were incorporated into the FRET quenched substrate with some modifications to the amino acid composition to include those which were previously shown to be compatible with BACE1 cleavage. The fluorophore/quencher FRET pair used was 7-dimethylamino-coumarin-4-acetic acid/4-(dimethylaminoazo)benzene-4-carboxylic acid (DMACA/DABCYL) since DMACA is functional at the desired lower pH 4.5 as well as pH 7.4, and DABCYL was small enough not to interfere with enzyme hydrolysis. The structure and description of the membrane anchored substrate ABP is shown in Figure 3.4.

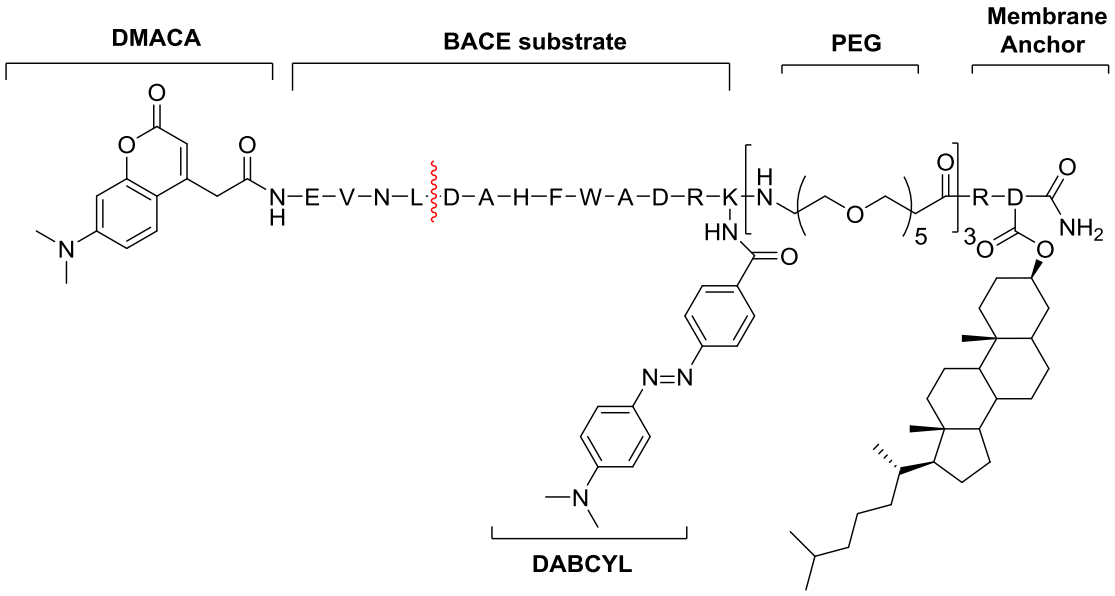


Figure 3.4. The β-MAP ABP substrate used for monitoring BACE1 activity in live cells.
The fluorophore used for fluorescence quantification is 7-dimethylamino-coumarin-4-acetic acid (DMACA) (λ_{ex}: 370 nm, λ_{em}: 480 nm) located at the N-terminus of the BACE substrate. The quencher, 4-(dimethylaminoazo)benzene-4-carboxylic acid (DABCYL) (λ_{ex}: 420-520 nm) is attached to a lysine residue at the opposite end. The cell membrane anchor is a dihydrocholesterol modified aspartic acid, which is linked to the rest of the probe with a poly(ethylene glycol) (PEG) linker, enabling the probe substrate to attach to lipid rafts within the cell membrane.[281]The cleavage site is shown in red, between the leucine and aspartic acid residues[282].

The β-MAP substrate (200 nM) was tested in live HeLa cells, and the fluorescence intensity stemming from the release of DMACA from β-MAP gradually increased over time within acidic intracellular vesicles, indicating that the substrate was turning over in the expected cellular location. Once the method was validated, the BACE1 activity was monitored in the presence of a known BACE1 inhibitor, Axon 1125 (0-500 nM), using 200 nM of β-MAP. A consistent dose-dependent decrease in fluorescence signal was observed with increasing inhibitor concentrations. β-MAP was useful in this study for concluding that the inhibitor Axon 1125 is a BACE1 inhibitor in a cellular setting, which is something that was not previously established due to the limitations of previous assays. Additionally, this study validated the potential for using β-MAP as an effective probe for screening future BACE1 inhibitors in cells, as well as for

visualizing BACE1 in real time while altering other cellular conditions which could play a role in Alzheimer disease-related studies related to BACE1.

Interestingly, from the perspective of this thesis, even though there are several excellent examples of FRET quenched substrates for visualizing protease activities in live cells, examples of such substrates for GHs in the literature are extremely sparse. There are examples displaying the use of fluorogenic substrates to monitor activity of bacterial enzymes such as *Escherichia coli* β-galactosidase[283] or mammalian senescence-associated β-galactosidase[284] but the only instance where normal endogenous levels of a GH are monitored in live mammalian cells using a quenched substrate is the example mentioned in Chapter 1 (Figure 1.22) for GBA1. The reason for this could be because proteases contain a cleft-like binding pocket, which makes it more simple to design substrates which could be accommodated, whereas most GHs are pocket-shaped.

3.3. Non-radiative energy transfer can also occur through Photoinduced Electron Transfer (PET)

Photoinduced electron transfer (PET) is a different method of non-radiative energy transfer that occurs when an electron rich donor absorbs light and donates an electron to an electron deficient acceptor, thereby creating a charge transfer complex. Quenching through a PET mechanism can occur through electron transfer from an excited donor to an acceptor (donor-excited PET, or d-PET), or from a ground state donor to an excited acceptor (acceptor-excited PET, or a-PET). During a d-PET mechanism, the excited fluorophore donates an electron from an excited singlet state (S_1) to the lowest unoccupied molecular orbital (LUMO) of the quencher. Alternatively, during an a-PET mechanism absorption of a photon prompts transfer of an electron from the highest occupied molecular orbital (HOMO) of the quencher to the lower energy semi-occupied orbital of the fluorophore.[285] Besides the fact that FRET involves long range dipole-dipole interactions while PET involves electron transfer between donor and acceptor, the nature of PET quenching differs from FRET since the direction of electron transfer can occur either from fluorophore to quencher or vice-versa, contrary to FRET, where during energy transfer the donor species is always the fluorophore donating

energy to the quencher acceptor. The most common direction of PET is the case where the fluorophore acts as an acceptor,[286] but regardless of the direction, there is always a resulting donor/acceptor radical ion pair that is formed. The mechanisms of PET quenching are summarized in Figure 3.5.

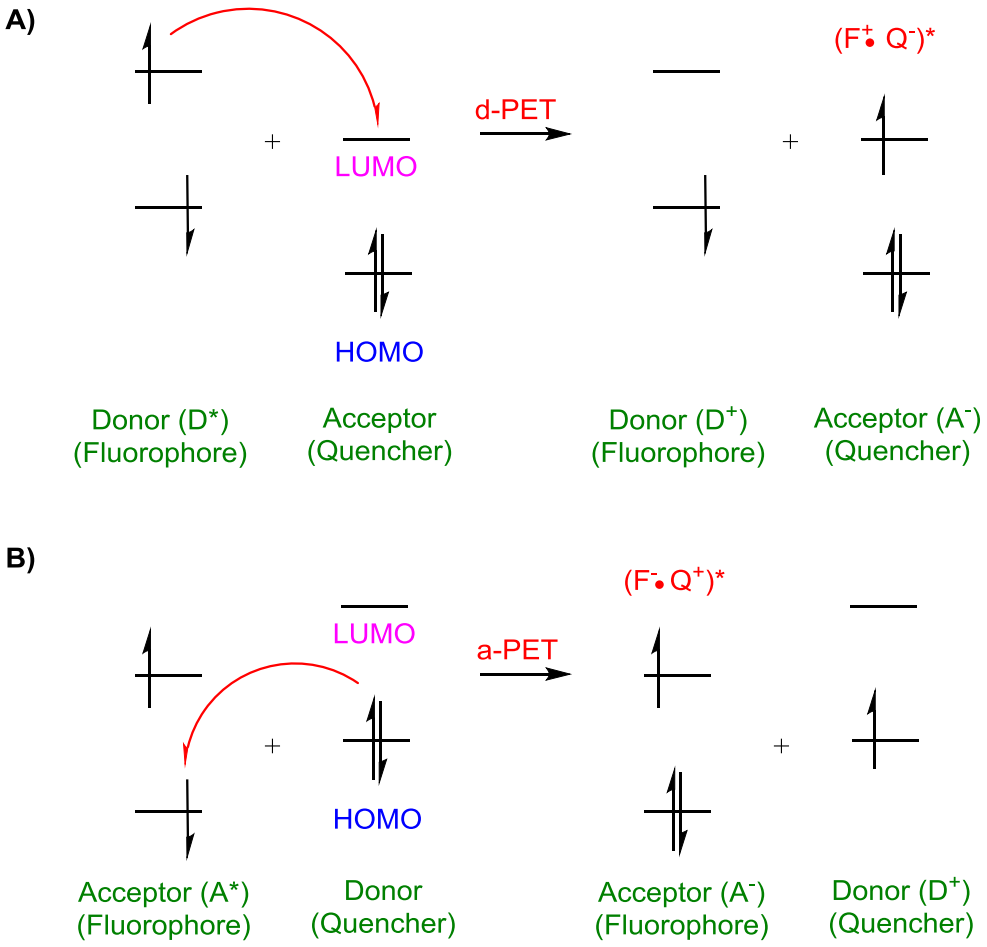


Figure 3.5. The mechanisms of PET quenching through d-PET (A) and a-PET (B).
A) The excited fluorophore donates an electron from S_1 to the lowest unoccupied molecular orbital (LUMO) of the quencher **B)** An electron is transferred from the highest occupied molecular orbital (HOMO) of the quencher donor which is higher in energy than the lower partially occupied molecular orbital (S_o) of the acceptor fluorophore. Ultimately, the result of PET is the formation of a donor/acceptor radical ion pair.[285]

Contrary to the conditions necessary for the occurrence of FRET, PET between a donor and acceptor is not dependent on the spectral overlap, which avoids one of the major limitations of FRET. However, both types of quenching still require proximity

between the donor and acceptor. In order for efficient PET quenching to occur, the donor and acceptor need to be in van der Waals contact, meaning that they must be within sub-nanometre distance from each other.[287] As a result, PET quenching probes have been developed for biomolecules that contain monomers close together, such as amino acids in peptides and proteins or nucleotides in DNA.[288] In addition to the strict distance requirement, PET efficiency is regulated by the redox potentials of the donor and the acceptor. In fact, the direction of electron transfer is governed by reduction and oxidation potentials of the excited and ground states. The Rehm-Weller model (Equation 3.2)[289] is used to define the energy change occurring during PET. The sign of the Gibbs free energy of electron transfer (ΔG_{ET}) is used to predict the spontaneity of electron transfer between a donor and acceptor and as a result dictates whether quenching can occur. In this equation, a negative ΔG_{ET} indicates that PET is favourable, whereas a positive value indicates that PET is unfavourable.[286]

$$\Delta G_{ET} = F\{E_{ox}(D) - E_{red}(A)\} - E_{0,0} + \frac{e^2}{\varepsilon d}$$

Equation 3.2. The Rehm-Weller model.

ΔG_{ET} represents the Gibbs free energy of electron transfer; F is the Faraday constant (9.65 x 10^4 C/mol); $E_{ox}(D)$ is the oxidation potential of the donor; E_{red} is the reduction potential of the acceptor; $E_{0,0}$ is the zero vibrational electronic excitation energy of the fluorophore (average of fluorophore absorption and emission wavelengths) and $e^2/\varepsilon d$ is the constant for Coulombic interactions occurring in the ion pair where ε is the solvent dielectric constant and d is the distance between the ions. This last term is relevant in cases where the ions are considered point charges, which typically applies in non-polar solvents. The value for this term in aqueous solution is negligible.[287, 290, 291]

Equation 3.2. stems from the free energy relationship in Equation 3.3 which defines the free energy change for moving charge in an electric potential, tying the relationship between the redox potential and free energy of electron transfer between donor and acceptor.

$$\Delta G = -nF\Delta E$$

Equation 3.3 The change in free energy for a moving charge in an electric potential.
∆E is the change in electric potential, n is the number of electrons, *F* is the Faraday constant (9.65 x 10^4 C/mol) and ∆G represents the energy change associated with moving 1 mol of electrons through an electric potential ∆E.

3.3.1. PET quenching can be categorized as dynamic or static

The mechanism of PET quenching can be carried out through dynamic (collisional) or static paths. In order for quenching to occur in solution between donor and acceptor molecules, the two species are required to collide. When these collisions are diffusion-limited, this translates into efficient quenching of all fluorophores in the excited state by the quencher, resulting in dynamic quenching. Experimentally, this type of mechanism can be detected by observing a decrease in the steady-state fluorescence intensity, since only the fluorophore excited state (F*) is affected by the quencher. This is reflected in reduced fluorescence lifetimes as well as reduced quantum yields (ϕ), since PET is occurring between the excited donor and quencher faster than emission. Dynamic quenching can be experimentally confirmed and graphically expressed by using the Stern-Volmer equation described by Equation 3.4, where the linear relationship between fluorescence intensity (*I*) and quencher concentration (*Q*) or fluorescence lifetime (τ) and Q can be correlated in a graph.

$$\frac{I^0}{I} = \frac{\tau^0}{\tau} = 1 + k_q\tau^0\,[Q]$$

Equation 3.4. The Stern-Volmer equation.
I^0 and I are the fluorescence intensities in the absence and presence of quencher (Q), respectively; τ^0 and τ are the lifetimes of the excited fluorophore in the absence and presence of (Q), respectively; k_q is the bimolecular rate constant of the excited fluorophore with Q. $k_q\tau^0$ can also be collectively expressed as K_D.

Alternatively, PET quenching can occur via a static mechanism if the acceptor and donor are in extremely close proximity, allowing for rates of PET in excess of those allowed by diffusion. In this case, the excited state of the fluorophore is not affected.

Instead, the ground state fluorophore interacts with the quencher, resulting in the formation of a non-emissive ground state complex which causes ϕ to be so low that it approaches zero.[287] Experimentally, this can be detected by the lack of change in the fluorescence lifetime with varying amounts of quencher, since only the excited state fluorophore contributes to the fluorescence lifetime, which in this case is not affected. Static quenching has been observed in cases where there is contact between molecules including van der Waals interactions, pi stacking, or other non-covalent interactions driving direct interaction between donor and acceptor. When considering rational designs for PET quenched substrates, bringing the fluorophore/quencher pair into close enough proximity on the same molecule would allow them to interact through both static and dynamic quenching and this could result in efficient quenching.

3.3.2. Many fluorescent probes are able to undergo PET quenching

PET quenching has been used to examine interactions stemming from conformational dynamics of biomolecules such as DNA and proteins. This has been successfully achieved by exploiting the strong electron donating properties and low oxidation potentials of tryptophan (Trp)[292] in peptides and proteins, and guanine (G)[293] in DNA, and using these attributes for quenching organic fluorophores tethered onto the appropriate biomolecules. Many common fluorophores have been explored in combination with Trp and G (Figure 3.6). Applications have included assaying proteolytic enzymes using fluorescent-labelled peptide probes which are self-quenched by Trp[294], and for studying the folding of a fluorophore/quencher labelled DNA hairpin molecule[295].

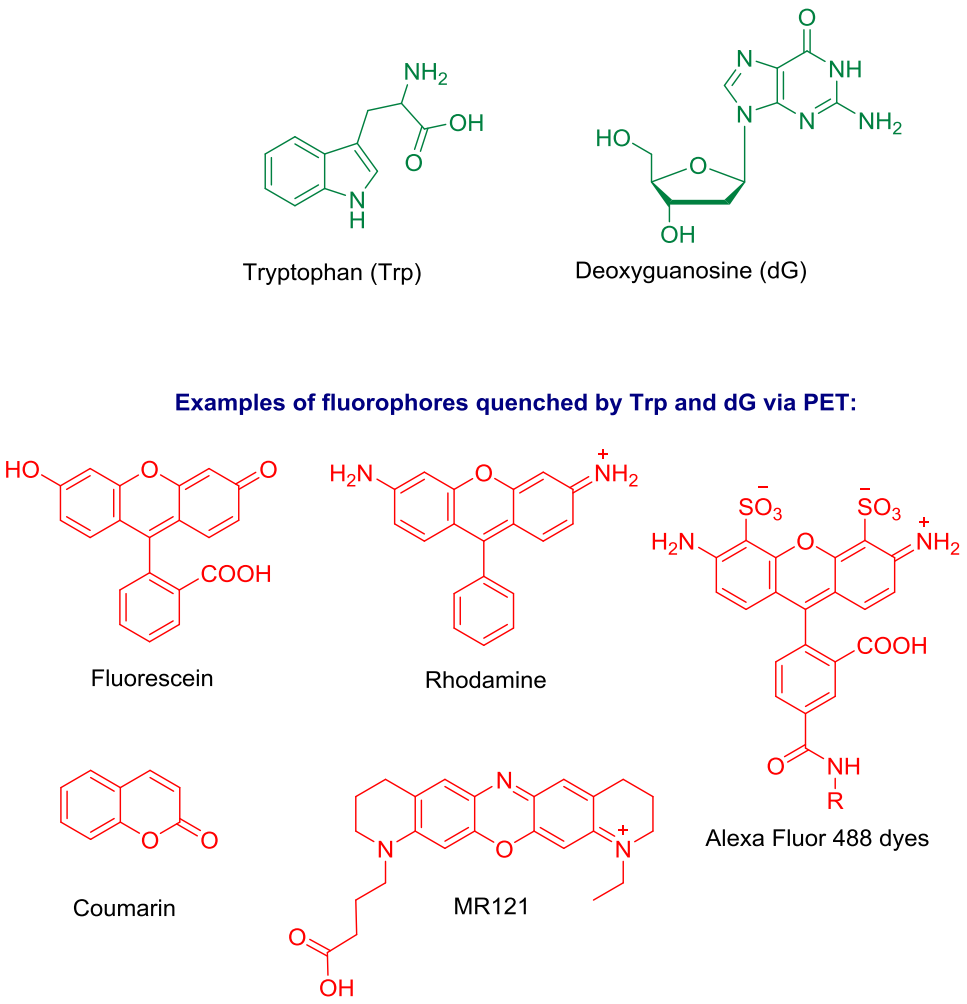
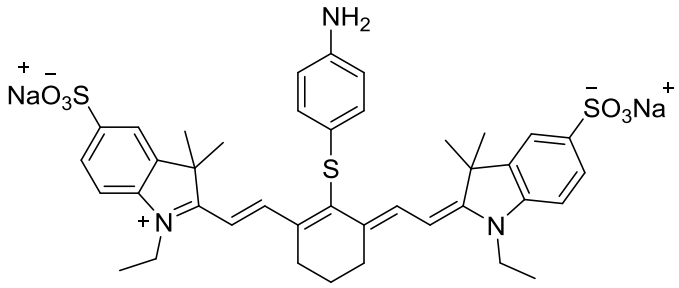


Figure 3.6. Tryptophan (Trp) and deoxyguanosine (dG) are useful natural electron donors which serve as useful PET quenchers for fluorophores such as the ones shown above.

Although this area of assessing biomolecule conformation by PET quenching systems has been continuously elaborated, there are few examples of fluorescent quenched substrates that are quenched by PET which have been used for assaying enzyme activities in live cells or tissues. A recent example of a PET quenched substrate used in this manner was published in 2013 by Wang *et al* [296], who described the selective probing of *N*-acetyl transferase 2 (NAT2) both *in vitro* and *in vivo* with a near-infrared (IR) fluorescent probe based on sulfonated-cyanine dyes (Figure 3.7).

NAT2 is a metabolic enzyme predominantly found in colorectal and liver tissues, which is responsible for transferring an acetyl group from acetyl CoA to aromatic amines

and arylhydroxylamines. It is of interest to monitor endogenous enzyme activity of NAT2 because the homozygous NAT2 genotype has been shown to be a bladder and lung cancer susceptibility factor[297]. The CYP1 substrate shown below in Figure 3.7 contains an arylamine, which fills a dual role as both the PET quencher as well as the recognition feature for the enzyme. In its native state, the CYP1 dye is quenched since the free amine acts as an electron donor. However, acetylation of the amine by NAT2 greatly reduces the electron donating ability of the nitrogen, preventing quenching and allowing the substrate to be detected by fluorescence.



CYP1

Figure 3.7. The structure of CYP1 used for *in vitro* and *in vivo* studies to examine the endogenous levels of NAT2.
The excitation and emission wavelengths are both in the NIR region, at 780 nm and 790 nm, respectively.

For the *in vivo* studies, mice were treated with the CYP1 probe (1.5 mg kg^{-1}) in parallel with a saline control via intravenous injection, and NIR fluorescence images and X-ray images were taken after 10, 30, 60 minutes and 24 hours. The X-ray images were used for contrasting purposes in order to detect the location of the tissues of interest based on the position of the bones. It was evident that there was a gradual increase in fluorescence over time in the liver, and after 30 minutes in the intestine, followed by complete clearance of the compound by 24 hours. To confirm the *in vivo* results, CYP1 was also quantified in tissue homogenates. A similar activity profile was observed in both cases, where the highest levels of CYP1 was detected in the liver and lungs. The levels of NAT2 were established in liver and compared to other tissues such as kidney, heart, spleen and brain. Although the endogenous levels of active NAT2 were not established in real time and the enzyme was quantified from tissue homogenates, this study was a useful example that illustrates the value of incorporating near-IR emitting fluorophores in

PET quenched substrates for *in vivo* use because of the low autofluorescence and good tissue penetration at these wavelengths.

This strategy of applying PET quenching *in vivo* is the long-term goal for making quenched substrates for hOGA. Interestingly, there are currently no examples of PET quenched substrates used to study the activity of glycosidases. The remainder of this chapter focuses on the development of PET quenched probe substrates for hOGA, using a thioamide as a small and convenient PET quencher.

3.4. Applications of thioamides as efficient PET quenchers

Currently, the most common PET quenchers of various fluorophores are large and bulky conjugated molecules. DABCYL (quenching range 400-550 nm) is a common collisional quencher, as well as a FRET quencher, which is compatible with many different fluorophores. The Black Hole Quencher (BHQ) dyes are also common static PET quenchers[298] and FRET quenchers, and are compatible with an even broader range of fluorophores since the different BHQ derivatives available have a broad absorption spectrum (450 to 750 nm) (Figure 3.8).

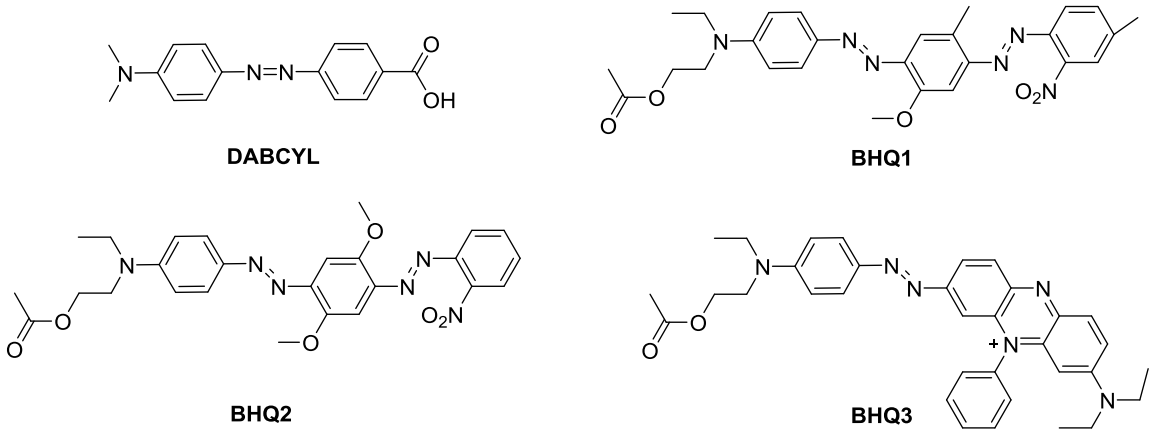


Figure 3.8. The structures of common dark quenchers DABCYL and BHQ derivatives.

Interestingly, thioamides have recently emerged as efficient PET quenchers of a wide spectrum of bright fluorophores excited with visible light, including Alexa Fluor 488, coumarin and Rhodamine 6G (Figure 3.6), amongst others.[291] This discovery was a

follow-up to previous work which was done by the Petersson group, involving the examination of thioamides incorporated into peptide structures as quenchers of Trp and tyrosine (Tyr) fluorescence.[299, 300] Thioamides are particularly attractive in this context because of their small size, which has a comparably small effect on the solubility of molecules to which they are tethered relative to many other hydrophobic, high molecular weight quenchers, which confers improved solubility in aqueous solutions. Additionally, owing to their small size, thioamides should contribute minimally to steric conflicts with enzyme active sites and may therefore be structurally compatible for use in quenched substrates that are processed by enzymes having a constrained binding pocket.

The distance-dependent quenching capabilities of fluorophores by Trp have been extensively exploited to examine conformational dynamics of macromolecules. However, the obvious drawback of using Trp in this capacity is having to incorporate such a bulky quencher within a protein, which may affect its dynamics. Goldberg *et al*. found that incorporating a thioamide instead of an amide directly into the polypeptide backbone quenches the fluorescence of both Trp and Tyr of chicken (CaM) protein calmodulin in aqueous buffer and in a distance-dependent manner.[299] A very small spectral overlap region was observed between the thioamide and Tyr (between 275 and 300 nm), and no spectral overlap between Trp and the thioamide. However, since in both cases similar quenching efficiencies were observed, it was evident that the principle quenching mechanism was likely mediated through a mechanism other than FRET. Since there is a large difference between the oxidation potential of oxoamides (3.75 eV) and thioamides (1.21 eV)[301] and there were previous observations of thioamides quenching fluorophores without any spectral overlap, it was predicted by Goldberg *et al* [299] that quenching was occurring through an electron transfer process.

An extension of this work assessed whether thioamides cause PET quenching of common fluorophores. In 2013, Goldberg *et al* evaluated the quenching effects of thioacetamide with 16 bright fluorophores.[291] The steady-state fluorescence of each fluorophore in solution was recorded at pH 7 in 100 mM sodium phosphate buffer containing either 50 mM acetamide or 50 mM thioacetamide. The quenching efficiency, $E_Q(SS)$, was then determined by comparing the fluorescence intensities between the

acetamide and thioacetamide solutions for each fluorophore and their results are summarized in Table 3.1.

Table 3.1. Thioamide quenching efficiencies of bright fluorophores. This table was extracted from the results obtained by Goldberg *et al.*[291]

Fluorophore	$\lambda_{ex}/\lambda_{em}$ (nm)	$E_Q(SS)$ (%)[a]
Coumarin102	393/487	0
NBD	467/538	27±1
Alexa Fluor 488	488/518	47±1
5-carboxyfluorescein	492/518	44±1
Fluorescein isothiocyanate	492/517	41±1
Fluorescein maleimide	492/516	44±1
Fluorescein click	494/521	44±1
BODIPY FL	502/510	61±1
Acridine Orange	502/525	3±1
Rhodamine R6G	526/556	16±1
BODIPY R6G	528/547	45±1
Cy3	550/570	0
Resorufin	571/585	51±1
Texas Red	582/602	0
Cy5	651/674	0
ATTO655	655/684	7±1

[a]The quenching efficiency was calculated using $EQ(SS) = 1-(F_{thioamide}/F_0)$; $F_{thioamide}$ = fluorescence at λ_{em} in 100 mM sodium phosphate buffer with 50 mM thioacetamide at pH 7.0; F_0 = fluorescence at λ_{em} in 100 mM sodium phosphate buffer at pH 7.

It can be seen from this data that red-shifted fluorophores that are commonly used for cellular imaging (500 nm and higher) undergo efficient intermolecular quenching by thioacetamide. This result brought to light the possibility of designing quenched substrates for hOGA, since we would be able to use these properties of thioamides and apply them to sugar-based glycosidase substrates by directly tethering a small thioamide quencher and fluorophore onto the same sugar residue.

3.5. Design of thioamide quenched hOGA substrates

The previous examples discussed in this chapter highlight the important qualities to consider when optimizing the structure of quenched substrates with the long-term goal of applying them to live cell or *in vivo* imaging studies. These desirable features include

using a dark quencher, a bright fluorophore that is preferably red-shifted and keeping the substrate small to enhance membrane permeability and processing efficiency by the enzyme of interest. The general substrate structure that we hypothesized may have all these traits is shown in Figure 3.9.

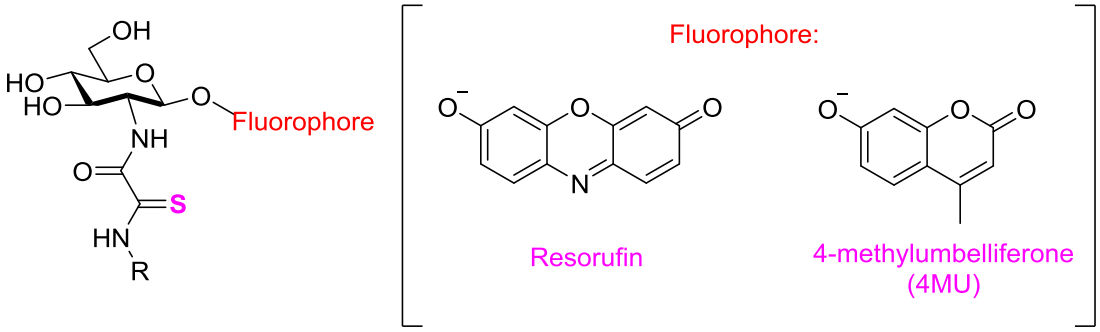


Figure 3.9. The general structure of the target probe substrate containing a fluorophore at the anomeric position and thioamide at the 2' position.
The fluorophores to be initially studied are 4-methylumbelliferone (4MU) and resorufin.

The rationale for the substrate design can be explained as follows:

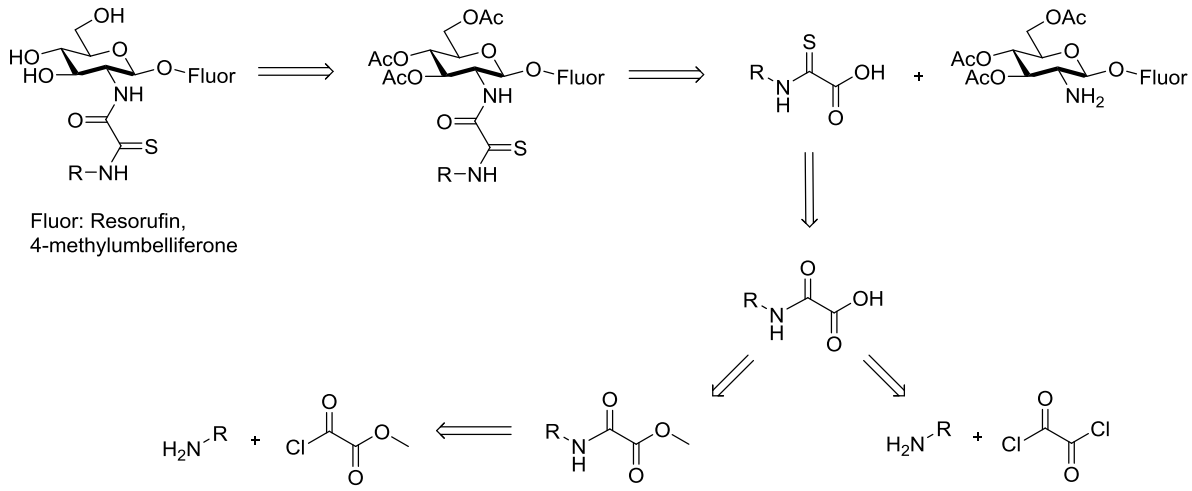
1) Using a thioamide as a small quencher should increase the likelihood of cell permeability as compared to using large, high molecular weight quenchers.
2) Since it was observed[291] that there is 51% quenching of resorufin (Figure 3.9) with thioacetamide in buffer, installing this red-shifted fluorophore at the anomeric position should result in even more efficient quenching since the effective concentration of the quencher would be greatly increased by tethering the fluorophore and quencher onto the same molecule. For synthetic simplicity, we chose to start with 4-methylumbelliferone (4MU) as the fluorophore, for a proof-of-concept.
3) Incorporation of resorufin as the fluorophore is also likely to maximize the cell permeability of the probe since resorufin-based substrates have often successfully been used in cellular imaging studies [302-304] and resorufin has been shown to be a cell-permeable inhibitor of protein kinase CK2[305].
4) Based on the structure-activity relationships observed in the inhibitor design study discussed in Chapter 2, by placing a chain which is at least 5 atoms in length at the 2 position of the sugar, we should be able to achieve high selectivity for hOGA over the lysosomal hexosaminidases.

3.6. Results and Discussion

3.6.1. Outline of substrate synthetic methods

The synthesis of quenched fluorogenic substrates of hOGA can be focused on two different structures: one containing 4MU and one containing resorufin as the fluorogenic leaving group. 4MU was chosen as the initial fluorophore for proof-of-concept studies since synthesis of the 4-methylumbelliferyl 3,4,6-*tri*-*O*-acetyl 2-amino-2-deoxy-β-D-glucopyranoside is known to be facile and high yielding.[84, 251] As mentioned previously, resorufin is a useful fluorophore for cell-based assays, and also showed 51% quenching by thioacetamide in buffer in the study by Goldberg *et al* (Table 3.1).[291] For this reason, resorufin was chosen as the fluorophore which would be used to make the quenched substrate to be assessed in cellular assays, after the proof-of-of concept was evaluated using the 4MU glycoside.

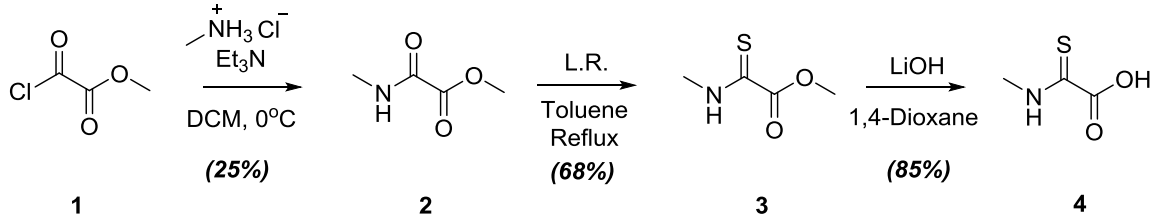
The synthesis of both quenched substrates was approached through a convergent synthesis by coupling 2-(methylamino)-2-thioxo-acetic acid (**4**) with the per-*O*-acetylated glucosamine sugar with the fluorogenic leaving group already installed (Scheme 3.1).



Scheme 3.1. Retrosynthetic outline for the synthesis of quenched fluorogenic substrates bearing resorufin and 4MU fluorophores.

3.6.2. Synthesis of the thioamide quencher

Both sugars containing fluorophores were coupled to acid **4**, which was synthesized conveniently in three steps. Starting from monomethyl oxalyl chloride reaction with methylamine hydrochloride in the presence of triethylamine provided methyl 2-(methylamino)-2-oxoacetate (**2**) (Scheme 3.2). This first step was modified from a literature procedure by Mecinovic *et al* who made a series of amino oxoacetates in reasonable yields.[306]



Scheme 3.2. The synthetic summary for 2-(methylamino)-2-thioxo-acetic acid (4).

The synthesis of thioamide **3** was done by refluxing amide **2** in toluene with Lawesson's reagent (Figure 3.10). Lawesson's reagent is commonly used to convert amides to thioamides, and more generally can be used to convert carbonyls to thiocarbonyl functionalities.[307, 308] Literature precedence showed cases where an amide in the presence of an ester was transformed to a thioamide without affecting the ester.[309]

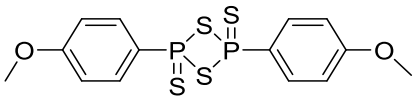
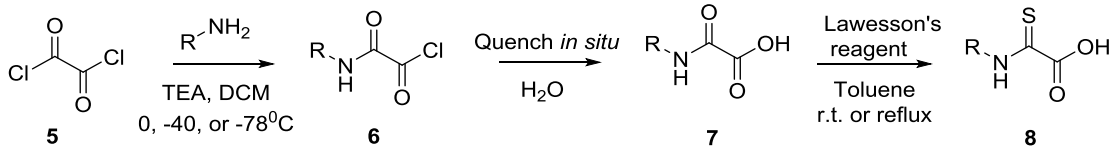


Figure 3.10. Lawesson's reagent: a common reagent for conversion of carbonyls to thiocarbonyls.

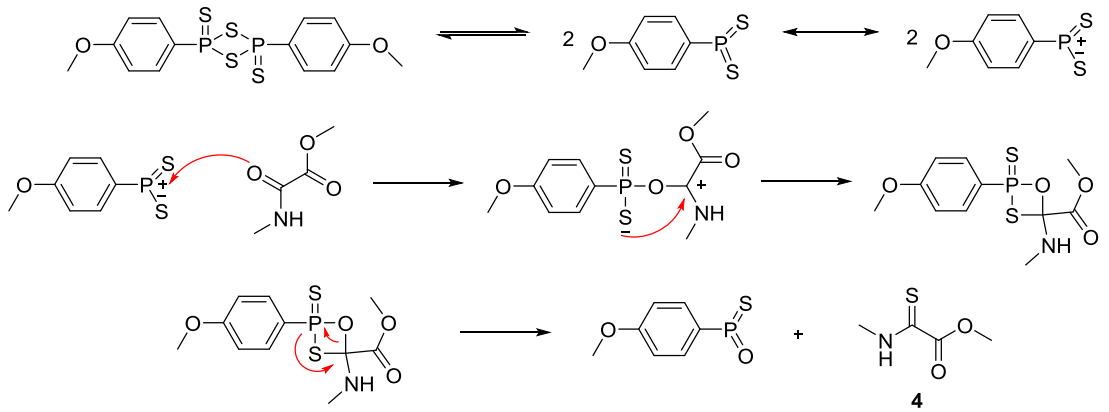
Initially, efforts were geared toward starting with oxalyl chloride **5** and using this to monoacylate an appropriate amine by using 0.5 equivalents of the amine and then hydrolyzing the remaining acyl chloride *in situ* with water to make acid **7** in one pot (Scheme 3.3). However, this approach proved problematic after observing that the majority of the product was generally the *di*-amide, making it difficult to obtain desired acid **7**. Decreasing the temperature of the reaction to -78 ^{o}C helped to obtain the desired mono-amide product, though some *di*-amide was still observed. Additionally, **7** is polar enough to be in the aqueous layer upon work-up and as a result was difficult to isolate and purify using column chromatography or crystallization. The solubility of the acid in water was problematic since the triethylammonium chloride salt, which was formed during the course of the reaction on converting **5** to **7** was also soluble in water, making them difficult to separate. Visualization of **7** was difficult during TLC analysis, even with carboxylic acid bromocresol green stain, making it difficult to monitor the progress of the reaction. Furthermore, after isolating **7**, its reaction with Lawesson's reagent to make the thioamide **8** resulted in degradation observed by TLC and the formation of multiple products. These difficulties prompted us to start with monomethyl oxalyl chloride.



Scheme 3.3. General initial approach to make the target thioamide coupling partner starting with oxalyl chloride.

The advantage of starting with the monomethyl oxalyl chloride to make ester **2** instead of making **7** from **5** was that the triethylammonium chloride salt, which forms a precipitate over the course of the reaction was simply eliminated by filtration prior to work-up, and residual salt partitions into the aqueous phase, while the desired ester **3** partitions into the organic phase enabling its convenient isolation. Additionally, the ester could be readily monitored by TLC using either ultraviolet (UV) light or $KMnO_4$ stain.

When amide **2** was reacted with Lawesson's reagent to generate compound **3**, it was important to reflux the reaction under an inert atmosphere, while closely monitoring the reaction. No conversion of starting material to product was observed at room temperature, in accord with the literature that indicates heating is required to activate Lawesson's reagent by fragmentation into reactive monomers (Scheme 3.4).[310] Carbonyl functional groups can then react with a dithiophospine ylide monomer to form a 4-membered thiaoxaphosphetane intermediate, which collapses to the resulting thiocarbonyl and oxothiophosphine. The driving force of this collapse is the formation of the thermodynamically favored P=O bond, similar to the Wittig reaction.[307, 310, 311] ^{1}H and ^{13}C NMR analysis of the isolated products revealed that thioamide **3** and not the thioester was formed based on the change in chemical shifts of the amide proton and carbon. This outcome is consistent with literature observations that esters are the least reactive functional groups with Lawesson's reagent[307] and amides can be smoothly converted to thioamides in their presence.

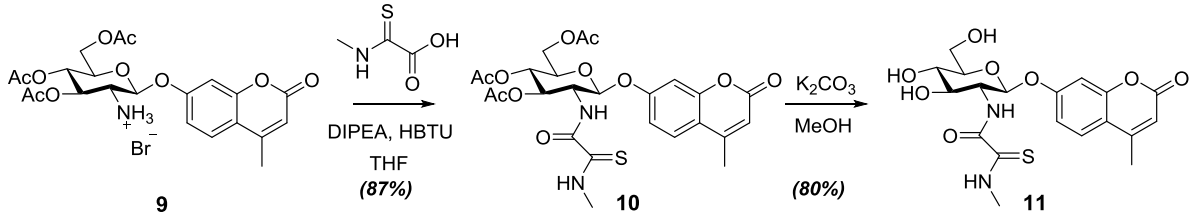


Scheme 3.4. The reaction of Lawesson's reagent to form 2-(methylamino)-2-thioxo-acetic acid (4).
The dissociation of Lawesson's reagent into reactive monomeric dithiophosphine ylides generates an electrophile for attack of a carbonyl nucleophile. The formation of 4-membered thiaoxaphosphetane intermediate follows, and this species collapses to form the thioamide. [307]

To obtain the desired acid **4** from ester **3**, ester hydrolysis was performed using 1N LiOH in 1,4-dioxane. For detailed procedures refer to methods as described in Section 3.7.1.

3.6.3. Synthesis of the 4-Methylumbelliferyl *2*-deoxy-2-[(2-methylamino)-2-thioacetamido]-β-D-glucopyranoside substrate

The fluorogenic substrate bearing a 4-methylumbelliferyl leaving group, 2-deoxy-2-[(2-methylamino)-2-thioxo-acetamido]-β-D-glucopyranoside **11**, was made by coupling 4-methylumbelliferyl 2-amino-2-deoxy-β-D-glucopyranoside **9**, (provided by Evan Perley Robertson), with **4** using standard 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) peptide coupling conditions.[312-316] De-protection using K_2CO_3 in methanol afforded target glycoside **11**. The overall synthetic scheme for the synthesis of substrate **11** is shown in Scheme 3.5. During the coupling to form compound **10**, it was important to react acid **4** and HBTU to pre-form the activated benzyltriazol-1-yl ester intermediate which serves as the electrophile for amine **9**. Diisopropyl ethyl amine (DIPEA) was added at least 30 minutes after the addition of amine **9** to ensure that the basic conditions required for the reaction to proceed to completion are maintained. The desired product **11** was purified by column chromatography, followed by HPLC purification to ensure that there was no free 4MU present that could complicate photophysical characterization of this substrate. The synthesis and purification conditions are described in Methods Section 3.7.2.

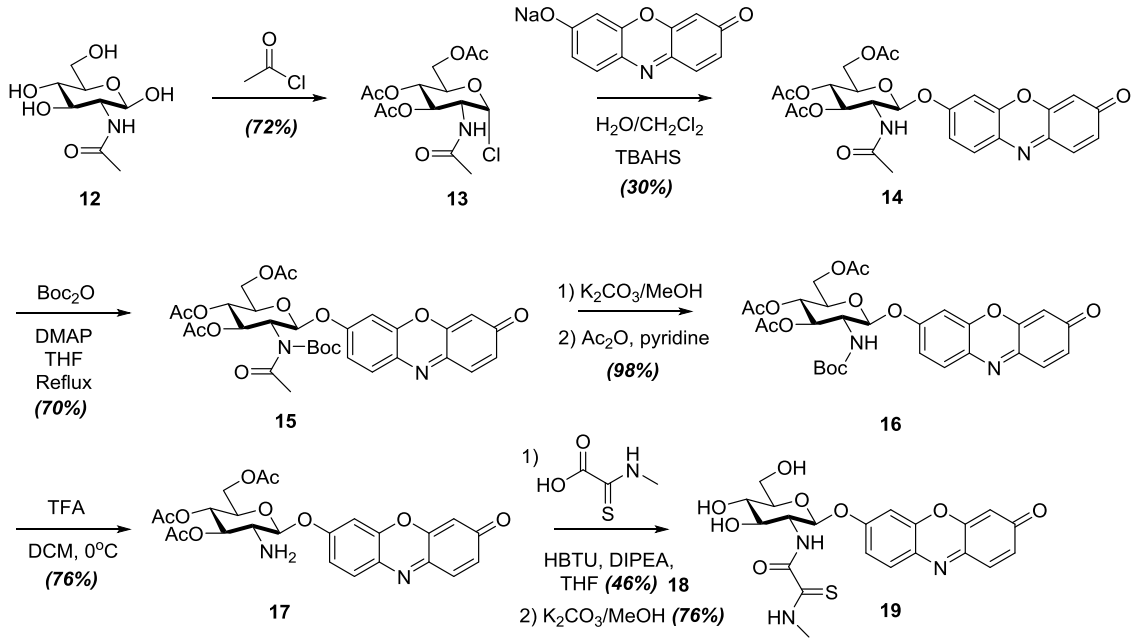


Scheme 3.5. Synthesis of fluorogenic substrate containing 4-MU as a fluorophore at the anomeric position.

3.6.4. Synthesis of the Resorufin 2-deoxy-2-(2-methylamino-2-thioxo-acetamido)-β-D-glucopyranoside substrate

To synthesize the resorufin fluorogenic PET-quenched substrate, the resorufin glycoside **17** was prepared in six steps and coupled to acid **4**, followed by the same coupling and deprotection sequence as described for glycoside **9** (Scheme 3.6). The starting material for making **19** was *N*-acetylglucosamine **12** which was per-*O*-acetylated

and then converted to the glycosyl chloride1,3,4,6-*tetra*-*O*-acetyl-2-acetamido-2-deoxy-α-D-glucopyranosyl chloride (**13**), using literature procedures.[317] Donor **13** was then reacted with the sodium salt of resorufin in a biphasic reaction of DCM and water using tetrabutylammonium hydrogen sulfate (TBAHS) as a phase-transfer catalyst. The phase transfer catalyst forms an organic soluble complex with the phenolate allowing it to react with the donor sugar. Vigorous stirring of the reaction mixture is required to ensure that the reactants come in contact. Extraction using EtOAc as the organic phase and extensive washing with saturated $NaHCO_3$, followed by washes of water, enabled isolation of the glycoside without contaminating resorufin. A series of protecting group manipulations was next required to obtain the 2-amino glycoside. Protection of the amide with a *tert*-butyloxy carbonyl (Boc) protecting group installed using di-*tert*-butyl dicarbonate (Boc-anhydride) enabled global deacetylation using K_2CO_3/MeOH. Subsequent *O*-acetylation using acetic anhydride in pyridine afforded *N*-Boc protected **16**. Treatment of compound **16** with TFA using standard deprotection conditions yielded amino glycoside **17**. The use of **17** for coupling was found to best be done as soon as possible in order to avoid decomposition of **17**. The coupling of **17** to **4** was performed using the same HBTU peptide coupling conditions with as used for glycoside **9**. De-*O*-acetylation of **18** to provide **19** was performed in the same way as for compound **10**. The final product was purified by column chromatography, followed by HPLC purification to ensure that there was no free resorufin present that might complicate subsequent photophysical characterization of this substrate. Detailed descriptions of the synthetic procedures for substrate **19** can be found in Methods Section 3.7.3.



Scheme 3.6. Synthesis of the resorufin fluorogenic substrate 19.

3.6.5. Photophysical properties of substrate 11

Following its preparation, substrate **11** was studied to establish some basic photophysical properties, and compared to the natural fluorogenic substrate, 4MU-GlcNAc (Figure 3.11) and free 4MU. The excitation and emission bands (Figure 3.11 **A** and **B**) for all 3 compounds were obtained by fluorescence spectroscopy to enable optimization of quenching and kinetic studies. The excitation maximum for all 3 compounds was 320 nm and the emission maxima were 375 nm for 4MU-GlcNAc and substrate **11**, and 460 nm for 4MU. Although the excitation and emission scans were evaluated at the same concentration (500 µM) of substrate **11** and 4MU-GlcNAc, there is a drastic difference in their fluorescence, indicating likely quenching of fluorescence in compound **11** in comparison to 4MU-GlcNAc.

The quenching efficiency of **11** was then evaluated at a range of concentrations by comparing fluorescence to that of 4MU-GlcNAc at the same concentrations. The overall % quenching was determined by taking the average of the quenching efficiencies between **11** and 4MU-GlcNAc at 5 different concentrations, with 2-fold dilutions starting from 500 µM to 31.25 µM (Figure 3.11, **C**) The average quenching efficiency of **11** was

determined to be 92% and no concentration dependence of quenching was observed, since this value hardly varied across all concentrations studied.

Typically, the wavelengths which are used for enzymatic assays with 4MU substrates such as 4MU-GlcNAc or 4MU-Glu are 355-365λ_{ex}/440-450λ_{em} for both screening assays[84, 179, 318-320] and cell-based assays[318, 321]. When considering potential applications of this substrate, it is important to note that liberation of the fluorescent phenolate anion is being detected, which is why the typical emission wavelength that is used with 4MU-substrates is 445 nm, the emission maximum for free 4MU. Since the excitation maximum of 4MU overlaps with that of substrate **11** and 4MU-GlcNAc at 320 nm, the quenching efficiency of substrate **11** was evaluated with an excitation wavelength of 320 nm and emission of 445 nm (Figure 3.11, C), which we observed to be over 90%. Accordingly, even though the emission spectra of substrate **11** and 4MU overlap at 445 nm, the background fluorescence of the intact substrate **11** is almost eliminated since the quenching efficiency at 445 nm is very high with an average quenching efficiency of 92%.

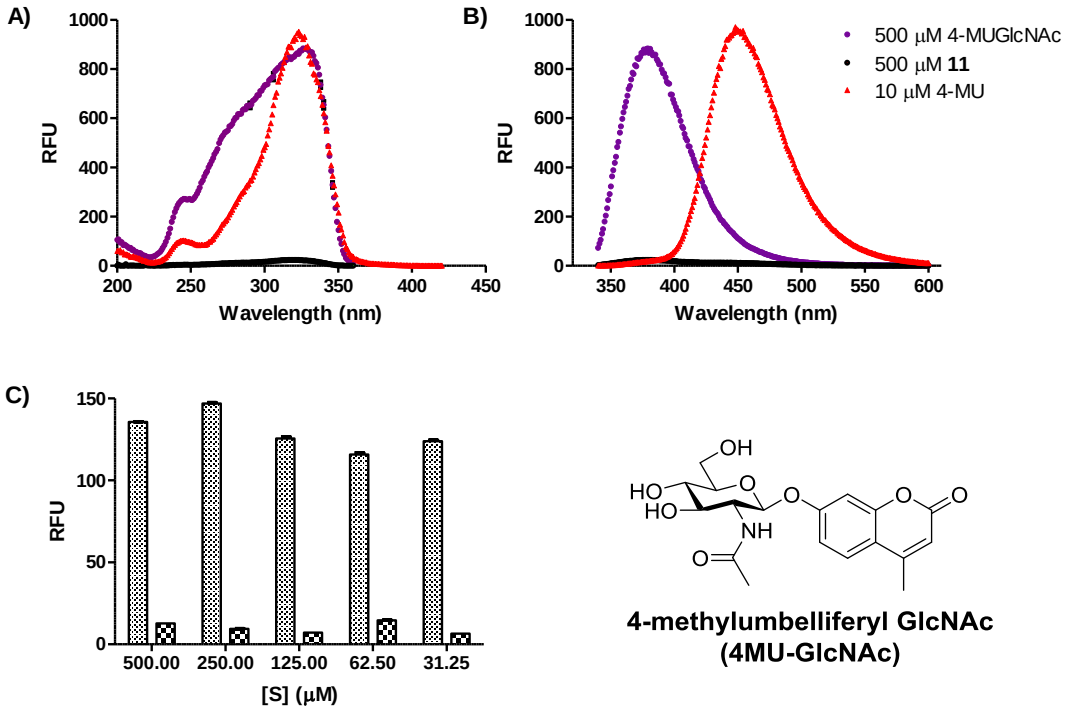


Figure 3.11. The photophysical properties and quenching efficiency of substrate 11.
A) The excitation spectra of 500 µM 4MU-GlcNAc, substrate **11** and 10 µM 4MU.The excitation maximum of all three compounds is 320 nm when the emission is measured at 445 nm. **B)** The emission spectra of 500 µM 4MU-GlcNAc, substrate **11** and 10 µM 4MU. The emission maximum for 4MU-GlcNAc and substrate **11** are at 375 nm, and 460 nm for 4-methylumbelliferone when excited at 320 nm. **C)** The quenching efficiency of substrate **11** was evaluated by comparison to fluorescence of 4MU-GlcNAc at a range of concentrations using an excitation wavelength of 320 nm and an emission wavelength of 445 nm. The average quenching efficiency is 92%, in 10% DMSO/H_2O.

3.6.6. Basic photophysical properties of substrate 19

Fluorescence spectra and quenching efficiency were also evaluated for substrate **19** in the same manner as was done for substrate **11**. The fluorescence spectra were evaluated in comparison to free resorufin and resorufin β-D-glucopyranoside (Res-β-Glu, Figure 3.12), a commercially available GH substrate bearing a resorufin leaving group. The absorption maxima (λ_{max}) measured by monitoring fluorescence emission at 600 nm for **19** and Res-β-Glu are 470 nm, and 570 nm for free resorufin by monitoring emission

at 640 nm (Figure 3.13, A). The λ_{max} values for commercially available Res-β-Glu and resorufin are in accord with literature reports.[177] The emission maximum determined for both substrate **19** and Res-β-Glu was 570 nm, while for resorufin the emission maximum was 590 nm, also in accord with the literature values.[177, 318] Both the absorption and emission spectra show differences in the intensities between substrate **19** and Res-β-Glu, indicating that fluorescence is being quenched, but not nearly as much as compared to substrate **11** and 4MU-GlcNAc (Figure 3.11 A and B). Since the absorption and emission spectra of **19** overlap with Res-β-Glu, this was determined to be an appropriate reference compound to use for comparison to substrate **19** in quenching efficiency assays.

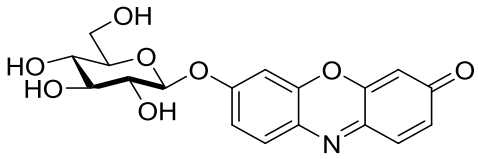


Figure 3.12. The structure of Resorufin β-D-glucopyranoside, Res-β-Glu.

The quenching efficiency of **19** was determined by measuring fluorescence at both 590 nm and 610 nm at a range of concentrations in a solution of 10% DMSO/H_2O and comparing these values to measurements made for Res-β-Glu at the same concentrations. It was predicted based on the studies mentioned by Goldberg *et al* that the quenching efficiency would be higher than that reported for free resorufin in buffer containing 50 mM thioacetamide [291] (Table 3.1). Measurements were made for both substrate **19** and Res-β-Glu with excitation at 470 nm or 570 nm, and emission at 590 nm or 610 nm. The quenching efficiencies varied, depending on the excitation and emission wavelengths. The highest quenching efficiency was observed at an excitation wavelength of 470 nm (λ_{max} of **19**), with a quenching efficiency of 53% and 54% for emission wavelengths of 590 and 610 nm, respectively. This was a surprising result, considering that with substrate **11**, there was consistently over 90% quenching across all substrate concentrations. Precautions were taken during HPLC purification to ensure there was no free resorufin in the commercial Res-β-Glu or substrate **19**. Since the retention times for the substrate and free fluorophore were 1.5 to 2 minutes apart (see Methods section 3.7.2.), care was taken not to collect the fractions before and after elution of resorufin. Additionally, if there was free resorufin in the samples, there would

have been an evident absorption peak observed at 570 nm when the absorption scan was performed for both substrate **19** and Res-β-Glu, but this was not the case.

Typically, in assays involving resorufin substrates such as Res-β-Glu, which is often used to study GBA1, the excitation and emission wavelengths used in enzymatic and cellular assays are 560-570$_{ex}$/590-610$_{em}$, since the liberation of the resorufin phenolate is being monitored as a detector of enzymatic activity.[176, 302, 322] As mentioned in Chapter 1, Section 1.4, the pK_a of resorufin (5.8) makes it an attractive fluorophore to use at lower pH and physiological pH, since the abundance of the phenolate is higher than the phenol ionization state, which means that more fluorescence can be detected when compared to 4MU substrates (pK_a of 7.8). However, inconsistent quenching was observed which ranges from 40-55% depending on the excitation/emission wavelengths used in the quenching assay (Figure 3.13). Although the highest quenching efficiency was observed when exciting at the maximum of 470 nm for both emissions of 590 and 610 nm (53 and 54%, respectively), if these wavelengths were used in a cellular assay there would still be a high fluorescent background detected from the intact substrate **19**. When substrate **19** and Res-β-Glu are excited at 570 nm, the excitation maximum of free resorufin, the quenching is 10% lower for both emissions at 590 nm and 610 nm. The maximum emission of Res-β-Glu and substrate **19** overlap at 570 nm and are close to the free resorufin emission maximum at 590 nm. In terms of applying this to a cellular assay, this could be problematic for result interpretation because there would be some quenching by substrate **19** of the fluorescence of free resorufin, since the emission spectra overlap, which could decrease the amount of signal that is being detected. Additionally, it would not be possible to determine whether the fluorescence detected comes from the intact substrate **19** fluorescence background or the phenolate fluorophore since there is only ~50% quenching at an emission of either 570 nm or 610 nm.

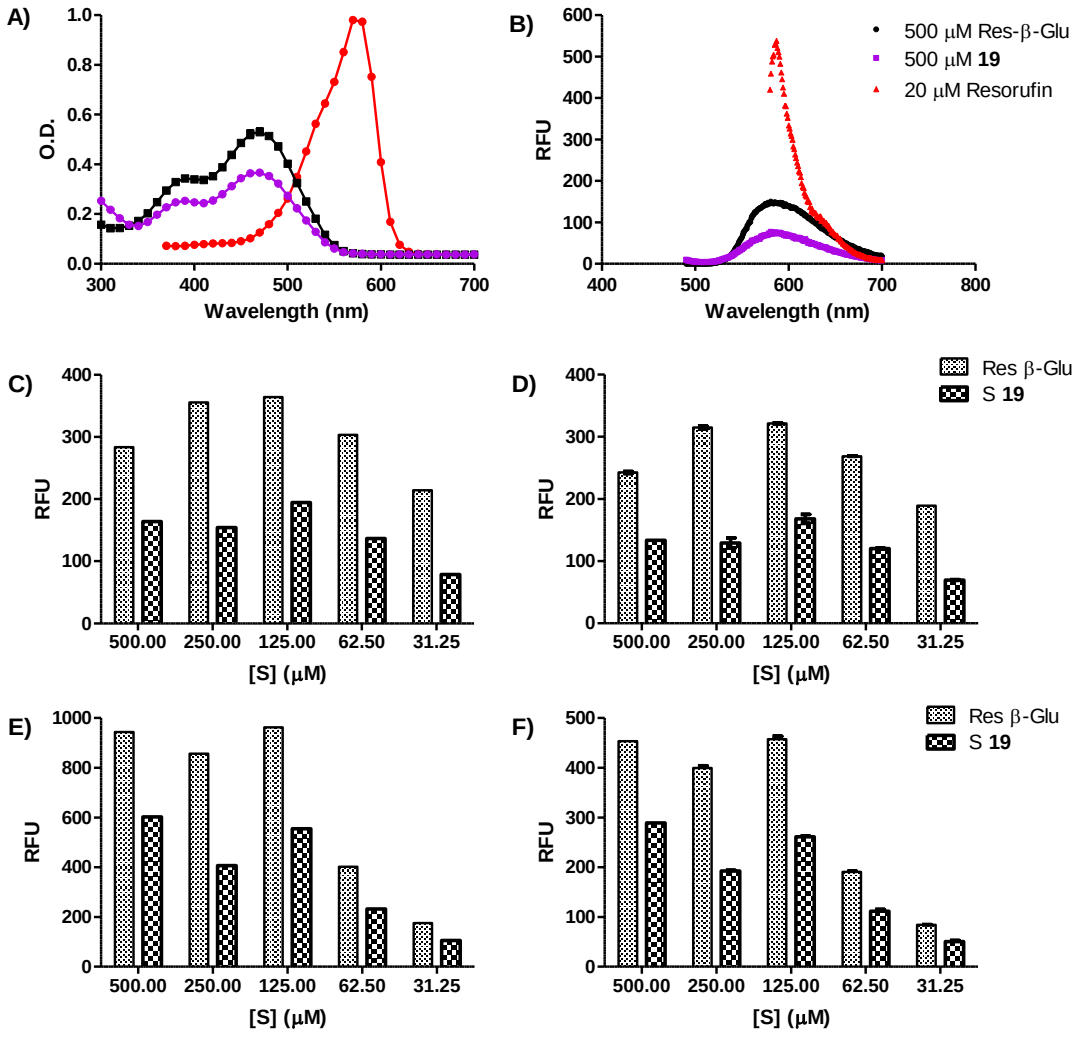


Figure 3.13. Summary of photophysical properties and quenching efficiency for substrate 19.

A) The absorption spectrum of 500 µM Res-β-Glu, substrate **19** (Em: 600 nm) and 20 µM resorufin (Em: 640 nm).The absorption maxima of Res-β-Glu and substrate **19** are 470 nm, and 570 nm for resorufin. **B)** Emission spectrum of 500 µM Res-β-Glu, substrate **19** (Ex: 470 nm) and 20 µM resorufin (Ex: 580 nm). Emission maxima for Res-β-Glu and substrate **19** are 570 nm and 590 nm for resorufin. **C)** Ex: 470 nm, Em: 590 nm, quenching = 53%; **D)** Ex: 470 nm, Em: 610 nm, quenching = 54%; E) Ex: 570 nm, Em: 590 nm, quenching = 43%; F) Ex: 570 nm, Em: 610 nm, quenching = 42% in 10% DMSO.

3.6.7. Kinetic properties of substrates 11 and 19

The kinetic parameters governing the hOGA catalyzed hydrolysis including K_m, k_{cat} and k_{cat}/K_m, were determined for both substrates **11** and **19** to assess their suitability as substrates for hOGA and their potential for their use in cellular assays. For substrate **11**, it was observed that the k_{cat}/K_m (1.49 x 10^{-5} µmol min^{-1} mg^{-1} $µM^{-1}$) was very low compared to the k_{cat}/K_m for 4MU-GlcNAc (9.77 x 10^{-1} µmol min^{-1} mg^{-1} $µM^{-1}$). The K_m values between **11** and 4MU-GlcNAc were comparable (435 and 548 µM, respectively) (Figure 3.14, A, Table 3.2). Although the quenching efficiency of **11** was high, the turnover rate being low limits the potential for this substrate to be used in cell-based assays because there might not be enough turnover by hOGA to be detected. Furthermore, the emission wavelength is not suitable for cell based studies in any event, since it is far into the visible region of the electromagnetic spectrum and the autofluorescence detected would be problematic.

In contrast, substrate **19** not only had a low K_m value of 8 µM, but the k_{cat}/K_m was comparable 4MU-GlcNAc (3.09 x 10^{-2} µmol min^{-1} mg^{-1} $µM^{-1}$), showing only a 30-fold decrease (Table 3.2). Although substrate **19** had a lower quenching efficiency than substrate **11**, it might still be a better contender than substrate **11** for cell-based studies since the K_m value is low and k_{cat}/K_m is comparable (Figure 3.14 B, Table 3.2). The second order rate constant is the important kinetic parameter in cases where there are competing substrates, such as in the cellular environment. Accordingly, less material would be needed to permeate into the cell to see enzymatic turnover. It was not entirely surprising that the better substrate contained resorufin as the fluorophore, since resorufin is a better leaving group than 4MU due to the more extensive conjugation across three ring systems as opposed to the two rings of 4MU. As mentioned previously, the pK_a of resorufin is 5.8, whereas for 4-MU it is 7.8. According to the Henderson-Hasselbalch equation, the dominant ionization state of resorufin at physiologial pH would be the phenolate, compared to 4MU which would be mainly the neutral phenol (Figure 3.15). As mentioned previously, this difference in pK_a is important for use in cellular systems because it is the phenolates that are the highly fluorescent species.

With regard to the physicochemical properties of these molecules, it is clear that their solubility is lower than that of 4MU-GlcNAc. The Michaelis Menten curve for **11** is

presented up until the highest concentration of substrate (500 µM) which did not precipitate out of solution during the 20 minute assay at 37 ^{o}C when using PBS buffer containing 6.5% DMSO. Substrate **19** was less soluble than substrate **11**, with the substrate precipitating from PBS containing 6.5% DMSO after 20 minutes at 37 ^{o}C at concentrations as low as 200 µM. However, because the K_m of the substrate is so low, full saturation kinetics can be determined.

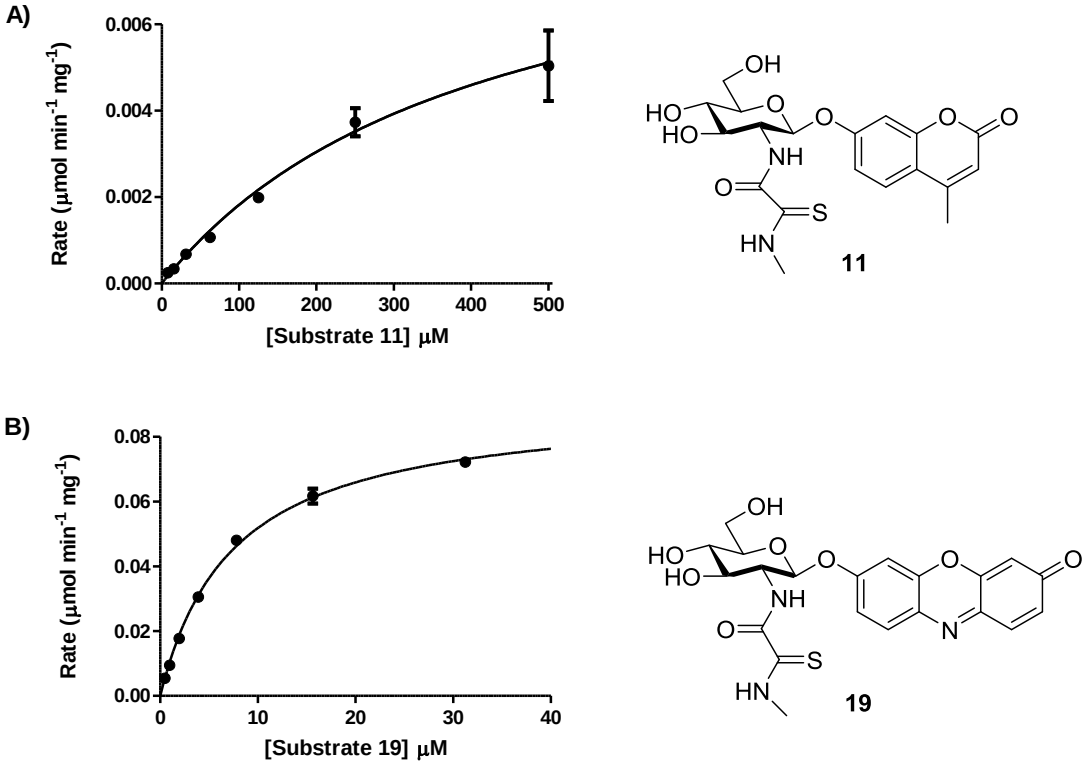


Figure 3.14. Summary of kinetic parameters for substrates 11 and 19.
A) The Michealis-Menten curve for substrate **11**. Substrate concentrations range from 500 µM to 8 µM with 2-fold dilutions throughout. The [hOGA] in this assay was 2.4 µM. **B)** The Michaelis-Menten curve for substrate **19.** Substrate concentrations range from 31.25 µM to 0.5 µM with 2-fold dilutions throughout. The [hOGA] used in this assay was 367 nM.

Table 3.2. Summary of kinetic parameters for quenched substrates.

Substrate	K_m (μM)	k_{cat} (μmol min^{-1} mg^{-1})	k_{cat}/K_m (μmol min^{-1} mg^{-1} μM^{-1})
4MU-GlcNAc	548 ± 74	[a]5.35 x 10^{2}	[b]9.77 x 10^{-1}
11	435 ± 173	[a]6.48 x 10^{-3}	[b]1.49 x 10^{-5}
19	8 ± 0.5	2.47 x 10^{-1}	3.09 x10^{-2}

[a]Values were estimated by nonlinear regression of the Michaelis-Menten data.
[b]Calculated by determination of the 2nd order rate constant in the linear portion of the Michaelis-Menten curve.

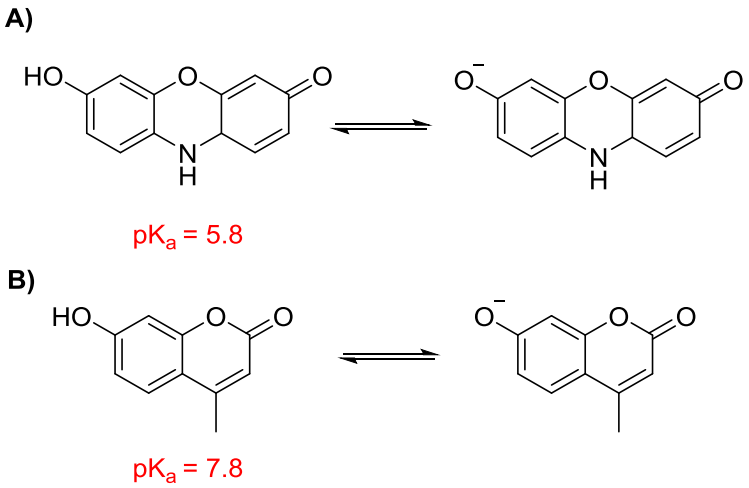


Figure 3.15. The difference between the pK_a for resorufin and 4MU.
At pH 7.4, resorufin exists mainly as the phenolate (A), while 4MU is mainly in its neutral state (B).

Lastly, substrates **11** and **19** were evaluated for turnover with human Hexosaminidase A and B from placenta. No turnover was observed at substrate concentrations ranging from 500 μM to 8 μM when using an enzyme concentration of 15 nM, indicating that these substrates are selective for hOGA over the human lysosomal hexosaminidases. This observation is in line with our predictions based on what was

observed with the amino-thiazoline inhibitor family (Chapter 2), where selectivity for hOGA over hHexB is observed on increasing the size of 2' position of the thiazoline.

3.6.8. Future directions

In future studies, substrate **19** will be tested for use in cells. As discussed above, problems which may arise are high background fluorescence from the intact substrate since quenching only ranges between 40-50% at the emission wavelength or not enough fluorescent signal being detected due to the overlap in the emission spectra of substrate **19** and free resorufin. Additionally, the solubility of substrate **19** could be problematic, depending on the concentration that would be used in cellular assays, since precipitation was observed to occur at concentrations as low as 200 µM after a 20 minute kinetic assay in buffer at 37^{o}C.

A possible way to overcome challenges which may be presented with resorufin is to consider using different phenolate fluorophores whose quenching efficiency with thioamides may be higher and which could be turned over efficiently by hOGA when conjugated to a sugar at the anomeric position. Commercial dyes such as fluorescein (λ_{ex} 494 nm/λ_{em} 521 nm) and a substituted fluorescein derivative with a high quantum yield described by Zhang *et al*[323] (Figure 3.16, A and B) could be a reasonable starting point, since PET quenching efficiencies for fluorescein derivatives were already observed to be around 45% in buffer containing 50 mM thioacetamide (Table 3.1)[291]. Additionally, the pK_a of the fluorescein phenolate is 6.35,[324] which is a value falling directly in the middle of free 4MU and resorufin. This could be beneficial because at physiological pH, the liberated fluorescein, like resorufin, would be mainly in its phenolate fluorescent form. Accordingly, it can be predicted that the cleavage by hOGA of leaving groups containing the fluorescein phenolate would occur more readily than 4MU substrates and with comparable efficiency as with resorufin substrates.

Interestingly, there has been precedence for the development of a fluorogenic substrate, fluorescein di(*N*-acetyl-β-D-glucosaminide (FDGlcNAc, Figure 3.16, C) for human hOGA which consists of fluorescein as the fluorophore conjugated to two *O*-GlcNAc moities.[325] This substrate was shown to have a higher affinity and catalytic

efficiency (K_m = 85 µM, k_{cat}/K_m = 1.6 x 10^4 $M^{-1}s^{-1}$ than one of the conventional substrates, *para*-nitrophenyl 2-acetamido-2-deoxy-β-D-glucopyranoside (*p*NP-GlcNAc, K_m = 1.1 mM, k_{cat}/K_m = 3.0 x 10^3 $M^{-1}s^{-1}$), and a higher affinity than 4MU-GlcNAc (K_m = 0.430 mM). Hydrolysis by OGA of FDGlcNAc forms fluorescein mono(*N*-acetyl-β-D-glucosaminidine, FMGlcNAc), which tautomerizes and generates fluorescence in its quinoid form (Figure 3.16, C). It was observed that FMGlcNAc can be further cleaved by OGA, if the enzyme concentration and reaction time is further increased.[325]

Overall, literature precedence showing that fluorescein conjugated to GlcNAc can serve as a good fluorogenic substrate for hOGA[325] and that thioamides are able to efficiently quench free fluorescein-derived fluorophores in buffer containing 50 mM thioacetamide as the source of thioamide[291], as well as the pK_a of fluorescein falling within a range that would make its phenolate form dominant at physiological pH are all strong reasons which support the proposed direction which this project could take.

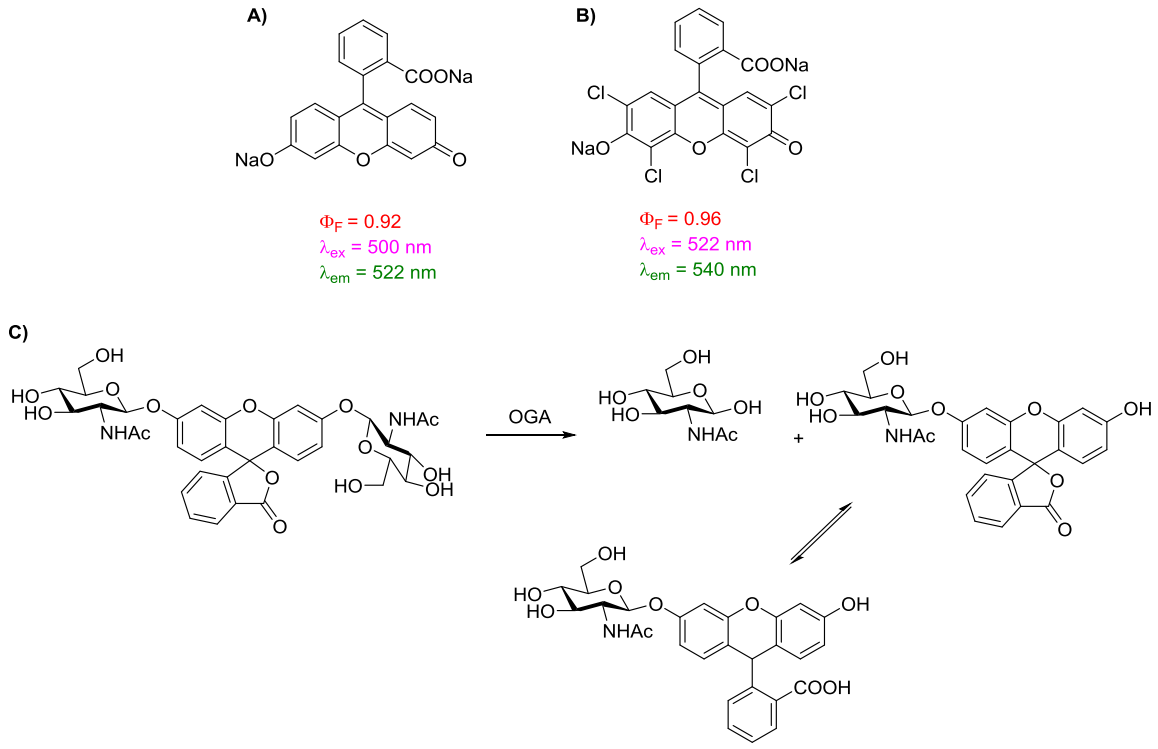


Figure 3.16. Possible future directions using fluorescein as a fluorophore. The structure and photophysical parameters of fluorescein **(A)** and its chlorinated derivative **(B)** which could both be potential fluorophores for future quenched substrates. **C)** An example of a fluorogenic substrate for hOGA, fluorescein di(*N*-acetyl-β-D-glucosaminide (FDGlcNAc, K_m = 85 µM, k_{cat}/K_m = 1.6 x 10^4 $M^{-1}s^{-1}$, λ_{ex} = 485 nm, λ_{em} = 535 nm).[325]

3.7. Methods

3.7.1. Preparation of thioamide acid (4)

Methyl-2-(methylamino)-2-oxoacetate (2): The method used was modified from Mecinovic *et al*.[306] To a round-bottom flask, methylamine hydrochloride (3.6 g, 53.3 mmol) and triethylamine (16.3 mL, 117.3 mmol) were added and dissolved in anhydrous DCM (180 mL) under argon at 0^0C, after which monomethyl oxalyl chloride (4.5 mL, 48.5 mmol) was added drop-wise to the mixture. The reaction mixture was stirred for 20 minutes at 0^oC and gradually warmed up to room temperature over one hour. The resulting yellow solid was then filtered and the filtrate was diluted with EtOAc (15 mL) and washed once with 10 mL of H_2O. The organic layer was dried over $MgSO_4$ and the

solvent evaporated under reduced pressure to afford the product as a yellow solid. Purification was performed using flash column silica gel gradient chromatography using DCM, to 1:1 Et_2O:DCM as the mobile phase (1.3 g, 25% yield). ^1H NMR (400 MHz, $CDCl_3$) δ 7.17 (bs, 1H), 3.91 (s, 3H), 2.95 (d, J = 5.2 Hz, 3H); ^{13}C NMR (101 MHz, CDCl3) δ 161.14 , 156.96 , 53.60 , 26.49. HRMS (m/z): [M+H]+ calculated for $C_4H_8NO_3$: 118.0459; found 118.0500; [M+Na]+ calculated for $C_4H_7NNaO_3$: 140.0324; found 140.0322.

2-(methylamino)-2-thioxo-acetate (3): Methyl-2-(methylamino)-2-oxoacetate (100 mg, 0.85 mmol) was dissolved in toluene. Lawesson's reagent was then added (190 mg, 0.47 mmol), and the reaction was heated to reflux for 30 minutes. The solvent was evaporated from the reaction mixture and purification was performed using column silica gel column gradient flash chromatography using DCM, to 5% Et_2O/DCM as the mobile phase to afford the pure product as a yellow solid (77 mg, 68% yield). ^1H NMR (400 MHz, $CDCl_3$) δ 9.05 (bs, 1H), 3.95 (s, 3H), 3.27 (d, *J* = 5.3 Hz, 3H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 183.47, 159.81, 54.53, 32.73. HRMS (m/z): [M+H]+ calculated for $C_4H_8NO_2S$: 134.0276; found 134.0271.

2-(methylamino)-2-thioxo-acetic acid (4): 2-(methylamino)-2-thioxo-acetate (300 mg, 2.3 mmol) was dissolved in 1,4 dioxane (25 mL), followed by the drop-wise addition of 1M LiOH (3.8 mL, 3.8 mmol). The reaction mixture was stirred for 1.5 hours at room temperature. After the reaction was judged complete by TLC analysis, the solvent was removed under reduced pressure and the product purified by silica gel column chromatography using 18% MeOH/DCM as the mobile phase to afford the pure yellow solid product. (216 mg, 80% yield). ^1H NMR (400 MHz, CD_3OD) δ 10.08 (bs, 1H), 3.13 (s, 3H); ^{13}C NMR (101 MHz, CD_3OD) δ 195.91, 166.22, 31.19. HRMS (m/z): [M-H]- calculated for $C_3H_4NO_2S$: 117.9968; found 117.9968.

3.7.2. Preparation of 4-Methylumbelliferyl *2*-deoxy-2-(2-methylamino-2-thioxo-acetamido)-β-D-glucopyranoside (11)

4-Methylumbelliferyl 3,4,6-*tri*-*O*-acetyl-*2*-deoxy-2-(2-methylamino-2-thioxo-acetamido)-β-D-glucopyranoside (10): 2-(methylamino)-2-thioxo-acetic acid (16 mg, 0.11 mmol) and 3-[Bis(dimethylamino)methyliumyl]-3*H*-benzotriazol-1-oxide

hexafluorophosphate (HBTU, 133 mg, 0.35 mmol) were dissolved in anhydrous THF (2 mL) under argon at 0°C. The resulting mixture was stirred for 3 hours. 4-methylumbelliferyl 3,4,6-*tri*-*O*-acetyl-2-amino-2-deoxy-β-D-glucopyranoside (50 mg, 0.10 mmol) was added to the reaction mixture and the combined solution was stirred for a further hour. N,N-Diisopropylethylamine (DIPEA, 0.06 mL, 0.35 mmol) was then added drop-wise at 0°C and the reaction mixture was stirred for 16 hours. Water (10 mL) and EtOAc (10 mL) were added to the reaction mixture, and the organic layer was collected, isolated, dried over $MgSO_4$ and the solvent removed under reduced pressure. The final product was isolated using silica gel flash column chromatography using 1:1 Et_2O:DCM as the mobile phase to isolate the product as a light yellow solid (48 mg, 86% yield). ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.47 (d, *J* = 5.5 Hz, 1H), 8.51 (d, *J* = 8.7 Hz, 1H), 7.45 (d, *J* = 8.8 Hz, 1H), 6.94 – 6.84 (m, 2H), 6.15 (d, *J* = 1.2 Hz, 1H), 5.63 – 5.53 (m, 2H), 5.19 – 5.12 (m, 1H), 4.31 (dd, *J* = 12.3, 5.8 Hz, 1H), 4.23 – 4.09 (m, 2H), 4.01 (ddd, *J* = 9.7, 5.8, 2.6 Hz, 1H), 3.18 (d, *J* = 5.3 Hz, 3H), 2.36 (d, *J* = 1.3 Hz, 3H), 2.11 (s, 3H), 2.07 (s, 3H), 2.04 (s, 3H). ^{13}C NMR (151 MHz, $CDCl_3$) δ 185.98, 170.74, 170.28, 169.62, 160.90, 159.33, 158.99, 154.91, 152.26, 125.78, 115.65, 114.11, 113.32, 104.43, 97.82, 72.64, 70.76, 68.44, 62.13, 55.74, 32.96, 20.87, 20.80, 20.73, 18.81. HRMS (m/z): [M+H]+ calculated for $C_{25}H_{29}N_2O_{11}S$: 565.1447; found 565.1496; [M+Na]+ calculated for $C_{25}H_{28}N_2NaO_{11}S$: 587.1312; found 587.1312.

4-Methylumbelliferyl *2*-deoxy-2-(2-methylamino-2-thioxo-acetamido)-β-D-glucopyranoside (11): 10 (20 mg, 0.035 mmol) was dissolved in anhydrous methanol (1.5 mL). A catalytic amount of dry K_2CO_3 (1.0 mg) was added to the reaction, and the resulting mixture was stirred for 3 hours. The solvent was evaporated under reduced pressure and the crude product was isolated using silica gel flash column chromatography using 3% MeOH/DCM as the mobile phase (11.6 mg, 77% yield). HPLC purification conditions: A 20 mM solution of compound in DMSO was manually injected into a semi-prep C^{18} column and eluted using a solvent gradient system of 40 to 55% MeOH in H_2O over 30 minutes at a flow rate of 2 mL/min. The elution was monitored by measuring absorbances at 254 nm, 320 and 365 nm and the retention times of **11** and free 4MU were 15.2 and 16.5 min, respectively. ^{1}H NMR (400 MHz, Methanol-d_4) δ 7.69 (d, *J* = 9.5 Hz, 1H), 7.05 – 6.99 (m, 2H), 6.22 (d, *J* = 1.3 Hz, 1H), 5.44 (d, *J* = 8.3 Hz, 1H), 4.05 – 3.92 (m, 2H), 3.79 (ddd, *J* = 20.2, 11.2, 7.2 Hz, 2H), 3.62 – 3.55 (m, 1H),

3.48 (dd, *J* = 9.8, 8.6 Hz, 1H), 3.17 (s, 3H), 2.46 (d, *J* = 1.2 Hz, 3H). ^{13}C NMR (151 MHz, Methanol-d_4) δ 186.61, 161.38, 159.91, 159.69, 154.12, 153.53, 125.42, 114.36, 113.26, 111.09, 103.21, 98.15, 76.56, 73.25, 69.99, 60.60, 56.57, 31.09, 16.73. HRMS (m/z): [M+H]+ calculated for $C_{19}H_{23}N_2O_8S$: 439.1130; found 439.1170; [M+Na]+ calculated for $C_{19}H_{22}N_2NaO_8S$: 461.0995; found 461.0992; [M+K]+ calculated for $C_{19}H_{22}N_2KO_8S$: 477.0734; found 477.0727.

3.7.3. Preparation of Resorufin 2-deoxy-2-[(2-methylamino)-2-thioxo-acetamido]-β-D-glucopyranoside (19)

2-acetamido-3,4,6-*tri*-*O*-acetyl-2-deoxy-α-D-glucopyranosyl chloride (12): N-Acetylglucosamine (10 g, 0.045 mol) was dissolved in acetyl chloride (70 mL, 0.9 mol) and the mixture was stirred for two days. The reaction mixture was then diluted with chloroform to 500 mL and washed 3 times with cold H_2O and then with saturated $NaHCO_3$. The organic layer was collected, dried over $MgSO_4$ and the solvent was removed under reduced pressure. The compound was isolated as a sticky solid with no further purification, and was used directly in the next step (12.5 g, 72% yield).

Resorufin 2-acetamido-3,4,6-*tri*-*O*-acetyl-2-deoxy-β-D-glucopyranoside (13): **12** (2 g, 5.5 mmol) was combined with resorufin sodium salt (2.57 g, 11 mmol) in a round-bottom flask under argon. The solids were dissolved in DCM (35 mL) and H_2O (35 mL) and *tetra*-butyl ammonium hydrogen sulfate (TBAHS, 1.86 g, 5.5 mmol) was added to the mixture. The reaction was vigorously stirred overnight at 700 rpm. Once the reaction was judged to be completed, the mixture was diluted with EtOAC and washed with saturated $NaHCO_3$ three times, then several times with H_2O until the resorufin salt was entirely removed from the organic layer. The organic layer was isolated, dried and the solvent removed under reduced pressure. Purification of the desired compound was performed using gradient silica gel flash column chromatography with 1:1 EtOAc:DCM, followed by EtOAc to elute the product out as a dark red solid (340 mg, 12% yield). ^{1}H NMR (600 MHz, $CDCl_3$) δ 7.77 (d, J = 8.8 Hz, 1H), 7.50 – 7.45 (m, 1H), 7.06 (dd, J = 8.8, 2.6 Hz, 1H), 7.01 (d, J = 2.6 Hz, 1H), 6.90 (dd, J = 9.8, 2.0 Hz, 1H), 6.35 (d, J = 2.0 Hz, 1H), 5.68 (d, J = 8.4 Hz, 1H), 5.54 – 5.49 (m, 2H), 5.20 (t, J = 9.6 Hz, 1H), 4.33 (dd, J = 12.3, 6.0 Hz, 1H), 4.27 (dd, J = 12.2, 2.6 Hz, 1H), 4.19 – 4.14 (m, 1H), 4.02 (ddd, J =

9.0, 6.0, 2.6 Hz, 1H), 2.17 (s, 3H), 2.14 (s, 3H), 2.13 (s, 3H), 2.04 (s, 3H). ^{13}C NMR (151 MHz, $CDCl_3$) δ 170.80, 170.48, 170.45, 169.42, 160.09, 146.83, 134.82, 134.69, 131.59, 115.06, 106.95, 103.42, 98.21, 91.83, 72.52, 71.62, 70.85, 68.33, 68.15, 67.85, 62.15, 54.89, 52.17, 23.40, 20.69, 20.66. HRMS (m/z): [M+Na]+ calculated for $C_{26}H_{26}N_2NaO_{11}$: 565.1434; calculated 565.1451.

Resorufin 2-[acetamido (*tert*-butyl-carbamoyl)]-3,4,6-*tri*-*O*-acetyl-2-deoxy-β-D-glucopyranoside (14): 13 (70 mg, 0.12 mmol) was dissolved in anhydrous THF (3 mL) under an atmosphere of argon. Dimethylamino pyridine (DMAP, 3 mg, 0.02 mmol) and di-*tert*-butyl dicarbonate (36 mg, 0.16 mmol) were then and the reaction mixture was heated at reflux for 3 hours. The solvent was then removed under reduced pressure and the product purified using silica gel column chromatography by gradient elution using DCM to 20% EtOAc:DCM as the mobile phase to afford the pure product as an orange powder (65 mg, 78% yield). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.73 (d, J = 8.4 Hz, 1H), 7.44 (d, J = 9.8 Hz, 1H), 7.01 – 6.92 (m, 2H), 6.86 (dd, J = 9.8, 2.0 Hz, 1H), 6.31 (d, J = 2.0 Hz, 1H), 6.12 (d, J = 8.2 Hz, 1H), 5.84 (d, J = 38.4 Hz, 2H), 5.19 (t, J = 9.5 Hz, 1H), 4.57 (s, 1H), 4.38 – 4.17 (m, 2H), 4.05 (d, J = 11.9 Hz, 1H), 2.19 (s, 3H), 2.15 (s, 3H), 2.09 (s, 3H), 2.04 (s, 3H), 1.64 – 1.51 (m, 9H). ^{13}C NMR (151 MHz, $CDCl_3$) δ 186.26, 170.48, 170.11, 169.62, 169.42, 160.00, 153.08, 149.56, 146.82, 145.13, 134.79, 134.69, 131.55, 129.52, 114.96, 106.93, 103.52, 97.08, 72.16, 70.19, 69.11, 62.22, 61.05, 56.35, 28.01, 20.72, 20.69, 20.68, 20.55. HRMS (m/z): [M+Na]+ calculated for $C_{31}H_{34}N_2NaO_{13}$: 665.1992; calculated 665.1974.

Resorufin 3,4,6-*tri*-*O*-acetyl-2-deoxy-2-[(*tert*-butyl)-carbamoyl]-β-D-glucopyranoside (16): 14 (245 mg, 0.39 mmol) was dissolved in anhydrous MeOH (10 mL) under argon. K_2CO_3 (12 mg) was then added and the reaction mixture was stirred at room temperature for one hour. After the reaction was judged complete, the solvent was removed under reduced pressure and the crude material was dissolved into dry pyridine (20 mL) under an atmosphere of argon. Acetic anhydride (0.4 mL, 10 eq) was added drop-wise to the mixture and stirred for another three hours. The pyridine was then removed under reduced pressure and the mixture diluted with EtOAc. The organic layer was collected and dried over $MgSO_4$ and the solvent removed under reduced pressure. The product was isolated using silica gel column chromatography using 1:1 Et_2O:DCM

as the mobile phase to afford the product as an orange powder (223 mg, 98% yield). ^{1}H NMR (601 MHz, $CDCl_3$) δ 7.64 (d, J = 8.8 Hz, 1H), 7.37 – 7.32 (m, 1H), 6.94 (dd, J = 8.8, 2.6 Hz, 1H), 6.90 (d, J = 2.6 Hz, 1H), 6.80 – 6.75 (m, 1H), 6.24 – 6.19 (m, 1H), 5.38 (s, 1H), 5.34 – 5.29 (m, 1H), 5.05 (t, J = 9.6 Hz, 1H), 4.75 (s, 1H), 4.21 (dd, J = 12.3, 6.1 Hz, 1H), 4.14 (dd, J = 12.3, 2.5 Hz, 1H), 3.90 – 3.84 (m, 1H), 3.76 (q, J = 9.0 Hz, 1H), 2.04 (s, 3H), 2.03 (s, 3H), 2.00 (s, 3H), 1.36 (s, 9H). ^{13}C NMR (151 MHz, $CDCl_3$) δ 186.30, 170.47, 170.43, 169.50, 160.31, 154.97, 149.57, 146.76, 145.13, 134.80, 134.65, 131.56, 129.50, 115.13, 106.92, 103.50, 98.62, 72.47, 71.61, 68.54, 62.20, 55.70, 29.73, 28.25, 20.69, 20.68, 20.66. HRMS (m/z): [M+Na]+ calculated for $C_{29}H_{32}N_2NaO_{12}$: 623.1853; found 623.1865.

Resorufin 3,4,6-*tri*-*O*-acetyl-2-amino-2-deoxy-β-D-glucopyranoside (17): 16 (200 mg, 0.34 mmol) was dissolved in cold, anhydrous DCM (10 mL) and trifluoroacetic acid (TFA, 2 mL, 5.2 mmol) was added drop-wise over a period of 20 minutes. The reaction mixture was stirred and the progress of the reaction was carefully monitored for 3 hours, at which it was judged complete and the product was isolated by using silica gel chromatography using a gradient elution system with 1:1 EtOAc:DCM, EtOAc, then 10% MeOH/DCM to afford the desired product as a dark red powder (126 mg, 74% yield). δ ^{1}H NMR (500 MHz, $(CD_3)_2CO$) δ 7.74 (t, *J* = 7.0 Hz, 1H), 7.46 (d, *J* = 9.8 Hz, 1H), 7.25 – 7.05 (m, 2H), 6.77 (dd, *J* = 9.8, 2.0 Hz, 1H), 6.18 (d, *J* = 2.0 Hz, 1H), 5.80 (d, *J* = 7.8 Hz, 1H), 5.54 (t, *J* = 9.6 Hz, 1H), 5.13 (t, *J* = 9.5 Hz, 1H), 4.41 – 4.15 (m, 4H), 3.49 – 3.40 (m, 1H), 2.07 – 2.05 (m, 9H); ^{13}C NMR (126 MHz, $(CD_3)_2CO$) δ 169.84 , 169.48 , 169.06 , 160.23 , 149.79 , 146.69 , 145.15 , 134.95 , 134.15 , 131.40 , 129.27 , 114.73 , 106.03 , 103.15 , 98.92 , 73.04 , 72.19 , 72.06 , 68.58 , 63.61 , 61.99 , 60.96 , 19.80 , 19.70; HRMS (m/z): [M+H]+ calculated for $C_{24}H_{25}N_2O_{10}$: 501.1464; found 501.1506.

Resorufin 3,4,6-*tri*-*O*-acetyl-2-deoxy-2-[(2-methylamino)-2-thioxo-acetamido)]- β-D-glucopyranoside (18): 4 (23 mg, 0.26 mmol) and HBTU, (264 mg, 0.70 mmol) were dissolved in anhydrous cold THF (8 mL) under an atmosphere of argon and the reaction mixture was stirred at 0^{o}C for 3 hours. Resorufin glucosamine salt (100 mg, 0.20 mmol) was added to the reaction mixture as a 63 mg/mL solution in dry THF under an atmosphere of argon via syringe and the resulting solution was stirred for one more hour at 0^{o}C. N,N-Diisopropylethylamine (DIPEA, 0.12 mL, 0.70 mmol) was then

added drop-wise and the reaction mixture was stirred for 16 hours. The crude reaction collected, dried with $MgSO_4$, and the solvent reduced under reduced pressure. The desired product was isolated using silica gel column chromatography 1:1 EtOAc:DCM as the mobile phase to afford the title compound as a dark red powder (60 mg, 42% yield). ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.41 (d, *J* = 5.8 Hz, 1H), 8.53 (d, *J* = 8.5 Hz, 1H), 7.73 (d, *J* = 8.7 Hz, 1H), 7.44 (d, *J* = 9.8 Hz, 1H), 7.04 – 6.96 (m, 2H), 6.87 (dd, *J* = 9.8, 2.0 Hz, 1H), 6.32 (d, *J* = 2.0 Hz, 1H), 5.68 – 5.56 (m, 2H), 5.18 (t, *J* = 9.3 Hz, 1H), 4.38 – 4.24 (m, 2H), 4.17 – 3.99 (m, 2H), 3.25 (d, *J* = 5.3 Hz, 3H), 2.14 (s, 3H), 2.13 (s, 3H), 2.08 (s, 3H). ^{13}C NMR (151 MHz, $CDCl_3$) δ 186.28 , 185.82 , 170.44 , 170.10 , 169.48 , 159.79 , 158.82 , 149.53 , 146.94, 145.10, 134.80, 134.75, 131.59, 129.62, 115.09, 106.99, 103.63, 97.53, 72.62, 70.37, 68.30, 62.12, 55.61, 32.87, 29.73, 20.68, 20.60; HRMS (m/z): [M+H]+ calculated for $C_{27}H_{28}N_3O_{11}S$: 602.1400; found 602.1434; [M+Na]+ calculated for $C_{27}H_{27}N_3NaO_{11}S$: 624.1264; found 624.1251.

Resorufin 2-deoxy-2-[(2-methylamino)-2-thioxo-acetamido]-β-D-glucopyranoside (19): 18 (15 mg, 0.025 mmol) was dissolved in anhydrous MeOH (2 mL). K_2CO_3 was added (1 mg) and the reaction mixture was stirred for 1.5 hours. The solvent was removed under reduced pressure and the desired material was isolated using silica gel column chromatography using a gradient of 5% MeOH/DCM to 18% MeOH/DCM as the mobile phase to yield the title compound as an orange powder (9 mg, 76% yield). HPLC purification conditions: A 20 mM solution of compound **19** in DMSO was manually injected into a semi-prep C^{18} column and eluted using a gradient of 50 to 60% MeOH in H_2O over 30 minutes at a flow rate of 2 mL/min. The elution of products was monitored by UV absorption at 254 nm, 470 and 610 nm and the retention times of **19** and free resorufin were 11.3 and 13.1 min, respectively. ^{1}H NMR (600 MHz, DMSO-d_6) δ 10.85 (t, *J* = 5.1 Hz, 1H), 8.67 (d, *J* = 9.8 Hz, 1H), 7.79 (d, *J* = 8.8 Hz, 1H), 7.55 (d, *J* = 9.8 Hz, 1H), 7.06 (d, *J* = 2.5 Hz, 1H), 7.00 (dd, *J* = 8.9, 2.5 Hz, 1H), 6.81 (dd, *J* = 9.8, 2.0 Hz, 1H), 6.29 (d, *J* = 2.0 Hz, 1H), 5.54 (d, *J* = 8.4 Hz, 1H), 5.23 (dd, *J* = 13.5, 5.6 Hz, 2H), 4.68 (t, *J* = 5.8 Hz, 1H), 3.90 (td, *J* = 10.1, 8.4 Hz, 1H), 3.77 (ddd, *J* = 11.9, 5.4, 2.1 Hz, 1H), 3.67 (ddd, *J* = 10.2, 8.6, 5.9 Hz, 1H), 3.51 (dd, *J* = 11.9, 6.1 Hz, 1H), 3.44 (ddd, *J* = 9.9, 6.0, 2.0 Hz, 1H), 3.28 – 3.22 (m, 1H), 3.04 (d, *J* = 4.9 Hz, 3H). ^{13}C NMR (151 MHz, DMSO-d_6) δ 186.99 , 185.38 , 160.62 , 159.57 , 149.62 , 146.03, 144.89, 134.98, 133.99, 131.34, 128.69, 114.78, 105.80, 102.80, 98.39, 77.86, 74.02,

70.28, 61.55, 56.40, 32.74. HRMS (m/z): [M+H]+ calculated for $C_{21}H_{21}N_3O_8S$: 476.1083; found 476.1127; [M+Na]+ calculated for $C_{21}H_{21}N_3NaO_8S$: 498.0947; found 498.0942; [M+K]+ calculated for $C_{21}H_{21}N_3KO_8S$: 514.0686; found 514.0668.

3.7.4. Analysis of substrates 11 and 19

Assessment of excitation and emission spectra: Excitation and emission spectra for substrates **11** and **19** were obtained using a Varian Cary Eclipse Fluorescence spectrophotometer and a 40 µL quartz cuvette. A slitwidth of 10 nm (**11**) and 5 nm (**19**) was used for both excitation and emission in conjuncition with a Xenon flash lamp with sensitivities set at 520 to 1000. Each substrate was evaluated in a solution of 10% DMSO in milli-Q H_2O in duplicate, at 5 different concentrations ranging from 500 µM to 32.5 µM in volume of 40 µL and temperature of 25^oC. The free fluorophores were evaluated at 10 µM (4-methylumbelliferone) or 20 µM (resorufin). A blank was obtained for 10% DMSO/H_2O at each sensitivity between readings and subtracted from the fluorescence spectra obtained.

Absorption scans for substrate 19, Res-β-Glu and resorufin (Figure 3.14 A): The maximum absorbance (λ_{max}) of the three compounds were assessed using a BioTek (Synergy^{TM} H4) plate reader in 96-well plates (Thermoscientific, FloroNunc, lot # 139825) in triplicate with the slitwidth set to 10 nm. The compounds were dissolved in 10% DMSO/phosphate buffered saline (PBS) buffer adjusted to pH 7.4 at a concentration of 500 µM for substrate **19** and Res-β-Glu and 20 µM for free resorufin.

Evaluation of quenching efficiency: Fluorescence quenching was evaluated using a Varian Cary Eclipse Fluorescence spectrophotometer with a 40 µL the slitwidth set to 5 or 10 nm in combination with a Xenon flash lamp set to sensitivities ranging from 520 to 1000. 10 mM solutions of each substrate in DMSO were diluted to 5 mM, 2.5 mM, 1.25 mM, 0.625 mM and 0.3125 mM in DMSO. Each resulting solution was then diluted 10-fold into milli-Q H_2O to result in a final concentration of 10% DMSO, immediately before taking a measurement, in order to ensure there was no substrate precipitation prior to measurement. The excitation and emission measurements were made 10 times for each compound at each concentration, and in duplicate. The fluorescence measurements were made at the same sensitivity setting for **19** and Res-β-Glu and the

measurement of a blank (10% DMSO in H_2O) was subtracted from each measurement at that sensitivity. The ratio of the fluorescence for **19/**Res-β-Glu was determined for each concentration to obtain the quenching efficiency. The overall quenching efficiency was defined as an average of quenching efficiencies determined at each concentration.

Michealis-Menten kinetic assessment: A 20 mM solution of **11, 19** or the appropriate control compound were prepared in 100% DMSO and diluted 10-fold to 10 mM on a serial dilution plate into 100% DMSO with a final volume of 200 µL. 2-fold serial dilutions were then generated from 5 mM to 0.005 mM in a final volume of 100 µL per well in 100% DMSO. On a black 96-well plate (Thermoscientific, FloroNunc, lot # 139825) 10 µL of each serial dilution was added into 90 µL of pH 7.4 PBS buffer, resulting in a final DMSO concentration of 10%. To the diluted compound in buffer at 10% DMSO was added 50 µL of a stock solution of enzyme that was maintained at 4^{o}C to yield the appropriate final enzyme concentration in a final volume of 150 µL of buffer containing 6.5% DMSO. The substrate hydrolysis was monitored for 20 minutes at 37^{o}C using a continuous fluorescence assay with excitation and emission wavelengths as follows: $365_{ex}/445_{em}$ for substrate **11** and $571_{ex}/610_{em}$ for substrate **19.** The rates in RFU for each substrate concentration were obtained, from the time period of 10 to 20 minutes. The k_{cat}/K_m was determined from the linear, second order region, of the Michaelis Menten plot for substrate **11**, since saturation kinetics were not observed. For substrate **19,** the k_{cat}/K_m was determined from the V_{max} value obtained. RFU values were converted to concentration of product by generating two standard curves using each free fluorophore that covered a range of concentrations from 0 to 0.3 µM.

References

1. Varki, A. *et al.*, *Essentials of Glycobiology, 2nd edition*, The Cons. of Glycobiol. Editors, 2009. Cold Spring Harbor Laboratory Press: Cold Spring Harbor (NY).

2. Gloster, T.M., *Development of inhibitors as research tools for carbohydrate-processing enzymes.* Biochem Soc Trans, 2012. **40**(5): p. 913-28.

3. Zachara, N.E. and G.W. Hart, *O-GlcNAc a sensor of cellular state: the role of nucleocytoplasmic glycosylation in modulating cellular function in response to nutrition and stress.* Biochim Biophys Acta, 2004. **1673**(1-2): p. 13-28.

4. Olszewski, N.E. *et al.*, *O-GlcNAc protein modification in plants: Evolution and function.* Biochim Biophys Acta, 2010. **1800**(2): p. 49-56.

5. Kelly, W.G., Dahmus, M.E. and Hart, G.W., *RNA Polymerase II is a Glycoprotein.* J Biol Chem, 1993. **266**(14): p. 10416-10424.

6. Kornfeld, R. and Kornfeld, S., *Assembly of Asparagine-Linked Oligosaccharides.* Ann Rev Biochem, 1985. **54**: p. 631-634.

7. Wells, L., Vosseller, K., and G.W. Hart, *Glycosylation of Nucleocytoplasmic Proteins: Signal Transduction and O-GlcNAc.* Science, 2001. **291**(5512): p. 2376-2378.

8. Love, D.C. and Hanover, J.A., *The Hexosamine Signaling Pathway: Deciphering the "O-GlcNAc Code".* Sci Signal, 2005. **312**(13): p. 1-14.

9. Hanover, J.A., *Glycan-dependent signaling: O-linked N-acetylglucosamine.* FASEB J., 2001. **15**: p. 1865-1876.

10. Comer, F.I. *et al*., *Characterization of a mouse monoclonal antibody specific for O-linked N-acetylglucosamine.* Anal Biochem, 2001. **293**(2): p. 169-77.

11. Shane, C.A. and Hart, G.W., *Dynamic Cytoskeletal Glycosylation and Neurodegenerative Disease.* Trends in Glycosci and Glycotech, 1999. **11**(62): p. 355-370.

12. Ngoh, G.A. *et al.*, *O-GlcNAc signaling attenuates ER stress-induced cardiomyocyte death.* Am J Physiol Heart Circ Physiol, 2009. **297**: p. H1711-H1719.

13. Sohn, K.C., *et al*., *OGT functions as a catalytic chaperone under heat stress response: a unique defense role of OGT in hyperthermia.* Biochem Biophys Res Commun, 2004. **322**(3): p. 1045-51.

14. Zachara, N.E., *et al., Dynamic O-GlcNAc modification of nucleocytoplasmic proteins in response to stress. A survival response of mammalian cells.* J Biol Chem, 2004. **279**(29): p. 30133-42.

15. Jang, I. *et al.*, *O-GlcNAcylation of eIF2α regulates the phospho-eIF2α-mediated ER stress response.* Biochim et Biophys Acta (BBA) - Mol Cell Res, 2015. **1853**(8): p. 1860-1869.

16. Arnold, S.C. *et al.*, *The Microtubule-associated Protein Tau is Extensively Modified with O-linked N-acetylglucosamine.* J Biol Chem, 1996. **271**(46): p. 28741-28744.

17. Chou, C.-F., Smith, J.A. and Omary, B.M., *Characterization and Dynamics of O-Linked Glycosylation of Human Cytokeratin 8 and 18.* J Biol Chem, 1992. **267**(6): p. 3901-3906.

18. Ramirez-Correa, G.A. *et al.*, *O-linked GlcNAc modification of cardiac myofilament proteins: a novel regulator of myocardial contractile function.* Circ Res, 2008. **103**(12): p. 1354-8.

19. Komura, K., Ise, H. and Akaike, T., *Dynamic behaviors of vimentin induced by interaction with GlcNAc molecules.* Glycobiol, 2012. **22**(12): p. 1741-59.

20. Lima, V.V. *et al.*, *O-GlcNAcylation: a novel pathway contributing to the effects of endothelin in the vasculature.* Am J Physiol Regul Integr Comp Physiol, 2011. **300**: p. R236-R250.

21. Bullen, J.W. *et al*., *Cross-talk between two essential nutrient-sensitive enzymes: O-GlcNAc transferase (OGT) and AMP-activated protein kinase (AMPK).* J Biol Chem, 2014. **289**(15): p. 10592-606.

22. Shi, J. *et al*., *Diverse regulation of AKT and GSK-3beta by O-GlcNAcylation in various types of cells.* FEBS Lett, 2012. **586**(16): p. 2443-50.

23. Lima, V.V. *et al., Vascular O-GlcNAcylation augments reactivity to constrictor stimuli by prolonging phosphorylated levels of the myosin light chain.* Brazilian J of Med and Biol Res, 2014. **47**(10): p. 826-833.

24. Li, B. and J.J. Kohler, *Glycosylation of the nuclear pore.* Traffic, 2014. **15**(4): p. 347-61.

25. Mizuguchi-Hata, C. *et al*., *Quantitative regulation of nuclear pore complex proteins by O-GlcNAcylation.* Biochim Biophys Acta, 2013. **1833**(12): p. 2682-9.

26. Miller, M.W. *et al.*, *Phosphorylation and Glycosylation of Nucleoporins.* Arch of Biochem and Biophys, 1999. **367**(1): p. 51-60.

27. Hart, C.R.T. and Hart, G.W., *Topography and Polypeptide Distribution of Terminal N-Acetylglucosamine Residues on the Surfaces of Intact Lymphocytes.* J. Biol. Chem. , 1984. **259**: p. 3308-3317.

28. Kreppel, L.K., *Dynamic Glycosylation of Nuclear and Cytosolic Proteins. CLONING AND CHARACTERIZATION OF A UNIQUE O-GlcNAc TRANSFERASE WITH MULTIPLE TETRATRICOPEPTIDE REPEATS.* J of Biol Chem, 1997. **272**(14): p. 9308-9315.

29. Lubas, W.A. *et al.*, *O-Linked GlcNAc Transferase Is a Conserved Nucleocytoplasmic Protein Containing Tetratricopeptide Repeats.* J of Biol Chem, 1997. **272**(14): p. 9316-9324.

30. Hart, G.W., Housley, M.P and Slawson, C., *Cycling of O-linked β-N-acetylglucosamine on nucleocytoplasmic proteins.* Nature, 2007. **446**(7139): p. 1017-22.

31. Dong, D. L-Y. and Hart, G.W., *Purification and Characterization of an O-GlcNAc Selective N-Acetyl-β-D-glucosaminidase from Rat Spleen Cytosol.* J Biol Chem, 1994. **269**(30):19321-19330

32. Rempel, B.P. and S.G. Withers, *Covalent inhibitors of glycosidases and their applications in biochemistry and biology.* Glycobiol, 2008. **18**(8): p. 570-86.

33. Banjeree, P.S., Ma, J. and Hart, G.W., *Diabetes-associated dysregulation of O-GlcNAcylation in rat cardiac mitochondria.* Proc Natl Acad Sci U S A, 2015. **112**(19): p. 6050-6055.

34. Tan, E.P. *et al*., *Altering O-linked beta-N-acetylglucosamine cycling disrupts mitochondrial function.* J Biol Chem, 2014. **289**(21): p. 14719-30.

35. Hart, G.W. *et al.*, *Cross talk between O-GlcNAcylation and phosphorylation: roles in signaling, transcription, and chronic disease.* Annu Rev Biochem, 2011. **80**: p. 825-58.

36. Zeidan, Q. *et al*., *O-GlcNAc cycling enzymes associate with the translational machinery and modify core ribosomal proteins.* Mol Biol Cell, 2010. **21**(12): p. 1922-36.

37. Marshall, S., Bacote, V. and Traxinger, R.R., *Discovery of a Metabolic Pathway Mediating Glucose-induced Desensitization of the Glucose Transport System: Role of Hexosamine Biosynthesis in the Induction of Insulin Resistance.* J Biol Chem, 1991. **266**(8): p. 4706-4712.

38. Buse, G.M. *et al.*, *Differential Effects of GLUT1 or GLUT4 Overexpression on Hexosamine Biosynthesis by Muscles of Transgenic Mice.* J Biol Chem, 1996. **271**(38): p. 23197-23202.

39. Robinson, A. *et al., Effects of Diabetes and Hyperglycemia on the Hexosamine Synthesis Pathway in Rat Muscle and Liver.* Diabetes, 1995. **44**: p. 1438-1446.

40. Hawkins, M. *et al*., *Role of the glucosamine pathway in fat-induced insulin resistance.* J Clin Invest, 1997. **99**(9): p. 2173-82.

41. McClain, D.A. and Crook, E.D., *Perspectives in Diabetes: Hexosamines and Insulin Resistance.* Diabetes, 1996. **45**: p. 1003-1009.

42. Hawkins, M. *et al.*, *The Tissue Concentration of UDP-N-acetylglucosamine Modulates the Stimulatory Effect of Insulin on Skeletal Muscle Glucose Uptake.* J Biol Chem, 1997. **272**(8): p. 4889-4895.

43. Watzele, G. and Tanner, W., *Cloning of the Glutamine:Fructose-6-phosphate Amidotransferase Gene from Yeast.* J Biol Chem, 1989. **265**(15): p. 8753-8758.

44. Manning, G., *et al., The Protein Kinase Complement of the Human Genome.* Science, 2002. **298**: p. 1912-1934.

45. Wells, L. *et al., O-GlcNAc transferase is in a functional complex with protein phosphatase 1 catalytic subunits.* J Biol Chem, 2004. **279**(37): p. 38466-70.

46. Hsieh-Wilson, L.C. and Vincent-Lamarre, N., *Dynamic Glycosylation of the Transcription Factor CREB: A Potential Role in Gene Regulation.* J Am Chem Soc, 2003. **125**: p. 6612-6613.

47. Chou, T.-Y., Hart, G.W. and Dang, C.V., *c-Myc is Glycosylated at Threonine 58, a Known Phosphorylation Site and Mutational Hot Spot in Lymphomas.* J Biol Chem, 1995. **270**(32): p. 18961-18965.

48. Kamemura, K., Hayes, B.K., Comer, F.I. and Hart, G.W., *Dynamic Interplay between O-Glycosylation and O-Phosphorylation of Nucleocytoplasmic Proteins.* J Biol Chem, 2002. **277**(21): p. 19229-19235.

49. Francisco, G.D. *et al.*, *Phenyl thiazolyl urea and carbamate derivatives as new inhibitors of bacterial cell-wall biosynthesis.* Bioorg Med Chem Lett, 2004. **14**(1): p. 235-238.

50. Chavan, M. and Lennarz, W., *The molecular basis of coupling of translocation and N-glycosylation.* Trends Biochem Sci, 2006. **31**(1): p. 17-20.

51. Chou, T.-Y., Hart, G.W. and Dang, C.V., *c-Myc is Glycosylated at Threonine 58, a Known Phosphorylation Site and a Mutational Hot Spot in Lymphomas.* J Biol Chem, 1995. **270**(22): p. 18961-18965.

52. Kamemura, K., *et al., Dynamic interplay between O-glycosylation and O-phosphorylation of nucleocytoplasmic proteins: alternative glycosylation/phosphorylation of THR-58, a known mutational hot spot of c-Myc in lymphomas, is regulated by mitogens.* J Biol Chem, 2002. **277**(21): p. 19229-35.

53. Zhu, Y., Shan, X., Yuzwa, S.A. and Vocadlo, D.J., *The Emerging Link between O-GlcNAc and Alzheimer Disease.* J Biol Chem 2014. **289**(50): p. 34472-34481.

54. Yuzwa, S.A. *et al.*, *A potent mechanism-inspired O-GlcNAcase inhibitor that blocks phosphorylation of tau in vivo.* Nat Chem Biol, 2008. **4**(8): p. 483-90.

55. Yuzwa, S.A. and Vocadlo, D.J., *O-GlcNAc and neurodegeneration: biochemical mechanisms and potential roles in Alzheimer's disease and beyond.* Chem Soc Rev, 2014. **43**(19): p. 6839-58.

56. Risso, G., *et al.*, *Akt/PKB: one kinase, many modifications.* Biochem J, 2015. **468**(2): p. 203-14.

57. Whelan, S.A., *et al.*, *Regulation of insulin receptor substrate 1 (IRS-1)/AKT kinase-mediated insulin signaling by O-Linked β-N-acetylglucosamine in 3T3-L1 adipocytes.* J Biol Chem, 2010. **285**(8): p. 5204-11.

58. Vosseller, K. *et al*., *Elevated nucleocytoplasmic glycosylation by O-GlcNAc results in insulin resistance associated with defects in Akt activation in 3T3-L1 adipocytes.* Proc Natl Acad Sci U S A, 2002. **99**(8): p. 5313-8.

59. Ma, Z. and Vosseller, K., *O-GlcNAc in cancer biology.* Amino Acids, 2013. **45**(4): p. 719-33.

60. Champattanachai, V., *et al*., *Proteomic analysis and abrogated expression of O-GlcNAcylated proteins associated with primary breast cancer.* Proteomics, 2013. **13**(14): p. 2088-99.

61. Yi, W. *et al., Phosphofructokinase 1 Glycosylation Regulates Cell Growth and Metabolism.* Science, 2012. **337**: p. 975-979.

62. Jensen, R.V. *et al.*, *Impact of O-GlcNAc on cardioprotection by remote ischaemic preconditioning in non-diabetic and diabetic patients.* Cardiovasc Res, 2013. **97**(2): p. 369-78.

63. Champattanachai, V. *et al., Glucosamine protects neonatal cardiomycocytes from ischemia-reperfusion injury via increased protein-associated O-GlcNAc.* Am J Physiol Cell Physiol, 2006. **292**: p. C178-C187.

64. Vibjerg Jensen, R. *et al.*, *Ischemic preconditioning increases myocardial O-GlcNAc glycosylation.* Scand Cardiovasc J, 2013. **47**(3): p. 168-74.

65. Dassanayaka, S. and S.P. Jones, *O-GlcNAc and the cardiovascular system.* Pharmacol Ther, 2014. **142**(1): p. 62-71.

66. Yu, Q. *et al*., *Effective glycaemic control critically determines insulin cardioprotection against ischaemia/reperfusion injury in anaesthetized dogs.* Cardiovasc Res, 2014. **103**(2): p. 238-47.

67. Jones, S.P. *et al.*, *Cardioprotection by N-acetylglucosamine linkage to cellular proteins.* Circulation, 2008. **117**(9): p. 1172-82.

68. Honda, H. and Ping, P., *Mitochondrial Permeability Transition in Cardiac Cell Injury and Death.* Cardiovasc Drugs and Therapy, 2006. **20**(6): p. 425-432.

69. Zechel, D.L. and Withers, S.G. *Glycosidase Mechanisms: Anatomy of a Finely Tuned Catalyst.* Acc of Chem Res, 2000. **33**: p. 11-18.

70. Wolfenden, R., Lu, X. and Young, G., *Spontaneous Hydrolysis of Glycosides.* J Am Chem Soc, 1998. **120**: p. 6814-6815.

71. White, A. and Rose, D.R., *Mechanism of catalysis by retaining beta-glycosyl hydrolases.* Curr Op in Struct Biol, 1997. **7**: p. 645-651.

72. Davies, G.J. and Henrissat, B., *Structures and mechanisms of glycosyl hydrolases.* Structure, 1995. **3**: p. 853-859.

73. Burns, M.D. and Touster, O. *Purification and Characterization of Glucosidase 11, an Endoplasmic Reticulum Hydrolase Involved in Glycoprotein Biosynthesis.* J Biol Chem, 1982. **257**(17): p. 9991-10000.

74. Yaoi, K. and Mitsuishi, Y. *Purification, characterization, cloning, and expression of a novel xyloglucan-specific glycosidase, oligoxyloglucan reducing end-specific cellobiohydrolase.* J Biol Chem, 2002. **277**(50): p. 48276-81.

75. Lombard, V. *et al.*, *The carbohydrate-active enzymes database (CAZy) in 2013.* Nucleic Acids Res, 2014. **42**(Database issue): p. D490-5.

76. Henrissat, B. and Bairoch, A. *Updating the sequence-based classification of glycosyl hydrolases.* Biochem J Lett, 1996. **316**: p. 695-696.

77. Rye, S.C. and Withers, S.G., *Glycosidase mechanisms.* Curr Op in Chem Biol, 2000. **4**: p. 573-580.

78. Sinnott, M.L., *Catalytic Mechanisms of Enzymic Glycosyl Transfer.* Chem Rev, 1990. **90**: p. 1171-1202.

79. Vasella, A., Davies, G.J. and Bohm, M., *Glycosidase mechanisms.* Curr Op in Chem Biol, 2002. **5**: p. 619-629.

80. Koshland, D.E., *Stereochemistry and the Mechanism of Enzymatic Reactions.* Biol Rev, 1953. **4**: p. 416-436.

81. Vocadlo, D.J. *et al.*, *Catalysis by hen egg-white lysozyme proceeds via a covalent intermediate.* Nature, 2001. **412**: p. 835-838.

82. Wang, Q. *et al., Changing Enzymatic Reaction Mechanisms by Mutagenesis: Conversion of a Retaining Glucosidase to an Inverting Enzyme.* J Am Chem Soc 1994. **116**: p. 11594-11595.

83. Wang, Q. and Withers, S.G.,*Substrate-assisted catalysis in glycosidases.* J Am Chem Soc, 1995. **117**: p. 10137-10138.

84. Macauley, M.S., *et al., O-GlcNAcase uses substrate-assisted catalysis: kinetic analysis and development of highly selective mechanism-inspired inhibitors.* J Biol Chem, 2005. **280**(27): p. 25313-22.

85. Gloster, T.M. and Vocadlo, D.J.V, *Developing inhibitors of glycan processing enzymes as tools for enabling glycobiology.* Nat Chem Biol, 2012. **8**.

86. Fenger, T.H. and Brumer, H., *Synthesis and analysis of specific covalent inhibitors of endo-xyloglucanases.* Chembiochem, 2015. **16**(4): p. 575-83.

87. Meloncelli, P.J. *et al., D-Glucosylated Derivatives of Isofagomine and Noeuromycin and Their Potential as Inhibitors of β-Glycoside Hydrolases.* Aust. J. Chem., 2007. **60**: p. 549-565.

88. Sumida, T., *et al.*, *Gaining insight into the inhibition of glycoside hydrolase family 20 exo-β-N-acetylhexosaminidases using a structural approach.* Org Biomol Chem, 2012. **10**(13): p. 2607-12.

89. Pluvinage, B., *et al.*, *Inhibition of the family 20 glycoside hydrolase catalytic modules in the Streptococcus pneumoniae exo-β-D-N-acetylglucosaminidase, StrH.* Org Biomol Chem, 2013. **11**(45): p. 7907-15.

90. Nerinckx, W., *et al*., *An elaboration on the syn-anti proton donor concept of glycoside hydrolases: electrostatic stabilisation of the transition state as a general strategy.* FEBS Lett, 2005. **579**(2): p. 302-12.

91. Pauling, L., *Molecular Architecture and Biological Reactions.* Biol Sci, 1946. **24**(10): p. 1375-1377.

92. Wolfenden, R., *Transition State Analogues for Enzyme Catalysis.* Nature, 1969. **223**: p. 704-705.

93. Kursula, I. and Wierenga, R.K., *Crystal structure of triosephosphate isomerase complexed with 2-phosphoglycolate at 0.83-A resolution.* J Biol Chem, 2003. **278**(11): p. 9544-51.

94. Lahiri, R., Ansari, A.A., and Vankar Y.D., *Recent developments in design and synthesis of bicyclic azasugars, carbasugars and related molecules as glycosidase inhibitors.* Chem Soc Rev, 2013. **42**(12): p. 5102-18.

95. Davies, G.J. *et al.*, *Mapping the conformational itinerary of beta-glycosidases by X-ray crystallography.* Biochem Soc Trans, 2003. **31**(3): p. 523-527.

96. Vocadlo, D.J. and G.J. Davies, *Mechanistic insights into glycosidase chemistry.* Curr Opin Chem Biol, 2008. **12**(5): p. 539-55.

97. Gloster, T.M. *et.al.*, *Glycosidase Inhibition: An Assessment of the Binding of 18 Putative Transition-State Mimics.* J Am Chem Soc, 2007. **129**: p. 2345-2354.

98. Inouye S, Tsuruoka., T. and Nida, T., *The structure of nojirimycin, a piperidinose sugar antibiotic.* J Antibiot (Tokyo), 1966. **19**(6): p. 288-292.

99. Inouye, S., Tsuruoka, T. Ito., T. and Niida, T., *Structure and synthesis of nojirimycin.* Tetrahedron, 1968. **24**(5): p. 2125-2144.

100. Peyrieras, N. *et al.*, *Effects of the glucosidase inhibitors nojirimycin and deoxynojirimycin on the biosynthesis of membrane and secretory glycoproteins.* The EMBO Journal, 1983. **2**(6): p. 823-832.

101. Glawar, A.F., *et al.*, *Scalable syntheses of both enantiomers of DNJNAc and DGJNAc from glucuronolactone: the effect of N-alkylation on hexosaminidase inhibition.* Chemistry, 2012. **18**(30): p. 9341-59.

102. Gradnig, G. *et al.*, *A novel approach to the 1-deoxynojirimycin system: synthesis from sucrose of 2-acetamido-1,2-dideoxynojirimycin, as well as some 2-N-modified derivatives.* Carb Res, 1996. **287**: p. 49-57.

103. L. Wells, K.V. and Hart, G.W., *Glycosylation of Nucleocytoplasmic Proteins: Signal Transduction and O-GlcNAc.* Science, 2001. **219**: p. 2376-2378.

104. Fleet, G.W. *et al.*, *Synthesis of 2-acetamido-1,5-imino-1,2,5-trideoxy-D-mannitol and of 2-acetamido-1,5-imino-1,2,5-trideoxy-D-glucitol, a potent and specific inhibitor of a number of β-N-acetylglucosaminidases.* Chem Lett, 1986. **15**(7): p. 1051-1054.

105. Best, D., *et al*., *Synthesis of 2-acetamido-1,2-dideoxy-D-galacto-nojirimycin [DGJNAc] from D-glucuronolactone: the first sub-micromolar inhibitor of α-N-acetylgalactosaminidases.* Tet Lett, 2010. **51**(17): p. 2222-2224.

106. de la Fuente, A. *et al.*, *Stereoselective Synthesis of 2-Acetamido-1,2-dideoxyallonojirimycin (DAJNAc), a New Potent Hexosaminidase Inhibitor.* Org Lett, 2013. **15**(14): p. 3638-3641.

107. Butters, T.D., Dwek, R.A., and Platt, F.M., *Imino sugar inhibitors for treating the lysosomal glycosphingolipidoses.* Glycobiol, 2005. **15**(10): p. 43R-52R.

108. Bowman, E.A., *et al*., *Longitudinal changes in cerebellar and subcortical volumes in adult-onset Niemann-Pick disease type C patients treated with miglustat.* J Neurol, 2015. **262**(9):2106-2114.

109. Brand, M., *et al., Results from a 9-year Intensive Safety Surveillance Scheme (IS(3)) in miglustat (Zavesca((R)))-treated patients.* Pharmacoepidemiol Drug Saf, 2015. **24**(3): p. 329-33.

110. Scott, L.J. and Spencer, M.C., *Miglitol: A review of its therapeutic potential in type II diabetes mellitus.* Drugs, 2000. **59**(3): p. 521-549.

111. Jespersen, T.M. *et al., Isofagomine, a potent, new glycosidase inhibitor.* Angew Chem Int Ed, 1994. **33**(17): p. 1778-1779.

112. M. Jespersen, T., *et al.*, *Synthesis of isofagomine, a novel glycosidase inhibitor.* Tetrahedron, 1994. **50**(47): p. 13449-13460.

113. Bols, M., Hazell, G.R. and Thomsen, B, *1-Azafagomine: A hydroxyhexahydropyridazine that potently inhibits enzymatic glycoside cleavage.* Chem: A Eur J, 1997. **3**(6): p. 940-947.

114. Dong, W. *et al.*, *Evaluation of isofagomine and its derivatives as potent glycosidase inhibitors.* Biochemistry, 1996. **35**: p. 2788-2795.

115. Steet, R.A., *et al*., *The iminosugar isofagomine increases the activity of N370S mutant acid β-glucosidase in Gaucher fibroblasts by several mechanisms.* Proc Natl Acad Sci U S A, 2006. **103**(37): p. 13813-8.

116. Dong, W. *et al., Evaluation of Isofagomine and Its Derivatives As Potent Glycosidase Inhibitors.* Biochem J, 1996. **35**: p. 2788-2795.

117. Lopez Lopez, O. and M. Bols, *Anomer-selective glycosidase inhibition by 2-N-alkylated 1-azafagomines.* Chembiochem, 2007. **8**(6): p. 657-61.

118. Mendes, R., *et al.*, *Synthesis and evaluation of α-, β-glucosidase inhibition of 1-N-carboxamide-1-azafagomines and 5-epi-1-azafagomines.* Carb Res, 2014. **395**: p. 52-7.

119. Welter, A., *et al., 2,5-Dihydroxymethyl 3,4-dihydroxypyrrolidine dans les feuilles de Derris elliptica.* Phytochem, 1976. **15**(5): p. 747-749.

120. Card, J.P. and Hitz, D.W., *Synthesis of 2(R), 5(R)-Bis(hydroxymethyl)-3(R)-di-hydroxypyrrolidine. A novel glycosidase inhibitor.* J Org Chem, 1985. **50**(6): p. 891-893.

121. Liang, P.H., *et al.*, *Novel five-membered iminocyclitol derivatives as selective and potent glycosidase inhibitors: new structures for antivirals and osteoarthritis.* Chembiochem, 2006. **7**(1): p. 165-73.

122. Rountree, J.S.S., *et al.*, *Efficient synthesis from d-lyxonolactone of 2-acetamido-1,4-imino-1,2,4-trideoxy-l-arabinitol LABNAc, a potent pyrrolidine inhibitor of hexosaminidases.* Tet Lett, 2007. **48**(24): p. 4287-4291.

123. Elbein, D.A., *Inhibitors of the biosynthesis and processing of N-linked oligosaccharide chains.* Ann. Rev. Biochem., 1987. **56**: p. 497-534.

124. Cutfield, M.S. *et al., The Structure of the Exo-β-(1,3)-Glucanase from Candida albicans in Native and Bound Forms: Relationship between a Pocket and Groove in Family 5 Glycosyl Hydrolases.* J Mol Bio, 1999. **294**(3): p. 771-783.

125. Macauley, M.S., *et al., Inhibition of O-GlcNAcase using a potent and cell-permeable inhibitor does not induce insulin resistance in 3T3-L1 adipocytes.* Chem Biol, 2010. **17**(9): p. 937-48.

126. Elbein, D.A. *et al., Swainsonine: An inhibitor of glycoprotein processing* Proc Natl Acad Sci U S A, 1981. **78**(12): p. 7393-7397.

127. Vasconcelos-Dos-Santos, A., *et al.*, *Biosynthetic Machinery Involved in Aberrant Glycosylation: Promising Targets for Developing of Drugs Against Cancer.* Front Oncol, 2015. **5**: p. 138.

128. Olden, K., *et al.*, *The potential importance of swainsonine in therapy for cancers and immunology.* Pharmacol Ther, 1991. **50**(3): p. 285-290.

129. Usuki, H., *et al.*, *Pochonicine, a polyhydroxylated pyrrolizidine alkaloid from fungus Pochonia suchlasporia var. suchlasporia TAMA 87 as a potent β-N-acetylglucosaminidase inhibitor.* Bioorg Med Chem, 2009. **17**(20): p. 7248-53.

130. Zhu, J.S., *et al.*, *Synthesis of eight stereoisomers of pochonicine: nanomolar inhibition of β-N-acetylhexosaminidases.* J Org Chem, 2013. **78**(20): p. 10298-309.

131. Whitworth, G.E. *et al., Analysis of PUGNac and NAG-thiazoline as Transition State Analogues for Human O-GlcNAcase: Mechanistic and Structural Insights into Inhibitor Selectivity and Transition State Poise.* J Am Chem Soc, 2007. **129**: p. 635-644.

132. Haltiwanger, R.S., Grove, K. and Philipsberg, G.A., *Modulation of O-Linked N Acetylglucosamine Levels on Nuclear and Cytoplasmic Proteins in Vivo Using the Peptide O-GlcNAc-β-N-acetylglucosaminidase Inhibitor O-(2-Acetamido-2-deoxy-D-glucopyranosylidene)amino-N-phenylcarbamate.* J Biol Chem, 1998. **273**(6): p. 3611-3617.

133. Aoyama, T. *et al.*, *The structure of nagstatin, a new inhibitor of N-Acetyl-β-D-Glucosaminidase.* J Antibiot, 1992. **45**(9).

134. Tatsuta, K. and Miura, S., *Total Synthesis of Nagstatin, an N-Acetyl·β-D glucosaminidase Inhibitor.* Tet Lett, 1995. **36**(37): p. 6721-6724.

135. Dorfmueller, C.H. *et al.*, *GlcNAcstatin: a Picomolar, Selective O-GlcNAcase Inhibitor that Modulates Intracellular O-GlcNAcylation Levels.* J Am Chem Soc, 2006. **128**: p. 16484-16485.

136. Dorfmueller, H.C. *et al.*, *GlcNAcstatins are nanomolar inhibitors of human O-GlcNAcase inducing cellular hyper-O-GlcNAcylation.* Biochem J, 2009. **420**(2): p. 221-7.

137. He, Y. *et al., Visualizing the Reaction Coordinate of an O-GlcNAc Hydrolase.* J Am Chem Soc, 2010. **132**: p. 1807-1809.

138. Knapp, S. *et al., NAG-thiazoline, An N-Acetyl-β-hexosaminidase Inhibitor That Implicates Acetamido Participation.* J Am Chem Soc, 1996. **118**: p. 6804-6805.

139. Macauley, M.S., *et al*., *Elevation of global O-GlcNAc levels in 3T3-L1 adipocytes by selective inhibition of O-GlcNAcase does not induce insulin resistance.* J Biol Chem, 2008. **283**(50): p. 34687-95.

140. Macauley, M.S., *et al*., *Elevation of Global O-GlcNAc in rodents using a selective O-GlcNAcase inhibitor does not cause insulin resistance or perturb glucohomeostasis.* Chem Biol, 2010. **17**(9): p. 949-58.

141. Graham, D.L., *et al.*, *Increased O-GlcNAcylation reduces pathological tau without affecting its normal phosphorylation in a mouse model of tauopathy.* Neuropharmacology, 2014. **79**: p. 307-13.

142. Yuzwa, S. *et al.*, *Increasing O-GlcNAc slows neurodegeneration and stabilizes tau against aggregation.* Nat Chem Biol, 2012. **8**: p. 393-399.

143. Yuzwa, S.A. *et al.*, *Pharmacological inhibition of O-GlcNAcase (OGA) prevents cognitive decline and amyloid plaque formation in bigenic tau/APP mutant mice.* Mol Neurodeg, 2014. **9**.

144. Clarke, A.J., Drummelsmith, J. and Yaguchi, M., *Identification of the catalytic nucleophile in the cellulase from Schizophyllum commune and assignment of the enzyme to Family 5, subtype 5 of the glycosidases.* FEBS Lett, 1997. **414**: p. 359-361.

145. Diaz Arribas, J.C. *et al., Differential mechanism-based labeling and unequivocal activity assignment of the two active sites of intestinal lactase/phlorizin hydrolase.* Eur. J. Biochem, 2000. **267**: p. 6996-7005.

146. Hinou, H., Kurogochi, M., and Nishimura, S.I. *Mechanism-Based Inhibitors to Probe Transitional States of Glycoside Hydrolases.* 2006. **415**: p. 202-212.

147. Kallemeijn, W.W., *et al., Chapter 4 - Mechanism-Based Inhibitors of Glycosidases: Design and Applications*, in *Advances in Carbohydrate Chemistry and Biochemistry*, H. Derek, Editor. 2014, Academic Press. p. 297-338.

148. Atsumi, S. *et al.*, *Production, Isolation and Structure Determination of a Novel β-Glucosidase Inhibitor, Cyclophellitol, from Phellinus sp.* J Antibiot, 1990. **43**(1): p. 49-53.

149. Tai, V. *et al.*, *Kinetic Studies on Cyclophellitol Analogues - Mechanism-Based Inhibitors.* Biochem and Biophys Res Comm, 1995. **213**(1): p. 175-180.

150. Withers, S.G. and K. Umezawa, *Cyclophellitol: A naturally occurring mechanism-based inactivator of β-glucosidases.* Biochem and Biophys Res Comm, 1991. **177**(1): p. 532-537.

151. Gloster, T.M., Madsen, R, and Davies, G.J. *Structural basis for cyclophellitol inhibition of a β-glucosidase.* Org Biomol Chem, 2007. **5**(3): p. 444-6.

152. Withers, S.G. *et al.*, *2-Deoxy-2-fluoroglucosides: A Novel Class of Mechanism-Based Glucosidase Inhibitors.* J. Am. Chem. Soc., 1987. **109**: p. 7530-7531.

153. Caron, G. and Withers, S.G., *Conduritol aziridine: A new mechanism-based glucosidase inactivator.* Biochem and Biophys Res Comm, 1989. **163**(1): p. 495-499.

154. Willems, L.I., *et al.*, *From covalent glycosidase inhibitors to activity-based glycosidase probes.* Chem: A Eur J, 2014. **20**(35): p. 10864-72.

155. Chakladar, S., *et al., A mechanism-based inactivator of glycoside hydrolases involving formation of a transient non-classical carbocation.* Nat Commun, 2014. **5**: p. 5590.

156. Grabowski, G.A. *et al., Human Acid β-Glucosidase.* J Biol Chem, 1986. **261**(18): p. 8263-8269.

157. Boot, R.G. *et al.*, *Identification of the non-lysosomal glucosylceramidase as β-glucosidase 2.* J Biol Chem, 2007. **282**(2): p. 1305-12.

158. Stephens, M.C., *et al.*, *Distribution of conduritol B epoxide in the animal model for Gaucher's disease (Gaucher mouse).* Biochim et Biophys Acta (BBA) - General Subjects, 1981. **672**(1): p. 29-32.

159. Atsumi, S., *et al., Inhibition of glucocerebrosidase and induction of neural abnormality by cyclophellitol in mice.* Arch of Biochem and Biophys, 1992. **297**(2): p. 362-367.

160. Farfel-Becker, T., Vitner, E.B. and Futerman, A.H., *Animal models for Gaucher disease research.* Dis Model Mech, 2011. **4**(6): p. 746-52.

161. Adams, B.T. *et al. N-Alkylated aziridines are easily prepared, potent, specific and cell-permable covalent inhibitors of human beta-glucocerebrosidase.* Chem Comm, 2015. **51**: p. 11390-11393.

162. Büchold, C., *et al*., *New cis-Configured Aziridine-2-carboxylates as Aspartic Acid Protease Inhibitors.* ChemMedChem, 2011. **6**(1): p. 141-152.

163. Alcaide, A. and Llebaria, A., *Aziridine ring opening for the synthesis of sphingolipid analogues: inhibitors of sphingolipid-metabolizing enzymes.* J Org Chem, 2014. **79**(7): p. 2993-3029.

164. Miao, S. *et al.*, *Identification of Glu^{340} as the active-site nucleophile in human glucocerebrosidase by use of electrospray tandem mass spectrometry.* J Biol Chem, 1994. **269**(15): p. 10975-10978.

165. Ferrer, M., *et al., A novel α-glucosidase from the acidophilic archaeon Ferroplasma acidiphilum strain Y with high transglycosylation activity and an unusual catalytic nucleophile.* Biochem J, 2005. **391**(Pt 2): p. 269-76.

166. Witte, M.D. *et al., Ultrasensitive in situ visualization of active glucocerebrosidase molecules.* Nat Chem Biol 2010. **6**(12): p. 907-913.

167. Kallemeijn, W.W., *et al.*, *Novel Activity-Based Probes for Broad-Spectrum Profiling of Retaining β-Exoglucosidases In Situ and In Vivo.* Angew Chem Int Ed, 2012. **51**(50): p. 12529-12533.

168. Vocadlo, D.J. and Bertozzi, C.R. *A strategy for functional proteomic analysis of glycosidase activity from cell lysates.* Angew Chem Int Ed Engl, 2004. **43**(40): p. 5338-42.

169. Phenix, C.P., *Imaging of enzyme replacement therapy using PET.* Proc. Natl. Acad. Sci. USA, 2010. **107**(24): p. 10842-10847.

170. Mead, J.A.R. and Williams, R.T. , *Studies in Detoxication: The biosynthesis of the glucuronides of umbelliferone and 4-methylumbelliferone and their use in fluorimetric determination of β-glucuronidase.* Biochem J, 1955. **61**: p. 569-574.

171. King, G.M., *Characterization of beta-glucosidase activity in intertidal marine sediments.* App and Env Microbiol, 1986. **51**(2): p. 373-380.

172. Shulman, M.L., Kulshin, V.A. and Khorlin, A.Y., *A continuous fluorimetric assay for glycosidase activity: Human N-acetyl-β-d-hexosaminidase.* Anal Biochem, 1980. **101**(2): p. 342-348.

173. Tropak, M.B., et al., *High-throughput screening for human lysosomal beta-N-Acetyl hexosaminidase inhibitors acting as pharmacological chaperones.* Chem Biol, 2007. **14**(2): p. 153-64.

174. Perry, C.C., *et al., Use of a coumarin-labeled hexa-arginine peptide as a fluorescent hydroxyl radical probe in a nanoparticulate plasmid DNA condensate.* J Phys Chem B, 2011. **115**(32): p. 9889-97.

175. Ahmed, V., *et al.*, *A fluorogenic substrate for the continuous assaying of aryl sulfatases.* Anal Biochem, 2005. **340**(1): p. 80-8.

176. Bueno, C., *et al.*, *The Excited-State Interaction of Resazurin and Resorufin with Aminesin Aqueous Solutions. Photophysics and Photochemical Reaction¶.* Photochem and Photobiol, 2002. **76**(4): p. 385-390.

177. Ibatullin, F.M., *et al., A real-time fluorogenic assay for the visualization of glycoside hydrolase activity in planta.* Plant Physiol, 2009. **151**(4): p. 1741-50.

178. Coleman, D.J., Studler, M.J. and Naleway, J.J. *A long-wavelength fluorescent substrate for continuous fluorometric determination of cellulase activity: resorufin-β-D-cellobioside.* Anal Biochem, 2007. **371**(2): p. 146-53.

179. Urban, J.D. *et al., Optimization and validation of two miniaturized glucocerebrosidase enzyme assays for high throughput screening.* Comb Chem and High Throughput Screening, 2008. **11**: p. 817-824.

180. Humphries, W.H.IV. and Payne, C.K. *Imaging lysosomal enzyme activity in live cells using self-quenched substrates.* Anal Biochem, 2012. **424**(2): p. 178-83.

181. Yadav, A.K., *et al.*, *Fluorescence-quenched substrates for live cell imaging of human glucocerebrosidase activity.* J Am Chem Soc, 2015. **137**(3): p. 1181-9.

182. Cottaz, S., Brasme, B. and Driquez, H. , *A fluorescence-quenched chitopentaose for the study of endo-chitinases and chitobiosidases.* Eur. J. Biochem, 2000. **267**: p. 5593-5600.

183. Butters, T.D., *Gaucher disease.* Curr Opin Chem Biol, 2007. **11**(4): p. 412-8.

184. Schapira, A.H., *Glucocerebrosidase and Parkinson disease: Recent advances.* Mol Cell Neurosci, 2015. **66**(Pt A): p. 37-42.

185. Deng, H., Xiu, X and Jankovic, J. *Genetic convergence of Parkinson's disease and lysosomal storage disorders.* Mol Neurobiol, 2015. **51**(3): p. 1554-68.

186. Khidekel, N., *et al., Exploring the O-GlcNAc proteome: direct identification of O-GlcNAc-modified proteins from the brain.* Proc Natl Acad Sci U S A, 2004. **101**(36): p. 13132-7.

187. Whelan, S.A. and Hart, G.W., *Proteomic approaches to analyze the dynamic relationships between nucleocytoplasmic protein glycosylation and phosphorylation.* Circ Res, 2003. **93**(11): p. 1047-58.

188. Myers, S.A., Panning, B. and Burlingame, A.L. , *Polycomb repressive complex 2 is necessary for the normal site-specific O-GlcNAc distribution in mouse embryonic stem cells.* Proc Natl Acad Sci U S A, 2011. **108**(23): p. 9490-9495.

189. Zhu, Y., *et al.*, *Post-translational O-GlcNAcylation is essential for nuclear pore integrity and maintenance of the pore selectivity filter.* J Mol Cell Biol, 2015.

190. Zhu, Y. *et al., O-GlcNAc occurs cotranslationally to stabilize nascent polypeptide chains.* Nat Chem Bio, 2015. **11**: p. 319-326.

191. Wang, Z., Gucek, M. and Hart, G.W. *Cross-talk between GlcNAcylation and phosphorylation: site-specific phosphorylation dynamics in response to globally elevated O-GlcNAc.* Proc Natl Acad Sci U S A, 2008. **105**(37): p. 13793-8.

192. Zhang, F. *et al., O-GlcNAc Modification is an Endogenous Inhibitor of the Proteasome.* Cell, 2003. **115**: p. 715-725.

193. Mishra, S., Ande, S.R. and Salter, N.W., *O-GlcNAc modification: why so intimately associated with phosphorylation?* Cell Commun Signal, 2011. **9**(1): p. 1.

194. Santoyo-Ramos, P., Cristina, M., and Robles-Flores, M., *The Role of O-Linked β-N-Acetylglucosamine (GlcNAc) Modification in Cell Signaling.* 2012. Chapter 3.

195. Zhu, Y., *et al., The emerging link between O-GlcNAc and Alzheimer disease.* J Biol Chem, 2014. **289**(50): p. 34472-81.

196. Liu, F., *et al., Reduced O-GlcNAcylation links lower brain glucose metabolism and tau pathology in Alzheimer's disease.* Brain, 2009. **132**(Pt 7): p. 1820-32.

197. Liu, F., *et al., O-GlcNAcylation regulates phosphorylation of tau: a mechanism involved in Alzheimer's disease.* Proc Natl Acad Sci U S A, 2004. **101**(29): p. 10804-9.

198. Gong, C.X., Liu, F., and Iqbal, K., *O-GlcNAc cycling modulates neurodegeneration.* Proc Natl Acad Sci U S A, 2012. **109**(43): p. 17319-20.

199. Ferrer, C.M., *et al*., *O-GlcNAcylation regulates cancer metabolism and survival stress signaling via regulation of the HIF-1 pathway.* Mol Cell, 2014. **54**(5): p. 820-31.

200. Singh, J.P., *et al.*, *O-GlcNAc signaling in cancer metabolism and epigenetics.* Cancer Lett, 2015. **356**(2 Pt A): p. 244-50.

201. Ferrer, C.M. and M.J. Reginato, *Sweet connections: O-GlcNAcylation links cancer cell metabolism and survival.* Mol & Cell Onc, 2014. **2**(1): p. e961809.

202. Gao, Y., *et al*., *Dynamic O-glycosylation of nuclear and cytosolic proteins: cloning and characterization of a neutral, cytosolic β-N-acetylglucosaminidase from human brain.* J Biol Chem, 2001. **276**(13): p. 9838-45.

203. L-Y Dong, and Hart., G.W., *Purification and Characterization of an O-GlcNAc Selective N-Acetyl-β-D-glucosaminidase from Rat Speen Cytosol.* J Biol Chem, 1994. **269**(30): p. 19321-19330.

204. Cantarel, B.L., *et al.*, *The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics.* Nucleic Acids Res, 2009. **37** (Database issue): p. D233-8.

205. Horsch, M., Hoesch, L., Vasella, A. and Rast, D.A., *N-Acetylglucosaminono-1,5-lactone oxime and the corresponding(phenylcarbamoy1)oxime.* Eur. J. Biochem, 1991. **197**: p. 815-818.

206. Beer, D. *et al., Synthesis of 2-Acetamido-2-deoxy-D-gluconhydroximolactone-and Chitobionhydroximolactone-Derived N-Phenylcarbamates, Potential Inhibitors of β-N-Acetylglucosaminidase.* Helv Chim Acta, 1990. **73**(7): p. 1918-1922.

207. Miller, D.J., Gong, X. and Shur, B.D., *Sperm require β-N-acetylglucosaminidase to penetrate through the egg zona pellucida.* Development, 1993. **118**: p. 1279-1289.

208. Triggs-Raine, B., *Naturally Occuring Mutations in G_{M2} Gangliosidosis: A Compendium.* Adv in Gen, 2001. **44**: p. 199-223.

209. Macauley, M.S., Stubbs, K.A. and Vocadlo, D.J., *O-GlcNacase Catalyzes Cleavage of Thioglycosides without General Acid Catalysis.* J Am Chem Soc, 2005. **127**(49): p. 17202-17203.

210. Bisel, B., Pavone, F.S., and Calamai, M. *GM1 and GM2 gangliosides: recent developments.* Biomol Concepts, 2014. **5**(1): p. 87-93.

211. Zhou, D. *et al.*, *Lysosomal Glycosphingolipid Recognition by NKT Cells.* Science, 2004. **306**(5702): p. 1786-1789.

212. Nojiri, H. *et al.*, *A Specific Type of Ganglioside as a Modulator of Insulin-dependent Cell Growth and Insulin Receptor Tyrosine Kinase Activity.* J Biol Chem, 1991. **266**(7): p. 4531-4537.

213. Matthews, J.A., *et al., Glucosamine-induced increase in Akt phosphorylation corresponds to increased endoplasmic reticulum stress in astroglial cells.* Mol Cell Biochem, 2007. **298**(1-2): p. 109-23.

214. Dennis, R.J., *et al., Structure and mechanism of a bacterial β-glucosaminidase having O-GlcNAcase activity.* Nat Struct Mol Biol, 2006. **13**(4): p. 365-71.

215. Naniye Cetinbas *et al., Identification of Asp^{174} and Asp^{175} as the Key Catalytic Residues of Human O-GlcNAcase by Functional Analysis of Site-Directed Mutants.* Biochemistry, 2006. **45**: p. 3835-3844.

216. Borghgraef, P., *et al., Increasing brain protein O-GlcNAc-ylation mitigates breathing defects and mortality of Tau.P301L mice.* PLoS One, 2013. **8**(12): p. e84442.

217. Dorfmueller, H.C., *et al., Cell-penetrant, nanomolar O-GlcNAcase inhibitors selective against lysosomal hexosaminidases.* Chem Biol, 2010. **17**(11): p. 1250-5.

218. Stubbs, K.A., Zhang, N. and Vocadlo, D.J.V, *A divergent synthesis of 2-acyl derivatives of PUGNAc yields selective inhibitors of O-GlcNAcase.* Org Biomol Chem, 2006. **4**(5): p. 839-45.

219. Kim, E.J. *et al., An O-GlcNAcase-Specific Inhibitor and Substrate Engineered by the Extension of the N-Acetyl Moiety.* J Am Chem Soc 2006. **128**: p. 4234-4235.

220. Avalos Gonzalez, M., *et al., Synthesis of 1,3,4,6-tetra-O-acetyl-2-[3-alkyl-(aryl)-thioreido]-2-deoxy-α-D-glucopyranoses and their transformation into 2-alkyl(aryl)amino-(1,2-dideoxy-α-D-glucopyrano)[2,1-D]-2-thiazolines.* Carb Res, 1986. **154**(1): p. 49-62.

221. Lineweaver, H. and Burk, D., *The Determination of Enzyme Dissociation Constants.* J Am Chem Soc, 1934. **56**: p. 658-666.

222. Morrison, J.F., *Kinetics of the Reversible Inhibition of Enzyme-Catalysed Reactions by Tight-Binding Inhibitors.* Biochim et Biophys Acta, 1969. **185**: p. 269-286.

223. Ribeiro, C. and Esteves da Silva, J.C., *Kinetics of inhibition of firefly luciferase by oxyluciferin and dehydroluciferyl-adenylate.* Photochem Photobiol Sci, 2008. **7**(9): p. 1085-90.

224. Williams, J.W. and Morrison, J.F., *The kinetics of reversible tight-binding inhibition*, in *Methods in Enzymology*, L.P. Daniel, Editor. 1979, Academic Press. p. 437-467.

225. Copeland, R.A., *Evaluation of Enzyme Inhibitors in Drug Discovery: A Guide for Medicinal Chemists and Pharmacologists*. Vol. Methods of Biochemical Analysis. 2005, Hoboken, New Jersey: John Wiley and Sons Inc.

226. Kuzmic, P., *et al., High-throughput screening of enzyme inhibitors: simultaneous determination of tight-binding inhibition constants and enzyme concentration.* Anal Biochem, 2000. **286**(1): p. 45-50.

227. Murphy, D.J., *Determination of accurate K_I values for tight-binding enzyme inhibitors: an in silico study of experimental error and assay design.* Anal Biochem, 2004. **327**(1): p. 61-7.

228. Rao, F.V., *et al.*, *Structural insights into the mechanism and inhibition of eukaryotic O-GlcNAc hydrolysis.* The EMBO Journal, 2006. **25**: p. 1569-1578.

229. Dixon, M., *The Determination of Enzyme Inhibitor Constants.* Biochem J, 1953. **55**: p. 170-171.

230. Mark, B.L., *et al., Crystal Structure of Human β-Hexosaminidase B: Understanding the Molecular Basis of Sandhoff and Tay–Sachs Disease.* J Mol Biol, 2003. **327**(5): p. 1093-1109.

231. Remko, M., Swart, M and Bickelhaupt, F.M., *Theoretical study of structure, pKa, lipophilicity, solubility, absorption, and polar surface area of some centrally acting antihypertensives.* Bioorg Med Chem, 2006. **14**(6): p. 1715-28.

232. Bacarella, A.L. *et al., The potentiometric measurement of acid dissociation constants and pH in the system methanol-water. pKa values for carboxylic acids and anilinium ions.* J org Chem, 1955. **20**(6): p. 747-762.

233. Harned, H.S. and Owen, B,B., *Determination of the Ionization and Thermodynamic Properties of Weak Electrolytes by Means of Cells without Liquid Junctions.* Chem Rev, 1939. **25**(1): p. 31-65.

234. Sakmar, T.P *et al., The role of the retinylidene Schiff base counterion in rhodopsin in determining wavelength absorbance and Schiff base pKa.* Proc Nat Acad Sci, 1991. **88**(8): p. 3079-3083.

235. Christ, P. *et al., pKa Values of Chiral Brønsted Acid Catalysts: Phosphoric Acids/Amides, Sulfonyl/Sulfuryl Imides, and Perfluorinated TADDOLs (TEFDDOLs).* Chem – A Eur J, 2011. **17**(31): p. 8524-8528.

236. Grunwald, E., Loewenstein, A., and Meiboom, S., *Application of Nuclear Magnetic Resonance to the Study of Acid-Base Equilibria.* J Chem Phys, 1957. **27**(3): p. 641.

237. Perrin, L.C. and Fabian, M.A. *Multicomponent NMR Titration for Simultaneous Measurement of Relative pKas.* Anal Chem, 1996. **68**: p. 2127-2134.

238. Shivapurkar, R. and Jeannerat, D., *Determination of the relative pKa's of mixtures of organic acids using NMR titration experiments based on aliased ^{1}H–^{13}C HSQC spectra.* Anal Meth, 2011. **3**(6): p. 1316.

239. Perrin, L.C., *et al., Solvation effect on steric bulk of ionic substituents: Imidazolium vs imidazole.* J Org Chem, 1994. **59**: p. 5246-5253.

240. Ellison, S.L.R. and Robinson, M.J.T., *A precise NMR method for measuring isotope effects on acid-base equilibria: effects of solvent, solvent isotopic composition and ionic strength on hydrogen and oxygen isotope effects in formic acid.* J Chem Soc Chem Commun, 1983: p. 745-746.

241. Perrin, L.C. and Thoburn, J.D., *Evidence for a Double Minimum Potential for Intramolecular Hydrogen Bonds of Aqueous Hydrogen Maleate and Hydrogen Phthalate Anions.* J Am Chem Soc, 1989. **111**: p. 8010-8012.

242. Rabenstein, D.L. and Mariappan, S.V.S., *Determination of ^{15}N Isotope Effects on the Acid-Base Equilibria of Amino Groups in Amino Acids by ^{13}C NMR.* J Org Chem, 1993. **58**: p. 4487-4489.

243. Pehk, T. *et al.*, *General approach to measurements of pK_a differences by ^{13}C NMR spectroscopy.* J Chem Soc Perkin Trans 2, 1996: p. 2351-257.

244. Roses, M. *et al., Dissociation constants of phenols in methanol-water mixtures.* J Chromatograph A, 2000. **867**: p. 45-56.

245. Reynolds, C.H., *Ligand efficiency metrics: Why all the fuss?* Future Med Chem, 2015. **7**(11): p. 1363-1365.

246. Kuntz, I.D. *et al., The maximal affinity of ligands.* Proc Natl Acad Sci U S A, 1999. **96**: p. 9997-10002.

247. Hopkins, A.L., Groom, C.R., and Alex, A., *Ligand efficiency: a useful metric for lead selection.* Drug Discovery Today, 2004. **9**(10): p. 430-431.

248. Lienhard, E.G., *Enzymatic Catalysis and Transition-State Theory.* Science, 1973. **180**(4082): p. 149-154.

249. Bartlett, P.A. and Mader, M.M., *Binding Energy and Catalysis: The Implications for Transitition-State Analogs and Catalytic Antibodies.* Cheml Rev, 1997. **97**: p. 1281-1301.

250. Bartlett, P.A. and Marlowe, C.K., *Phosphonamidates as Transition-State Analogue Inhibitors of Thermolysin.* Biochemistry, 1983. **22**: p. 4618.

251. Roeser, K.-R and Legler, G., *Role of sugar hydroxyl groups in glycoside hydrolysis. Cleavage mechanism of deoxyglucosides and related substrates by β-glucosidase A3 from Aspergillus wentii.* Biochim et Biophys Acta (BBA) - Enzymol, 1981. **657**(2): p. 321-333.

252. Wang, Z., *Zemplén Deacetylation*, in *Comprehensive Organic Name Reactions and Reagents*. 2010, John Wiley & Sons, Inc.

253. Wicki, J. *et al., Transition-State Mimicry by Glycosidase Inhibitors: A Critical Kinetic Analysis.* J Am Chem Soc, 2007. **129**: p. 4530-4531.

254. Mosi, R. *et al., Reassessment of Acarbose as a Transition State Analogue Inhibitor of Cyclodextrin Glycosyltransferase.* Biochem J, 1998. **37**: p. 17192-17198.

255. Frye, S.V., *The art of the chemical probe.* Nat Chem Biol, 2010. **6**(3): p. 159-161.

256. Cunha, A. *et al., Use of Protecting Groups in Carbohydrate Chemistry: An Advanced Organic Synthesis Experiment.* J. Chem. Edu., 1999. **76**(1): p. 79-80.

257. Frangioni, J., *In vivo near-infrared fluorescence imaging.* Curr Opin Chem Biol, 2003. **7**(5): p. 626-634.

258. Sakabe, M., *et al., Rational design of highly sensitive fluorescence probes for protease and glycosidase based on precisely controlled spirocyclization.* J Am Chem Soc, 2013. **135**(1): p. 409-14.

259. Forster, T., *Zwischenmolekulare Energiewanderung und Fluoreszenz.* Ann. Phys., 1948. **6**(2): p. 55-75.

260. Shrestha, D., *et al*., *Understanding FRET as a research tool for cellular studies.* Int J Mol Sci, 2015. **16**(4): p. 6718-56.

261. Hatzakis, N.S., *et al., How curved membranes recruit amphipathic helices and protein anchoring motifs.* Nat Chem Biol, 2009. **5**(11): p. 835-41.

262. Kajihara, D., *et al., FRET analysis of protein conformational change through position-specific incorporation of fluorescent amino acids.* Nat Methods, 2006. **3**(11): p. 923-9.

263. Johansson, L.B.A., *et al., Photophysics, molecular reorientation in solution and X-ray structure of a new fluorescent probe, 1,7-diazaperylene.* J Am Chem Soc, Faraday Transactions, 1993. **89**(1): p. 49-54.

264. Koglin, P.K.F., *et al.*, *Determination of micelle aggregation numbers by energy transfer.* J Phys Chem, 1981. **85**(16): p. 2363-2366.

265. Kremers, G.-J., *et al*., *Cyan and Yellow Super Fluorescent Proteins with Improved Brightness, Protein Folding, and FRET Förster Radius†,‡.* Biochemistry, 2006. **45**(21): p. 6570-6580.

266. Berney, C. and Danuser, G., *FRET or no FRET: A quantitative comparison.* Biophys J, 2003. **84**: p. 3992-4010.

267. Yaron, A., *et al., Intramolecularly quenched fluorogenic substrates for hydrolytic enzymes.* Anal Biochem, 1979. **95**(1): p. 228-235.

268. Meldal, M. *et al., Portion-mixing peptide libraries of quenched fluorogenic substrates for complete subsite mapping of endoprotease specificity.* Proc Nat.Acad Sci USA, 1994. **91**: p. 3314-3318.

269. Meldal, M. and Breddam, K., *Anthranilamide and nitrotyrosine as a donor-acceptor pair in internally quenched fluorescent substrates for endopeptidases: Multicolumn peptide synthesis of enzyme substrates for subtilisin carlsberg and pepsin.* Anal Biochem, 1991. **195**(1): p. 141-147.

270. Stennincke, H.R. *et al., Internally quenched fluorescent peptide substrates disclose the subsite preferences of human caspases 1, 3, 6, 7 and 8.* Biochem J, 2000. **350**: p. 563-568.

271. Angliker, H. *et al.*, *Internally Quenched Fluorogenic Substrate for Furin.* Anal Biochem, 1995. **224**: p. 409-412.

272. Mittoo, S., Sundstrom, L.E. and Bradley, M., *Synthesis and evaluation of fluorescent probes for the detection of calpain activity.* Anal Biochem, 2003. **319**(2): p. 234-238.

273. Edgington, L.E., Verdoes, M., and Bogyo, M., *Functional imaging of proteases: recent advances in the design and application of substrate-based and activity-based probes.* Curr Opin Chem Biol, 2011. **15**(6): p. 798-805.

274. Rudin, M. and Weissleder, R., *Molecular imaging in drug discovery and development.* Nat Rev Drug Discov, 2003. **2**(2): p. 123-31.

275. Watzke, A., *et al., Selective activity-based probes for cysteine cathepsins.* Angew Chem Int Ed Engl, 2008. **47**(2): p. 406-9.

276. Turk, V. *et al.*, *Lysosomal cysteine proteases: facts and opportuninites.* The EMBO Journal, 2001. **20**(17): p. 4629-4633.

277. Backes, B.J., *et al., Synthesis of positional-scanning libraries of fluorogenic peptide substrates to define the extended substrate specificity of plasmin and thrombin.* (1087-0156 (Print)).

278. Frisardi, V. *et al., Towards disease-modifying treatment of Alzheimer's disease: drugs targeting β-amyloid.* Curr Alzheimer Res, 2010. **7**(1): p. 40-55.

279. Evin, G., Barakat, A., and Masters, C.L., *BACE: Therapeutic target and potential biomarker for Alzheimer's disease.* Int J Biochem Cell Biol, 2010. **42**(12): p. 1923-6.

280. Stachel, S.J. *et al., Discovery of aminoheterocycles as a novel β-secretase inhibitor class: pH dependence on binding activity part 1.* Bioorg Med Chem Lett, 2009. **19**(11): p. 2977-80.

281. Folk, S.D. *et al., Monitoring β-Secretase Activity in Living Cells with a Membrane-Anchored FRET Probe.* Angew Chem Int Ed, 2012. **51**: p. 10795-10799.

282. Folk, S.D. and Franz, K.J., *A Prochelator Activated by β-Secretase Inhibits Aβ Aggregation and Suppresses Copper Induced Reactive Oxygen Species Formation.* J Am Chem Soc, 2010. **132**: p. 4994-4995.

283. Komatsu, T. *et al., Design and Synthesis of an Enzyme Activity-Based Labeling Molecule with Flurescence Spectral Change.* J Am Chem Soc, 2006. **128**: p. 15946-15947.

284. Debacq-Chainiaux, F. *et al., Protocols to detect senescence-associated β-galactosidase (SA-β gal) activity, a biomarker of senescent cells in culture and in vivo.* Nat Prot, 2009. **4**(12): p. 1798-1806.

285. Chen, Y., Tsao, K., and Keillor, J.W., *Fluorogenic protein labelling: a review of photophysical quench mechanisms and principles of fluorogen design.* Can J of Chem, 2015. **93**(4): p. 389-398.

286. Lakowicz, J.R., *Principles of Fluorescence Spectroscopy*, 2006, Springer Science + Business Media LLC: 233 Spring Street, New York, USA.

287. Doose, S., Neuweiler, H., and Sauer, M., *Fluorescence quenching by photoinduced electron transfer: a reporter for conformational dynamics of macromolecules.* Chemphyschem, 2009. **10**(9-10): p. 1389-98.

288. Doose, S., *et al*., *Probing polyproline structure and dynamics by photoinduced electron transfer provides evidence for deviations from a regular polyproline type II helix.* Proc Natl Acad Sci U S A, 2007. **104**(44): p. 17400-5.

289. Rehm, D. and Weller, A., *Kinetics of Fluorescence Quenching by Electron and H-Atom Transfer.* Israel J Chem, 1970. **8**(2): p. 259-271.

290. Adams, D.M. *et al., Charge Transfer on the Nanoscale: Current Status.* J. Phys. Chem. B, 2003. **107**: p. 6668-6697.

291. Goldberg, J.M., *et al., Thioamide quenching of fluorescent probes through photoinduced electron transfer: mechanistic studies and applications.* J Am Chem Soc, 2013. **135**(49): p. 18651-8.

292. Vaiana, A.C. *et al., Fluorescence Quenching of Dyes by Tryptophan: Interactions at Atomic Detail from Combination of Experiment and Computer Simulation.* J Am Chem Soc, 2003. **125**: p. 14564-14572.

293. Seidel, C.A.M., Shulz, A. and Sauer, M.H.M., *Nucleobase-Specific Quenching of Fluorescent Dyes. 1. Nucleobase One Electron Redox Potentials and Their Correlation with Static and Dynamic Quenching Efficiencies.* J. Phys. Chem. B, 1996. **100**: p. 5541-5553.

294. Marme, N., *et al., Highly sensitive protease assay using fluorescence quenching of peptide probes based on photoinduced electron transfer.* Angew Chem Int Ed Engl, 2004. **43**(29): p. 3798-801.

295. Jung, J. and van Orden, A., *A Three-State Mechanism for DNA Hairpin Folding Characterized by Multiparameter Fluorescence Fluctuation Spectroscopy.* J Am Chem Soc, 2006. **128**: p. 1240-1249.

296. Wang, X., *et al., A highly selective and sensitive near-infrared fluorescence probe for arylamine N-acetyltransferase 2 in vitro and in vivo.* Chem Sci, 2013. **4**(7): p. 2936.

297. Cascorbi, I. *et al., Homozygous rapid arylamine N-acetyltransferase (NAT2) genotype as a susceptibility factor for lung cancer.* Can Res, 1996. **56**: p. 3961-3966.

298. Johansson, M.K. *et al.*, *Intramolecular Dimers: A New Strategy to Fluorescence Quenching in Dual-Labeled Oligonucleotide Probes.* J Am Chem Soc, 2002. **124**: p. 6950-6956.

299. Goldberg, J.M., *et al.*, *Thioamide quenching of intrinsic protein fluorescence.* Chem Commun (Camb), 2012. **48**(10): p. 1550-2.

300. Batjargal, S., *et al., Native chemical ligation of thioamide-containing peptides: development and application to the synthesis of labeled α-synuclein for misfolding studies.* J Am Chem Soc, 2012. **134**(22): p. 9172-82.

301. Bordwell F.G., *et al., The Relative Ease of Removing a Proton, a Hydrogen Atom, or an Electron from Carboxamides versus Thiocarboxamides.* J Am Chem Soc, 1988. **110**: p. 5903-5904.

302. Qian, Y. *et al., A resorufin-based colorimetric and fluorescent probe for live-cell monitoring of hydrazine.* Biosens Bioelectron, 2014. **58**: p. 282-6.

303. Li, Z. *et al., Nitroreductase detection and hypoxic tumor cell imaging by a designed sensitive and selective fluorescent probe, 7-[(5-nitrofuran-2-yl)methoxy]-3H-phenoxazin-3-one.* Anal Chem, 2013. **85**(8): p. 3926-32.

304. Zhang, Y. *et al., A spectroscopic off-on probe for simple and sensitive detection of carboxylesterase activity and its application to cell imaging.* Analyst, 2012. **137**(3): p. 716-21.

305. Sandholt, I.S. *et al.*, *Resorufin: a lead for a new protein kinase CK2 inhibitor.* Anticancer Drugs, 2009. **20**(4): p. 238-48.

306. Mecinovic, J., *et al., 2-Oxoglutarate analogue inhibitors of prolyl hydroxylase domain 2.* Bioorg Med Chem Lett, 2009. **19**(21): p. 6192-5.

307. Ozturk, T., Erdal, E., and Mert, O., *Use of Lawesson's Reagent in Organic Syntheses.* Chem Rev, 2007. **107**: p. 5210-5278.

308. Coats, S.J., Link, S.J. and Hlasta, J.D., *Alternative Solvents for Elevated-Temperature Solid-Phase Parallel Synthesis. Application to Thionation of Amides.* Org Lett, 2003. **5**(5): p. 721-724.

309. Hilty, M.F., Brun, A.K. and Heimgartner, H., *A New 2H-Azirin-3-amine as a Synthon for alpha-Methyl Glutamate.* Helv Chim Acta, 2004. **87**: p. 2539-2548.

310. Cava, M.P. and Levinson, M.I., *Thionation reactions of lawesson's reagents.* Tetrahedron, 1985. **41**(22): p. 5061-5087.

311. Oshida, H. *et al.*, *First isolation of 1,3,2-dithiaphosphetane 2-sulfide.* Tet Lett, 2004. **45**: p. 1331-1334.

312. Knorr, R., *et al., New coupling reagents in peptide chemistry.* Tet Lett., 1989. **30**: p. 1927.

313. Pon, R.T. and Yu, S., *Rapid automated derivatization of solid phase supports for oligonucleotide synthesis using uronium or phosphonium coupling reagents.* Tet Lett., 1997. **38**: p. 3331.

314. Jadhav, S., *2-(1H-Benzotriazol-1-yl)-1,1,3,3-tetramethyluronium Hexafluoro-phosphate (HBTU).* Synlett, 2010. **2010**(08): p. 1287-1288.

315. Sureshbabu, V.V., *et al., New and simple synthesis of acid azides, ureas and carbamates from carboxylic acids: application of peptide coupling agents EDC and HBTU.* Org Biomol Chem, 2010. **8**(4): p. 835-40.

316. Mali, S.M., *et al., HBTU mediated 1-hydroxybenzotriazole (HOBt) conjugate addition: synthesis and stereochemical analysis of beta-benzotriazole N-oxide substituted gamma-amino acids and hybrid peptides.* Org Biomol Chem, 2014. **12**(42): p. 8462-72.

317. Sieber, S., Orth, R., and Pitscheider, M., *Chemical Probes for Labeling of the Bacterial Glucosaminidase NagZ via the Huisgen Cycloaddition.* Synthesis, 2010. **2010**(13): p. 2201-2206.

318. Zheng, W., *et al., Three classes of glucocerebrosidase inhibitors identified by quantitative high-throughput screening are chaperone leads for Gaucher disease.* Proc Natl Acad Sci U S A, 2007. **104**(32): p. 13192-7.

319. Olivova, P., *et al., An improved high-throughput dried blood spot screening method for Gaucher disease.* Clin Chim Acta, 2008. **398**(1-2): p. 163-4.

320. Molinari, L.M., *et al., Identification and partial characterisation of a chitinase from Nile tilapia, Oreochromis niloticus.* Comp Biochem Physiol B Biochem Mol Biol, 2007. **146**(1): p. 81-7.

321. Mauri, V., *et al., A rapid and sensitive method for measuring N-acetylglucosaminidase activity in cultured cells.* PLoS One, 2013. **8**(6): p. e68060.

322. Goldin, E. *et al., High Throughput Screening for Small Molecule Therapy for Gaucher Disease Using Patient Tissue as the Source of Mutant Glucocerebrosidase.* PLoS One, 2012. **7**(1).

323. Zhang, X.F., Zhang, J., and Liu, L., *Fluorescence properties of twenty fluorescein derivatives: lifetime, quantum yield, absorption and emission spectra.* J Fluoresc, 2014. **24**(3): p. 819-26.

324. Lavis, D.L. *et al.*, *Tuning the pKa of Fluorescein to Optimize Binding Assays.* Anal Chem, 2007. **79**: p. 6775-6782.

325. Kim, E.J., *et al., Enzymatic characterization of O-GlcNAcase isoforms using a fluorogenic GlcNAc substrate.* Carb Res, 2006. **341**(8): p. 971-982.

Appendix A.

Supporting information for Chapter 2

Kinetic, pK_a and NMR data

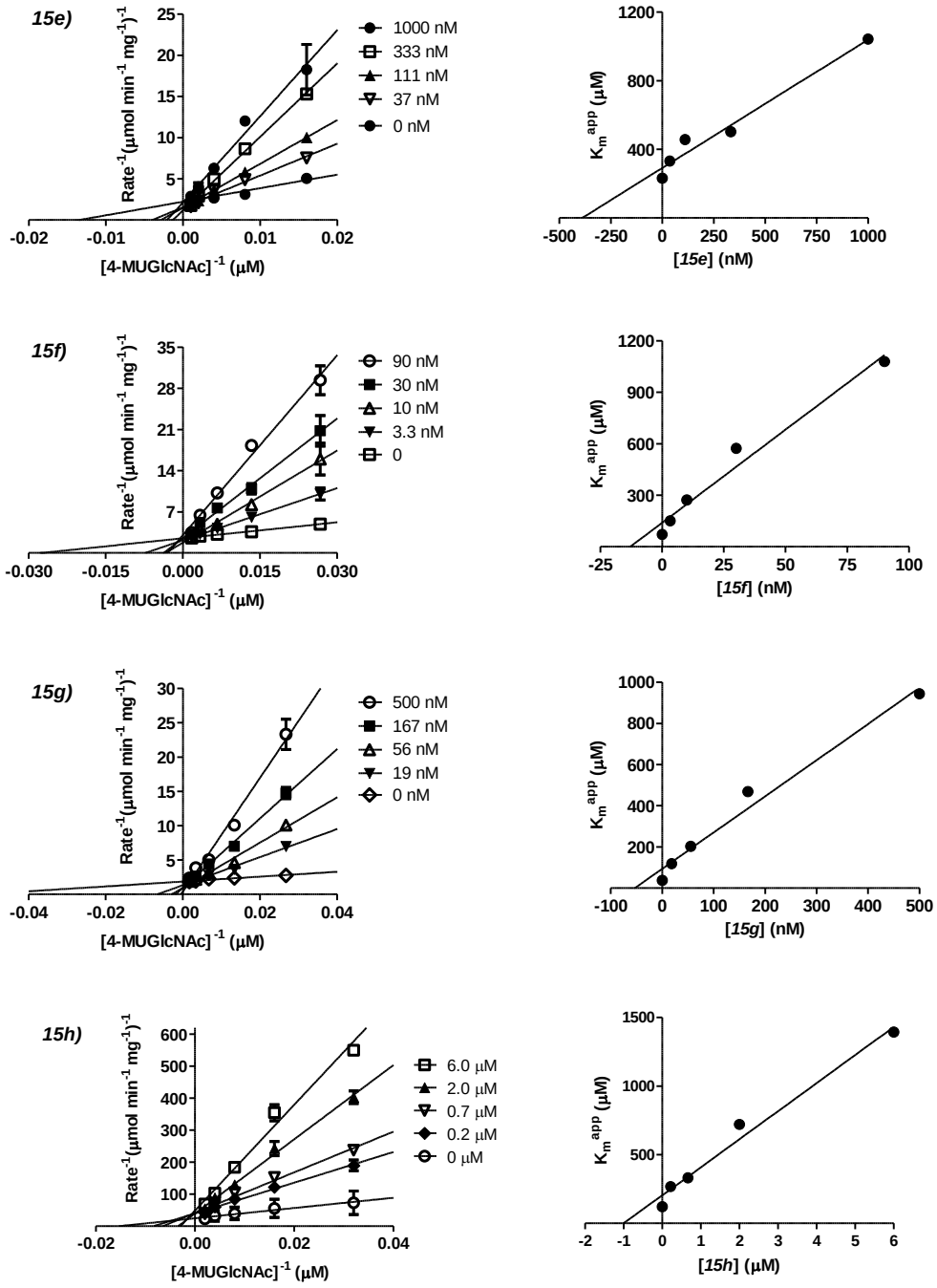


Figure A1. Lineweaver-Burke plots for inhibition of hOGA by compounds ***15e-15h*** (left panels) and the K_m^{app} vs. [I] plots for each inhibitor (right panels). The inhibitor concentrations used for each experiment are indicated in the legend of the K_i plots and the [hOGA] used in all cases was 1 nM.

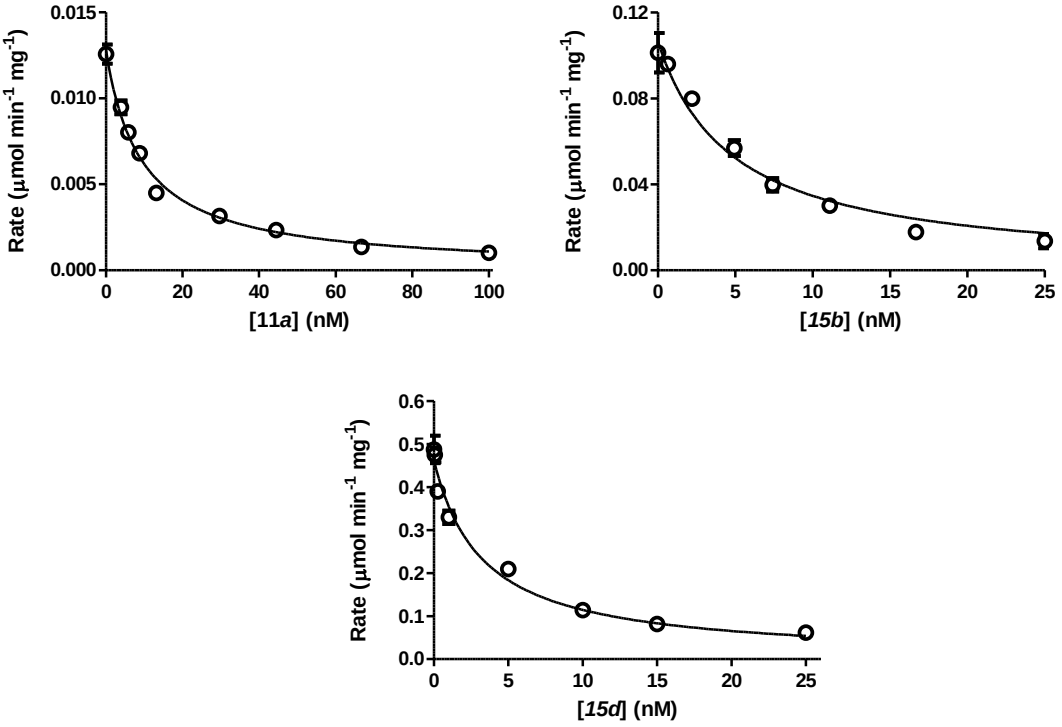


Figure A2. Morrison K_i plots for compounds ***11a, 15b*** and ***15d***. ***11a***: [hOGA] = 15 nM, [4MU-GlcNAc] = 170 µM; ***15b***: [hOGA] = 10 nM, [4MU-GlcNAc] = 60 µM; ***15d***: [hOGA] = 10 nM, [4MU-GlcNAc] = 180 µM.

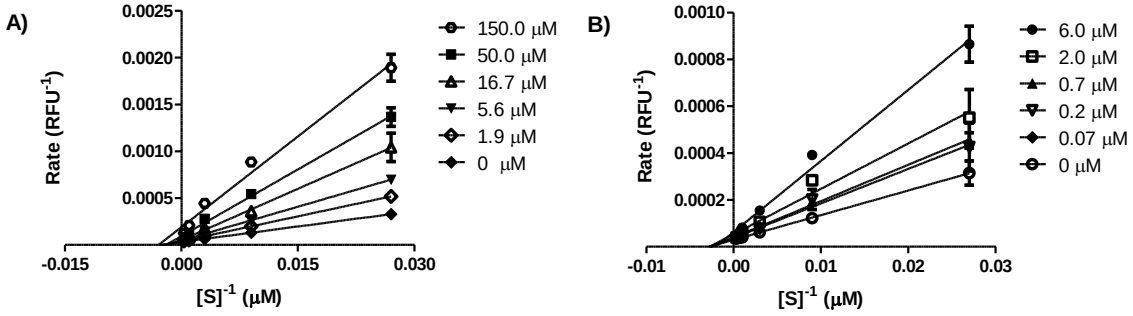


Figure A3. Lineweaver-Burke plots for inhibition of hHexB by compounds ***11a*** (panel A) and ***15a*** (panel B) using 4MU-GlcNAc as the substrate. Inhibitor concentrations are presented in each graph.

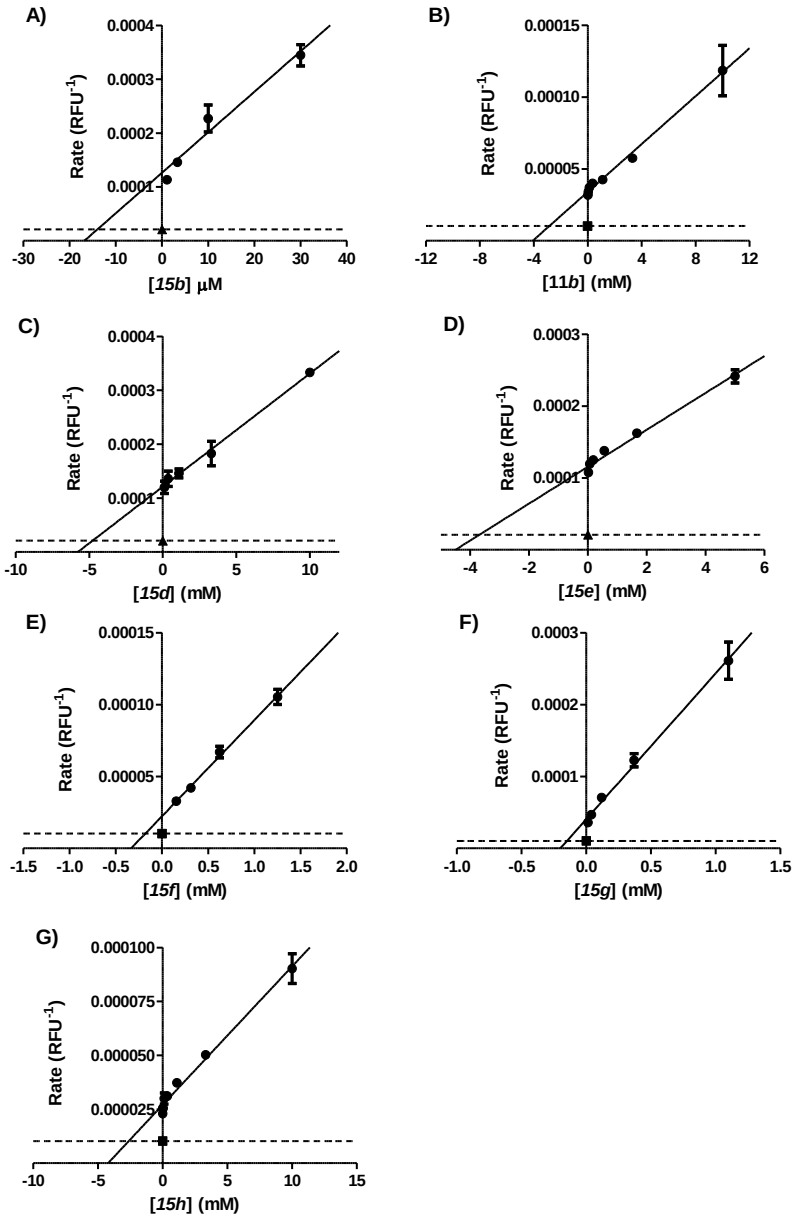


Figure A4. Dixon plot analysis of inhibition of hHexB by compounds ***11b***, ***15d-15h*** using 4MU-GlcNAc as a substrate. A) [E] = 5 nM, [S] = 230 µM; B) [E] = 5 nM, [S] = 500 µM; C) [E] = 2 nM, [S] = 60 µM; D) [E] = 2nM, [S] = 60 µM; E) [E] = 5 nM, [S] = 220 µM; F) [hHexB] = 5 nM, [4MU-GlcNAc] = 220 µM; G) [E] = 5 nM, [S] = 220 µM. Inhibitors were serially diluted 3-fold from the highest concentration. The K_i values were determined by assessing the point of intersection between the $1/V_{max}$ (dotted) line and the linear regression curve.

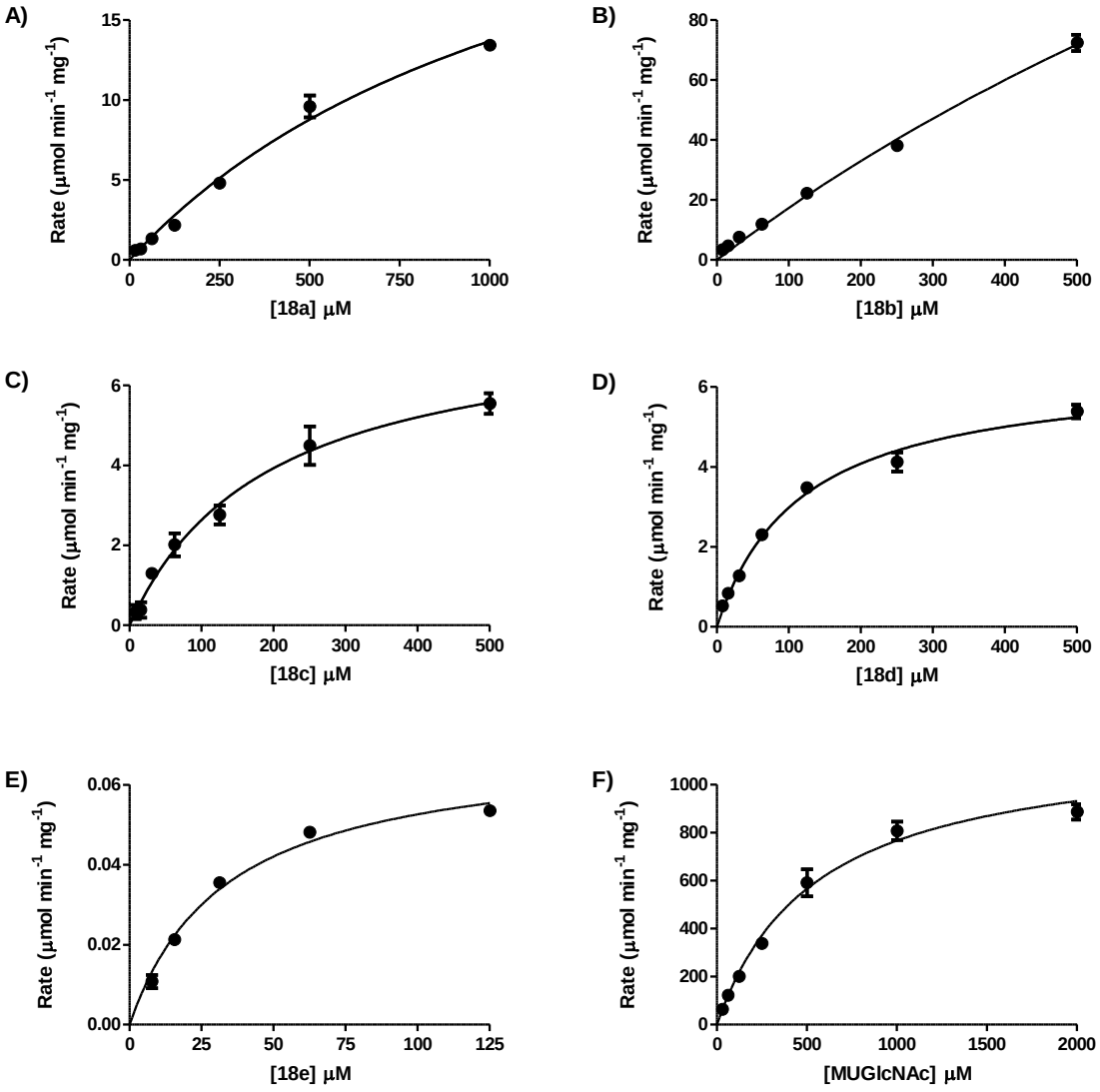


Figure A5. Michaelis-Menten plots obtained for urea substrates ***18a***-***18e*** and 4MU-GlcNAc with hOGA. All assays were performed in triplicate using pH 7.4 PBS buffer (0.03% BSA) containing 2.5% DMF. A) [E] = 0.1 µM B) [E] = 0.1 µM C) [E] = 0.2 µM D) [E] = 0.5 µM E) [E] = 12 µM F) [E] = 0.010 µM.

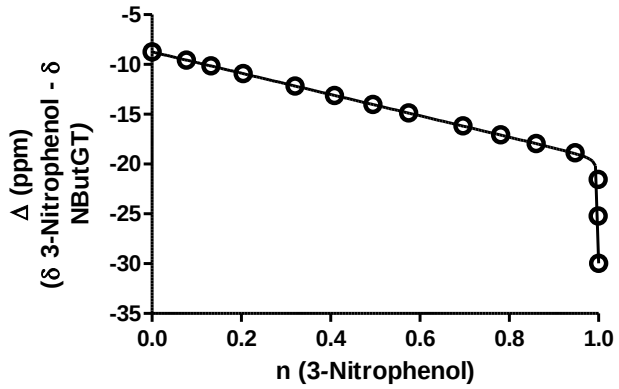
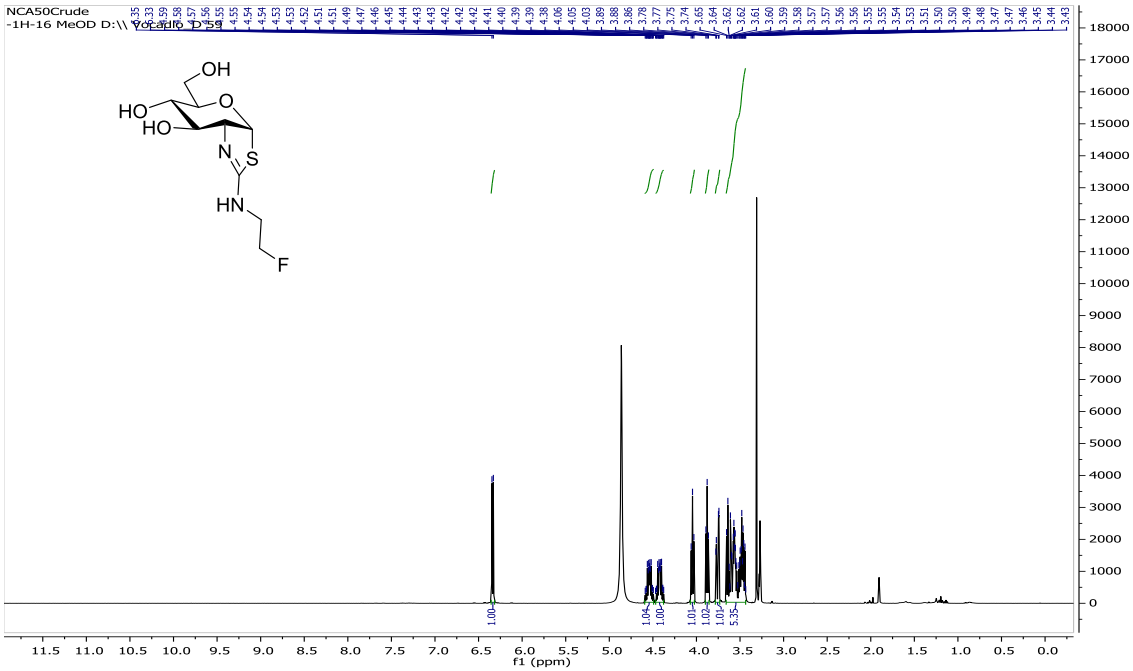


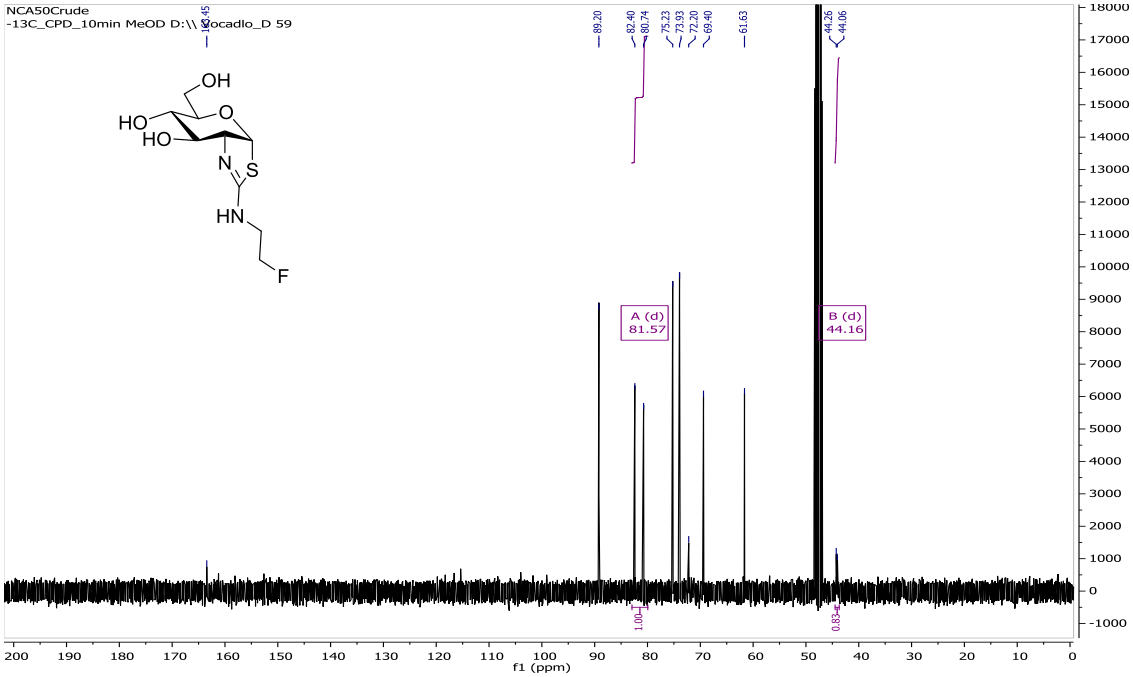
Figure A6. The difference between the chemical shifts of the thiazoline carbon resonances (Δ ppm) for NButGT as a function of the fractional protonation (*n*) of 3-nitrophenol, which has a pK_a of 8.42. The chemical shift changes (Δ ppm) of the **C**-O carbon in 3-nitrophenol and the 2' carbon of NButGT were monitored to define the fractional protonation of 3-nitrophenol, where $n = (\delta-\delta_d)/(\delta_p-\delta_d)$. The data was fit to equation 3. R (K_a^{NButGT}/K_a^r) = 0.0006 ± 0.0001; R^2 = 0.9865.

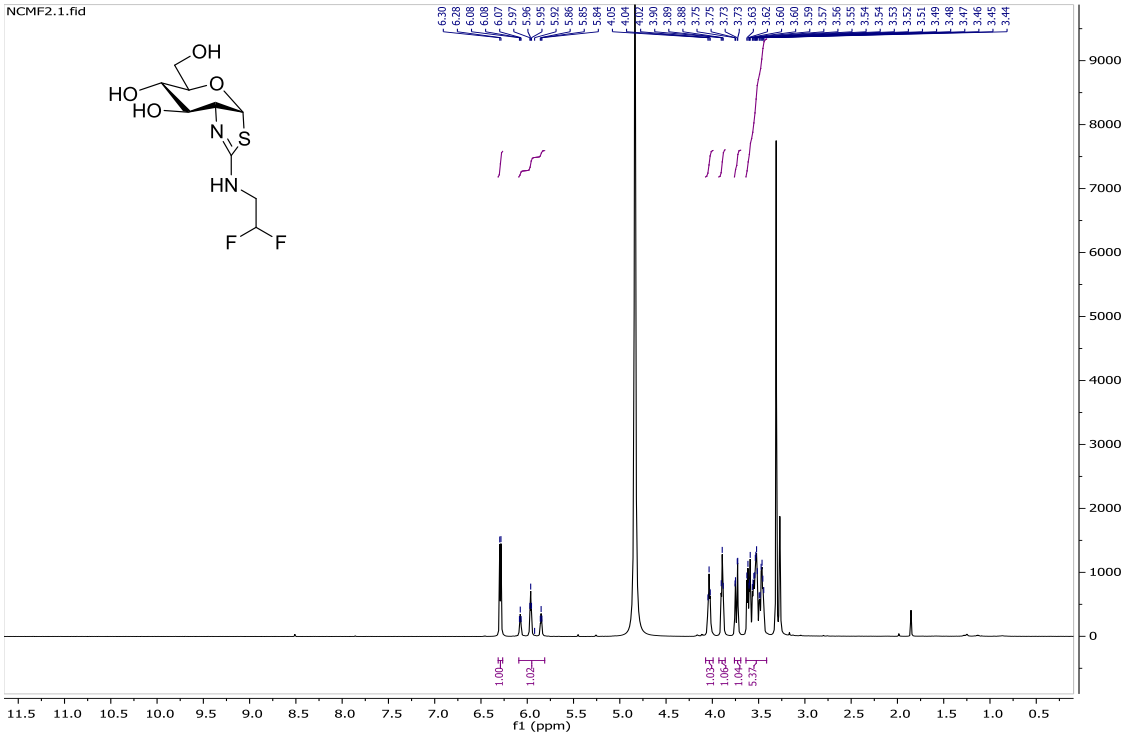
15c		15f		15g		15h		NButGT	
N-C-N	C-O	N-C-N	C-O	N-C-N	C-O	N-C-N	C-O	C-C-N	C-O
162.77	167.24	162.89	167.25	163.11	167.26	163.20	167.26	176.01	167.27
162.83	166.65	162.90	167.25	163.11	167.25	163.20	167.25	176.01	166.45
162.94	165.66	162.90	167.24	163.11	167.25	163.20	166.50	176.01	165.86
163.04	164.85	162.92	166.85	163.11	167.05	163.20	164.55	176.01	165.09
163.19	163.94	162.97	164.16	163.12	165.51	163.20	162.57	176.01	163.85
163.31	163.31	163.20	160.75	163.12	163.12	163.20	160.90	176.01	162.91
163.49	162.48	163.78	158.42	163.18	160.55	163.23	158.87	176.01	161.99
163.69	161.61	165.79	157.07	163.35	158.12	163.39	157.17	176.01	161.14
163.95	161.09	167.34	156.75	164.34	156.94	165.10	156.65	176.01	159.84
164.22	160.69	168.31	156.63	166.28	156.71	166.79	156.62	176.01	158.94
164.50	159.99	170.01	156.62	168.54	156.64	168.93	156.62	176.02	158.09
164.90	159.36	170.02	156.62	171.02	156.61	170.94	156.61	176.03	157.16
165.02	159.23			171.26	156.61	172.00	156.61	178.14	156.61
165.50	158.82							181.83	156.61
166.19	158.12							186.57	156.6
166.76	157.76								
167.20	157.48								
167.74	157.22								
168.34	157.1								
168.93	156.77								

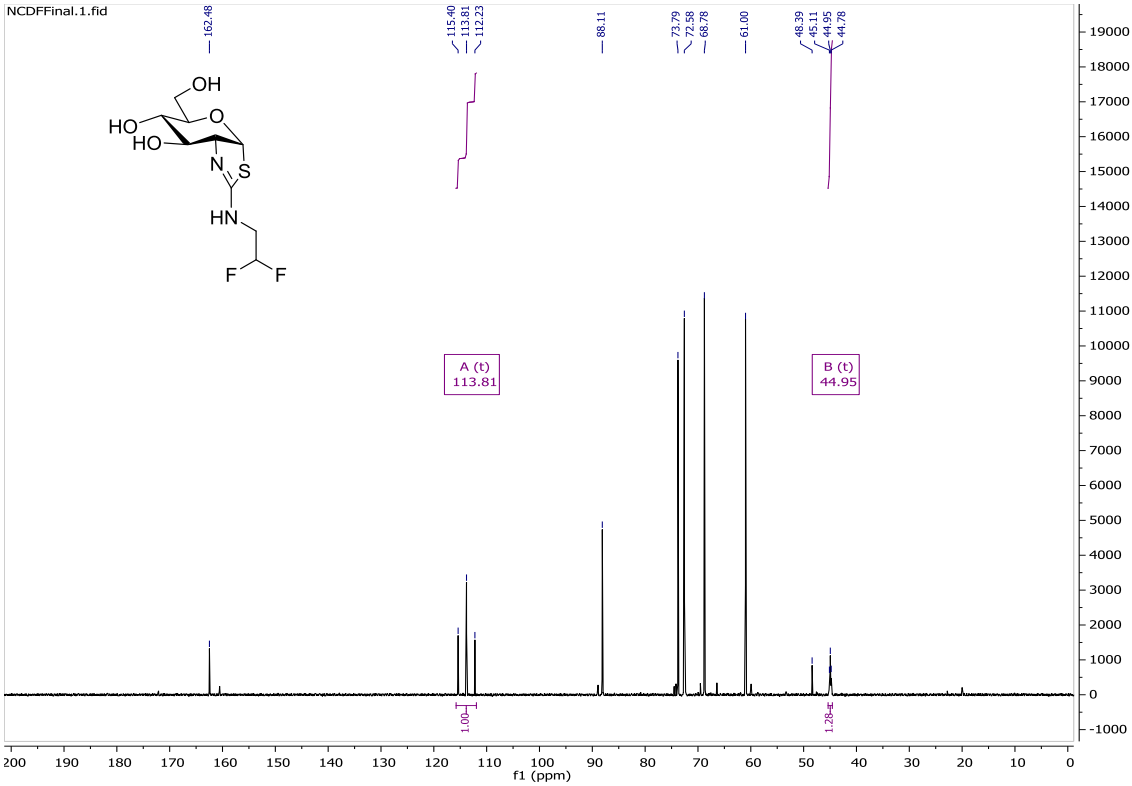
Figure A7. ^{13}C chemical shifts (δ, ppm) for compounds ***15c-15h*** and NButGT which were used to measure the parameters necessary for ΔpK_a determination. The chemical shifts which were directly compared were those of the amino-thiazoline carbon (N-**C**-N) for inhibitors ***15c-15h*** and those of the **C-**O carbon of 3-nitrophenol. ***15c*****:** 0.161 mmol of each compound titrated with 1M HCl in 10 µL increments (20 injections); ***15f*****:** 0.0262 mmol (12 injections), ***15g:*** 0.0302 mmol (13 injections), ***15h:*** 0447 mmol (13 injections) of each compound, titrated with 0.5 M HCl in 10 µL increments. **NButGT:** 0.161 mmol of each compound titrated with 1M HCl in 10 µL increments for a total of 15 injections.

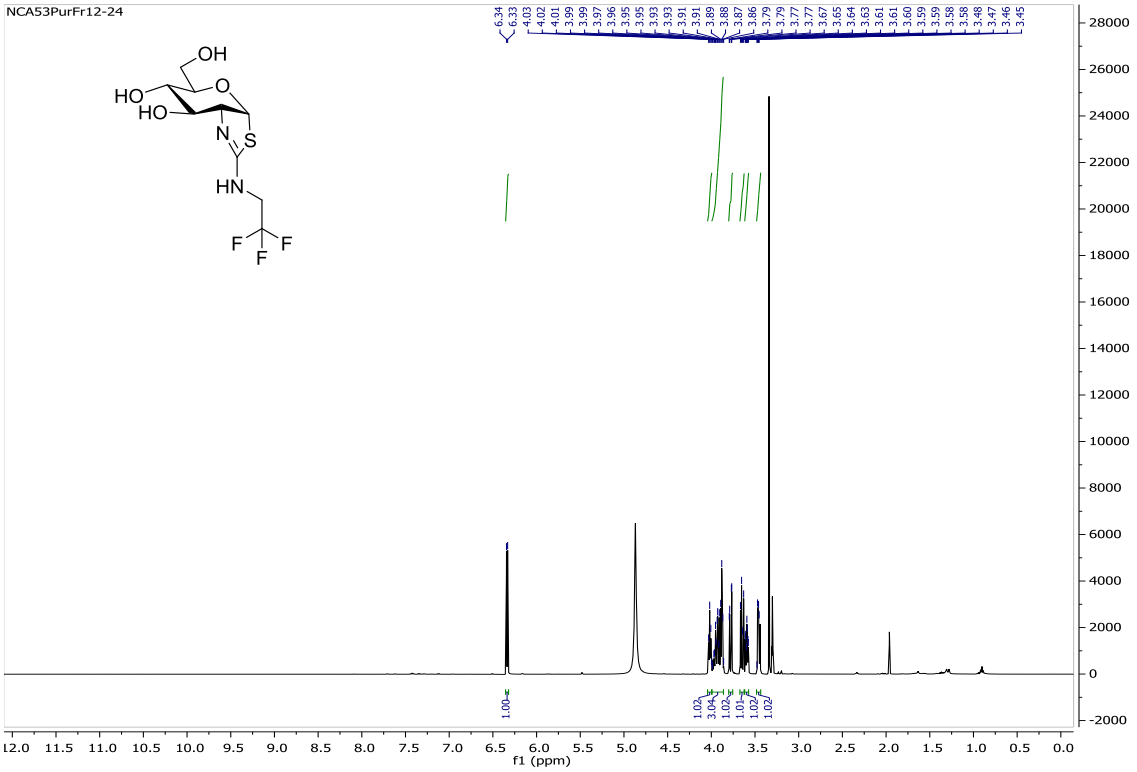
NMR spectra for 2'-aminothiazoline inhibitors and urea substrates:

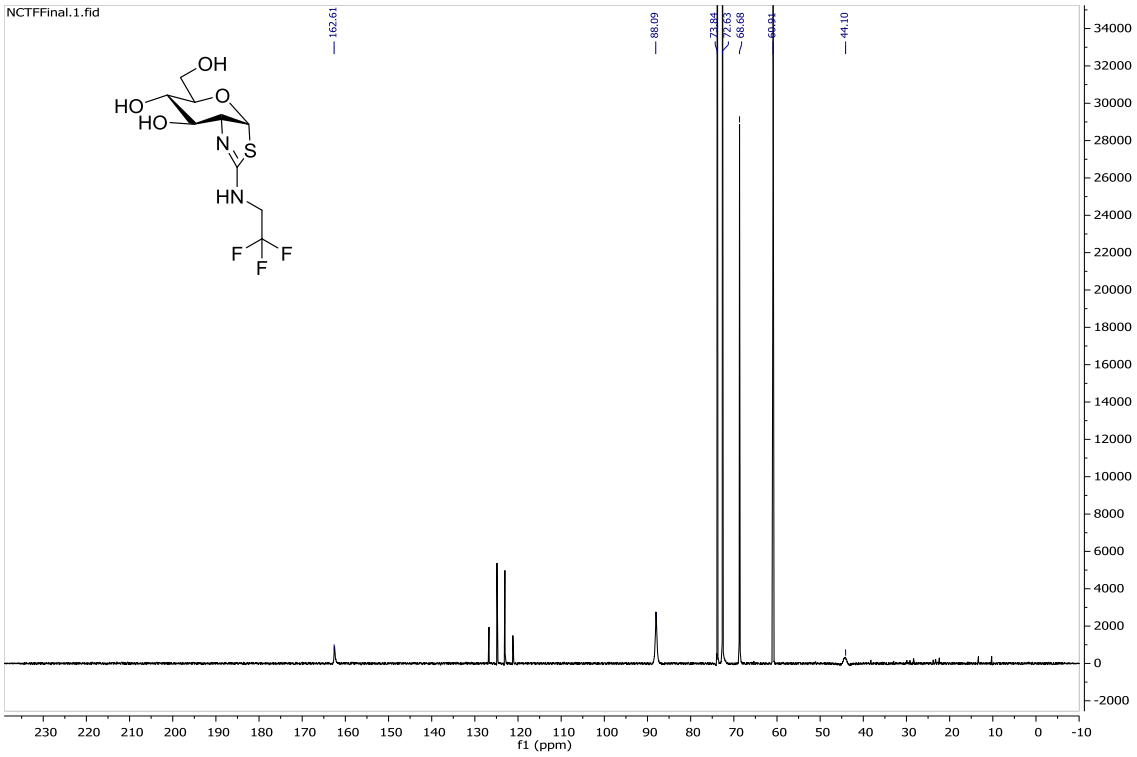


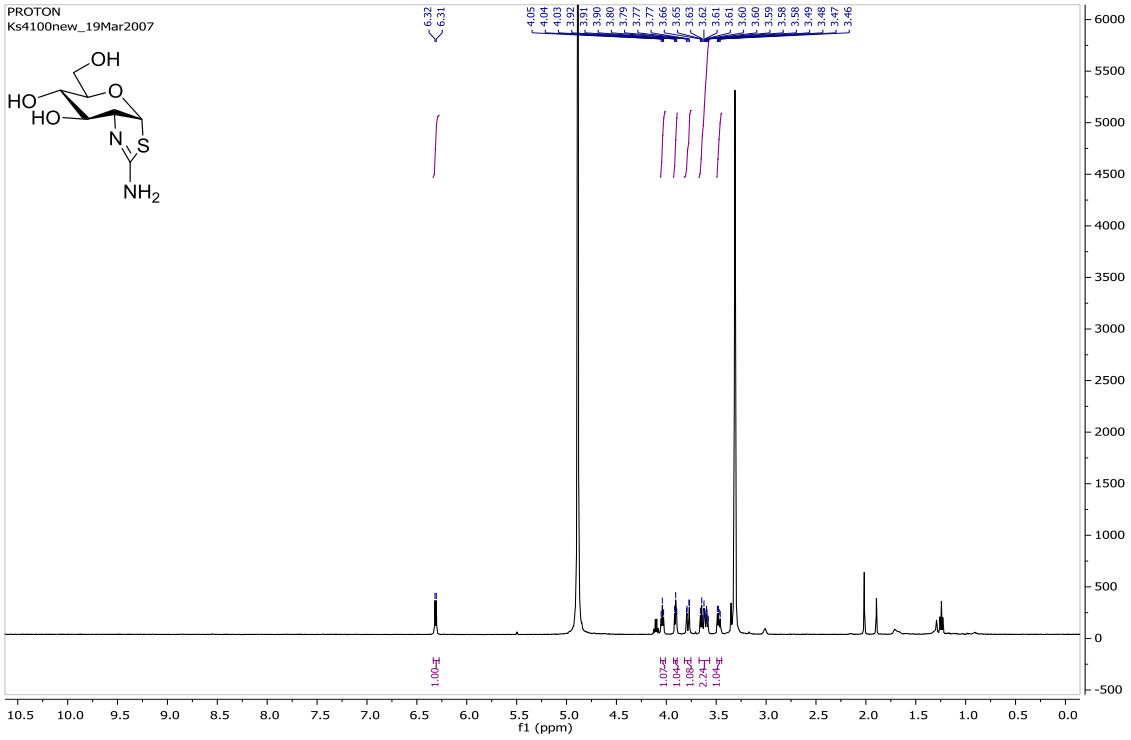


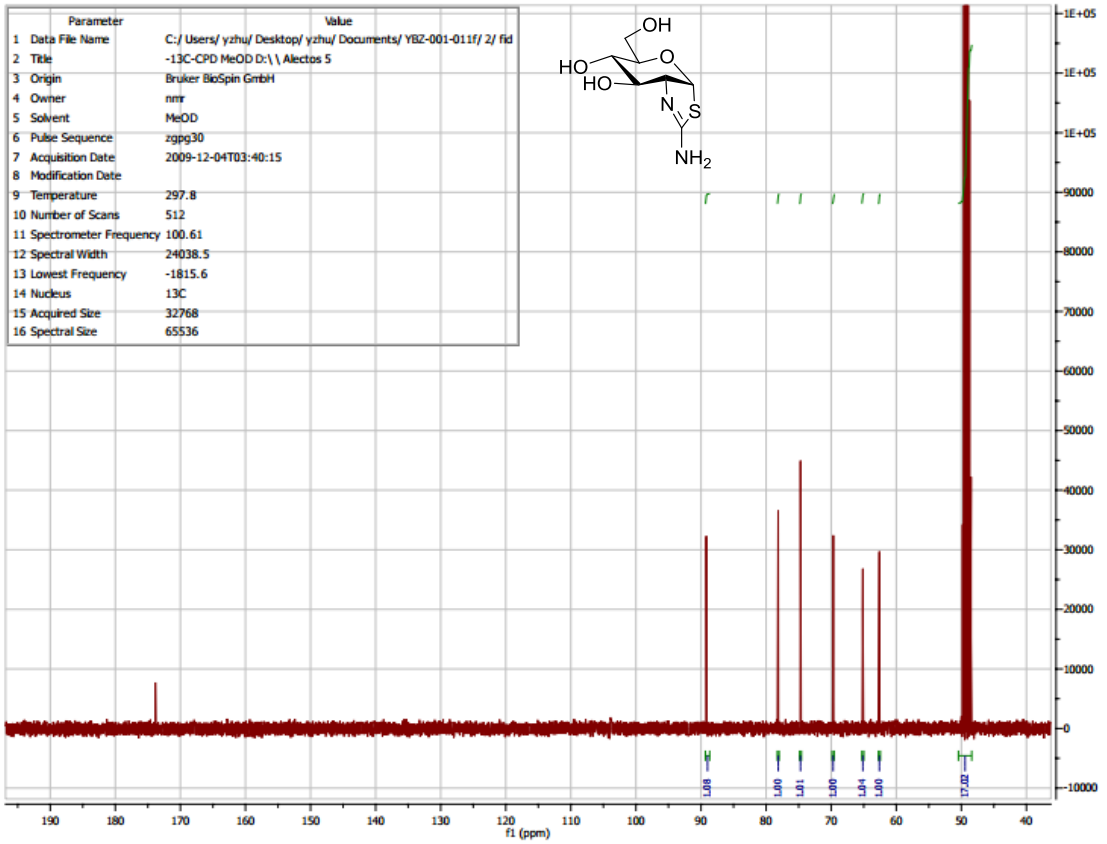

NCMF2.1.fid
OH
HO
HO
O
N
S
HN
F
F
f1 (ppm)

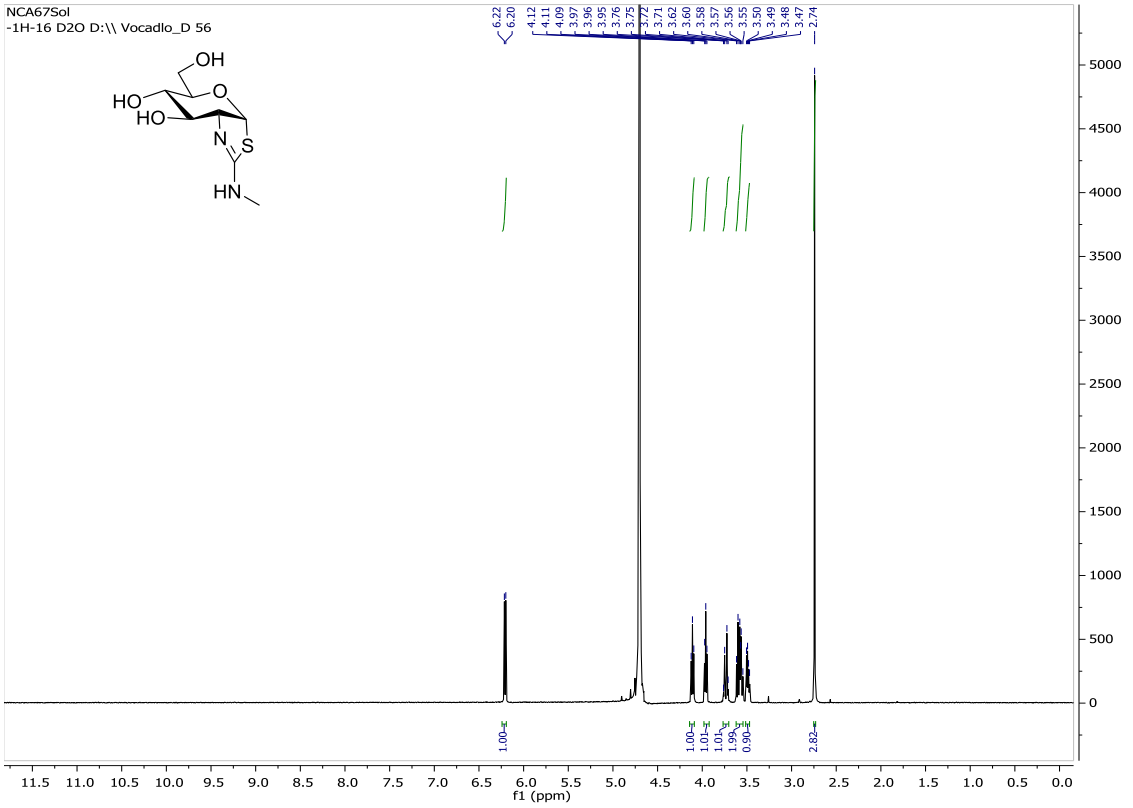

NCDFFinal.1.fid
162.48
115.40
113.81
112.23
88.11
73.79
72.58
68.78
61.00
48.39
45.11
44.95
44.78
A (t)
113.81
B (t)
44.95
OH
HO
HO
O
N
S
HN
F
F
f1 (ppm)

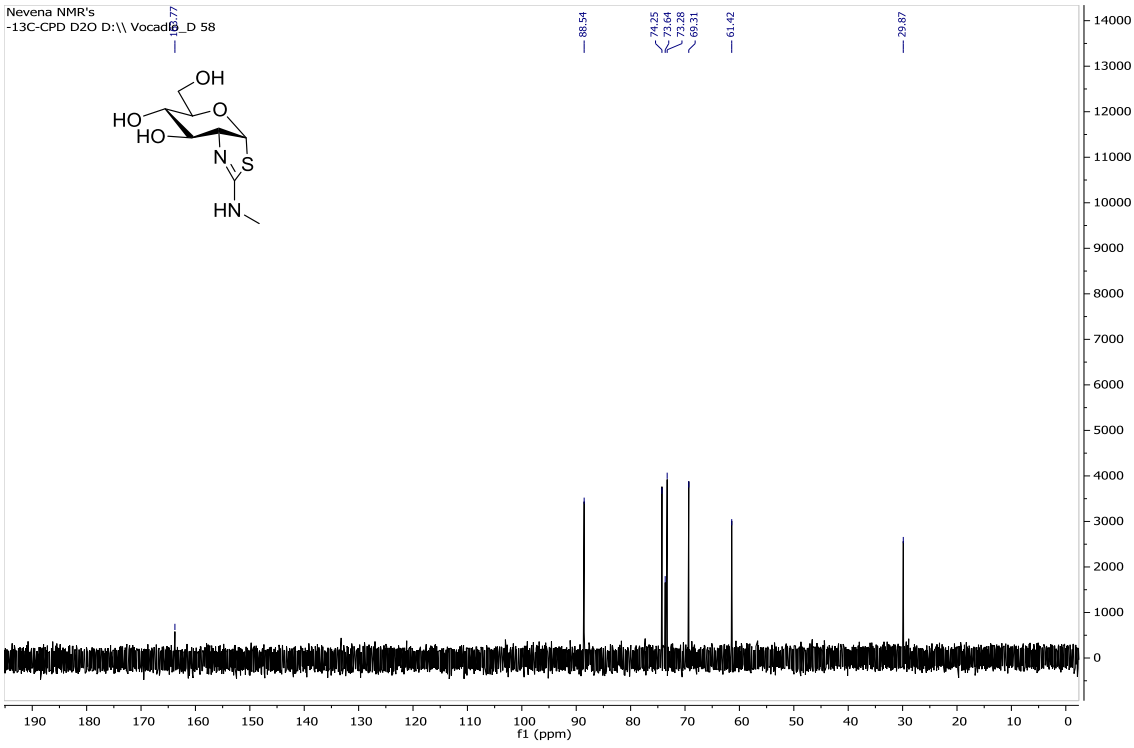

NCA53PurFr12-24
OH
HO
HO
O
N
S
HN
F
F
F
f1 (ppm)

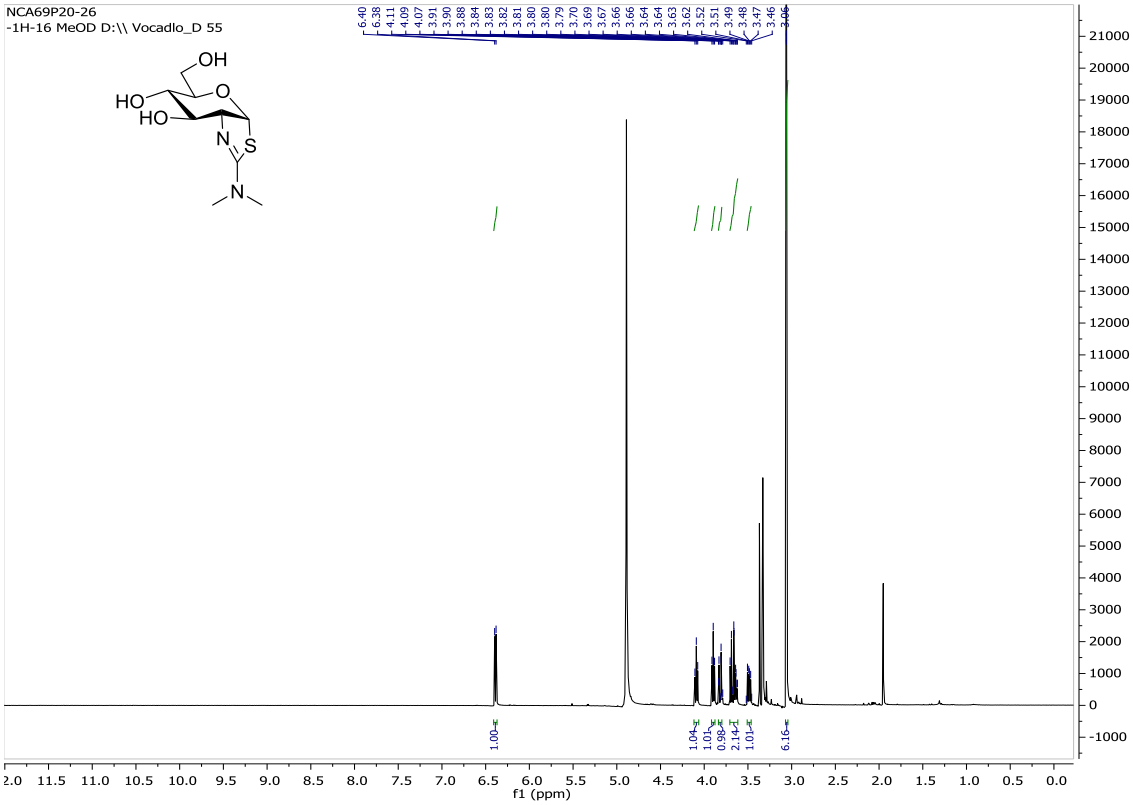

NCTFFinal.1.fid
162.61
88.09
73.84
72.63
68.68
60.01
44.10
OH
HO
HO
O
N
S
HN
F
F
F
f1 (ppm)

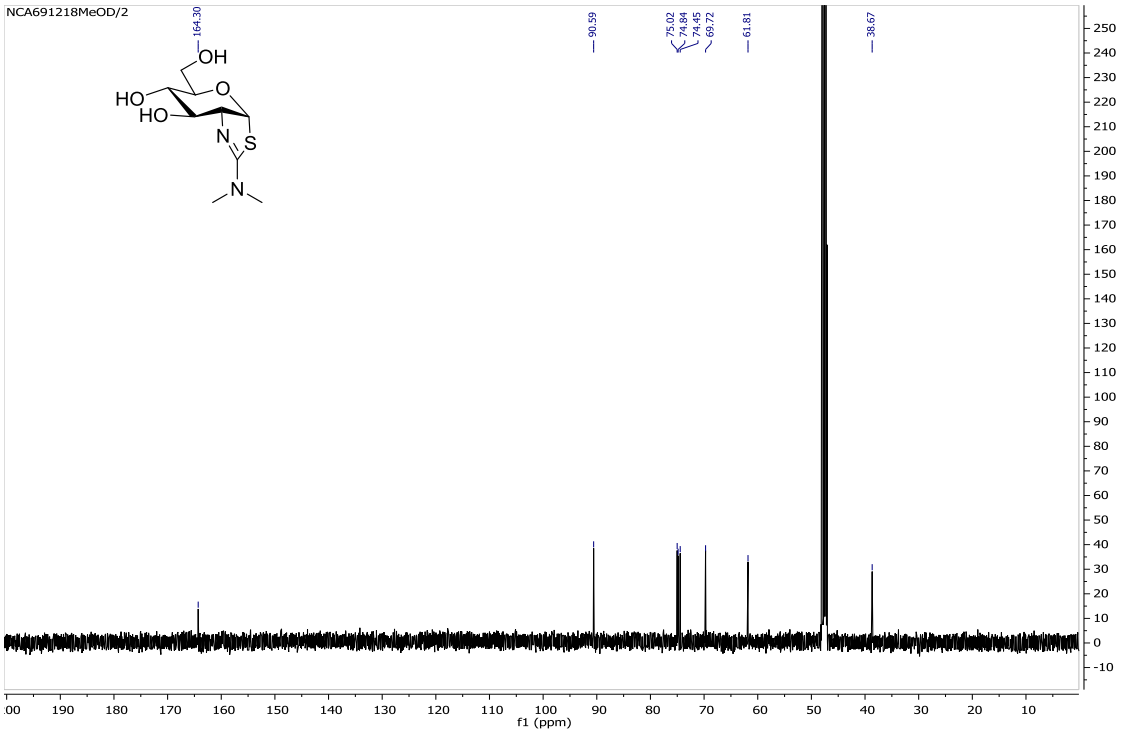

PROTON
Ks4100new_19Mar2007
OH
HO
HO
O
N
S
NH2
f1 (ppm)

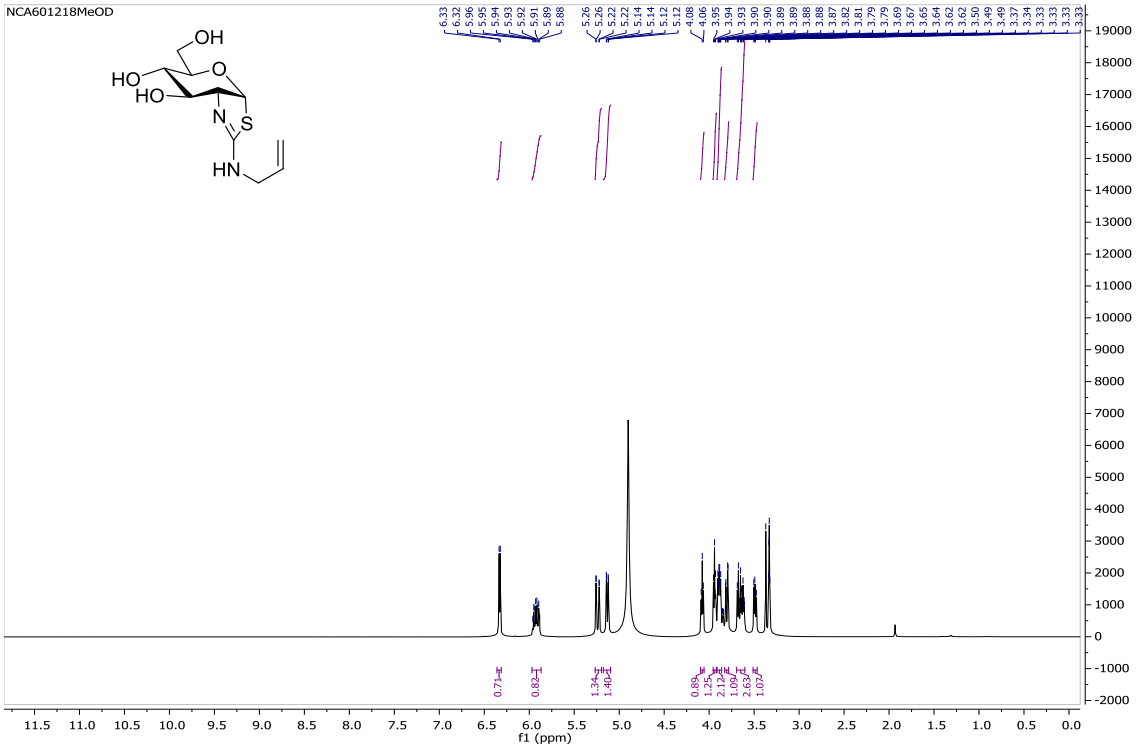

Parameter Value
1 Data File Name C:/ Users/ yzhu/ Desktop/ yzhu/ Documents/ YBZ-001-011f/ 2/ fid
2 Title -13C-CPD MeOD D:\ \ Alectos 5
3 Origin Bruker BioSpin GmbH
4 Owner nmr
5 Solvent MeOD
6 Pulse Sequence zgpg30
7 Acquisition Date 2009-12-04T03:40:15
8 Modification Date
9 Temperature 297.8
10 Number of Scans 512
11 Spectrometer Frequency 100.61
12 Spectral Width 24038.5
13 Lowest Frequency -1815.6
14 Nucleus 13C
15 Acquired Size 32768
16 Spectral Size 65536
OH
HO
HO
O
N
S
NH2
f1 (ppm)

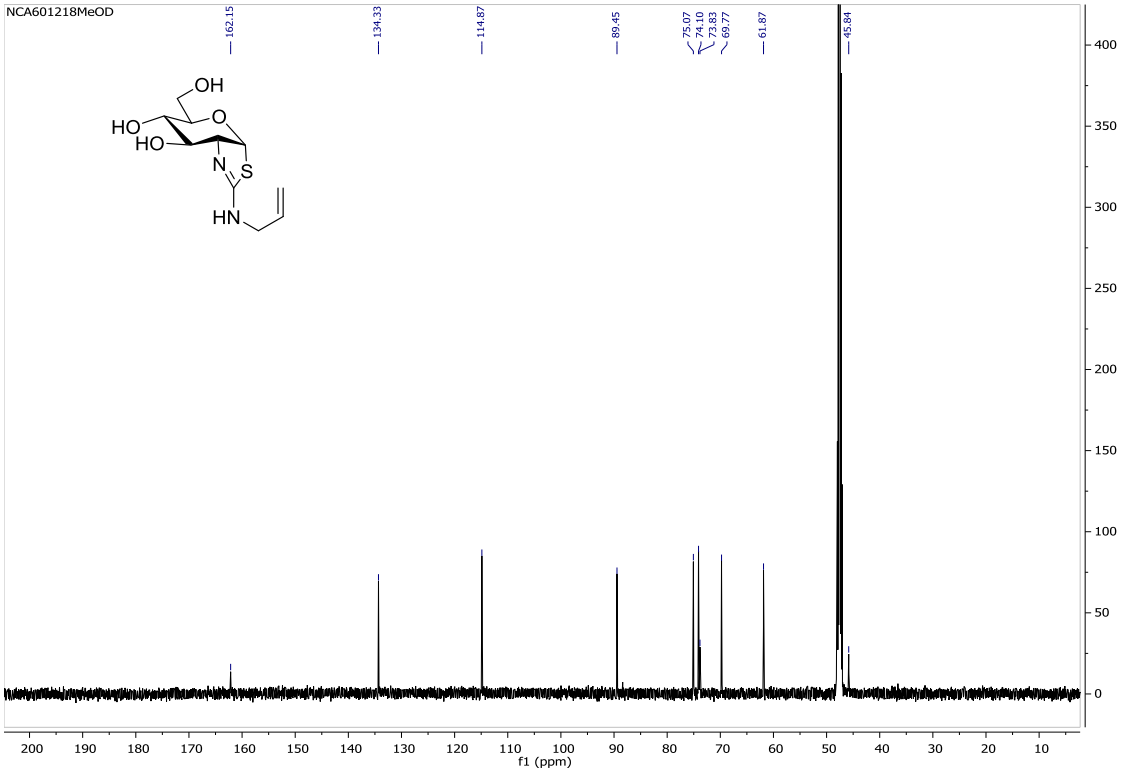

NCA67Sol
-1H-16 D2O D:\\ Vocadlo_D 56
OH
HO
HO
O
N
S
HN
6.22
6.20
4.12
4.11
4.09
3.97
3.96
3.95
3.76
3.75
3.72
3.71
3.62
3.60
3.58
3.57
3.56
3.55
3.50
3.49
3.48
3.47
2.74
1.00
1.00
1.01
1.01
1.99
0.90
2.82
f1 (ppm)

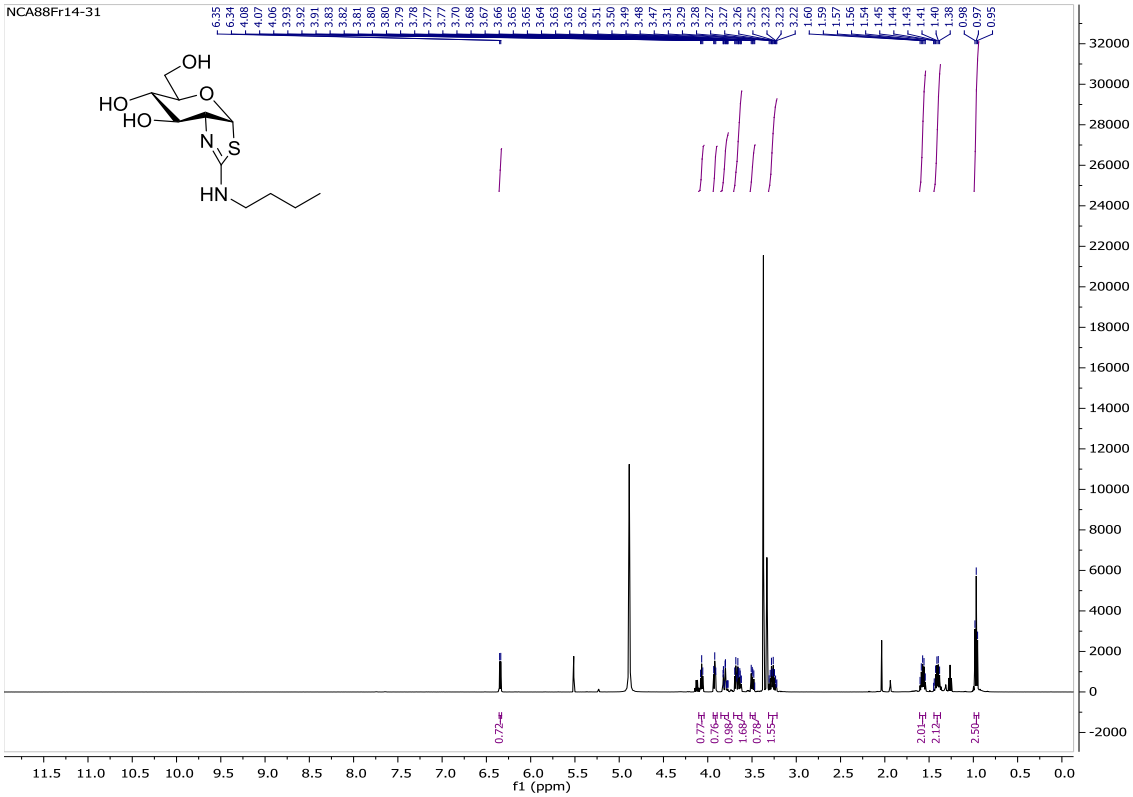

Nevena NMR's
-13C-CPD D2O D:\\ Vocadlo_D 58
OH
HO
HO
O
N
S
HN
156.77
88.54
74.25
73.64
73.28
69.31
61.42
29.87
f1 (ppm)

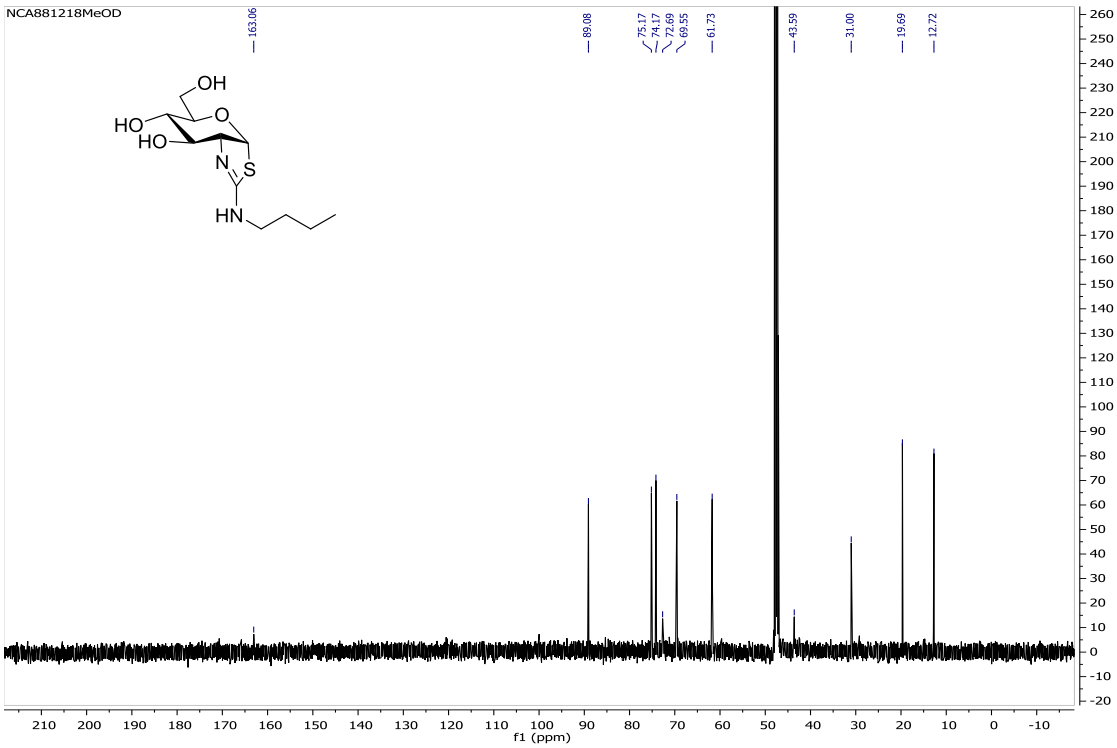

NCA69P20-26
-1H-16 MeOD D:\\ Vocadlo_D 55
OH
HO
HO
O
N
S
N
6.40
6.38
4.11
4.09
4.07
3.91
3.90
3.88
3.84
3.83
3.82
3.81
3.80
3.80
3.79
3.70
3.69
3.67
3.66
3.66
3.64
3.64
3.63
3.62
3.52
3.51
3.49
3.48
3.47
3.46
3.06
1.00
1.04
1.01
0.98
2.14
1.01
6.16
f1 (ppm)

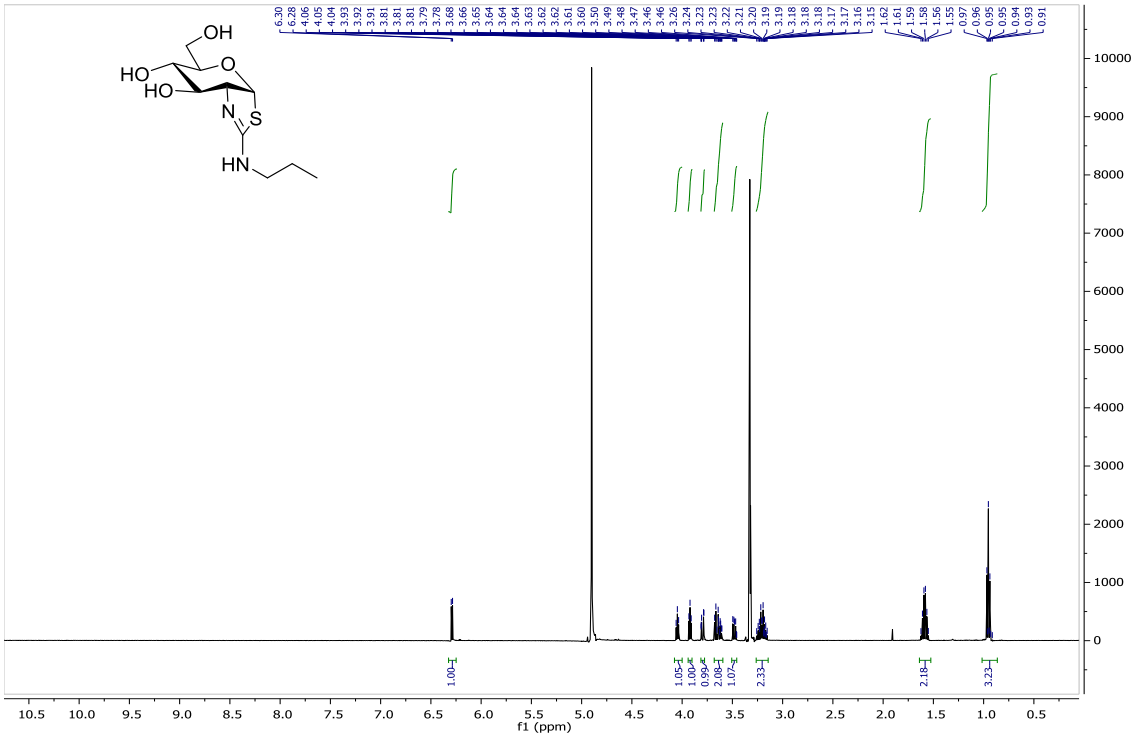

NCA691218MeOD/2
164.30
OH
HO
HO
O
N
S
N
90.59
75.02
74.84
74.45
69.72
61.81
38.67
f1 (ppm)

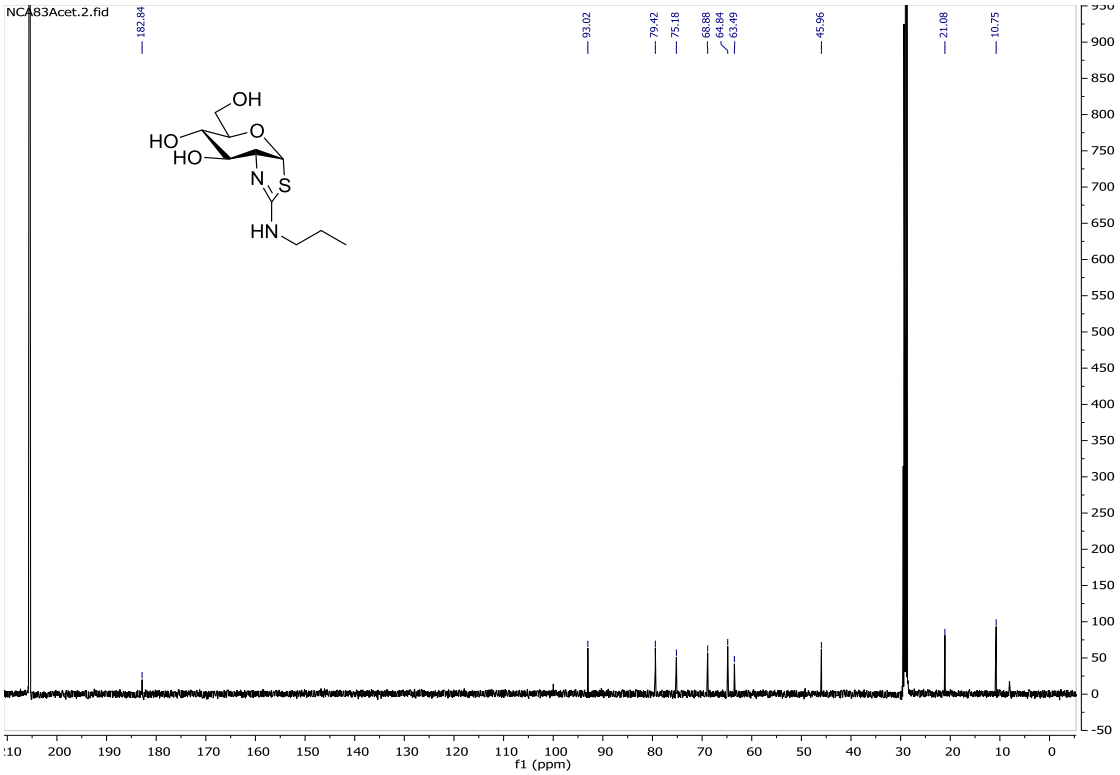

NCA601218MeOD
OH
HO
HO
O
N
S
HN
f1 (ppm)

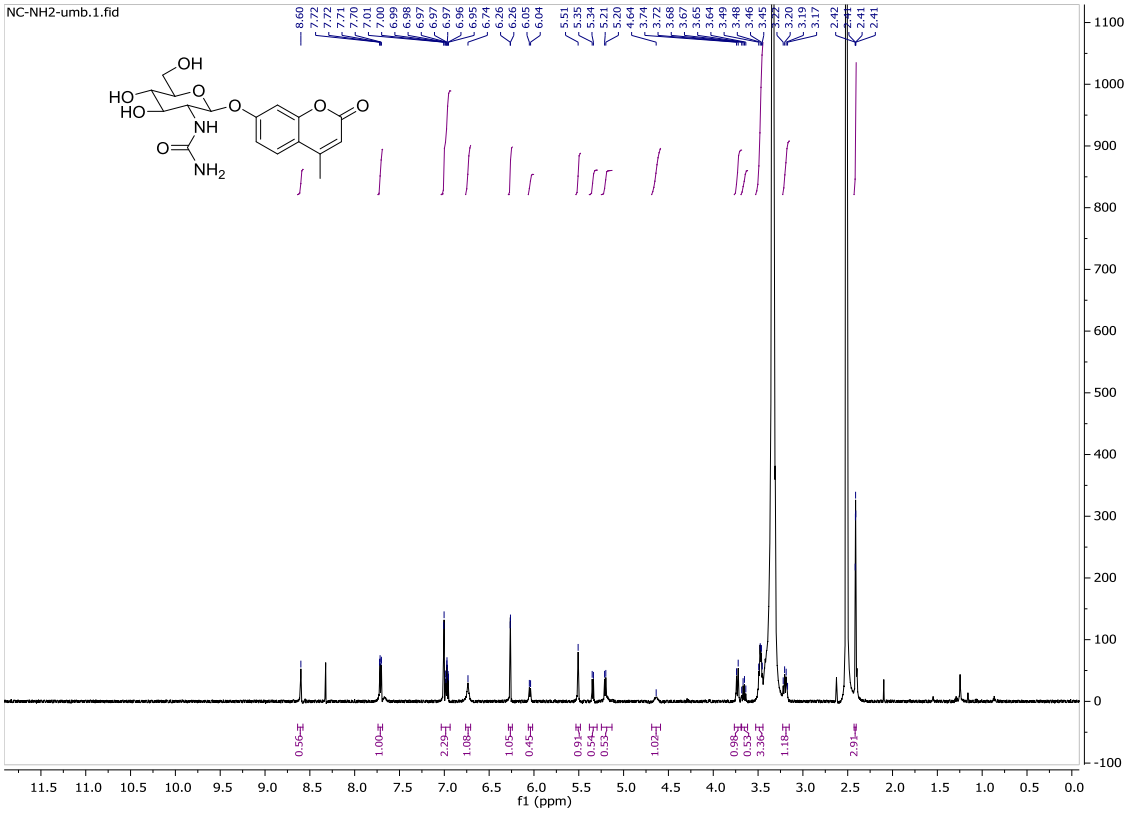

NCA601218MeOD
162.15
134.33
114.87
89.45
75.07
74.10
73.83
69.77
61.87
45.84
OH
HO
HO
O
N
S
HN
f1 (ppm)

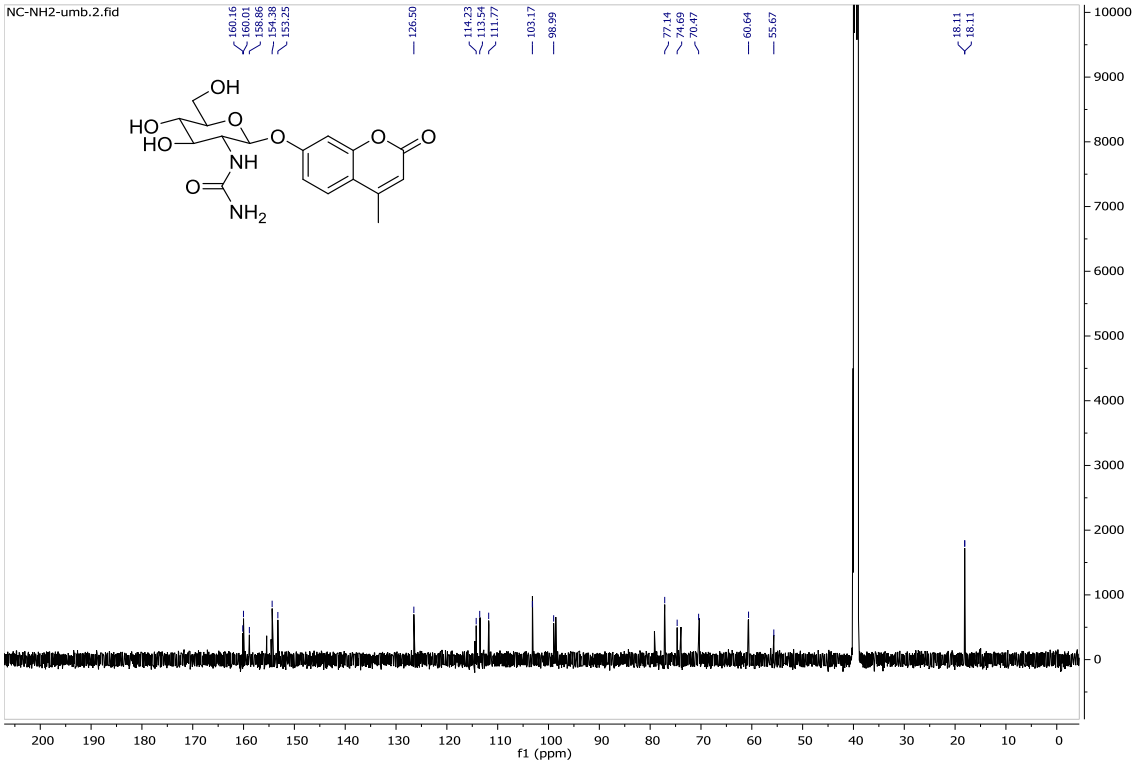

NCA88Fr14-31
OH
HO
HO
O
N
S
HN
f1 (ppm)

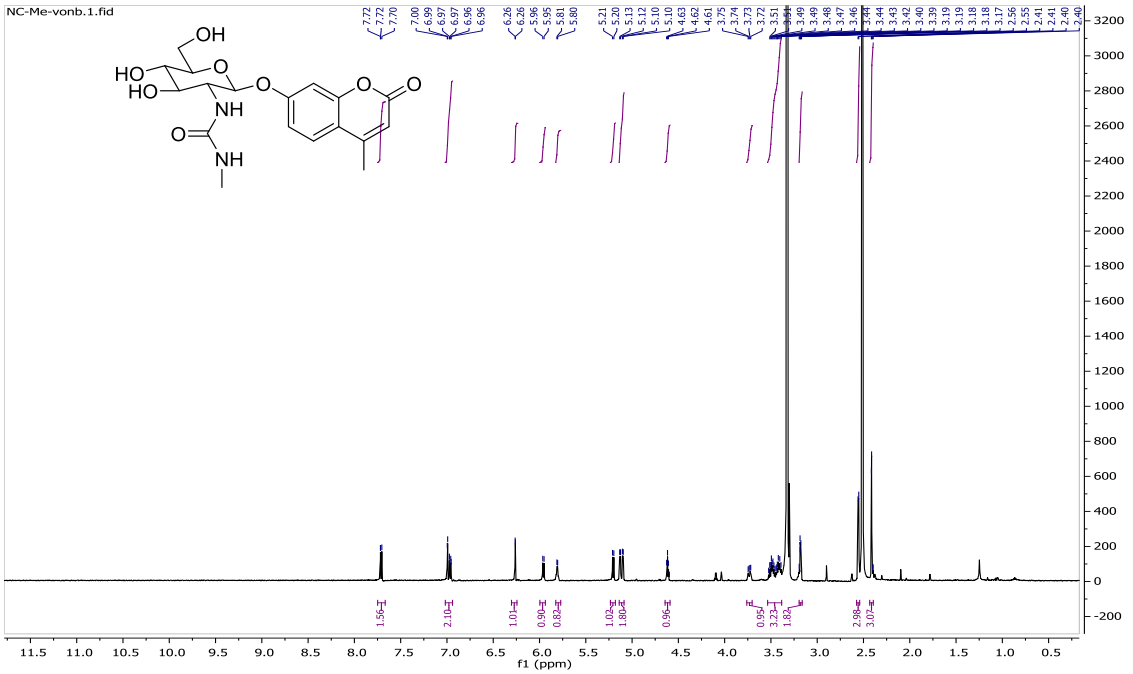

NCA881218MeOD
163.06
89.08
75.17
74.17
72.69
69.55
61.73
43.59
31.00
19.69
12.72
OH
HO
HO
O
N
S
HN
f1 (ppm)

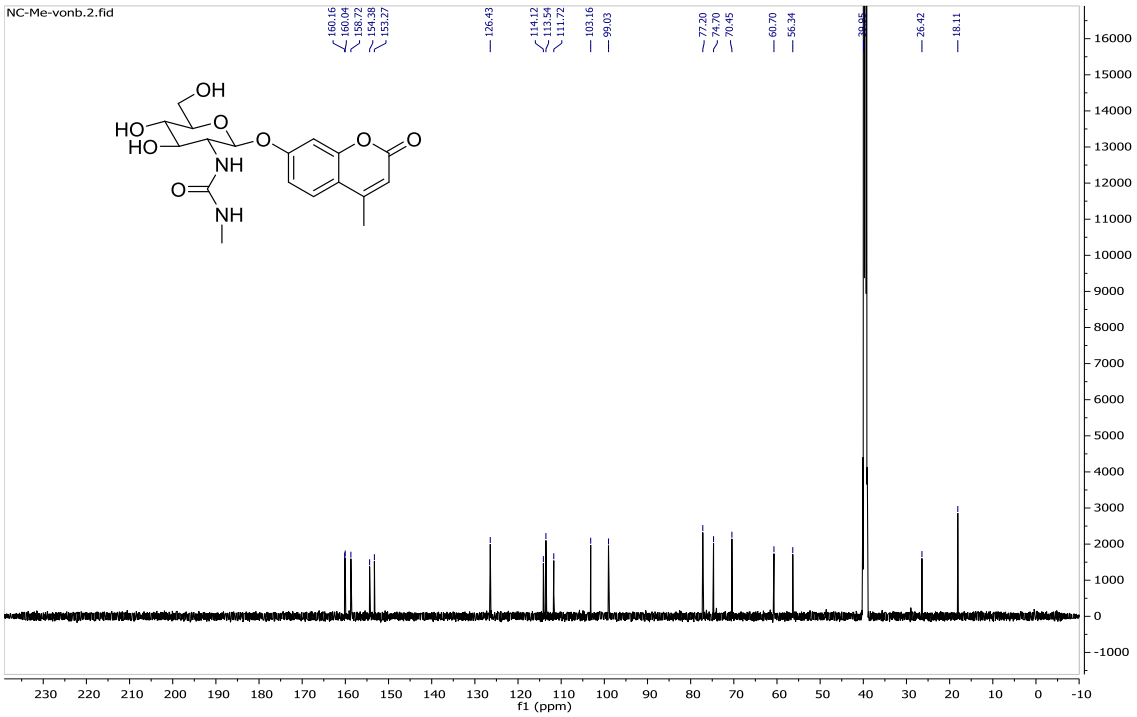

OH
HO
HO
O
N
S
HN
f1 (ppm)

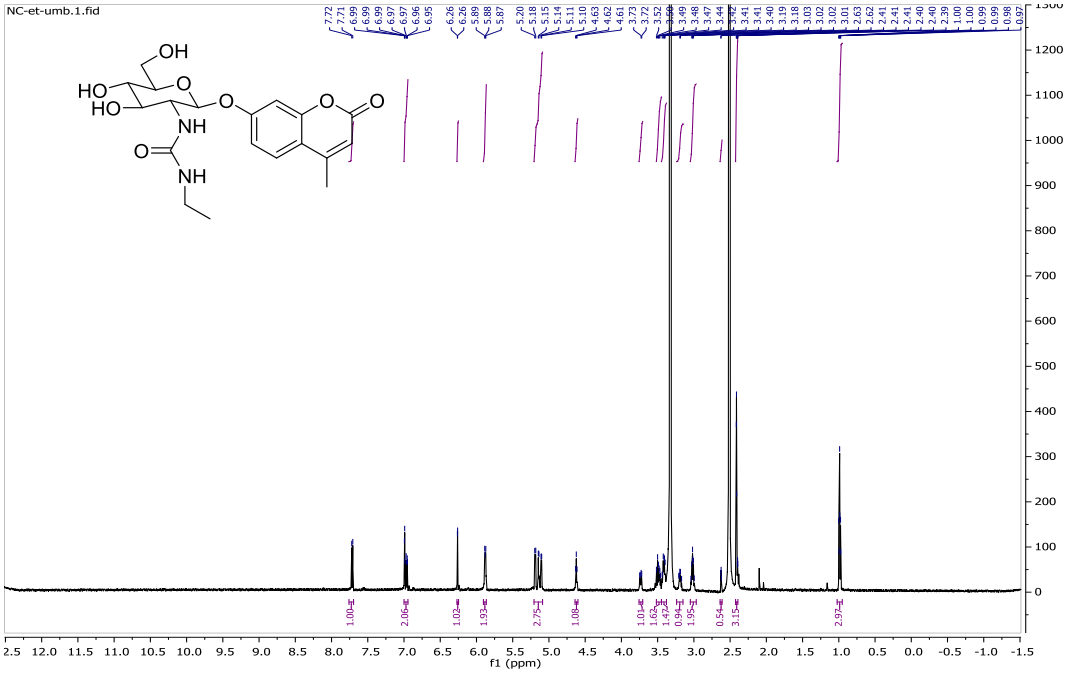

NCA83Acet.2.fid
182.84
93.02
79.42
75.18
68.88
64.84
63.49
45.96
21.08
10.75
OH
HO
HO
O
N
S
HN
f1 (ppm)

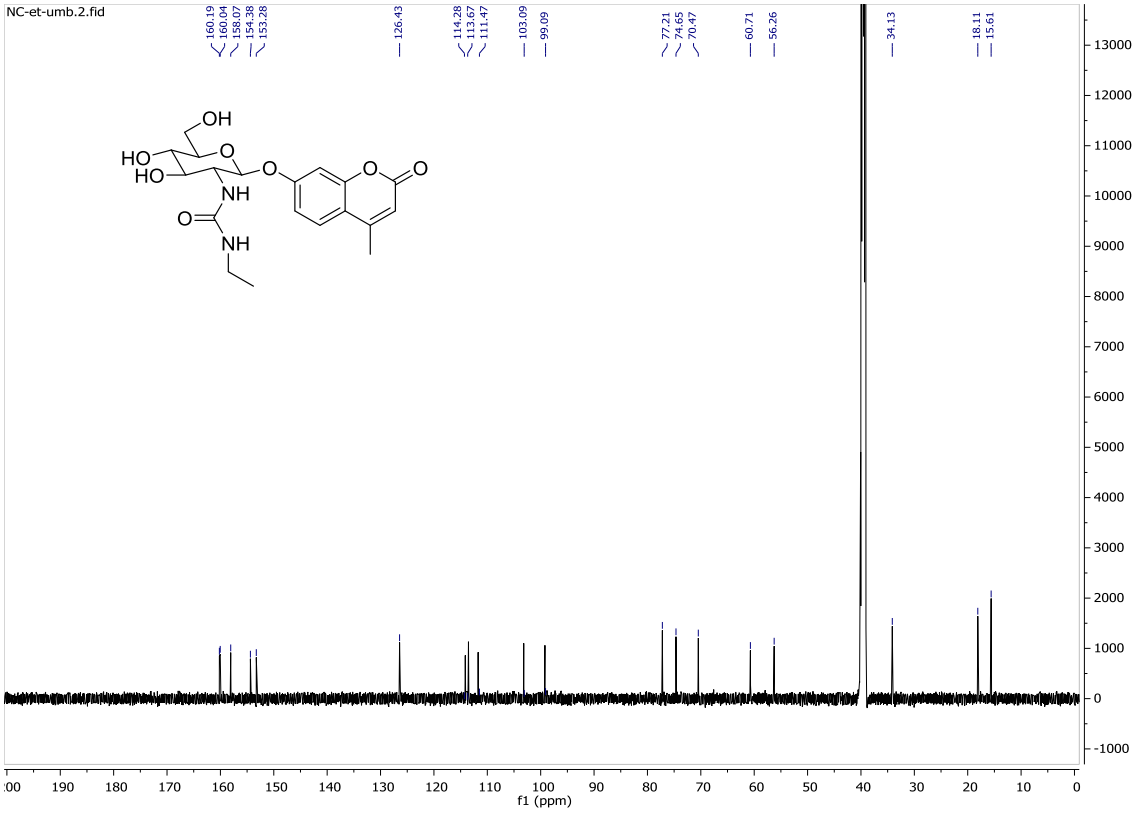

NC-NH2-umb.1.fid
f1 (ppm)

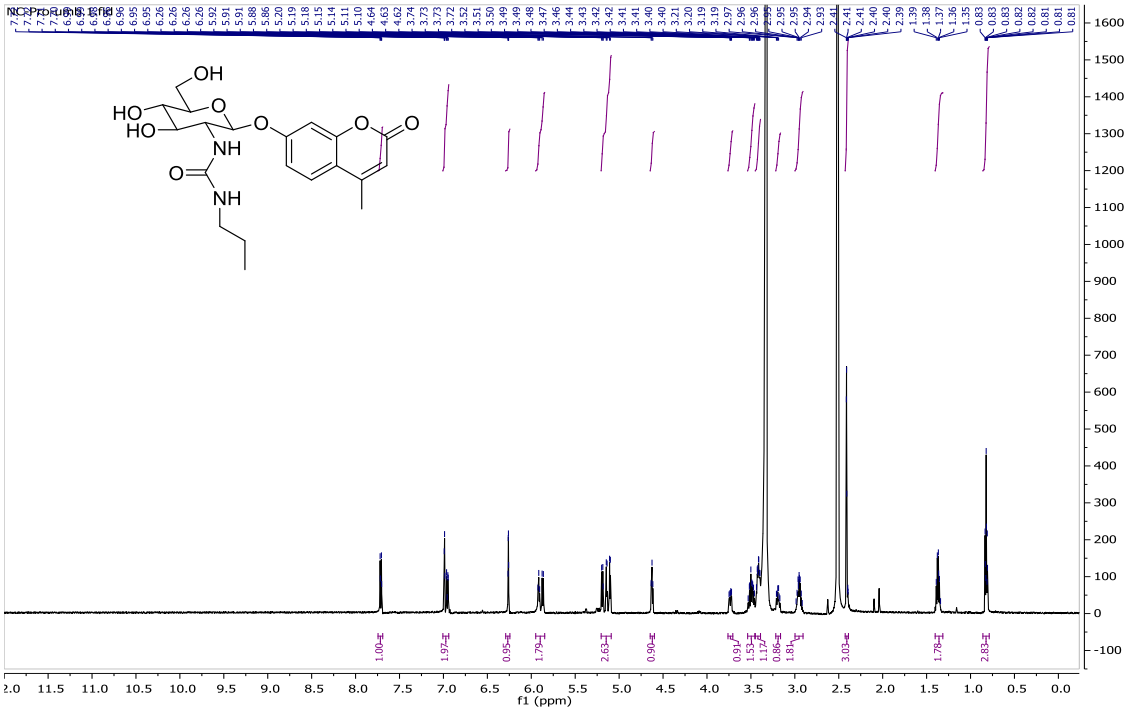

NC-NH2-umb.2.fid
f1 (ppm)

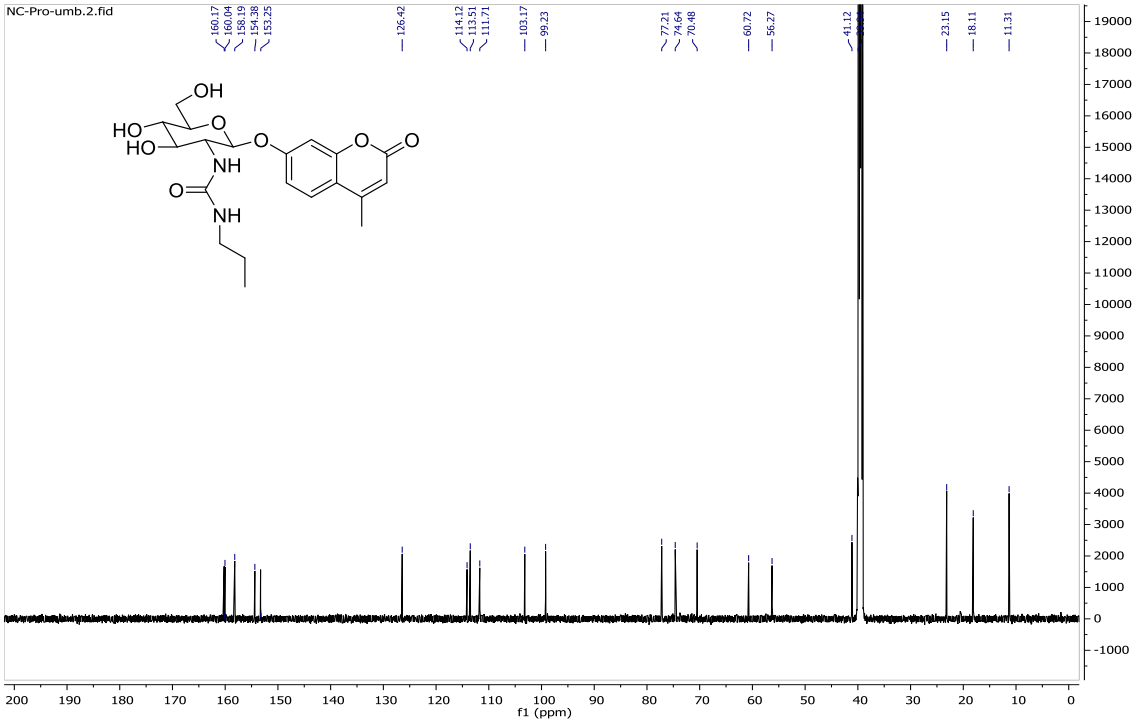

NC-Me-vonb.1.fid
f1 (ppm)

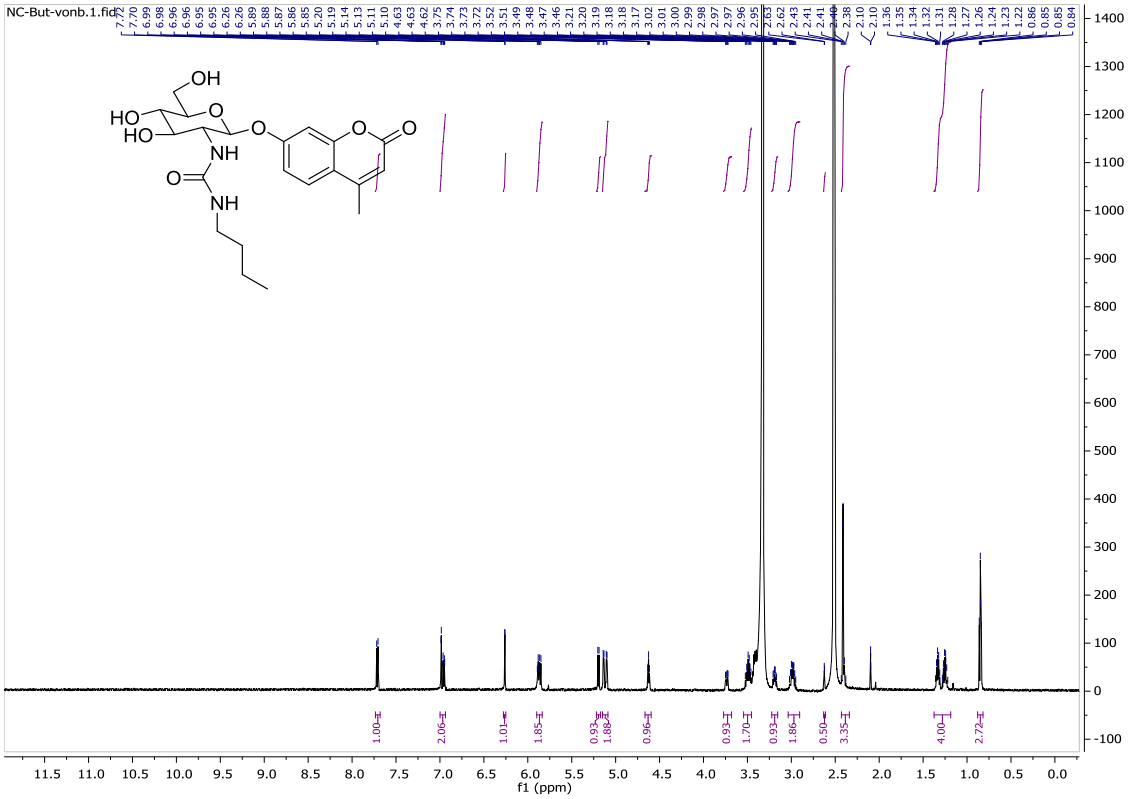

NC-Me-vonb.2.fid
f1 (ppm)

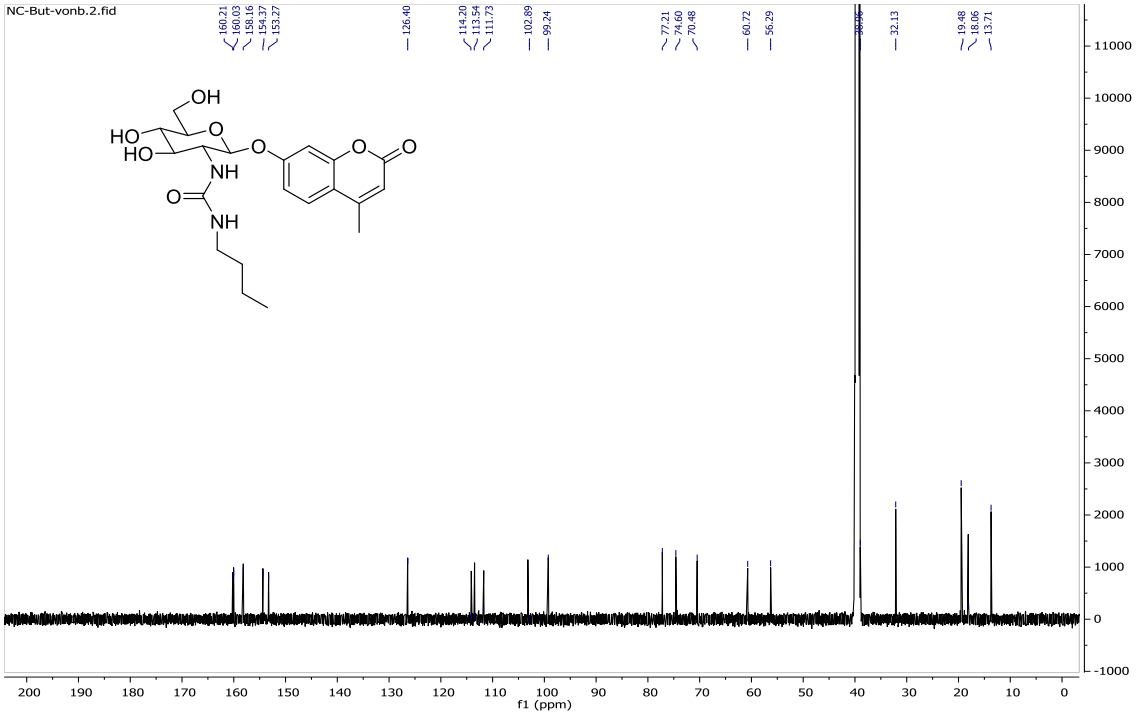

NC-et-umb.1.fid
OH
HO
HO
O
O
NH
O
NH
O
O
f1 (ppm)


NC-et-umb.2.fid
160.19
160.04
158.07
154.38
153.28
126.43
114.28
113.67
111.47
103.09
99.09
77.21
74.65
70.47
60.71
56.26
34.13
18.11
15.61
OH
HO
HO
O
O
NH
O
NH
O
O
f1 (ppm)


NC-Pro-umb.2.fid
OH
HO
HO
O
O
NH
O
NH
O
O
f1 (ppm)


NC-Pro-umb.2.fid
160.17
160.04
158.19
154.38
153.25
126.42
114.12
113.51
111.71
103.17
99.23
77.21
74.64
70.48
60.72
56.27
41.12
23.15
18.11
11.31
OH
HO
HO
O
O
NH
O
NH
O
O
f1 (ppm)

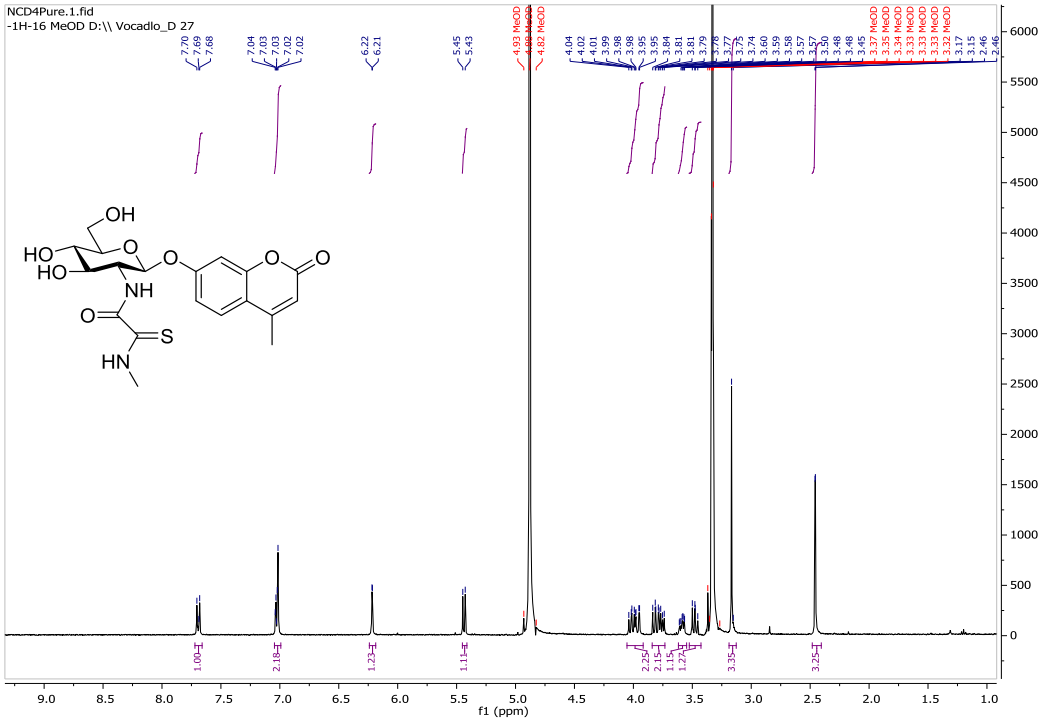

NC-But-vonb.1.fid
OH
HO
HO
O
O
NH
O
NH
O
O
f1 (ppm)

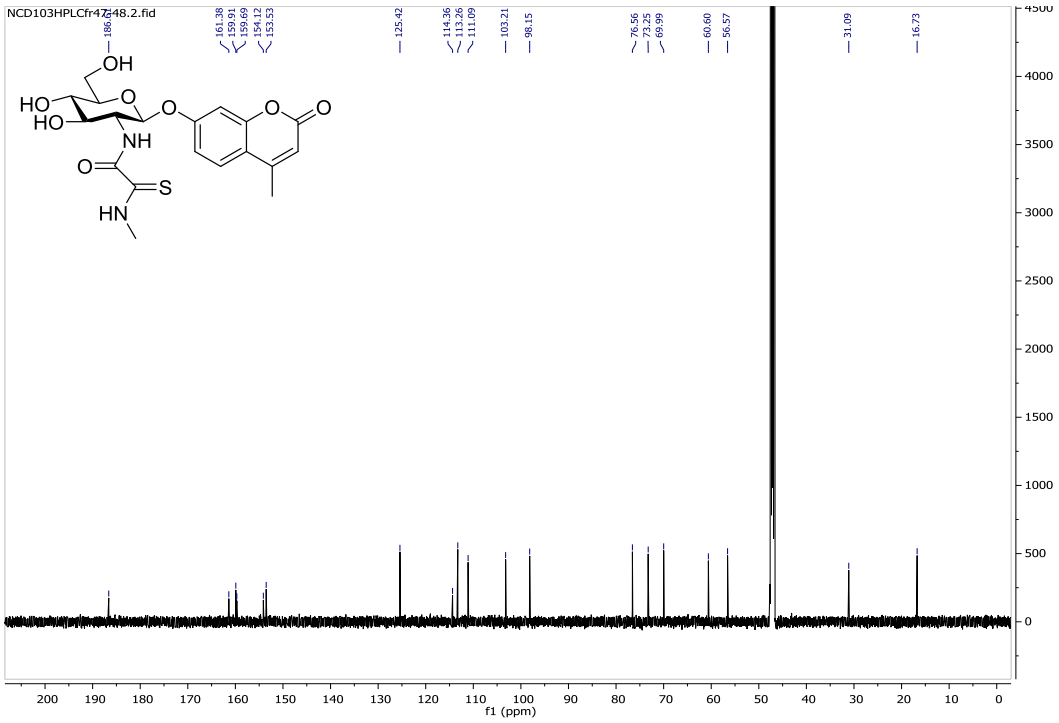

NC-But-vonb.2.fid
160.21
160.03
158.16
154.37
153.27
126.40
114.20
113.54
111.73
102.89
99.24
77.21
74.60
70.48
60.72
56.29
32.13
19.48
18.06
13.71
OH
HO
HO
O
O
NH
O
NH
O
O
f1 (ppm)

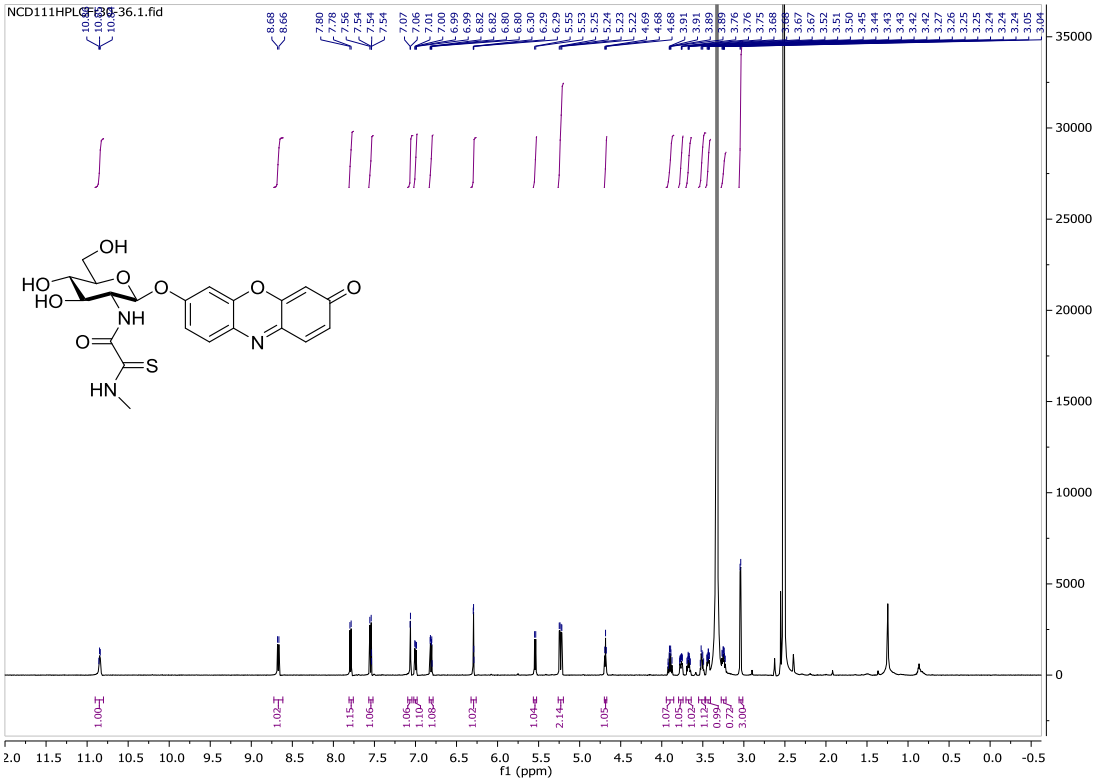
Appendix B.

Supporting information for Chapter 3

NMR spectra of substrates 11 and 19


NCD4Pure.1.fid
-1H-16 MeOD D:\\ Vocadlo_D 27
f1 (ppm)


NCD103HPLCfr47_48.2.fid
f1 (ppm)


NCD111HPLCF3G-36.1.fid
f1 (ppm)