Development of Tools and Methods for Studying Glycan Processing Proteins in Living Systems

by

George Evan Perley-Robertson

B.Sc., University of Ottawa, 2013

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

in the
Department of Chemistry
Faculty of Science

© George Evan Perley-Robertson 2016

SIMON FRASER UNIVERSITY

Summer 2016

All rights reserved.
However, in accordance with the *Copyright Act of Canada*, this work may be reproduced, without authorization, under the conditions for Fair Dealing. Therefore, limited reproduction of this work for the purposes of private study, research, education, satire, parody, criticism, review and news reporting is likely to be in accordance with the law, particularly if cited appropriately.

Approval

Name: **George Evan Perley-Robertson**

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

Title: ***Development of Tools and Methods for Studying Glycan Processing Proteins in Living Systems***

Chair: Dr. Krzysztof Starosta
Associate Professor

Examining Committee:

Dr. David J. Vocadlo
Senior Supervisor
Professor

Dr. Andrew J. Bennet
Supervisor
Professor

Dr. Robert N. Young
Supervisor
Professor

Dr. Roger Linington
Internal Examiner
Associate Professor

Date Defended/Approved: May 12, 2016

Abstract

Carbohydrates are a class of biomolecules present in all domains of life that provide energy for cellular processes, afford structural support, and take part in molecular recognition and signalling. Given the ubiquity of carbohydrates in living systems, gaining an improved understanding of the proteins that process them – glycosyl transferases, glycoside hydrolases, lectins, and sugar transporters – is of key interest. Compared to *in vitro* assays, few live-cell or *in vivo* assays of carbohydrate-processing proteins have been developed, despite the wealth of knowledge that they provide. This discrepancy is largely due to the difficulties associated with live-cell and *in vivo* examination of protein function, namely issues of substrate selectivity, sensitivity, reactivity, and cell permeability. This thesis aims to develop substrates and methods to study two carbohydrate-processing proteins: human *O*-GlcNAcase, a glycoside hydrolase involved in Alzheimer's disease, cancer, and the stabilization of nascent proteins; and bacterial AmpG, a sugar transporter implicated in β-lactam antibiotic resistance. In doing so, I hope to not only provide insight into the function of these proteins, but to also lay a foundation for live-cell or *in vivo* study of these and related proteins in the years to come.

Keywords: glycoside hydrolase *O*-GlcNAcase; sugar transporter AmpG; live-cell and *in vivo*; spheroplast transport assay; PET quenched substrates; β-lactam antibiotic resistance

To my family and friends for their boundless support.

Acknowledgements

Firstly I thank my senior supervisor David Vocadlo for his guidance throughout the course of my degree, and for helping me to realize my potential as a scientist. I also acknowledge the support of my colleagues and collaborators, who have all been integral in the realization of this work. In particular, I thank Samy Cecioni for sharing his knowledge and experience in chemical synthesis and enzymology; Nevena Cekic for working with me to tackle the design of quenched OGA substrates; Anuj Yadav, Judith Winogrodzki, Brian Mark, and Keith Stubbs for their enormous contributions in bringing the AmpG transport assay to fruition; and my committee members Andrew Bennet and Robert Young for their constructive input into my projects and, along with Margo Moore and the researchers at Alectos Therapeutics, for granting me the use of their instruments. Moreover, I would like to thank all of my labmates for harbouring such a positive atmosphere and for being such uplifting people to work with throughout the highs and lows that are often a part of scientific research.

There are many people not directly associated with my research without whom this work would not have been possible. Thank you to all of my friends, both here in the Greater Vancouver Area and across the country, for many a laugh and for plenty of skiing excursions, mountain climbs, game nights, wine & cheeses, camping trips, and general adventures. Thank you to my mother and sister for your unconditional love, support, and understanding. Thank you to my partner Adam, for keeping me grounded throughout it all and for never failing to be there for me, no matter the distance.

Finally, I would like to thank my father, George Perley-Robertson, for raising me to be the person I am today – I know that he would be proud to see how far I've come.

Table of Contents

List of Tables

List of Figures

List of Schemes

Abbreviations

$(Bu_3Sn)_2O$	bis(tributyltin) oxide
$(D_P^+A_P^-)$*	charge-transfer complex
$(NH_4)_6Mo_7O_{24}$	ammonium molybdate
[X]	concentration of substance X
°C	degrees Celsius
3-*O*-MG	3-*O*-methyl-D-glucose
4MU	4-methylumbelliferone
$4MU^+Na^-$	4-methylumbelliferone sodium salt
4MUGlc	4-methylumbelliferyl β-D-glucopyranoside
4MUGlcNAc	4-methylumbelliferyl 2-acetamido-2-deoxy-β-D-glucopyranoside
Å	angstrom
AB	combined blood type A and B
ABC	ATP-binding cassette
ABO	a common system of blood group classification
ABP	activity-based probe
Ac_2O	acetic anhydride
AcBr	acetyl bromide
A_F	FRET acceptor
A_F*	excited state FRET acceptor
AF350	Alexa Fluor 350
AgOTf	silver triflate
AmpC	an inducible β-lactamase found in clinically relevant Gram-negative bacteria
AmpG	a bacterial inner membrane permease responsible for transporting GlcNAc-anhMurNAc from the periplasm to the cytoplasm for recycling
ampG$^-$	AmpG knockout
AmpR	a transcriptional regulator of AmpC β-lactamase
A_P	PET acceptor
A_P*	excited state PET acceptor
Ar	argon

Asp	aspartic acid
ATP	adenosine triphosphate
$Ba(OH)_2$	barium hydroxide
BaO	barium oxide
BHQ-2	black hole quencher 2
BnBr	benzyl bromide
Boc_2O	di-*tert*-butyl dicarbonate
BODIPY	boron dipyrromethene
BW25113	a wild-type *E. coli* strain
C. freundii	*Campylobacter freundii*
CAM	cerium ammonium molybdate
CAZy	carbohydrate active enzyme
CD_3OD	deuterated methanol
$CDCl_3$	deuterated chloroform
$Ce(SO_4)_2$	cerium sulfate
CFTR	cystic fibrosis transmembrane conductance regulator
$CH_2(N_3)_2$	diazidomethane
CH_3COOH	acetic acid
Cl_2Ind	2,6-dichlorophenol-indophenol
CMP	cytidine monophosphate
CO_2	carbon dioxide
CPP	cell-penetrating peptide
$CuSO_4$	copper (II) sulphate
d	doublet
DABCYL	4-(4-dimethylaminophenylazo) benzoic acid
DCM	dichloromethane
dd	doublet of doublets
ddd	doublet of doublets of doublets
ddt	doublet of doublets of triplets
deGlc	2-deoxy-D-glucose
D_F	FRET donor
D_F*	excited state FRET donor
DIC	differential interference contrast

DIPEA	diisopropylethylamine
DMF	dimethylformamide
DMSO	dimethyl sulfoxide
DMSO-d6	deuterated dimethyl sulfoxide
DNA	deoxyribonucleic acid
D_P	PET donor
D_P*	excited state PET donor
dt	doublet of triplets
E	enzyme
E. coli	*Escherichia coli*
E_c	efficiency of collisional quenching
Ec	*Escherichia coli*
*Ec*BW	wild-type *E. coli* strain BW25113
*Ec*MG	wild-type *E. coli* strain MG1655
EDANS	5-[(2-aminoethyl)amino]naphthalene-1-sulfonic acid
EDC·HCl	*N*-ethyl-*N*'-(3-dimethylaminopropyl)-carbodiimide hydrochloride
EDTA	ethylenediaminetetraacetic acid
E_f	free enzyme OR efficiency of FRET quenching
eq	equivalent
ER	endoplasmic reticulum
ERAD	endoplasmic reticulum associated degradation
ES	enzyme-substrate complex
Et_2O	diethyl ether
EtOAc	ethyl acetate
EtOH	ethanol
EWG	electron withdrawing group
EXTL1	exostosin glycosyltransferase-like 1
FRET	Förster resonance energy transfer
g	grams
GAA	α-glucosidase A
GalNAc	*N*-acetylgalactosamine
GCase	glucocerebrosidase

GDP	guanosine diphosphate
GH	glycoside hydrolase
Glc	glucose
GlcNAc	*N*-acetylglucosamine
GlcNAc-anhMurNAc	a peptidoglycan fragment that is a precursor to inducers of AmpC β-lactamase
Glu	glutamic acid
GLUT-4	human glucose transporter 4
Gly	glycine
GPH	galactoside-pentoside-hexuronide
GT	glycosyl transferase
GTP	guanosine triphosphate
H_2	hydrogen
H_2O	water
H_2SO_4	sulfuric acid
HBr	hydrogen bromide
HBTU	*N*,*N*,*N*′,*N*′-tetramethyl-*O*-(1*H*-benzotriazol-1-yl)uronium hexafluorophosphate
HCl	hydrogen chloride
HIV	human immunodeficiency virus
hOGA	human *O*-GlcNAcase
HOSu	*N*-hydroxysuccinimide
HPLC	high performance liquid chromatography
hr	hours
HRMS	high resolution mass spectrometry
HTS	high-throughput screening
human *O*-GlcNAcase	a human enzyme responsible for removing *N*-acetylglucosamine from serine and threonine residues of nucleocytoplasmic proteins
I_2	iodine
IC_{50}	concentration causing 50% of maximum inhibition
ITC	intramolecular charge transfer
$k_{\#}$	a rate constant for a given step (#)
K_2CO_3	potassium carbonate

k_{cat}	turnover number
k_{cat}/K_M	second-order rate constant
K_i	inhibition constant
$K_{i,app}$	apparent inhibition constant
K_M	Michaelis constant
Km	kanamycin resistant
$K_{M,app}$	apparent Michaelis constant
$KMnO_4$	potassium permanganate
K_S	thermodynamic dissociation constant
LB	lysogeny broth
LiOH	lithium hydroxide
LRMS	low resolution mass spectrometry
M	molar
m	multiplet
Man	mannose
MeCN	acetonitrile
MeOH	methanol
MFS	major facilitator superfamily
mg	milligram
MG1655	a wild-type *E. coli* strain
$MgCl_2$	magnesium chloride
MHz	megahertz
min	minutes
mL	millilitre
mm	millimeter
mM	millimolar
mmol	millimol
MS	mass spectrometry OR molecular sieves
MurNAc	*N*-acetylmuramic acid
N_2H_4	hydrazine
Na_2SO_4	sodium sulfate
NAD+	oxidized nicotinamide adenine dinucleotide
NADH	reduced nicotinamide adenine dinucleotide

NADPH	reduced nicotinamide adenine dinucleotide phosphate
NagZ	a bacterial β-glucosaminidase responsible for cleaving the glycosidic linkage of GlcNAc-anhMurNAc
NaH	sodium hydride
$NaHCO_3$	sodium bicarbonate
NaN_3	sodium azide
NaOH	sodium hydroxide
NaOMe	sodium methoxide
NBD-Glc	2-(*N*-(7-nitrobenz-2-oxa-1,3-diazol-4-yl)amino)-2-deoxyglucose
NIR	near infrared
nm	nanometer
NMR	nuclear magnetic resonance
O_2	oxygen
OD600	optical density at 600 nm
OGA	*O*-GlcNAcase
OGT	*O*-GlcNAc transferase
OHS	oligosaccharide:H^+ symporter
OPA	organophosphate:phosphate antiporter
P	product
P. aeruginosa	*Pseudomonas aeruginosa*
Pa	*Pseudomonas aeruginosa*
Pa-AmpG	a *P. aeruginosa* protein with sequence similarity to *E. coli* AmpG
Pa-AmpP	a *P. aeruginosa* protein with sequence similarity to *E. coli* AmpG, does not appear to have a similar function
pBAD	a promotor
PBS	phosphate buffered saline
PBu_3	tributylphosphine
PBu_3O	tributylphosphine oxide
PCR	polymerase chain reaction
Pd/C	activated palladium on charcoal
PET	photoinduced electron transfer
PG	peptidoglycan

pH	negative logarithm of the proton concentration
PhMe	toluene
pK_a	negative logarithm of the acid dissociation constant
PMSF	phenylmethanesulfonyl fluoride
PQQ	pyrroloquinoline quinone
PTS	phosphoenolpyruvate:carbohydrate phosphotransferase system
PTSA	*para*-toluenesulfonic acid
p-TsCl	*para*-toluenesulfonyl chloride
Py	pyridine
q	quartet
rcf	relative centrifugal force
Res-α-Glu	resorufin α-glucoside
R_f	retention factor
RFU	relative fluorescence units
RT	room temperature
s	singlet
S	substrate
$S_{\#}$	singlet electronic energy level #
SAC	substrate-assisted catalysis
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
S_f	free substrate
sGDH	soluble quinoprotein glucose dehydrogenase
t	triplet
$T_{\#}$	triplet electronic energy level #
TBAF	tetrabutylammonium fluoride
TBS	tris(hydroxymethyl)aminomethane buffered saline
TBSCl	*tert*-butyldimethylsilyl chloride
TCEP	tris(2-carboxyethyl)phosphine
TEA	trimethylamine
TFA	trifluoroacetic acid
THF	tetrahydrofuran
TLC	thin layer chromatography

Trp	tryptophan
UDP	uridine diphosphate
UGGT	UDP-glucose:glycoprotein glucosyltransferase
UV	ultraviolet
UV-vis	ultraviolet-visible
v_d	rate of decay
v_f	rate of formation
V_{max}	maximum rate
δ	chemical shift
Δ*ampG*	AmpG knockout
µL	microliter
τ	lifetime of an excited state

Chapter 1.

Introduction

1.1. Carbohydrates and Their Roles in Living Systems

It is without question that the chemistry that occurs in living systems is immensely complicated. However, aspects of the chemistry of life can, largely, be thought of as the interactions between four molecular classes: nucleic acids, proteins, lipids, and carbohydrates. While the structure and function of nucleic acids and proteins are relatively well characterized, carbohydrates remain arguably the least well understood of the four major classes of biological molecules. The importance of carbohydrates in living systems is highlighted by estimates that they make up 50% of the dry weight of Earth's biomass.[1] Not only do carbohydrates provide a critical energy source for cells, they also comprise polymers that afford structural integrity and act as key protein modifications and effectors of numerous processes including cell adhesion, recognition, and signalling.[1] The plethora of functional roles that can be fulfilled by carbohydrates stems from their potential for structural complexity. In fact, for a chain composed of six monosaccharides the number of all linear and branched isomers is on the order of 10^{12}, many orders of magnitude greater than the number of possible isomers for a similar length chain of any other biological molecule.[2] It is for this reason that, while many techniques exist for the characterization and high-throughput construction of nucleic acid and amino acid polymers, progress in this regard for carbohydrates has lagged behind. In this section, a brief history and a description of carbohydrate structure will be given to provide the reader a foundation in carbohydrate terminology. A discussion of some of the biological functions of carbohydrates across the domains of life will also be described, as this will offer insight into the ubiquity of carbohydrates and will emphasise the importance of continued research into their structures and functions.

1.1.1. Carbohydrate Structure

The term ‘carbohydrate’ was coined as a descriptive term, as at the time of its inception all known molecules within this class were thought of as simple hydrates of carbon, with the generic formula $(CH_2O)_n$. Today the definition has broadened slightly, and carbohydrates are defined to be polyhydroxylated aldehydes/ketones of at least three carbons, or compounds which can be derived from these by oxidation or reduction of the carbonyl group, or by replacement of the hydroxyl groups with other substituents.[3] Much of the first work on carbohydrates was carried out by Emil Fischer who, following the works of Van’t Hoff and Le Bel who postulated the asymmetry of carbon atoms, elucidated the stereochemical configurations of many sugars including glucose, mannose, and gulose. As an accurate and descriptive way to represent these chiral molecules in two dimensions, Fischer developed the projection which bears his name and which still sees use to this day as a way of representing carbohydrates. Although the popularity of the Fischer projection has waned in recent years in favour of the more conformationally informative ‘chair’ representation, the Fischer projection remains the basis for defining the absolute stereochemical configuration of a sugar. To represent a sugar using the Fischer projection, the carbon backbone is drawn as a vertical line oriented such that the aldehyde/ketone is positioned as close as possible to the top of the chain, and substituents are indicated by horizontal lines perpendicular to the carbon backbone. The atoms of the backbone are numbered sequentially, with the topmost carbon being designated as C-1 (Figure 1.1). The vertical lines comprising the carbon backbone project into the page, and can be imagined to curve around to form the shape of a ‘C’ looked at edge-on, while the horizontal lines project out of the page toward the reader.

Every sugar has an enantiomer, a nonsuperimposible mirror image, resulting from the inversion of all the stereocentres along the carbon backbone. These two possible configurations are termed ‘D’ or ‘L’ based on the stereochemistry of the chiral carbon furthest away from the carbonyl group in the Fischer projection, which itself is known as the configurational atom. If the substituent bound to the configurational atom is positioned on the right-hand side of the carbon backbone, the sugar is said to be the ‘D’ configuration (from Latin ‘dexter’ for ‘right’), while a sugar with a left-hand positioned

hydroxyl substituent is designated as the 'L' configuration (from Latin 'laevus' for 'left') (Figure 1.1).[3] The 'D' and 'L' configurations should not be confused with the case where only some of the stereocentres are inverted, such as in D-glucose and D-gulose where the C-3 and C-4 stereocentres are inverted, or D-galactose and L-altrose where the C-5 stereocentre is inverted; these molecules are diastereomers. In particular for single atom stereochemical inversion (as in the case of D-galactose and L-altrose), the resulting isomers are known as epimers (Figure 1.1).

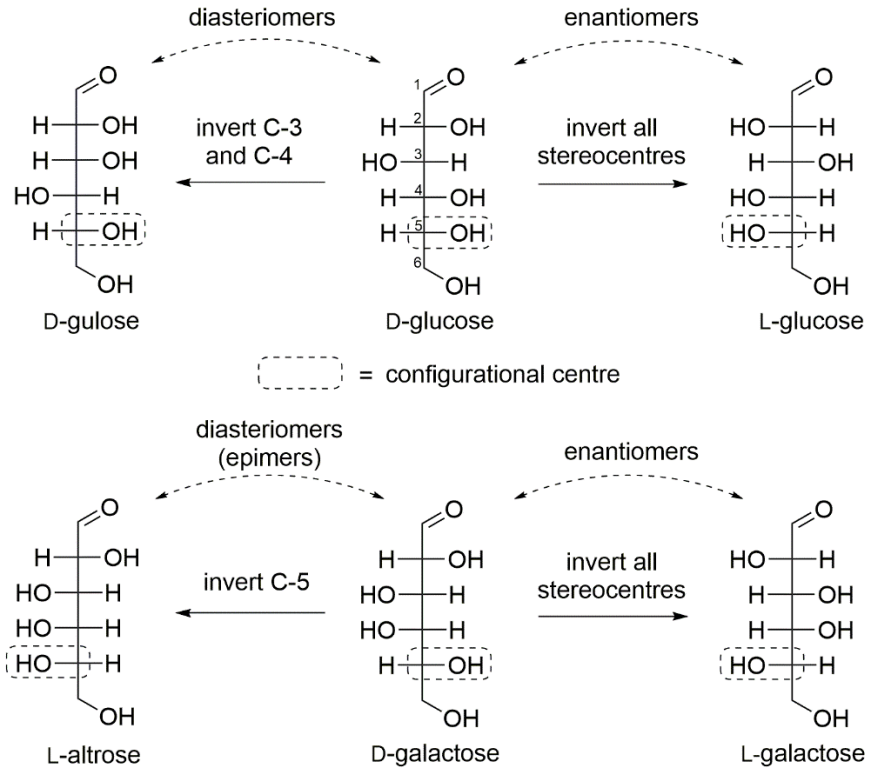


Figure 1.1. Fischer projections of various sugars indicating stereochemical relationships, atom numbering, and configuration.

In the case of cyclic sugars an additional naming convention exists when referring to the stereochemistry of the anomeric carbon – the carbon bearing the aldehyde or ketone in the open-chain form. For example, the most common form of D-glucose is the 6-membered ring (glucopyranose, the term derived from its structural similarity to the oxygen heterocycle pyran) formed when the C-5 hydroxyl of the open-chain form attacks the aldehyde at C-1 to produce a hemiacetal. If the newly formed hydroxyl group at C-1 is on the same side of the carbon backbone in the Fisher projection as is the substituent of the configurational atom, the prefix 'α' is given and the

isomer is said to be the α-anomer (a name specific to epimers at the anomeric carbon, Figure 1.2). Alternatively, if the two are on opposite sides of the carbon backbone in the Fisher projection, the isomer is termed the β-anomer (Figure 1.2). The atom designated as the configurational atom does not change, and when the attacking hydroxyl group is a substituent of the configurational atom, the relative positions in the Fischer projection of the cycle and the newly formed hydroxyl are used to determine the 'α' or 'β' designation (as is the case for glucopyranose). Although the β-pyranose isomer of D-glucose is the most thermodynamically favoured isomer, a solution of D-glucose actually exists as an equilibrium mixture of several possible isomers. This is a result of the lability of the hemiacetal. In fact, regardless of the isomer initially dissolved, a room temperature solution of D-glucose in water will always consist of 66% β-D-glucopyranose and 33% α-D-glucopyranose at equilibrium, with the remaining 1% being a mixture of the open-chain form and the α- and β-anomers of the 5-member cycle (glucofuranose, the term derived from its structural similarity to the oxygen heterocycle furan, Figure 1.2).[3] This is indeed the case for all sugars, with thermodynamics dictating the proportions of each isomer at equilibrium. While pyranose rings are almost universally depicted in their low-energy 'chair' representation, the analogous 'envelope' or 'twist' representations are not as commonly used for furanose rings, which are usually depicted using the Haworth representation (Figure 1.2).

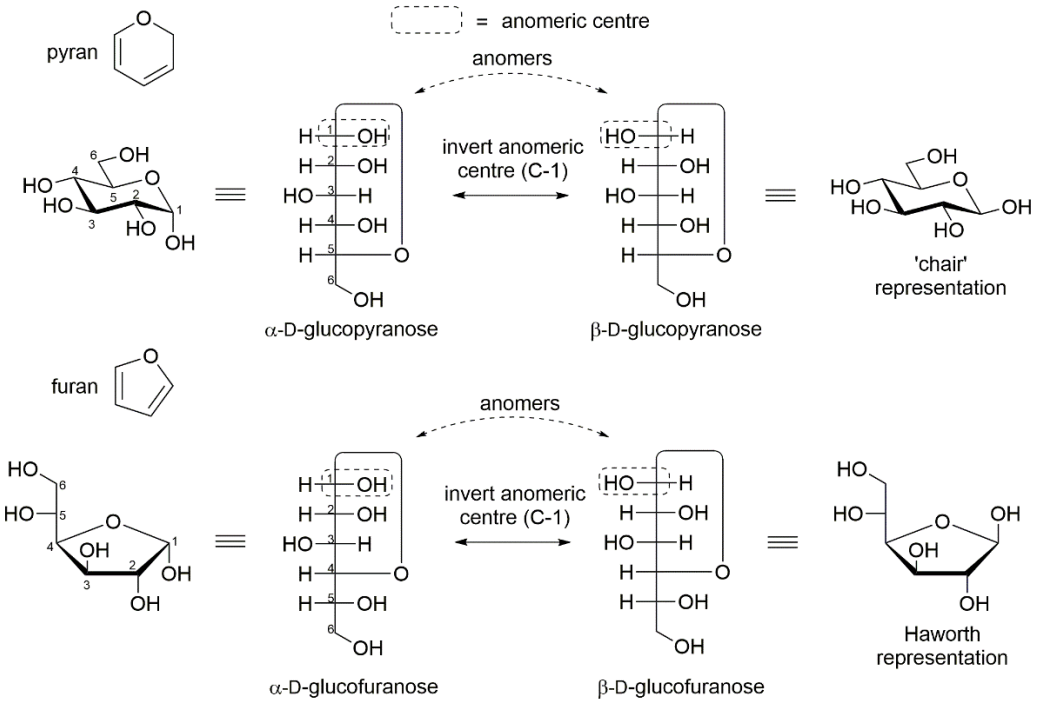


Figure 1.2. Fischer projections of the α- and β-anomers of D-glucopyranose and D-glucofuranose, along with the appropriate modern representation.

After undergoing intramolecular reaction to form a hemiacetal, a sugar can then be reacted with an alcohol in the presence of strong acid to form an acetal. The resultant bond between the anomeric carbon of the sugar and the oxygen of the alcohol is called the glycosidic linkage. If two monosaccharides are linked via a glycosidic linkage, the resultant molecule is called a disaccharide. A typical case is the formation of a glycosidic linkage between D-galactose and the C-4 hydroxyl of D-glucose to give the disaccharide lactose, which is commonly found in milk (Figure 1.3). Lactose (β-D-galactopyranosyl-(1-4)-D-glucopyranose), like monosaccharides with a free hydroxyl at the anomeric centre, is a reducing sugar because it still bears a hemiacetal (in this case on the glucose residue) giving it the potential to act as a reducing agent in subsequent reactions. This is in contrast to a non-reducing disaccharide such as sucrose (table sugar, α-D-glucopyranosyl-(1,2)-β-D-fructofuranoside), formed when the anomeric positions of D-glucopyranose and D-fructofuranose react to form a bis-acetal, in which case no hemiacetal remains (Figure 1.3). The sequential reaction of reducing sugars forms oligo- (up to 10 residues) or poly- (potentially thousands of residues) saccharides; and branching, which results from the formation of additional glycosidic linkages with an

internal sugar residue in the chain, can produce highly complex structures. It is these linear and branched polysaccharides that nature uses to solve many of the structural and energy-storage requirements of cells.

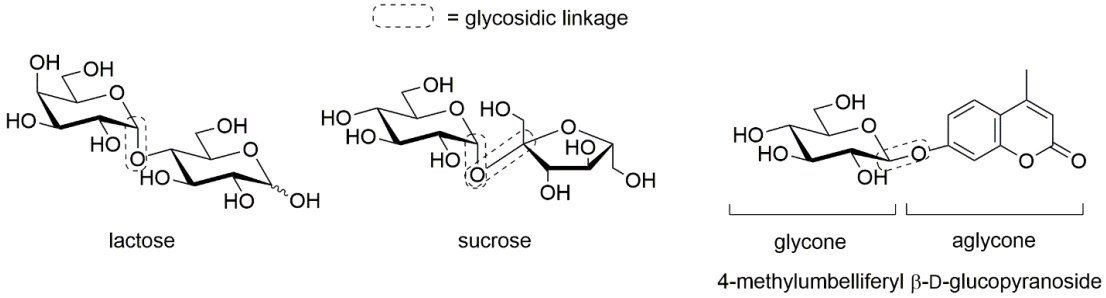


Figure 1.3. Disaccharides lactose and sucrose, as well as the *O*-glycoside 4-methylumbelliferyl β-D-glucopyranoside, with glycosidic linkages indicated.

When a sugar and a non-sugar alcohol are linked via a glycosidic linkage the resultant molecule is called a glycoside, wherein the sugar is called the glycone and the non-sugar is called the aglycone (Figure 1.3). Like the disaccharide sucrose, glycosides are generally non-reducing because of the formation of an anomeric acetal. Although the term glycoside typically refers to *O*-linked glycosides in which the glycone and aglycone are linked by an acetal, *N*-linked, *S*-linked, and *C*-linked glycosides; in which the glycone and aglycone are linked via a hemiaminal ether, thioacetal, or ether, respectively, also exist.[4] All classes of glycosides fall under the broader category of glycoconjugates, which include glycopeptides, glycoproteins, and glycolipids, and which are responsible for critical roles in living systems including, for example, providing barriers against infection,[5] facilitating cell signaling,[6] and aiding in the quality control of protein folding.[7]

1.1.2. Carbohydrates as a Source of Energy

At the forefront of most people's minds is the picture of carbohydrates as a key energy source, and this view is not without merit since much of the energy produced in human cells is derived from carbohydrate metabolism. By virtue of our dependence on carbohydrates, so too are we dependant on plants, which are responsible for producing most of the planet's carbohydrates. All species of plants take advantage of the daily and essentially inexhaustible supply of energy provided by the sun in order to carry out

photosynthesis. In this endothermic multistep process, water is photolysed to liberate O_2 and produce ATP and NADPH, which are then used to reduce CO_2 to give glyceraldehyde-3-phosphate.[8,9] Further biosynthesis consumes glyceraldehyde-3-phosphate to produce more complex carbohydrates, chief among them being glucose. Sugars which are not immediately required by the plant are stored as two D-glucose polymers which together make up starch: amylose, a shorter and mostly linear polymer joined by α(1-4) glycosidic linkages; and amylopectin, a much longer and highly branched polymer with abundant α(1-6) linkages (Figure 1.4). Storage of excess carbohydrates as a polysaccharide rather than as the monosaccharide is necessary, as polysaccharides are generally water insoluble and therefore osmotically inactive in contrast to monosaccharides, which would otherwise exert a tremendous osmotic pressure on the cells.[3] Furthermore, the carbonyl moieties of open-chain monosaccharides have a propensity to react with proteinaceous amines to form toxic adducts, but the occurrence of these reactions is precluded by polymerization, which ensures the saccharides remain in their more inert cyclized form.[10] Interestingly, it has been postulated that the evolutionary emergence of glucose as the universal metabolic fuel is due to the superior stability of its cyclic forms (which constitute 99% of its isomers under physiological conditions) over its open chain form (less than 1%) in comparison to other sugars.[11]

Animals derive a large portion of their energy from the consumption of starch, which is metabolized to produce glucose monomers that are in turn used to fuel the cells. In humans, this process begins with the breakdown of starch by α-amylases and α-dextrinases in the mouth and small intestine to give D-glucose monosaccharides, which can then be taken up by the intestinal mucosa.[8] After transport via the blood stream to cells throughout the body, glycolysis begins by first converting D-glucose into two molecules of glyceraldehyde-3-phosphate by five steps using two molecules of ADP.[8] From here, each glyceraldehyde-3-phosphate molecule is transformed through five steps into pyruvate to produce one molecule of NADH and two of ATP.[8] Pyruvate can be metabolized further via the citric acid cycle to give three molecules of NADH.[8] Glucose that is not required immediately can be stored as the polysaccharide glycogen, whose biosynthesis has been shown to be seeded on the protein glycogenin by *O*-glycosidic linkages to two tyrosine residues.[12] Glycogen shares the α(1-4) linkages and α(1-6)

branching points of the amylopectin component of starch (Figure 1.4), but is very large (10^8 g/mol) and much more highly branched than amylopectin, with branching points every dozen or so residues.[3] The storage of glucose as the polysaccharide glycogen instead of one of the starch polysaccharides is of critical importance for animal cells, which have much more acute energy requirements than plants – the size and dendritic structure of glycogen affords a large number of non-reducing ends, which in turn facilitate rapid degradation of the polymer when energy demands are high.

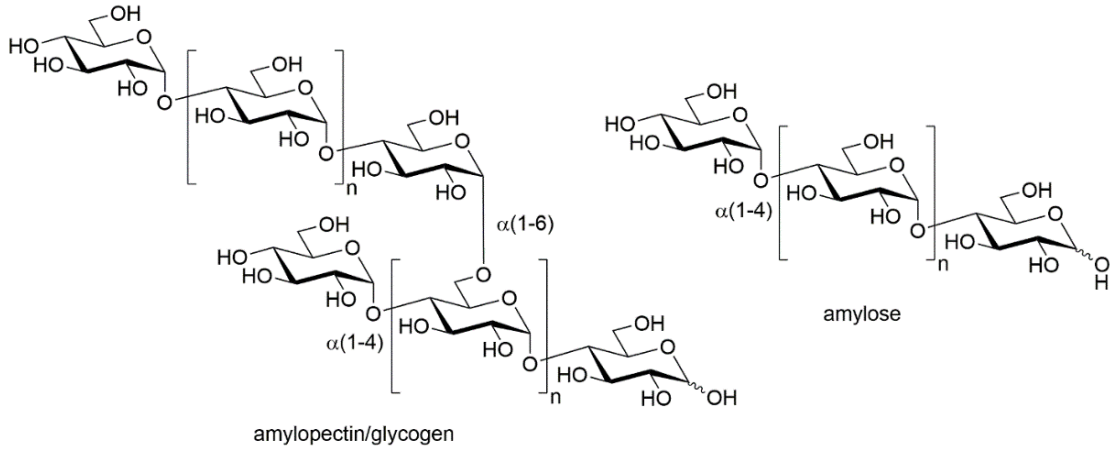


Figure 1.4. Basic structures of the energy storage polysaccharides amylose, amylopectin, and glycogen, with the α(1-4) and α(1-6) glycosidic linkages indicated.

1.1.3. Carbohydrates as Structural Elements

Not only are polymers a convenient form in which to store carbohydrates for later use, they also act to provide physical barriers or afford rigidity, thereby protecting cells from the extracellular environment and providing structural support in multicellular organisms. The most abundant structural polysaccharide, and indeed the most common organic polymer found on earth,[13] is cellulose, which is a major component of the cell wall in all plants and accounts for approximately 30% of the dry weight of plant tissues.[14] In cellulose, D-glucose units are joined by β(1-4) linkages to form linear chains (Figure 1.5), which then assemble into microfibrils whose crystalline nature, water impermeability, and tensile strength contribute to the insolubility of cellulose and its inertness towards chemical and enzymatic degradation.[1] Cellulose is the major carbohydrate polymer responsible for the rigidity of plant cells, and affords plants with

the structural scaffold required for vertical growth. While most animals gain little nutritional benefit from the cellulose they consume, some, such as the ruminants, have developed a symbiotic relationship with microbial gut flora that express β(1-4) D-glucosidase (cellulase), and so acquire most of their energy from cellulose itself.

The structural polysaccharide chitin is perhaps the second most abundant, after cellulose, and is found in invertebrates and fungi. In the case of invertebrates chitin is found almost exclusively in the exoskeleton, which serves as both a barrier against the environment and as a frame to maintain the organisms shape. Fungi, on the other hand, employ chitin as a major structural component of their cell walls, akin to cellulose for plants, leading to a relatively even distribution throughout the organism.[1] Chitin is composed of β(1-4) linked 2-acetamido-2-deoxy-D-glucopyranose residues (GlcNAc) (Figure 1.5), and the macroscopic structure of chitin is quite similar to that of cellulose, with individual chitin polymers coming together to form microfibrils with macroscopic properties.[1] The *N*-acetyl functionality of GlcNAc allows for increased hydrogen bonding within chitin, and thus provides increased strength when compared to cellulose. To provide further strength, in the shells of crustaceans for example, chitin is usually present as a composite along with proteins and calcium carbonate.

Other prominent structural polysaccharides are found in bacteria, which make up a considerable portion of the Earth's biomass. Chief among them is peptidoglycan (PG), a variable length heteropolymer consisting of alternating D-GlcNAc and 2-acetamido-2-deoxy-3-*O*-[(R)-1-carboxyethyl]-D-glucopyranose (*N*-acetylmuramic acid, D-MurNAc) residues joined by β(1-4) linkages, and with adjacent chains intermittently cross-linked by peptides 5-10 residues in length (Figure 1.5). The PG surrounds both Gram-positive and Gram-negative bacteria, although the layer is much thicker in Gram-positive bacteria, and acts primarily to counteract a strong osmotic pressure brought about by the relatively high solute concentration within the cytosol which would otherwise burst the cell.[9] Additionally, the PG layer acts as a barrier preventing passive permeation of molecules larger than 2 nm, thereby retaining nutrients and preventing internalization of extracellular toxins.[15] The importance of the PG layer in affording the survival of bacteria is highlighted by the fact that β-lactam antibiotics, a class of drugs which interfere with PG synthesis, are the most widespread antimicrobial agents in clinical use.[16]

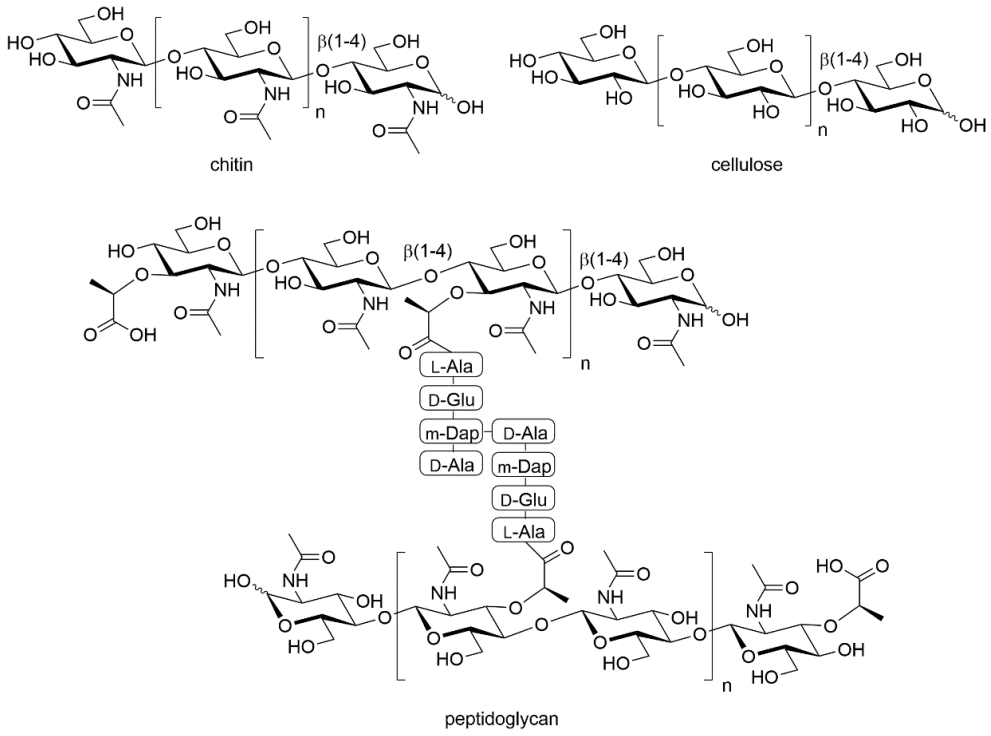


Figure 1.5. Basic structures of the structural polysaccharides chitin, cellulose, and peptidoglycan, with the β(1-4) linkages indicated.

1.1.4. Carbohydrates in Recognition and Signalling

While carbohydrates play critical roles as structural and energy-providing molecules, perhaps their most diverse functions stem from their serving as modulators of protein function, as recognition features in macromolecular interactions, and as signalling molecules. Outlined are a few choice illustrations of the broad scope of glycans in the functioning of living systems. A fascinating example is the pathway governing the quality control of protein folding in the endoplasmic reticulum (ER), which ensures that newly synthesized proteins reach their active conformations in a timely fashion, and that those proteins which fail to do so are sent for recycling (Figure 1.6).[7] Within the crowded ER, where the protein concentration can be in excess of 100 mg/mL, newly synthesized proteins are greatly outnumbered by folding chaperone proteins such as membrane-bound calnexin and its soluble homologue calreticulin.[7] In order to recognize proteins and assist in their folding, calnexin and calreticulin associate with the oligosaccharide modification $Glc_1Man_9GlcNAc_2$, which is attached to proteins by the reducing end GlcNAc residue.[17] After folding is attempted, the terminal glucose is

removed by action of the glycoside hydrolase (GH) α-glucosidase II, yielding the $Man_9GlcNAc_2$-protein, which is then released from the chaperone.[17] Despite the abundance of folding chaperones, there are inevitably proteins which are unable to achieve their native conformations due to issues such as mistranslation. If the protein is not properly folded after release from the chaperone, it is recognized as unfolded by the glycosyl transferase (GT) UDP-glucose:glycoprotein glucosyltransferase (UGGT), which reinstalls the terminal glucose residue to facilitate recognition of the unfolded protein by chaperones so that folding can be attempted again. If folding efforts repeatedly fail, α-mannosidase I will remove a mannose residue from the oligosaccharide.[18] Removal of the first mannose residue, and indeed subsequent demannosylation, causes iterative decreases in the affinity of the glycopeptide for UGGT, and therefore makes it increasingly likely that the unfolded protein will be targeted to the ER-associated degradation (ERAD) pathway via lectins that recognize a partially demannosylated feature such as $Man_6GlcNAc_2$.[18,19] In this way, the degree of mannosylation on newly synthesized ER proteins acts as a molecular clock, indicating how long a protein has spent in the calnexin/calreticulin cycle and determining when it should be sent for degradation.

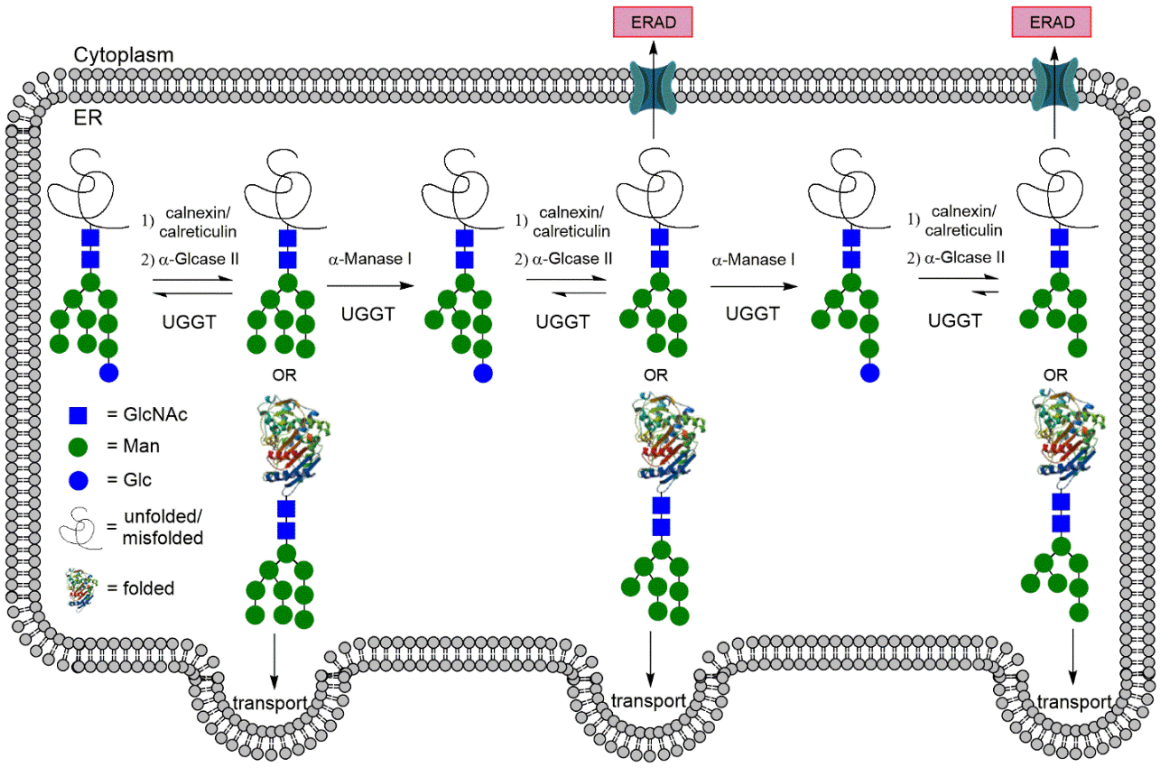


Figure 1.6. Simplified schematic of the quality control pathway for protein folding within the endoplasmic reticulum (ER).

Particularly significant in medicine is the role of carbohydrates in making up the blood group antigens in the ABO blood typing system. Of the 300 blood groups defined thus far, most of which are based on proteins, the carbohydrate-based ABO antigens are perhaps the most important as they have proven to be among the most immunogenic.[1,20] As a result, defining a person's ABO blood group is critical for the purposes of blood transfusions and organ transplantation. All antigens in the ABO blood group are glycoproteins or glycolipids in which the terminal saccharides are responsible for the immunogenicity. Specifically, O-type (also called H-type) arises from a terminal fucose with an α(1-2) linkage to galactose; while A-type and B-type are based on the O-type template but contain an additional *N*-acetylgalactosamine (GalNAc) or galactose residue, respectively, α(1-3) linked to the galactose (Figure 1.7).[1] Furthermore, individuals can be of blood type AB, in which case their erythrocytes express both the A-type and B-type antigens. Since the only differences between blood antigens under the ABO classification lay in the identity of the terminal saccharide residues, significant effort has been invested in developing methods to convert A-type and B-type antigens to the

universal donor O-antigen.[21] Much of the focus of this area of research has been on using α-glactosidase enzymes to trim the terminal GalNAc/galactose residues of the immunogenic blood groups to yield O-type erythrocytes. However, issues with the efficiency of enzymes employed in this process have hampered advancement of this technology to a level suitable for clinical use.[21]

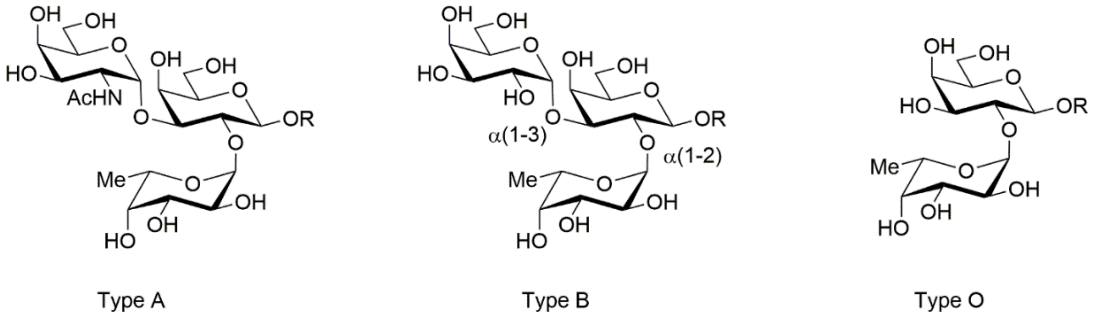


Figure 1.7. Blood group antigens A, B, and O, with α(1-2) and α(1-3) linkages indicated.
The blood group antigens are bonded to a Gal-Glc-ceramide (R) that is embedded in the erythrocyte cell membrane. Subtypes with varying degrees of branching and with sugar residues located between the antigen and the R group also exist.

1.2. Glycan Processing Proteins

Given the prevalence and complexity of glycans in living systems, there is naturally a diverse set of proteins involved in their synthesis, recognition, and degradation. Transmembrane proteins are responsible for transporting sugars, either simple monosaccharides or more complicated oligosaccharides or glycoconjugates, into the cell from the extracellular environment or from one cellular compartment to another. Using activated sugar donors, glycosyl transferases (GTs) synthesize oligo/polysaccharides or modify other biomolecules to form glycoconjugates. Once synthesized, these more complex glycans can interact with carbohydrate binding proteins (lectins) to mediate signalling within a cell, or to communicate with other cells within an organism. Ultimately, glycans which are no longer needed, or that require structural modification to tune their function, are either completely or partially degraded by glycoside hydrolases (GHs). The resulting liberated monosaccharides are then recycled. In this section, the superfamilies of enzymes which install (GTs) and hydrolyse

(GHs) glycosidic linkages will be discussed, as will the proteins responsible for transporting carbohydrates across biomembranes.

1.2.1. Glycosyl Transferases

In general terms, proteins and polysaccharides are quite similar – they are both constructed from a pool of possible monomers, and are linked together in variable length chains to produce macromolecules with properties that serve particular functions. However, this is where the similarities between these two classes of molecules end, and examination of their structure and how they are assembled sheds light on their differences. Proteins consist of a linear chain of amino acids linked together by amide bonds, and whose sequence is predetermined by the genetic code. Polysaccharides on the other hand often have complex branching patterns, are linked together by glycosidic bonds with variable regio- and stereochemistry, and have no pre-determined template for their synthesis. The complex task of oligo- and polysaccharide synthesis is undertaken by a class of enzymes known as glycosyl transferases (GTs), which either install sugars on protein or lipid scaffolds, or elongate pre-existing carbohydrate chains.[4] In order for the reactions they catalyze to be thermodynamically favourable, GTs use activated (high-energy) glycosyl donors. Most commonly, glycosyl donors are nucleotide mono- or di-phosphates (UDP- GDP-, and CMP-sugars) known as Leloir donors. However, non-nucleotide glycosyl donors (phosphate/pyrophosphate-sugars, or lipid-phosphate/pyrophosphate-sugars) also exist, and these are termed non-Leloir donors.[22] In the majority of cases, a given GT is responsible for forming one specific type of glycosidic linkage, an idea first popularized as the so-called ‘one enzyme one linkage hypothesis’. What is meant by this hypothesis is that GTs are unique regarding the stereo- (α or β) and regio- (1-2, 1-3, 1-4, etc.) chemistry of the glycosidic linkage they form, and the acceptor/donor saccharide they recognize. This is in stark contrast to protein synthesis, in which all the peptide bonds in a polypeptide are installed by one, albeit complex, protein – the ribosome. Exceptions to the one enzyme-one linkage hypothesis do exist, an example being EXTL2 which exhibits the ability to transfer both GalNAc and GlcNAc in the initiation of heparin sulfate synthesis.[23]

The enzymes which are responsible for the assembly, modification, and breakdown of oligo- and poly-saccharides, whether or not they are part of a glycoconjugate, have been classified into families based on their sequence in the carbohydrate active enzymes (CAZy) database.[24] As of 2015, the CAZy database classifies 97 families of GTs based on sequence similarity, and new families are identified regularly. However, despite the relatively large number of GT families, all GTs whose structures have been solved to date adopt one of three canonical folds: GT-A, GT-B, or GT-C.[25] Enzymes belonging to the GT-A superfamily are composed of two abutting β/α/β Rossmann-like domains, a structural motif found in many nucleotide-binding proteins.[26] GT-A superfamily proteins are generally dependant on a divalent metal ion to stabilize the charged phosphate of the nucleotide donor and aid in leaving group departure.[25] The first member of the GT-A superfamily to be structurally characterized was SpsA, which is responsible for the synthesis of a structural polysaccharide found in *Bacillus subtilis* (Figure 1.8).[27] Converse to the GT-A enzymes, those belonging to superfamily GT-B, such as the β-glucosyltransferase from T4 bacteriophage (Figure 1.8),[28] are composed of two facing Rossmann-like domains and are metal ion independent, with leaving group departure instead facilitated by enzymatic residues.[25] While the solved structures of GT-A and GT-B enzymes have indicated that these superfamilies use Leloir (nucleotide) donors, the more recently discovered GT-C superfamily is predicted to prefer lipid-activated donors. Indeed, a recent crystal structure of oligosaccharyltransferase, a GT-C enzyme from *Campylobacter lari* (Figure 1.8), in complex with an acceptor peptide was shown to have a hydrophobic cleft into which a lipid-anchored glycosyl acceptor might bind.[29] In addition to classification based on sequence or fold, GTs can also be classified according to the mechanism by which they act – either with retention or inversion of anomeric stereochemistry. Although correlations can be drawn between the catalytic mechanisms of GTs and those used by the more thoroughly understood glycoside hydrolases (GHs), the mechanistic characterization of GTs lags behind, especially in the case of retaining GTs.[22]

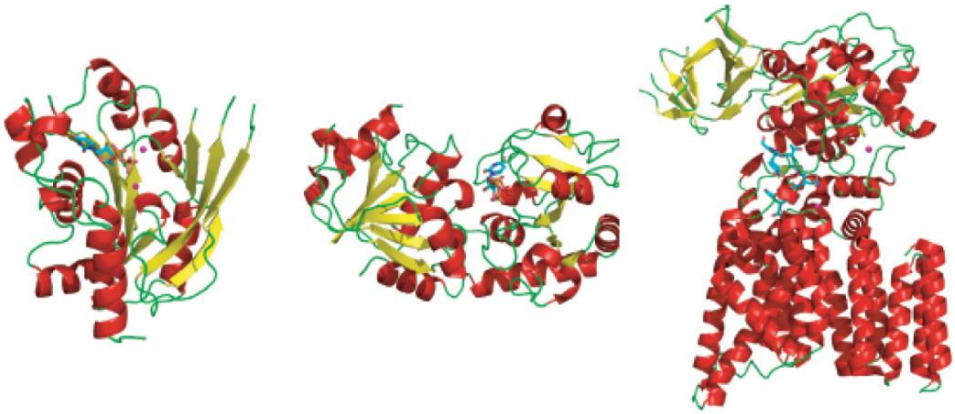


Figure 1.8. Structures representing GT-A, GT-B, and GT-C glycosyltransferase folds.
Adapted, with permission, from Gloster 2014.[25]

1.2.2. Glycoside Hydrolases

Converse to the GTs, which install glycosidic linkages, the glycoside hydrolases (GHs) catalyze their hydrolytic cleavage to liberate a sugar hemiacetal/hemiketal from a polysaccharide or glycoside. GHs are remarkably proficient, affording rate enhancements of up to 10^{17} times as compared to the uncatalyzed hydrolysis reaction.[30] Among the enzymes classified within the CAZy database, GHs are the most numerous and the best understood, with 135 families existing as of 2015. Since the three dimensional structures of proteins are better conserved and often more informative than their sequence, the GHs are also grouped on this basis into 12 superfamilies (clans).[31] Separate from the CAZy family and clan classifications, GHs can be categorized into two groups depending on the location of the glycosidic bond on which they act. Endoglycosidases catalyze the cleavage of internal glycosidic linkages to produce two smaller saccharide chains, while exoglycosidases hydrolyze terminal glycosidic linkages to liberate a monosaccharide, usually but not always at the non-reducing end of the chain (Figure 1.9).[32] Despite the many GH families, only two main active-site topologies exist: pocket- and cleft-shaped.[32] The pocket-shaped active site can be thought of as a crater in the surface of the enzyme, and is usually only able to accommodate one sugar residue. As such, pocket-shaped active sites are characteristic of exoglycosidases such as glucocerebrosidase.[33] Conversely, multiple residues of a polysaccharide can dock

within a cleft-like active site, and so these are the mark of endoglycosidases such as chitinases.[34] A third topology, tunnel, is found far less commonly and is essentially a cleft-like active site which is closed over on the substrate. It is thought that these active-sites allow the enzyme to maintain an association with the substrate after the hydrolytic event to allow further processing, and examples include cellobiohydrolase and carrageenase.[32,35,36]

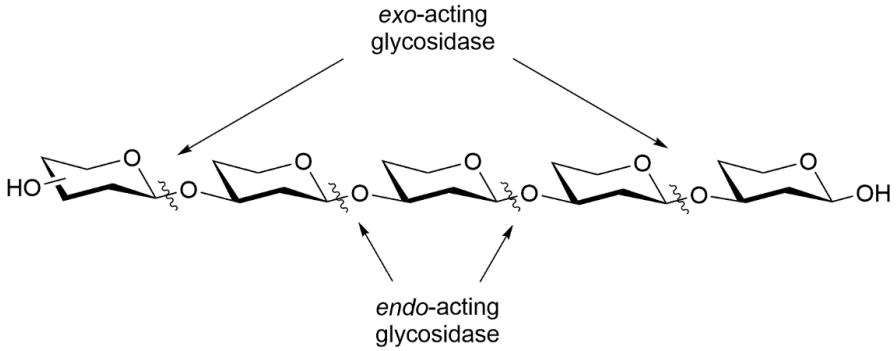


Figure 1.9. Positions of glycosidic linkages cleaved by *endo*- and *exo*-acting glycosidases.

Despite the great variety in the structure of the substrates they process, GHs nevertheless act by a limited set of rather similar mechanisms. The basics of these mechanisms were first postulated by Koshland in 1953,[37] and since that time a great deal of progress has been made towards their understanding. Succinctly, glycoside hydrolases act by either an inverting or retaining mechanism, wherein the anomeric stereochemistry (α/β) of the resultant reducing sugar either changes or remains the same, respectively, relative to its stereochemistry before cleavage. The inverting mechanism is the more simple of the two, and consists of a one-step single-displacement involving an oxocarbenium ion-like transition state (Figure 1.10).[4,38] In this process, a general base (Asp/Glu) enhances the nucleophilicity of a water molecule to facilitate its attack on the anomeric centre. Concomitantly, leaving group departure is aided by an enzymatic general acid (Asp/Glu) that stabilizes the building negative charge on the anomeric oxygen. The resulting dissociative transition state in which positive charge develops at the anomeric centre is stabilized by electron donation from the endocyclic oxygen to form an oxocarbenium ion-like species. At the resolution of this process, the resulting sugar hemiacetal/hemiketal has an anomeric stereochemistry that is inverted relative to the substrate.

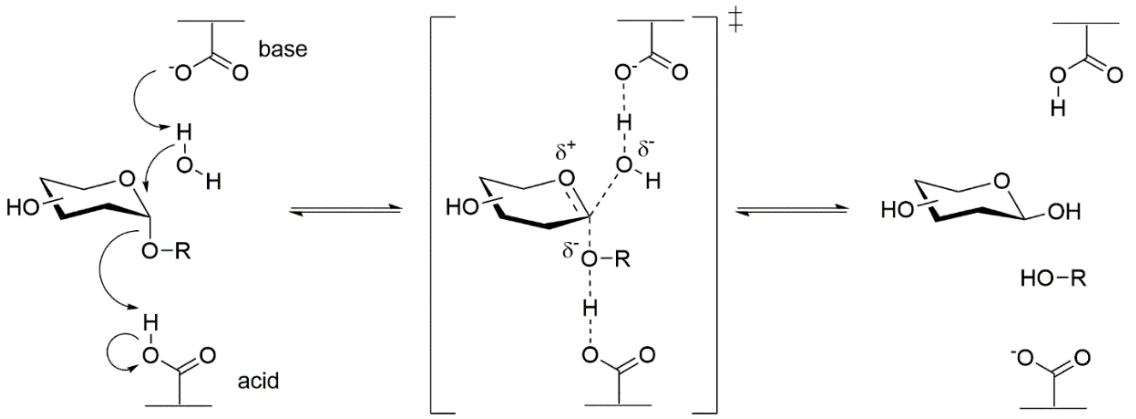


Figure 1.10. Generalized catalytic mechanism for an inverting α-glycosidase.

The mechanism for retaining glycosidases is somewhat more complicated, occurring by a two-step double-displacement wherein a covalent glycosyl-enzyme intermediate is flanked in the reaction coordinate by two oxocarbenium ion-like transition states (Figure 1.11).[4,38] Akin to inverting enzymes, retaining enzymes feature two catalytic residues, which are either aspartic or glutamic acid in the majority of cases. However, these residues serve different roles in retaining enzymes, with one acting as a general acid/base while the other acts as a catalytic nucleophile/leaving group.[38] In the general catalytic cycle for a retaining GH, attack on the anomeric centre by the catalytic nucleophile occurs in concert with leaving group departure, which is facilitated by a residue serving as a general acid as for inverting enzymes. This results in formation of an oxocarbenium ion-like transition state, with the endocyclic oxygen facilitating charge delocalization, which collapses with the catalytic nucleophile to give a covalent glycosyl-enzyme intermediate with inversion of anomeric stereochemistry relative to the substrate. The short-lived glycosyl-enzyme intermediate is then hydrolysed by general base-facilitated nucleophilic attack of water on the anomeric centre, and the catalytic leaving group concomitantly departs through a dissociative oxocarbenium ion-like transition state. The formation of the glycosyl-enzyme intermediate with inversion of stereochemistry is key, as it allows the nucleophilic attack of water to be directed such that overall retention of stereochemistry is achieved. Interestingly, the distance between the catalytic residues for retaining glycosidases is usually around 5 Å.[38,39] This is in contrast to inverting glycosidases for which the distance is closer to 9 Å so that the active site can simultaneously accommodate the intact substrate as well as the nucleophilic water molecule.[38] Accordingly, the distance between the catalytic residues

can give some indication of whether a newly discovered GH acts by an inverting or retaining mechanism.

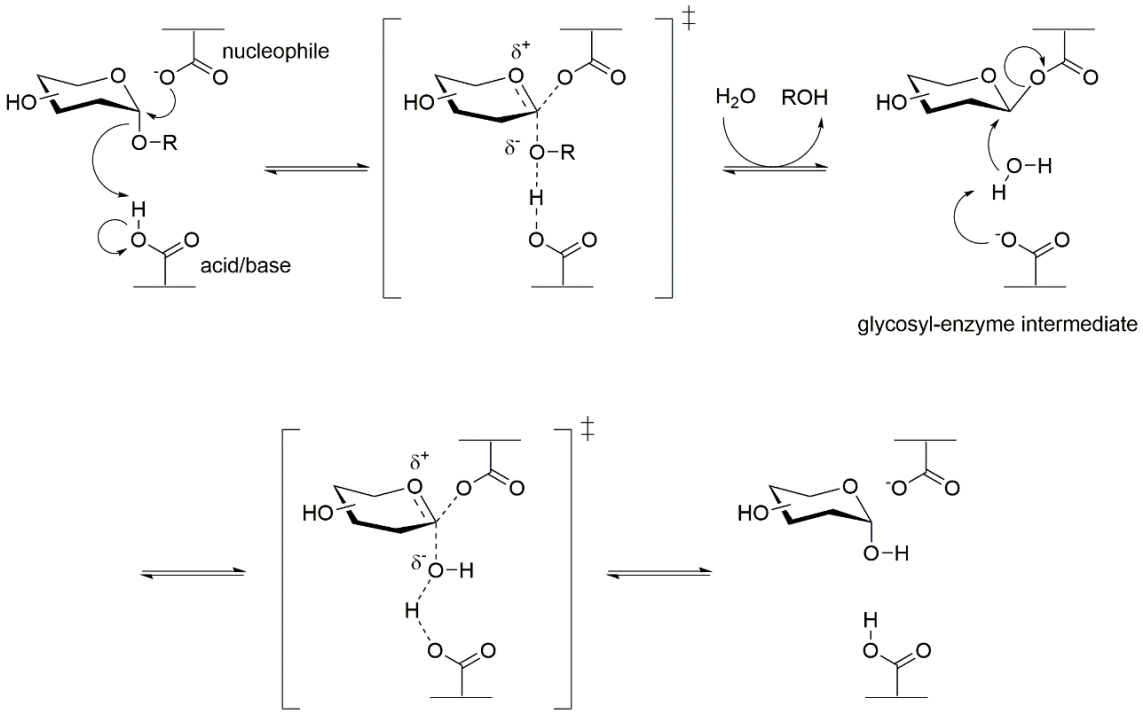


Figure 1.11. Generalized catalytic mechanism for a retaining α-glycosidase involving the formation of a glycosyl-enzyme intermediate.

A second type of retaining GH exists in which no catalytic nucleophile/leaving group is required. Such enzymes rely on substrate-assisted catalysis (SAC), in which an *N*-acetyl moiety facilitates departure of the anomeric substituent via anchimeric assistance (neighbouring group participation). Despite the fact that relatively few explicit examples of SAC exist in the literature for glycosidases, this mechanism is well established for a variety of other enzyme classes including serine proteases and GTPases.[40] The fact that a given GH processes an *N*-acetyl-bearing substrate does not obligate that enzyme to act by SAC; many such enzymes, including the GH-3 β-*N*-acetylglucosaminidases like NagZ,[41] act by the more classical retaining mechanism. All known GHs which do engage in SAC can be placed in the GH families 18, 20, 25, 56, 84, or 85. The catalytic mechanism for GHs using SAC is similar to that used for typical retaining GHs: a two-step double-displacement involving oxocarbenium ion-like transition states and a relatively low energy intermediate. However, attack on the anomeric centre to displace the anomeric substituent is performed not by an enzymatic

residue, but rather intramolecularly by the C-2 *N*-acetyl moiety to yield an oxazoline intermediate (Figure 1.12**Error! Reference source not found.**). An enzymatic carboxylate usually aids in this process, although it remains unclear whether its role is to electrostatically stabilize the developing positive charge on an oxazolinium ring, or to act as a general base to generate an oxazoline intermediate.[42] Subsequent to its formation, the oxazoline is hydrolyzed by general base-facilitated nucleophilic attack of water to yield the reducing sugar with overall retention of anomeric configuration. A particularly prominent GH which uses SAC is the human nucleocytoplasmic GH-84 enzyme β-*N*-acetylglucosaminidase,[43] also known as human *O*-GlcNAcase (hOGA), which is discussed in more detail in Chapter 3. Interestingly, several GHs exist whose mechanism of hydrolysis differs from those discussed. Two notable examples include the sialidases and trans-sialidases of GH families 33, 34, and 89 which use a tyrosine residue as the catalytic nucleophile,[44] and GH family 4 and 109 GHs which require NAD+ as a co-factor.[45]

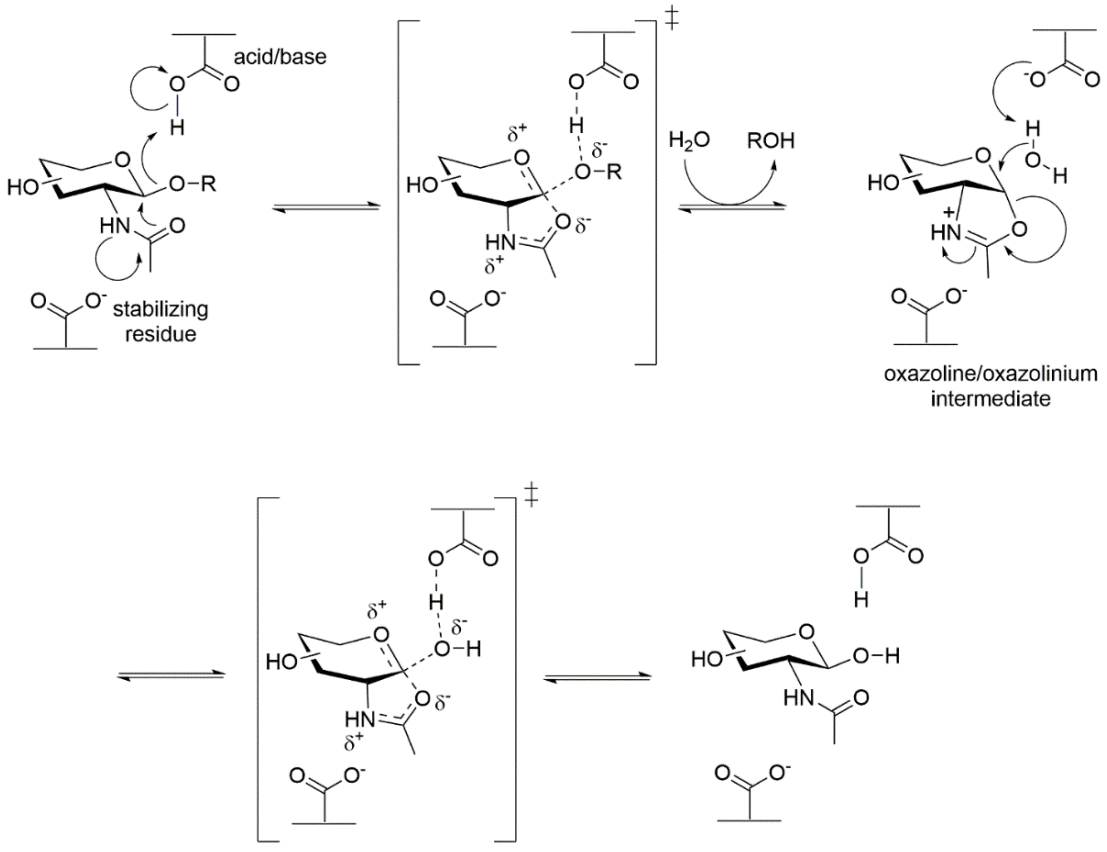


Figure 1.12. Generalized catalytic mechanism for substrate-assisted catalysis involving the formation of an oxazoline/oxazolinium intermediate.

1.2.3. Sugar Transport Proteins

Cell and organelle membranes are in place to exclude unwanted foreign substances and to compartmentalize cellular processes. Despite this requirement, a mechanism must also exist to allow compounds to permeate membranes so that substrates can reach the cellular compartments in which they are processed, and so that nutrients and metabolic waste products can be gathered and discarded, respectively. Although some materials can permeate biomembranes with relative ease via passive transport (diffusion), this is limited to small non-polar molecules which are not repelled by the hydrophobic environment of the lipid bilayer.[9] On the other hand, large molecules, or those which are highly polar or bear formal charges, must be assisted in crossing membranes by transport proteins. The breadth of molecules whose transport must be aided in this way is quite large and, as such, all organisms express a large repertoire of transport proteins. Indeed, a significant number of diseases in humans result from aberrantly functioning transport proteins. Prominent examples include the glucose transporter GLUT4 which plays a role in obesity and diabetes,[46,47] and the chloride channel CFTR whose mutation causes cystic fibrosis.[48]

Two major subsets of transport proteins exist: those belonging to the major facilitator superfamily (MFS) which rely on electrochemical gradients to facilitate transport, and those belonging to the ATP-binding cassette (ABC) superfamily which drive transport with the energy derived from ATP hydrolysis.[49] Originally, proteins in the MFS were thought to primarily function in the transport of sugars,[50] but work in recent decades has shown that many other classes of molecules, including amino acids, vitamins, nucleobases, enzyme cofactors, and organic and inorganic ions, are also transported by members of this superfamily.[51] As of 2012, the MFS consists of 74 subfamilies, many of which are still of unknown function.[51] MFS proteins can operate by three general transport reactions.[52] The first is uniport, or facilitated diffusion, which transports a recognized substrate along its concentration gradient and therefore does not require the use of cellular energy (Figure 1.13). The second, symport, is a type of active transport that can be used to move a substance against its concentration gradient. This is accomplished by coupling the transport of the substrate of interest with the transport of a monovalent cation (usually H^+ or Na^+) which moves along its concentration

gradient and in the same direction as the substrate (Figure 1.13). The third transport reaction is antiport, which is also a type of active transport and is similar to symport except that the substrate of interest and the co-transported species travel in opposite directions. Moreover, in the case of antiport the coupled species may either be a cation, or a compound whose structure is similar to that of the substrate of interest (Figure 1.13). Among the MFS subfamilies, eight are involved in the transport of sugars and their derivatives, with the largest of these being the sugar porter (SP) subfamily.[52] This subfamily is ubiquitous in nature, with member proteins being distributed through all domains of life: bacteria, archaea, and eukaryota.[52] The members of the SP subfamily are mainly responsible for the transport of monosaccharides including glucose, fructose, mannose, galactose, arabinose, and xylose; but can also transport other saccharide species including, but not limited to, maltose, lactose, and α-glucosides.[52]

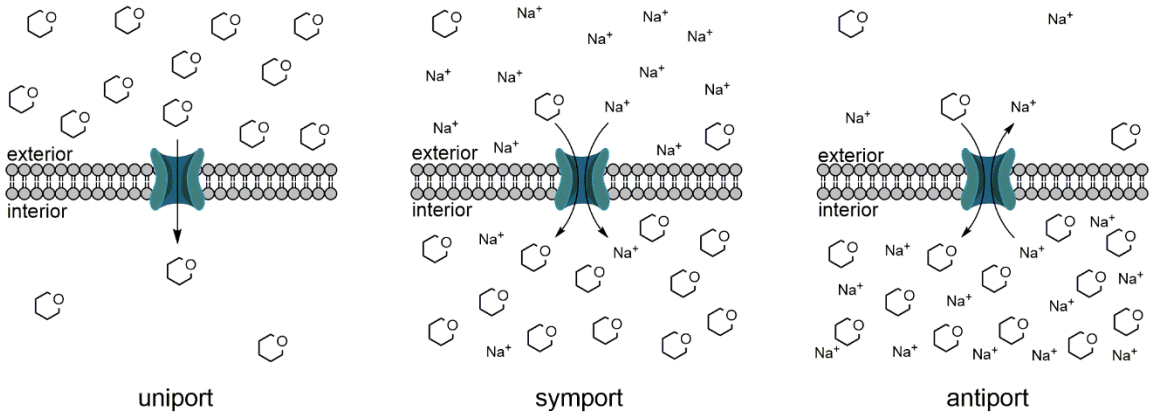


Figure 1.13. Illustrations of carbohydrate transporters acting by uniport, symport, and antiport.

Two particularly well characterized sugar transporters belonging to the MFS are LacY and GlpT, both of which are native to *E. coli* and act to scavenge nutrients from the environment. LacY is a member of the oligosaccharide:H^+ symporter (OHS) subfamily and is responsible for the proton-coupled internalization of lactose, which is then metabolized to produce energy for the cell.[53] On the other hand, GlpT belongs to the organophosphate:phosphate antiporter (OPA) subfamily and acts to internalize glycerol-3-phosphate which serves as an energy source and as a precursor to the biosynthesis of phospholipids.[54] The 12 transmembrane helices that comprise LacY and GlpT, known as the canonical MFS fold, are arranged to form a cylindrical bundle.[53,54] The helices are amphipathic and are carefully arranged to create a hydrophilic binding pocket for the

saccharide at the bundle's core, while maintaining a hydrophobic surface on the exterior of the bundle for interfacing with the alkyl chains of the membrane phospholipids. Although the directionality of substrate transport differ for LacY (a symporter) and GlpT (an antiporter), the data for both suggest that they act by an alternating-access mechanism. This mechanism is particularly well supported in the case of LacY, for which biochemical and biophysical data exist (Figure 1.14).[55] For LacY, the alternating-access transport mechanism begins with an outward facing V-shaped conformation such that the binding pocket can be accessed from the periplasmic space. Protonation of this conformation results in lactose having a higher affinity for the binding site, and binding of lactose triggers a conformational transition of LacY to the inward (cytoplasmic) facing V-shaped conformation. Release of lactose to the cytoplasm, followed by release of the proton, allows the protein to again adopt the periplasmic-facing conformation such that the process can be repeated. Since the elucidation of the structures of LacY and GlpT, only a handful of other MFS proteins have been crystallized, and for each not more than one conformation has been structurally characterized. However, together these crystal structures paint a picture supporting the alternating-access mechanism, as each of the structures obtained thus far happen to represent distinct conformations in the proposed transport cycle.[56,57] Moreover, all of the crystal structures obtained so far exhibit the canonical MFS fold seen in LacY and GlpT, offering further support that the alternating-access mechanism is shared among all members of this family of transporters. Indeed it is believed that the majority of MFS proteins act by this alternating-access mechanism,[56,57] consistent with bioinformatic analysis which predicted that most family members contain the 12 transmembrane helices characteristic of the MFS fold.[51]

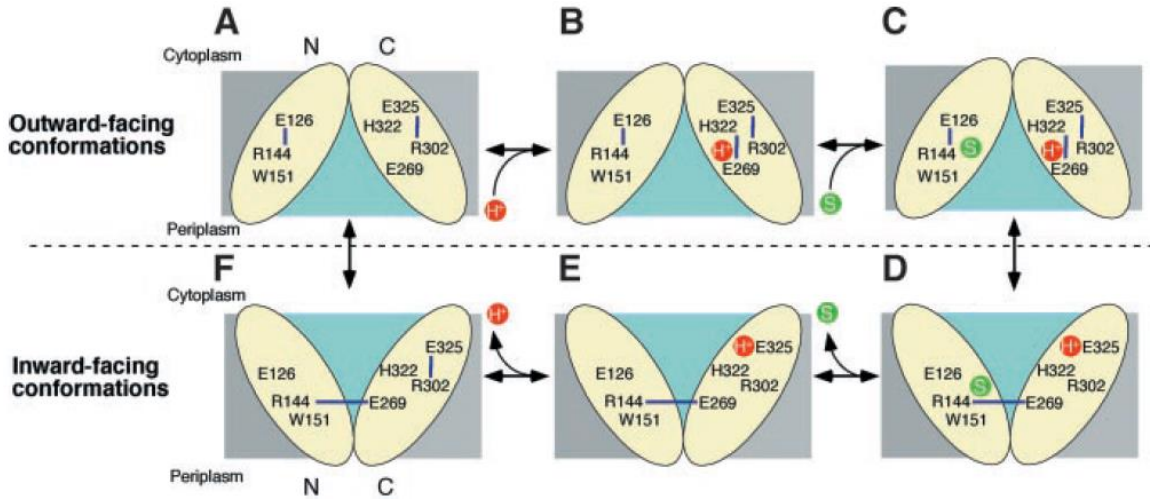


Figure 1.14. Alternating-access mechanism for LacY.
Yellow ovals are the N- and C-terminal domains, blue lines are hydrogen bonds, red and green circles are the proton and substrate, respectively. Beginning in the open periplasm-facing conformation (**A**), a proton binds (**B**) which primes the transporter for substrate binding (**C**). Substrate binding triggers a conformational transition to the cytoplasm-facing conformation (**D**), and release of substrate (**E**) and proton (**F**) to the cytoplasm allows transition back to the periplasm-facing conformation. Adapted, with permission, from Abramson *et al*. 2003.[53]

1.3. Monitoring the Activity of Glycan Processing Proteins

Many approaches have been developed to study enzyme-catalyzed reactions and the various aspects of catalysis. Using genetic techniques, an enzyme's expression can be reduced or entirely eliminated, and the resulting phenotypic effects can be observed to gain an understanding of the role of the enzyme within an organism. Further, information on the catalytic role of enzymatic residues can be obtained by using site-directed mutagenesis to alter residues of interest. The three-dimensional structure of a crystallized enzyme can be determined using X-ray crystallography, and more in depth information can be gleaned by co-crystallization of the enzyme with a substrate, product, or inhibitor. Moreover, through rational substrate design or mutation of enzymatic residues, certain steps of the catalytic process can be sufficiently slowed to provide a snapshot of catalysis. Multidimensional NMR techniques have also been employed to determine enzymatic structures and to study enzyme dynamics.[58,59] However, the most prevalent and easily applied technique to study enzymes is kinetic analysis, in which natural enzyme substrates or labelled analogues are used to monitor enzymatic activity.

Further, structural modification of labelled substrates to alter the kinetics of their processing can give insight into the important structural elements required for substrate recognition and catalysis. Not only can kinetic techniques offer insight into the mechanism of an enzyme, they can also be used to screen for modulators of enzyme activity which may ultimately be used as probes for further research or even as therapeutics.[60] This section will examine how enzymes, in particular glycoside hydrolases and sugar transporters, are characterized using kinetic techniques. Special attention will be given to fluorescent substrates, and specifically to those whose fluorescence is activatable and are thus used to study enzymes in the more physiologically relevant environment of a living system.

1.3.1. Michaelis-Menten Kinetics

Humanity has unwittingly been putting enzymes to use for thousands of years, particularly in the preservation or flavour enhancement of foods (cheese, wine, bread), but also for other purposes such as the softening of leather. Despite this, only in the previous two centuries did we learn of the existence of enzymes and gain an appreciable understanding of their function. Toward the turn of the 20th century, researchers discovered that the rate of product formation in an enzyme-catalyzed reaction exhibits a non-linear dependence on the concentration of substrate. British scientist Adrian Brown put forth a qualitative description to explain this phenomenon, which was supported by growing evidence that enzymes form complexes with the substrates on which they act. Brown proposed that free enzyme (E_f) and free substrate (S_f) first combine (rate constant k_1) to form an enzyme-substrate complex (ES), and that this complex either dissociates (rate constant k_{-1}) to reform the enzyme and substrate, or reacts (rate constant k_2) to furnish the product (P) and regenerate E_f (1.1).[61] In the scheme proposed by Brown (1.1), product formation depends only on the existence of the enzyme-substrate complex, and thus the rate of product formation (v) can be written in terms of a rate constant (k_2) and the concentration of the complex ([ES]) (1.2).

$$\mathrm{E_f + S_f} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \mathrm{ES} \xrightarrow{k_2} \mathrm{E_f + P} \quad (1.1) \qquad \mathrm{v = k_2[ES]} \quad (1.2)$$

Based on these equations, it is understandable how v can have a non-linear dependence on the free substrate concentration ($[S_f]$). At sufficiently low $[S_f]$, when not all the enzyme is bound to form ES, an increase in $[S_f]$ causes a proportional increase in [ES], and thus v appears to depend on $[S_f]$ in a first-order manner. However, when $[S_f]$ is high such that essentially all E is bound to form ES, increasing $[S_f]$ has a negligible effect on [ES], and so v appears independent of $[S_f]$. A graphical depiction of this relationship is given below (Figure 1.15). Although Brown's model provided a qualitative understanding of enzyme saturation, a mathematical approach was developed in 1903 by French chemist Victor Henri. In 1913 the works of Brown and Henri were built upon by German biochemist Leonor Michaelis and Canadian physician Maud Menten to study the kinetics of invertase, a GH which hydrolyses sucrose to yield glucose and fructose.[62] Although Michaelis and Menten were standing on the shoulders of giants, their formalization of previous research into a general procedure which was readily applicable to other cases has cemented them as the most prominent historical figures in the field of enzyme kinetics.

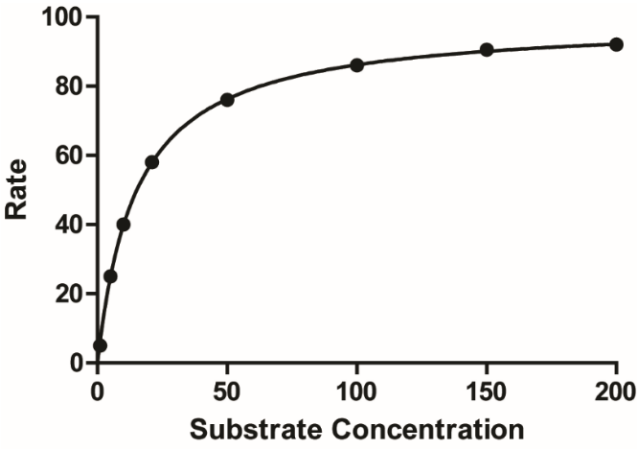


Figure 1.15. A depiction of the non-linear dependence of the rate of product formation on the substrate concentration for an enzyme-catalyzed reaction.

The model of enzyme kinetics first followed by Michaelis and Menten is known as the rapid equilibrium model. This model assumes that E_f and S_f are in rapid equilibrium with the ES complex, and that the complex reacts relatively slowly to give E_f and P. In terms of rate constants, this can be described as the situation where $k_2 << k_{-1}$ in (1.1). However, this assumption is not always accurate, as some systems may convert the ES

complex relatively rapidly to E_f and P, resulting in a significant value for k_2. Therefore, to arrive at a more generalizable theory, Briggs and Haldane proposed the steady-state model in which the rate of formation of the ES complex is exactly matched by its rate of decay.[63] This model does not require $k_2 << k_{-1}$, so long as any ES complex consumed in the formation of P is rapidly replenished, thus holding [ES] at a steady-state. This situation arises during the initial stages of the reaction, when $[S_f] >> [E_f]$ and the total concentration of product ([P]) is insignificant. In this case, the abundant S_f rapidly complexes with any E_f to reform the ES complex, and the low [P] means that this process is negligibly impacted by product inhibition. Due to the conditions required to achieve steady-state, some assumptions can be made about the initial stages of the reaction.[64] Firstly, the enzyme molecules in the system are assumed to have two possible states: either free or bound in the ES complex. Thus, the total enzyme concentration ([E]) can be written as the sum of $[E_f]$ and [ES] (1.3). Secondly, since $[S_f] >> [E_f]$, the amount of S_f consumed to form the ES complex is assumed to be negligible as compared to the total substrate concentration ([S]). As a result, $[S_f]$ can be reasonably approximated by [S] (1.4). Lastly, since it's assumed in the steady-state model that the rate of ES complex formation is exactly matched by its rate of decay, the rate of change in [ES] is zero (1.5).

$$[E] = [E_f] + [ES] \quad (1.3) \qquad [S_f] \simeq [S] \quad (1.4) \qquad \frac{d[ES]}{dt} = 0 \quad (1.5)$$

Using these assumptions to simplify the mathematics, an equation can be derived to describe steady-state kinetics. Looking at (1.1), the [ES] at a given moment is affected by two processes, the rate of ES complex formation (v_f) from S_f and E_f (1.6), and the rate of ES complex decay (v_d) into reactants or products (1.7). Since the total rate of change in [ES] is negligible from (1.5), the rate of formation and rate of decay must be equal (1.8).

$$v_f = k_1[S_f][E_f] \quad (1.6) \qquad v_d = (k_{-1} + k_2)[ES] \quad (1.7)$$

$$k_1[S_f][E_f] = (k_{-1} + k_2)[ES] \quad (1.8)$$

Rearranging this equation, an expression can be obtained for [ES] (1.9).

$$[\mathrm{ES}] = \frac{\mathrm{k_1}[\mathrm{S_f}][\mathrm{E_f}]}{\mathrm{k_{-1}} + \mathrm{k_2}} \quad (1.9)$$

The rate constants in (1.9) are abbreviated as the Michaelis constant (K_M), which represents the apparent dissociation constant for the ES complex (1.10). This results in a simplified expression for [ES] (1.11).

$$K_M = \frac{k_{-1} + k_2}{k_1} \quad (1.10) \qquad [\mathrm{ES}] = \frac{[\mathrm{S_f}][\mathrm{E_f}]}{\mathrm{K_M}} \quad (1.11)$$

Further substitutions can be made to (1.11) such that [ES] is expressed in terms of total, as opposed to free, species concentrations (1.12). These substitutions include the assumptions made earlier: that [E] can be written as a sum of the concentrations of the enzyme species (1.3), and that substrate depletion is negligible (1.4).

$$[\mathrm{ES}] = \frac{[\mathrm{S}][\mathrm{E}]}{K_M + [S]} \quad (1.12)$$

Substitution of (1.12) with the relation from Brown's model, that the rate of product formation (v) is governed by [ES] and k_2 (1.2), results in a description for v in terms of total enzyme and substrate concentrations (1.13).

$$\mathrm{v} = \frac{k_2[\mathrm{S}][\mathrm{E}]}{K_M + [S]} \quad (1.12)$$

From a plot of v against [S], it is clear that as [S] approaches infinity the rate of product formation asymptotically approaches some maximum value (V_{max}, Figure 1.15). When substituting an arbitrarily large [S] into (1.12) such that [S] >> K_M, the K_M becomes negligible and so the [S] terms cancel, resulting in an expression for V_{max} (1.13). This can then be inserted into (1.12) to obtain a final expression for v in terms of V_{max}, [S], and K_M (1.14).

$$\mathrm{V_{max}} = \mathrm{k_2}[\mathrm{E}] \quad (1.13) \qquad \mathrm{v} = \frac{\mathrm{V_{max}}[\mathrm{S}]}{K_\mathrm{M} + [\mathrm{S}]} \quad (1.14)$$

Interestingly, the final equation of the steady-state model is almost identical to the one found earlier by Michaelis and Menten using the rapid equilibrium model, with the only difference being the identity of the equilibrium constants. The steady-state model more accurately takes into account the value of k_2 and uses K_M (1.10), the apparent dissociation constant for the ES complex. On the other hand, the rapid equilibrium model assumes $k_2 << k_{-1}$ and uses K_S (k_{-1}/k_1), the thermodynamic dissociation constant for the ES complex. The similarity between the final equations of the two models has led to (1.14) being commonly known as the Michaelis-Menten equation.[64]

1.3.2. Fluorescent Substrates

Although Michaelis and Menten developed an excellent procedure to monitor enzyme activity based on initial rates, a variety of detection methods had to be developed to generalize this procedure to other enzymes. In the case of invertase, the enzyme studied by Michaelis and Menten, polarimetry was used to distinguish between reactant and products.[62] Although polarimetry was an advanced technique for its day, a plethora of techniques with improved sensitivity and adaptability exist today. One simple technique is colourimetry, in which the appearance or disappearance of a coloured compound is monitored. This can easily be used as a qualitative measurement of enzyme activity, as the change in colour can be observed by eye, however quantitative measurements can also be obtained using a spectrophotometer. Extending this approach to monitor changes in UV absorption is particularly useful since many substances of interest absorb in this region of the electromagnetic spectrum.[65] Enzymes, such as chitinases, that can degrade insoluble material to produce soluble metabolites can be studied by turbidimetry, which measures decreases in transmitted light intensity due to scattering by suspended particles.[66] Radioactive assays are often used to monitor the incorporation of small molecules into larger biomolecules, such as the phosphorylation of proteins using radiolabelled ATP.[67] Simple assays using radiolabelled substrates require the labelled product to be separated from the excess substrate for quantification. This complication can be surmounted using scintillation proximity assays in which the target is bound to a scintillant, and only substrates that come into proximity of the scintillant by association with the target will result in a signal.[68]

In recent years there has been a surge in the use of fluorescence techniques to study enzyme activity, in part as a desire to move away from radiolabelling which can pose safety concerns, be inconvenient and cumbersome, and for which proper use and waste disposal can be complicated and costly. Moreover, the sensitivity of fluorescence measurements, which are made directly with respect to a dark background, over other optical techniques, which are made with respect to a bright reference beam background, has increased the attractiveness of fluorescent substrates.[69] There are two general designs for fluorescent substrates: those that are intrinsically fluorescent, and those whose fluorescence is activatable. The more conceptually simple of the two are the intrinsically fluorescent substrates, in which an appended fluorescent compound results in the substrate itself being fluorescent. These compounds are akin to simple radiolabelled compounds in that the signal is always detectable. As such, in order to employ intrinsically fluorescent substrates a method must be used to separate the excess substrate and the product resulting from enzyme processing. Commonly, separation techniques such as filtration, chromatography, immunoprecipitation, and centrifugation are used. Intrinsically fluorescent substrates are often employed for the study of transporters, as these proteins do not effect a chemical change on their substrates. An exemplary case is that of 2-(*N*-(7-nitrobenz-2-oxa-1,3-diazol-4-yl)amino)-2-deoxyglucose (NBD-Glc), an intrinsically fluorescent glucose analogue that is widely used to monitor glucose uptake activity in cells as a measure of cell viability (Figure 1.16).[70] The fact that transport proteins act to translocate their substrates into compartments affords a basis on which to remove any excess material – cells/vesicles can be washed so that only the internalized material is retained for quantification. Another class of substrate-like molecules which can be intrinsically fluorescent are activity-based probes (ABPs). These mimic an enzyme's natural substrate in order to be recognized, but during one of the catalytic steps they form a reactive species which covalently bonds with an enzymatic residue, resulting in labelling of the enzyme. A separation technique such as SDS-PAGE or immunoprecipitation can then be used so that the labelled protein can be identified. While a large number of ABPs have been designed for proteases,[71] examples also exist for GHs, such as a BODIPY labelled substrate for examination of α-fucosidase activity (Figure 1.16).[72]

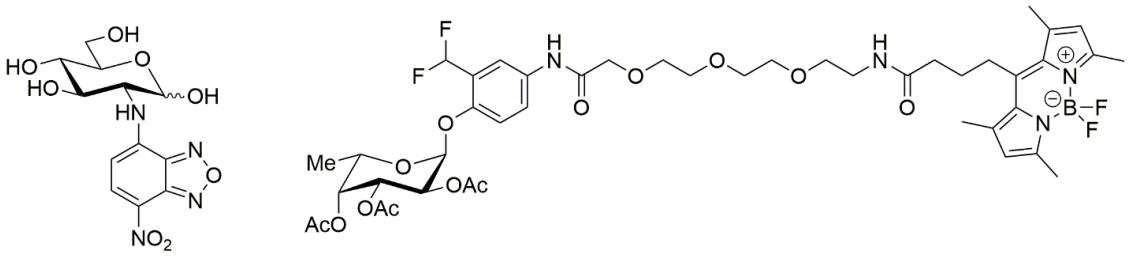


Figure 1.16. Examples of intrinsically fluorescent probe substrates: NBD-Glc and an α-fucosidase activity-based probe.

Separate from intrinsically fluorescent substrates are the activatable substrates which, while intact, exhibit minimal fluorescence at the desired emission wavelength. However, processing of the substrate by the target enzyme gives rise to a highly fluorescent product. As a result, the background signal arising from the unprocessed compound is often negligible for activatable substrates, allowing for real-time measurements and increased detection sensitivity as compared to intrinsically fluorescent substrates.[73] Generally, activatable substrates require the targeted enzyme to have hydrolytic activity in order to release a fluorescent moiety, though this is not universally true.[74] Activatable substrates can be further classified as either fluorogenic or quenched. Fluorogenic substrates are somewhat simpler in design and application than are quenched substrates, and because of this they are widely used to examine enzyme activity *in vitro* when the researcher has access to a relatively large amount of enzyme and to carefully controlled conditions. Fluorogenic substrates employ an array of fluorophores including, for example, coumarins (umbelliferones), phenoxazines (resorufins), and xanthenes (fluoresceins) (Figure 1.17).[73] These fluorophores exhibit a significant red-shift in their fluorescence when free compared to when they are covalently linked to the recognition feature of the substrate, and hence allow for turn-on fluorescence. In the case of the coumarins and phenoxazines, the red-shift is due to the phenolate's increase in intramolecular charge transfer (ITC) relative to the intact substrate,[75] whereas in the xanthenes an increase in conjugation due to lactone cleavage is responsible for the red-shift in fluorescence.[76]

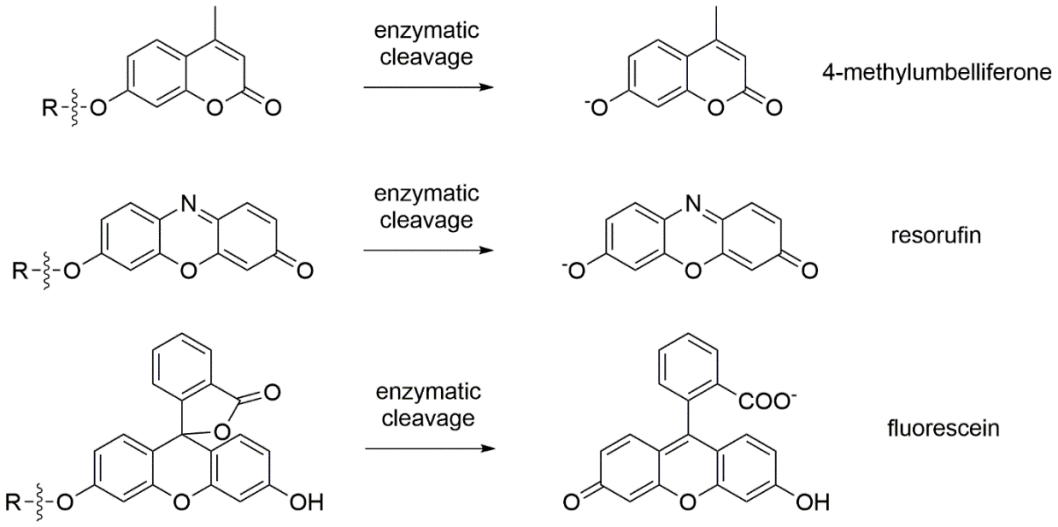


Figure 1.17. Fluorophores commonly used in the design of fluorogenic substrates.

While even the most simple fluorogenic substrates can be powerful tools to study enzyme activity, numerous advancements have been made to improve their utility. One particular challenge for the coumarin- and phenoxazine-based substrates is that the fluorescence intensity is dependent on the protonation state of the phenol; the phenolate ion is highly fluorescent while the phenol is not. One could foresee solving this issue by running enzymatic assays under basic conditions such that the vast majority of the free fluorophore molecules exist as the phenolate. However, enzymatic activity is highly pH dependent, with the majority of enzymes operating at the physiological pH of 7 which, for example, is significantly lower than the pH 10 optimum for coumarin fluorescence.[73] While the pH-dependence of fluorescence may not have a significant impact when monitoring highly active enzymes which produce a large quantity of product molecules, it can become an impedance when assaying low activity enzymes where measurements rely on the relatively few product molecules being highly fluorescent. Aside from circumventing this problem by using inconvenient stopped assays, in which a basic solution is added to ionize the free fluorophore, synthetic fluorophore derivatives can be designed such that the optimum pH for fluorescence coincides with the optimum pH for enzyme activity. Specifically, coumarin and phenoxide derivatives have been designed in which modification of the aromatic rings with electron-withdrawing groups (EWGs) stabilizes the buildup of negative charge on the phenolic oxygen, and thus increases the acidity of the phenol to afford pK_a values below the pH used in the assay.[77] In fact, even

in the simple case of EWG-substituted coumarins versus unsubstituted coumarins, measurement of the signal intensity found a 10-fold increase for the substituted derivative at physiological pH.[78] A side-effect of increasing phenolate stability in the coumarins and phenoxazine fluorophores is that they are made better leaving groups, which increases the rate of non-enzymatic hydrolytic cleavage. This can raise the background so much that the enzymatic rate of hydrolysis can be obscured in some cases.

To counter the increased propensity for non-enzymatic hydrolysis, self-immolative linkers have been designed which bridge the enzymatically cleavable bond and the fluorophore via a bond with increased stability towards acid-catalyzed hydrolysis. Upon cleavage of the enzymatically-labile bond, a cascade begins which results in rapid degradation of the self-immolative linker to yield the free fluorophore. Moreover, these linkers have the potential to prevent steric congestion from impacting the rate of enzymatic hydrolysis by placing the bulky fluorophore at a distance from the active site.[73] Recent work has applied self-immolative linkers as a technique to study β-galactosidases, with a p-benzoxycarbonyl-based linker used to join galactose and a synthetic phenoxazine-derived fluorophore (Figure 1.18).[79] Upon enzymatic cleavage of the glycosidic linkage, the adjoined fluorophore and linker are released and the linker undergoes rapid hydrolysis to yield the fluorophore, 4-hydroxy-3-nitrobenzyl alcohol, carbon dioxide, and glycine. Interestingly, hydrolysis of the released linker to yield glycine appeared to occur rapidly in water without the aid of another enzyme, as a mixture of the substrate and β-galactosidase in buffer showed glycine by HPLC. In addition to being fluorogenic this substrate is also chromogenic, with the intact substrate appearing yellow while the released 4-hydroxy-3-nitrobenzyl alcohol is blue in colour, thus providing a method to quantify the rate of linker breakdown.

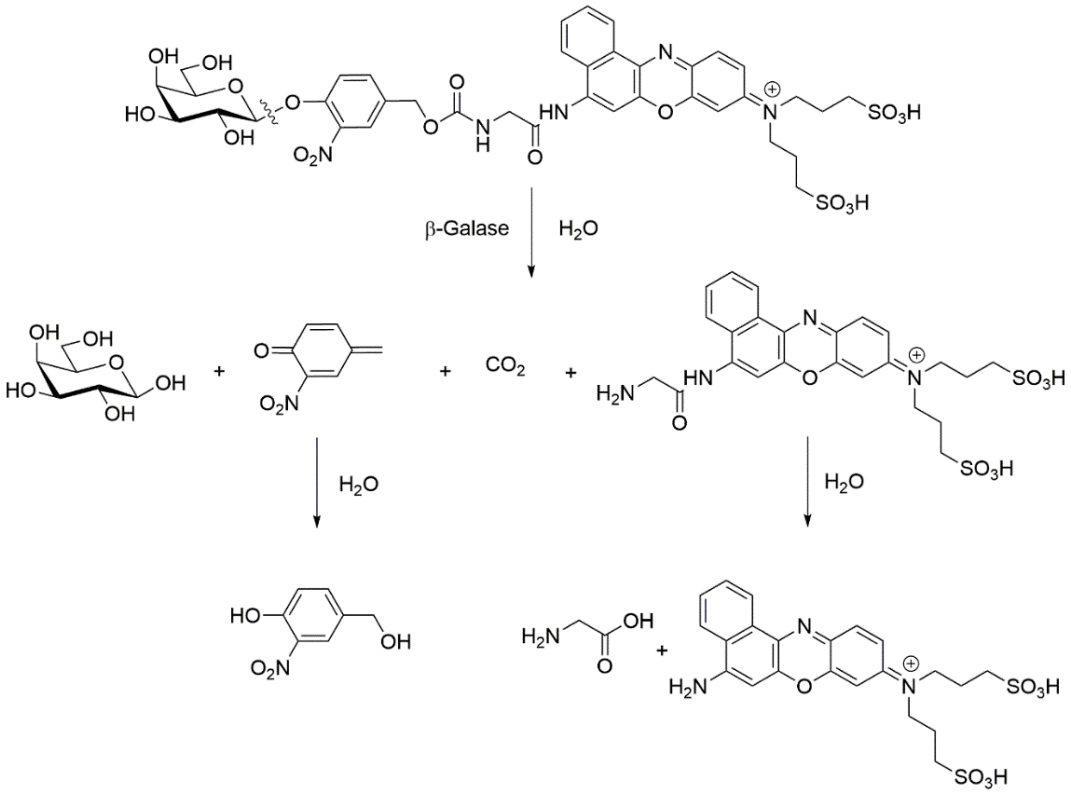


Figure 1.18. A self-immolative substrate for β-galactosidase, along with the proposed mechanism of linker breakdown.

Somewhat more complex than fluorogenic substrates are quenched substrates, which require a substrate to be substituted with a complementary fluorophore-quencher pair. While the substrate is intact, any energy which goes into excitation of the fluorophore is dissipated by the quencher, thereby preventing fluorescence. It is only once the fluorophore and quencher are no longer in contact that fluorescence is restored. As such, the locations at which these groups are installed is critical – hydrolysis of the enzymatically labile bond must result in separation of the fluorophore-quencher pair, allowing the two to diffuse apart so that the fluorophore can be detected. Many of the first examples of quenched substrates took advantage of the fluorescent properties of the amino acid tryptophan to study peptidases. One early example consisted of a dipeptide (Gly-Trp) to which a 5-(dimethylamino)-1-sulfonyl quencher was attached.[80] Upon cleavage of the peptide bond by carboxypeptidase A, tryptophan fluorescence was restored and so the activity of the enzyme could be monitored in real-time. However, the early examples of quenched substrates shared a limitation: their quenching mechanisms (mostly collisional) were very short-range processes,

necessitating nearly adjacent positioning of the fluorophore/quencher pair.[73] Since the majority of early fluorophores/quenchers were large aromatic molecules, the resulting steric bulk near the enzymatically cleavable linkage placed limitations on which enzymes could be studied using these substrates.[73] However, the design of substrates quenched by a longer range process, Förster resonance energy transfer (FRET), meant that a much greater diversity of enzymes could be approached. The first of such designs was for an HIV-1 protease and used the synthetic fluorophore/quencher pair, 5-[(2-aminoethyl)amino]naphthalene-1-sulfonic acid (EDANS)/4-(4-dimethylaminophenylazo) benzoic acid (DABCYL) (Figure 1.19), separated by an eight amino acid peptide.[81] In the years following, a great variety of fluorophore/quencher pairs have been described for the design of FRET quenched substrates: coumarin/p-nitroaniline,[82] 5-carboxyfluorescein/DABCYL,[83] and tetramethylrhodamide/black hole quencher-2,[84] to name a few (Figure 1.19). Although the short range of collisional quenching can be a detriment, it can also offer an advantage in that the proximity of the fluorophore and quencher can be known with high precision. This has particular use in examining protein dynamics, for which the range of FRET quenching is too large to precisely study the structure of proteins.[85] Moreover, identification of small collisional quenchers, such as thioamides which quench by photoinduced electron transfer (PET),[86] may open the possibility of collisionally quenched substrates being used to study hydrolytic enzymes that have sterically congested active sites. As a result of their potential for exceptionally low background fluorescence and the superior photophysical properties of the fluorophores used, quenched substrates are generally preferable over fluorogenic substrates for live-cell and *in vivo* assays. Further details about quenching mechanisms and the application of quenched substrates to assays in living systems will be described in following sections.

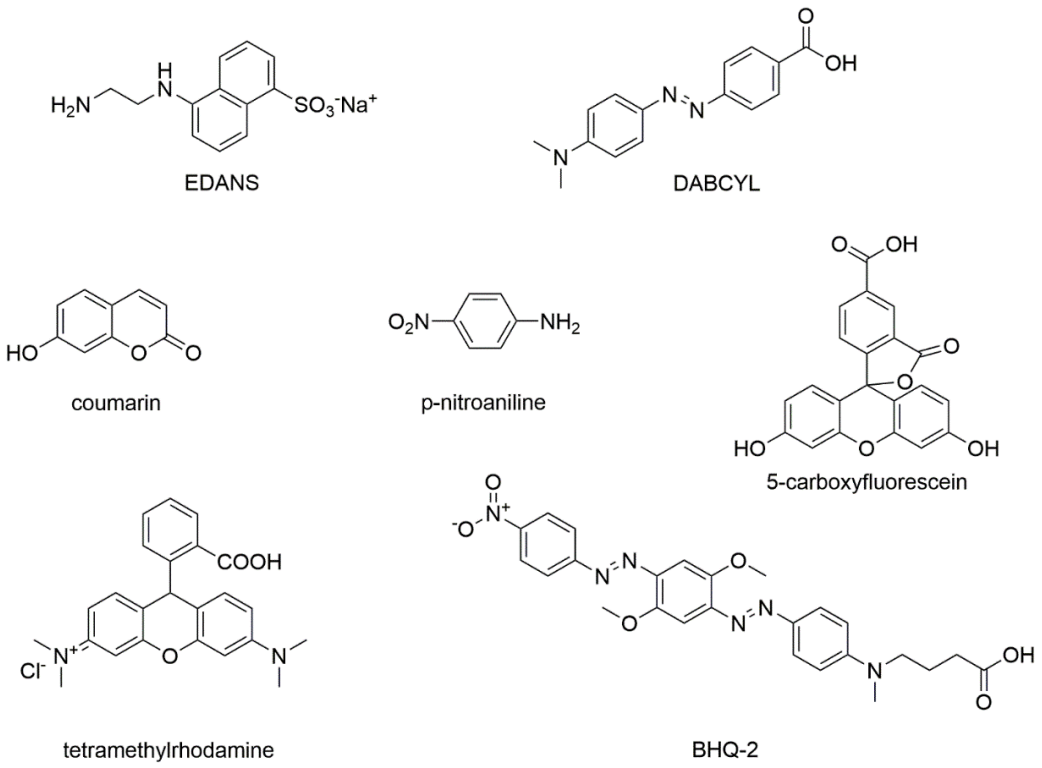


Figure 1.19. Examples of fluorophores and quenchers commonly used in the design on FRET quenched substrates.

1.3.3. Mechanisms of Fluorescence and Quenching

Fluorescence, along with phosphorescence, is a type of luminescence – emission of light from a substance resulting from relaxation of excited-state electrons. One requirement for fluorescence is that the excited-state electron and its paired ground-state electron have opposite quantum mechanical spin states. As such, return of the excited-state electron to the ground state is spin allowed as dictated by the Pauli Exclusion Principle, and so relaxation by the release of a photon happens rapidly with a lifetime (τ, the average time between excitation and relaxation) on the order of 10^{-8} seconds.[69] Phosphorescence is similar, except that the excited and ground-sate electrons have same quantum mechanical spin, and so the Pauli Exclusion Principle forbids the excited-state electron from relaxing to occupy the same quantum state as the ground-state electron. As such, the τ for phosphorescence is many orders of magnitude greater than for fluorescence, and gives rise to applications such as glow-in-the-dark toys.[69]

The processes of luminescence can be described pictorially using a Jablonski diagram, named after Polish physicist Aleksander Jablonski who was a pioneer in the field of fluorescence spectroscopy (Figure 1.20).[69] The electronic energy levels are depicted as line groupings, while individual lines within a grouping represent vibrational energy levels. The singlet electronic energy levels (S_0, S_1, S_2) represent the case where the paired electrons have opposite spin, while the triplet level (T_1) indicates paired electrons with the same spin. At room temperature and without an external stimulus, only the lowest electronic energy level (S_0) is occupied. Within S_0, the ground state vibrational level is by far the most populated, with thermal energy allowing only a small number of electrons to populate the higher vibrational energy levels.[69] Upon exposure to light, an S_0 electron (which is statistically likely to be in the ground state vibrational level) absorbs a photon and is promoted to a higher electronic energy level. The cumulative energy of the electronic and vibrational level to which an electron is promoted corresponds to the energy of the photon that was absorbed. After absorption, the excited electron can lose energy to vibrational relaxation within an excited electronic state, or to internal conversion from higher to lower excited electronic states. Internal conversion occurs rapidly (typically within 10^{-12} s), and vibrational relaxation even more so, and thus there is usually time for the excited electron to reach the lowest vibrational level of the first excited electronic state (S_1) before luminescence occurs. After sufficient time, an excited singlet-state electron can relax to S_0, and usually to an excited vibrational level therein, emitting a photon (fluorescence) in the process. The wavelength (colour) of the photon emitted is inversely proportional to its energy, which is simply the difference between the energy of the excited electronic/vibrational level from which the electron comes, and the energy of the electronic/vibrational level to which the electron relaxes. The photon emitted is universally of lower energy (higher wavelength) then the absorbed photon, since some of the absorbed energy is lost to vibrational relaxation and internal conversion after absorption, and because the relaxation usually occurs to an excited vibrational energy level within S_0. After emission, further vibrational relaxation brings the electron back to the ground vibrational state of S_0 from which it is likely to have originated. In an alternative process, after excitation an electron may undergo relaxation by spin-conversion to a triplet excited state in a process called intersystem crossing. From the lowest energy triplet state (T_1), phosphorescent relaxation to S_0 occurs

similarly to fluorescence, but over a much longer timescale due to the relaxation process being spin forbidden as mentioned previously.

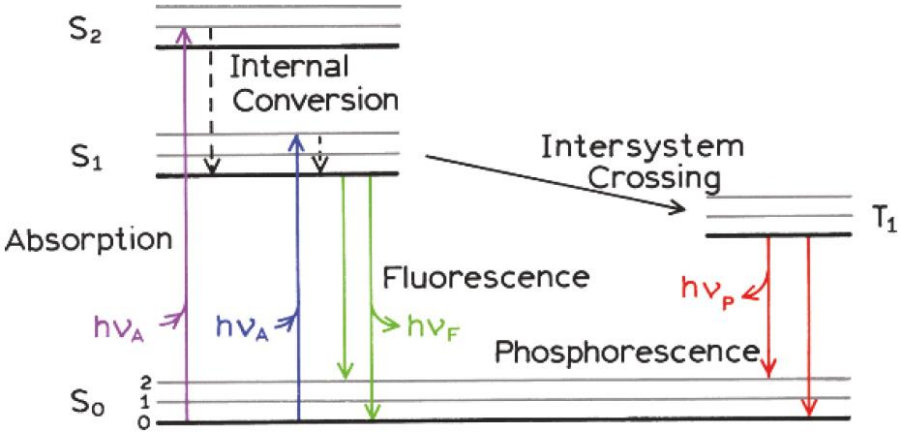


Figure 1.20. A Jablonski diagram indicating singlet (S_0, S_1, S_2) and triplet (T_1) electronic energy levels, associated vibrational levels, and various relaxation mechanisms.
Adapted, with permission, from Lakowicz 2006.[69]

When the energy of a fluorophore's excited-state electron is dissipated by a non-luminescent process, thereby causing a decrease in the fluorescence intensity, the fluorescence is said to be quenched.[69] Many mechanisms of fluorescence quenching exist, including excited-state reactions, molecular rearrangements, energy transfer, complex formation, and molecular collisions. Today, the most common quenching mechanism taken advantage of in the design of quenched enzyme substrates is energy transfer, and specifically Förster resonance energy transfer (FRET). In a FRET process, irradiation with an appropriate wavelength of light first causes an electron in the highest occupied molecular orbital of the ground-state fluorophore (FRET donor, D_F) to be promoted to the lowest unoccupied molecular orbital (Figure 1.21). Upon relaxation of the excited-state fluorophore (D_F*), the energy which would otherwise be converted to a photon is instead transferred to the ground-state acceptor (FRET quencher, A_F), causing an electron in its highest occupied molecular orbital to be promoted to the lowest unoccupied molecular orbital. The efficiency of the energy transfer depends on the amount of overlap between the emission spectrum of the donor and the excitation spectrum of the acceptor – greater overlap means an increased likelihood that the quanta of energy released upon relaxation of the donor will be appropriate to cause excitation of the acceptor. If the quencher is also fluorescent, relaxation of its excited

state (A_F*) can then occur to emit a photon. Alternatively, if the quencher is non-fluorescent then the excess energy of A_F* is dissipated as heat to return to A_F. If a fluorescent quencher is used, then the emission wavelength of the fluorophore should be significantly different from that of the quencher to enable distinguishing an intact substrate from the products of enzymatic cleavage.

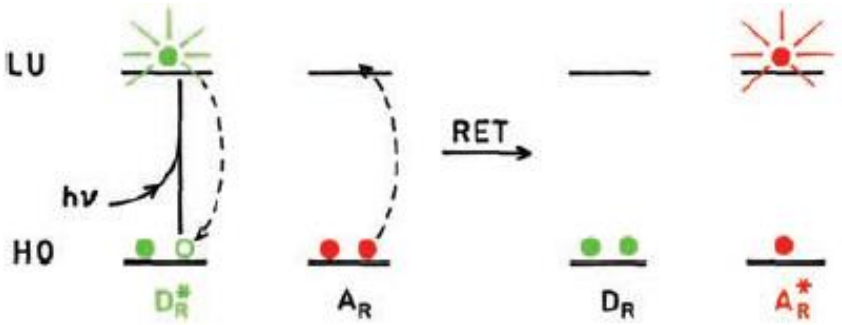


Figure 1.21. A molecular orbital depiction of FRET quenching.
Relaxation of a donor (fluorophore) excited-state electron from the lowest unoccupied molecular orbital to the highest occupied molecular orbital results in non-radiative resonance energy transfer (RET) to the ground-state acceptor (quencher), promoting an electron. The quencher can then relax either through luminescence or through a non-radiative process such as collisions. Adapted, with permission, from Lakowicz 2006.[69]

FRET is a non-radiative through-space process, meaning that it does not rely on interactions between the electron clouds of the fluorophore and quencher.[69] As such, FRET efficiency is mostly independent of factors that affect the propensity of close-range interactions, such as steric shielding and electrostatics. The efficiency of FRET (E_F) for a particular fluorophore-quencher pair is described by equation (1.15).

$$\mathrm{E_F} = \frac{1}{1 + \left(\frac{r}{R_0}\right)^6} \quad (1.15)$$

In this equation, r is the centre-to-centre distances between the electron clouds of the fluorophore and quencher, and R_0 is the Förster distance – the distance at which FRET is 50% efficient. The value of R_0 varies depending on the overlap of the emission spectrum of the fluorophore and the excitation spectrum of the quencher, and is typically between 20 Å and 60 Å.[87] As can be seen, the FRET efficiency decreases with the 6th power of r, and so FRET quenching is rapidly negated as the fluorophore and quencher

diffuse beyond R_0. In fact, the strong distance-dependence of the FRET plays a major role in making FRET-quenched substrates such powerful tools, since the increase in average distance between the fluorophore and quencher upon enzymatic processing, even if only modest, yields a dramatic increase in signal.

Collisional quenching processes were the first employed in the development of quenched enzyme substrates. However, the short-range of collisional quenching mechanisms, requiring interaction of the electron clouds of the fluorophore and quencher,[69] meant that the only enzymes amenable to this technique were those having active sites that could tolerate steric crowding.[73,80,88,89] Three collisional quenching mechanisms exist: intersystem crossing, electron-exchange, and photoinduced electron transfer (PET).[69] The short-range nature of collisional quenching means that the rate of diffusion, sterics, charge-charge interactions, and the lifetime of the fluorophore all impact the likelihood of quenching.[69] The efficiency of collisional quenching (E_C) is given by equation (1.16).

$$\mathrm{E_c} = \frac{1}{1 + \frac{1}{\tau A e^{-\beta(r - r_c)}}} \qquad (1.16)$$

In equation (1.16), τ is the lifetime of the fluorescent state, β and A are constants relating to the distance and time dependencies of orbital interactions, respectively, r is the centre-to-centre distance between the fluorophore and quencher, and r_c is the distance of closest approach at molecular contact.[69] From the equation, it is clear that E_c approaches unity as τ increases, which is logical since a longer lifetime means that the fluorophore and quencher have more time to encounter one another so that quenching can occur. Moreover, as r approaches the value of r_c the exponential term tends toward unity. In such a case E_c depends only on τ and A, and approaches unity if the fluorescent state is long-lived (large τ) and orbital interactions occur frequently (large A).[69]

Although collisional quenching is a short-range process, the technique gains applicability for quenched enzyme substrates if small quenchers are used that will not adversely influence enzyme function. Thioamides have been shown to quench

fluorescence of many fluorophores, and are anticipated to do so by a PET mechanism.[86] As for FRET, PET quenching can be depicted using a simple molecular orbital diagram (Figure 1.22). The fluorophore (PET donor, D_P), has an electron promoted from the highest occupied molecular orbital to the lowest unoccupied molecular orbital upon excitation with a wavelength of light corresponding to the energy difference between the orbitals. If, during the lifetime of the excited state (τ), the excited fluorophore (D_P*) comes into contact with the quencher (PET acceptor, A_P), a charge-transfer complex ($D_P^+A_P^-$)* is formed by transfer of the fluorophore's excited state electron to the lowest unoccupied molecular orbital of the quencher. The complex can then return to the ground state with or without emission of a photon, although if emission does occur the wavelength is shifted with respect to the emission wavelength of the fluorophore. Once in the ground state, the complex dissociates with transfer of an electron from the quencher back to the fluorophore. The case that the fluorophore acts as the electron donor and the quencher as the electron acceptor is not universal, as the direction of electron transfer depends on oxidation and reduction potentials, however the principles remain the same regardless.[69]

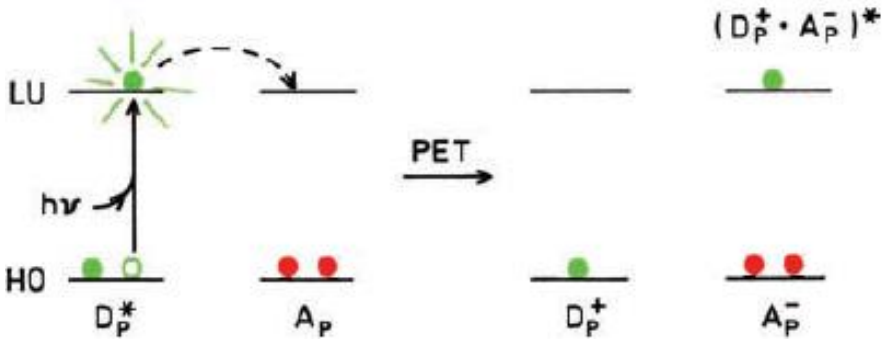


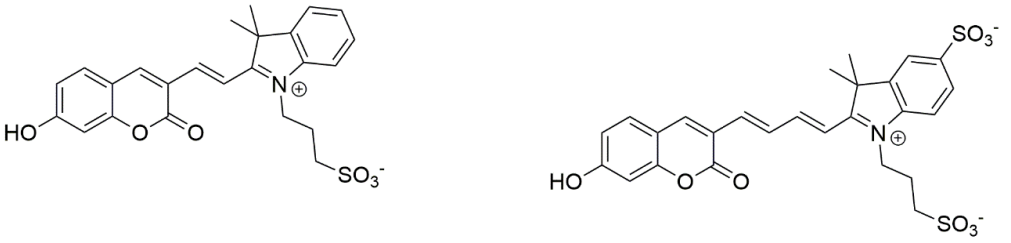
Figure 1.22. A molecular orbital depiction of PET quenching.
Collision of the excited donor (fluorophore) with the ground-state acceptor (quencher) during the lifetime of the excited state can result in a transfer of an electron from the lowest unoccupied molecular orbital of the donor to that of the acceptor, forming a charge-transfer complex. Transfer of the electron back to the donor may result in emission, but at a wavelength distinct from that of donor due to energy losses from the collision and from relaxation of the charge-transfer complex. Adapted, with permission, from Lakowicz 2006.[69]

1.3.4. Properties of Live-Cell and *in vivo* Fluorescent Substrates

While the *in vitro* study of enzymes is certainly invaluable, providing a convenient and rigorous method to identify kinetic parameters, screen substrates/inhibitors, and better understand catalytic mechanisms, the information gleaned is nevertheless limited. This arises from the simple fact that no matter how well suited an *in vitro* technique may be, it will still fall short from capturing the true complexity present in an enzyme's endogenous environment. While the function of an enzyme is tied to its primary sequence and tertiary structure, a host of other factors play essential roles in its activity, including: spatiotemporal regulation of enzyme expression, interaction of the enzyme with molecular effectors, and posttranslational modifications to the enzyme.[90] As a result, methods that allow the study of enzymes within the physiologically relevant environment of whole cells or a living organism are invaluable, and ultimately offer novel approaches towards understanding, diagnosing, monitoring and treating diseases caused by enzyme deficiencies.[73] However, the interior of a cell, be it part of a single- or multi-cellular organism, is a highly complex mixture of molecules, and so a number of factors must be considered when approaching the design of substrates for live-cell or *in vivo* assays. In brief, ideal fluorescent substrates for such applications must: (a) be dark when intact and bright when processed, (b) have significantly red-shifted product fluorescence, (c) be selective for the enzyme of interest, and (d) be cell permeable.

The stipulation that a live-cell or *in vivo* substrate should be dark when intact (have a low background signal), while being highly fluorescent when processed, ensures that even a small signal can be detected. This is necessary because unlike *in vitro* assays, in which the enzyme concentration can be greatly increased in order to obtain adequate signal, endogenous enzyme levels may be comparatively low and are not easily manipulated. Even in the case of transfected cells, in which enzyme expression is induced by the researcher, oftentimes overexpression is limited by toxicity or strength of the promotor system used.[91] Ensuring that the fluorescent product of the enzymatic cleavage is bright increases the sensitivity of the substrate by allowing the signal to be detected over any residual fluorescence of the intact substrate and over cellular autofluorescence. Further gains in sensitivity and an increase in tissue penetration can be achieved by using a fluorophore whose emission is shifted towards the near-infrared

(NIR) region of the electromagnetic spectrum. Common fluorophores, which generally emit in the ultraviolet (UV) or visible region of the spectrum, suffer a decreased signal-to-noise ratio when used *in vivo* due to scattering of visible light and the propensity of biomolecules to absorb and emit UV-visible light.[92] Further, other major biological components such as water and lipids strongly absorb in the infrared.[93] Thus, the NIR region provides an optimum window for observing fluorescence in a living system. Notably, conventional dyes can have their emission profiles shifted towards the NIR region by extension of the π-conjugation system, although this can lead to complications such as increased sensitivity to photobleaching and decreased water solubility. However, such problems can be circumvented by further modification of the fluorophore, as was done in the case of a 7-hydroxycoumarin derivative for which the water solubility issue was solved by introducing sulfonate moieties (Figure 1.23).[94]



orange fluorescence, low water solubility NIR fluorescence, high water solubility

Figure 1.23. Extension of the π-conjugation system and installation of sulfonate moieties shifts fluorescence to the NIR region and improves water solubility.

Separate from the photochemical traits preferred for *in vivo* substrates are the requirements of selectivity cell permeability. When an enzyme is studied *in vitro*, it is often the case that a pure or nearly pure sample of the enzyme is being used. As a result, the researcher can be fairly certain that any observed signal arises from catalysis by the enzyme of interest, assuming that non-enzymatic hydrolysis was also tested for using a deactivated enzyme preparation as a control. The same assumption cannot be made as easily in an *in vivo* assay however, as the substrate may be recognized and processed by other enzymes within the cell or tissue. Even in the case that the substrate is processed most efficiently by the enzyme of interest, minor off-target activities from other enzymes can still obscure the results of the assay. Circumventing this issue often involves engineering a compound to have increased affinity for the target enzyme while

having decreased affinity for off-target enzymes. For example, protease substrates containing unnatural amino acids can be used to find substrates which are selective for particular members of the same protease family.[95] The use of unnatural amino acids, as opposed to the 20 classically natural amino acids, provides a huge increase in the number of structures that can be evaluated, and thus greatly increases the likelihood that a sequence specific for one enzyme can be found.

Most fluorophores in use today are highly conjugated organic molecules, and so it would seem that their hydrophobic character would lend to their cell permeability. However, their overall bulk, as well as installation of charged moieties to increase water solubility,[90] tends to detract from their ability to passively diffuse across cell membranes. Accordingly, methods to increase the cell permeability of substrates are desired. Cell-penetrating peptides (CPPs), short peptide sequences that are recognized and internalized by cells, often by endocytosis, can be used in this regard.[96] By attaching CPPs to fluorescent substrates their cell permeability has been dramatically increased, facilitating imaging of, for example, caspase activity in live cells.[97,98] Interestingly, the attachment of moieties to increase cell-penetration is not limited to CPPs, other tags such D-glucosamine in the case of a TEM-1 β-lactamase substrate have also been examined (Figure 1.24).[99] In situations where solubility is less of an issue or can be recovered in another way, the 'prodrug' approach can be taken in which polar groups are chemically modified to afford an increase in cell permeability of the prodrug. On permeation, these modifications are removed by the action of endogenous enzymes, unmasking the substrate and allowing it to be processed by the enzyme of interest. Moreover, the difference in permeability between the masked an unmasked substrate results in accumulation within the cell and a concomitant increase in signal. A β-galactosidase substrate has been designed in this way – the acetylated galactose provides an increase in cell permeability leading to a buildup of substrate within the cell, and endogenous esterases act to yield the active substrate (Figure 1.24).[100]

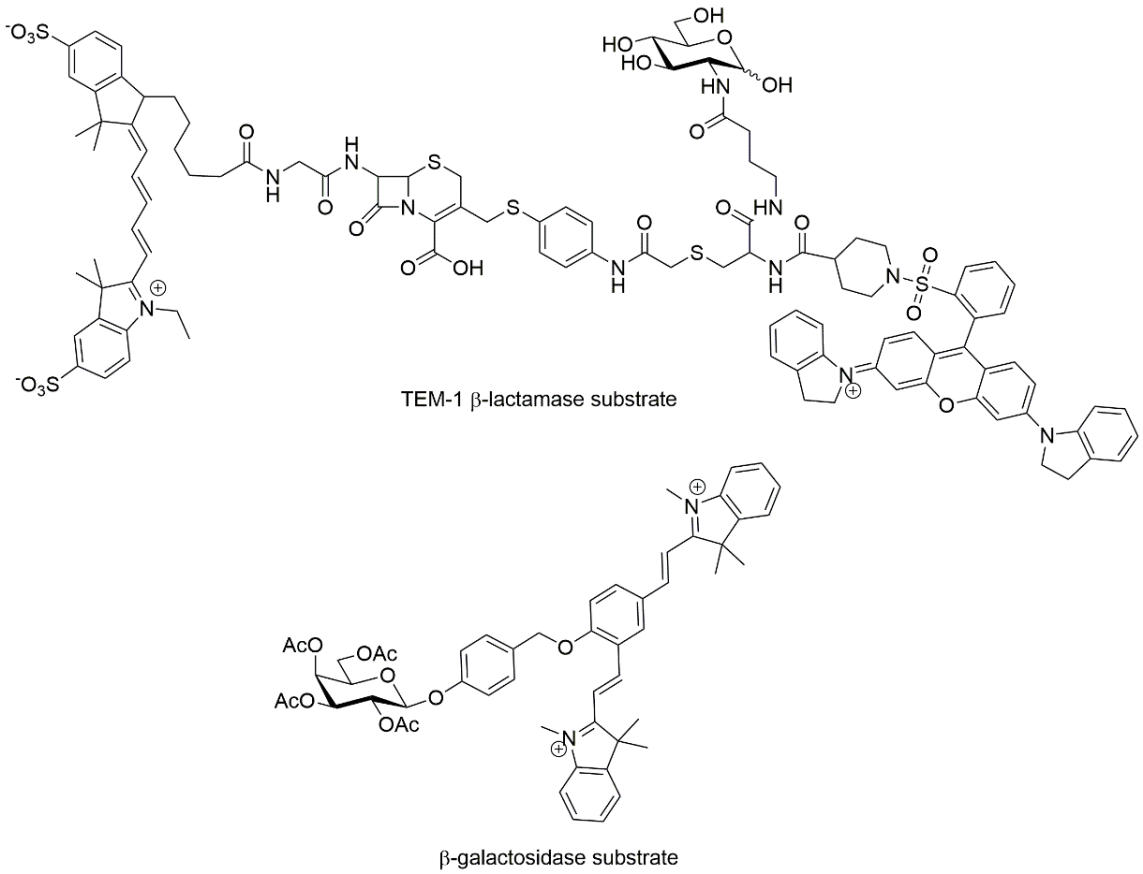


Figure 1.24. Cell permeable fluorescent substrates.
A TEM-1 β-lactamase substrate which makes use of glucosamine transporters by attachment to the *N*-acetyl moiety. A β-galactosidase substrate which uses the 'prodrug' approach to increase cell permeability by acetylation of the sugar.

1.3.5. Assays of Glycoside Hydrolases

Since glycoside hydrolases (GHs) catalyze a hydrolytic reaction, they make ideal candidates for study using activatable substrates. Moreover, the fact that GHs are usually specific for a given recognition saccharide and anomeric stereochemistry should facilitate their study in the complex environment of a living system. Simple *in vitro* assays of glycoside hydrolases typically use a fluorogenic substrate, such as 4-methylumbelliferyl 2-acetamido-2-deoxy-β-D-glucopyranoside (4MUGlcNAc) which is typically used to study β-hexosaminidases (Figure 1.25). In fact, *in vitro* techniques for GHs have become so refined that many high-throughput screening (HTS) methods have been developed to identify potential therapeutics. One such assay using a fluorogenic substrate has been developed for the human enzyme α-glucosidase A (GAA), the

deficient enzyme in Pompe disease.[101] GAA is a lysosomal enzyme belonging to GH family 31, and is responsible for hydrolyzing the terminal α-1,4 and α-1,6 linkages of glycogen. Those afflicted with Pompe disease, also called glycogen storage disease type II, suffer from progressive muscle weakening, and do not usually survive beyond two years of age.[102] Recently, small-molecule chaperones have been shown to rescue unstable proteins, and in particular the aminosugar N-butyldeoxynojirimycin has been shown to have this effect for mutant GAA.[103] Accordingly, development of a HTS to identify additional GAA activators was thought to be a promising approach to drug development for Pompe disease. To this end, resorufin α-glucoside (Res-α-Glu) was synthesized (Figure 1.25), with resorufin chosen for its red-shifted fluorescence in comparison to the previously trialed fluorophore 4-methylumbelliferone.[101] Red-shifted fluorescence is important not only for live-cell or *in vivo* assays, but also for *in vitro* HTS assays due to the propensity for many compounds from chemical libraries to fluoresce in the blue region of the spectrum, and because small particles such as dust tend to add to the background in this region.[104] Moreover, the increased acidity of the phenol moiety of resorufin as compared to 4-methylumbelliferone enables use of a continuous kinetic assay, which is much more amenable to HTS than are stopped assays. Using Res-α-Glu in a 1536-well plate format, a standard library of 1280 pharmaceutically active compounds was screened with GAA, from which two compounds were identified as GAA inhibitors.[101] In comparison, an assay using the blue-fluorescent substrate 4MU-α-Glu appeared to have also found three GAA activators. However these compounds were found to be false positives due to inherent fluorescence in the blue region of the spectrum, again highlighting the concerns of using such fluorescent substrates for a HTS application.

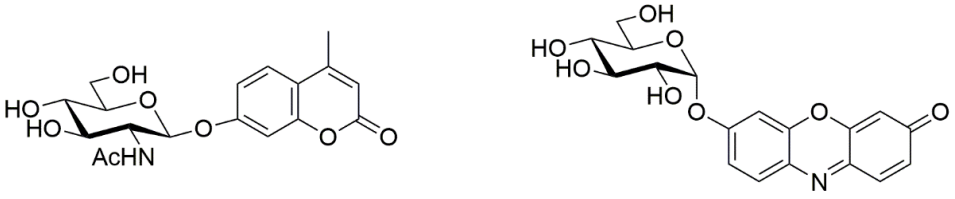


Figure 1.25. Substrates for *in vitro* assays of β-hexosaminidases (4MUGlcNAc), and for high-throughput screening of α-glucosidase A (Res-α-Glu).

While the optical properties of fluorogenic substrates are not always well suited to live-cell or *in vivo* experimentation, this is not universally the case, as is exemplified by a fluorescein-based β-galactosidase substrate (HMRef-βGal) that has been used for *in vivo* imaging of tumors (Figure 1.26).[105] Traditionally, *E. coli* β-galactosidase, expressed from the gene *LacZ*, has been extensively used as a marker enzyme to study gene expression. Interestingly, human β-galactosidase has been found to be overexpressed in ovarian[106] as well as breast and colon cancer cells,[107] and thus could act as a marker to identify metastases originating from these tissues. The design of HMRef-βGal improved upon previous β-galactosidase substrates, which either were cell impermeable, had their fluorescent products exported by overexpressed organic anion transporters, or whose low fluorescence enhancement upon cleavage prevented specific visualization of cancerous tissue.[105] Although HMRef-βGal fluorescence occurs in the blue region of the spectrum, fluorophore red-shift is not critical for detection of surface tumors by laparotomy or laparoscopy, since the light does not need to penetrate tissues.[105] Moreover, the fluorescence enhancement and brightness of HMRef-βGal proved to be significantly greater than the previously reported blue-fluorescent compound which was used in cultured cells, suggesting usefulness for *in vivo* imaging of surface tumors. The *in vivo* testing of HMRef-βGal was carried out using a tumor mouse model, which was prepared by intraperitoneal injection of cancer cells into live mice followed by two to four week incubation. After adequate time was allowed for tumors to develop, mice were injected intraperitoneally with HMRef-βGal, which was allowed time to incubate before imaging. Importantly, an HMRef-βGal dose 10-times higher than was used for imaging caused neither significant body weight change nor lethality in the mice. Even after 5 min incubation, tumors as small as 1 mm were detectible in sacrificed mice using an imaging system, and after 1 hr incubation the fluorescence was intense enough to be detected by the unaided eye and enable real-time fluorescence-guided tumor resection (Figure 1.26).[105] The short incubation period required before detection could allow rapid diagnosis of suspicious lesions encountered during surgery. Moreover, tumors of anesthetized mice were even able to be visualized through a small abdominal incision using an endoscope equipped with a fluorescence detection system.

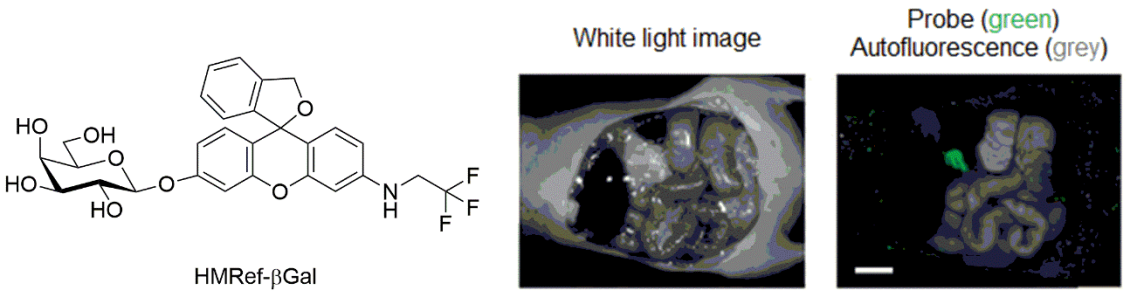


Figure 1.26. HMRef-βGal, a substrate used for *in vivo* imaging of β-galactosidase which is overexpressed in ovarian tumor metastases, and images of abdominal surgery on mice treated with HMRef-βGal.
Adapted, with permission, from Asanuma *et al.* 2015.[105]

Owing to their generally lower background and the improved optical properties of the fluorophores used, quenched substrates are often the preferred choice for live-cell or *in vivo* enzyme assays. However, using quenched substrates to study glycoside hydrolases, especially exo-acting enzymes with pocket shaped active-sites, can be a challenge since the fluorophore and quencher can be sterically bulky and often have to be appended directly to the features of the substrate important for recognition by the enzyme. However, the approach can be tractable if the enzyme active-site is relatively open or if a flexible linker can be installed as a bridge to orient the bulky substituent away from the active-site. A recent paper elegantly applied a FRET quenched substrate to study the activity of glucocerebrosidase (GCase) in live human cells (Figure 1.27).[108] A lysosomal glycolipid hydrolase classified in GH family 30, GCase is responsible for the catabolism of glucosylceramide to yield glucose and ceramide.[33] Not only is GCase the defective enzyme in patients with the most prevalent lysosomal storage disorder, Gaucher's disease,[109] recent work has also showed a link between GCase and Parkinsonism.[110] To allow optimization substrate properties, a modular synthesis was designed in which installation of the fluorophore and quencher were the final steps. While a substrate employing an EDANS/DABCYL fluorophore/quencher pair was synthesized as a proof-of-concept for *in vitro* use, the fluorescence maxima (490 nm) precluded live-cell use and so an alternate substrate using BODIPY and BHQ-2 as the fluorophore and quencher, respectively, was prepared.[108] Rather than take the traditional approach of appending the fluorophore to the sugar by the enzymatically labile bond, these substrates instead placed the quencher at this position (C-1). By designing the substrates in this way, it was anticipated the resolution and sensitivity of the assay would

be improved, since the fluorophore would remain appended to the polar sugar after enzymatic activity and so would likely be unable to diffuse out of the lysosome. Since GCase has been shown to tolerate modification to C-6,[111] it was at this position that fluorophores were installed. Design of the substrate also drew parallels from the natural substrate, in which the aliphatic ceramide moiety is accommodated at the anomeric centre, by installing an aliphatic linker between glucose and the anomeric quencher to prevent steric congestion in the enzyme's active site. Quenching efficiencies of 97.6% and 99.9% were determined for the EDANS and BODIPY substrates, respectively, and their catalytic efficiencies (k_{cat}/K_M) with respect to GCase were determined to be within an order of magnitude of the widely used GCase substrate 4-methylumbelliferyl β-D-glucopyranoside (4MUGlc).[108] When used in human cells, signal arising from the BODIPY substrate was localized to the lysosome, increased in a time-dependant manner (Figure 1.27), and showed no toxicity after up to 20 hr incubation. Pre-treatment of human cells with conduritol β epoxide, which inactivates GCase but not other β-glucosidases, abolished processing of the BODIPY substrate, indicating selectivity for GCase.

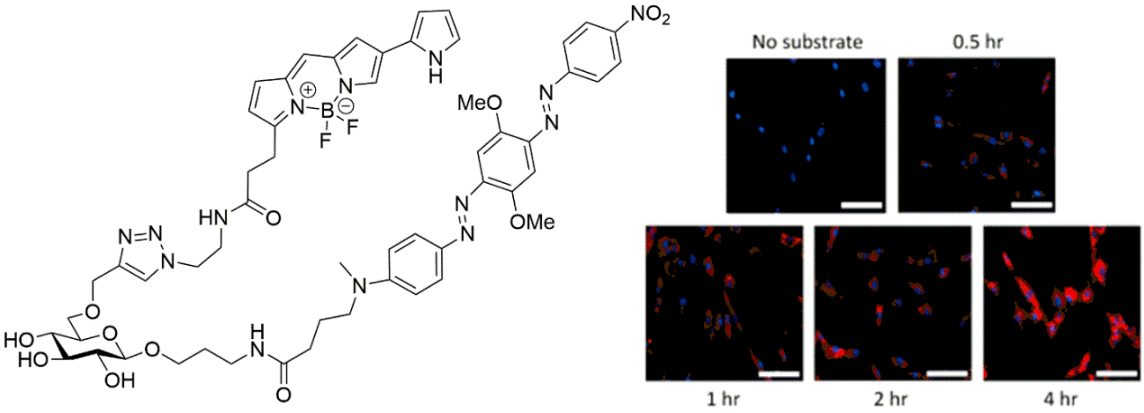


Figure 1.27. A FRET quenched substrate used to monitor GCase activity in live human cells. Images show a time-dependant increase in signal.
Adapted, with permission, from Yadav *et al*. 2015.[108]

1.3.6. Assays of Sugar Transporters

The functional analysis of transport proteins is more challenging and laborious than analysis of glycoside hydrolases. This is due in large part to the fact that transporters do not effect a chemical change on their substrates, and so the visualization

strategies used for GHs, such as fluorescence enhancement on catalysis, are not readily applicable for transporters. Moreover, being native to biological membranes, transport proteins are highly hydrophobic and so are intrinsically more difficult to handle in aqueous media.[112] Accordingly, assays of transporters require that the protein be imbedded in a natural or artificial membrane,[112] and usually monitor the amount of a labelled substrate entering or exiting a vesicle as a function of time. Also possible is to measure changes in vesicle volume or pH.[113] Much of the pioneering work on carrier-mediated transport was done on the *E. coli* lactose permease, the founding member of the major facilitator superfamily (MFS), which was first identified as a specific system for lactose transport and metabolism in 1956.[112,114] It was quickly recognized that the lactose permease operates by the protonmotive force, with protons being transported down their electrochemical gradient. However, it was not until 1979 that a lactose transport assay using radioactively labelled lactose was used to identify that the permease acted by symport, with a 1:1 proton:lactose stoichiometry, and was independent of extracellular pH.[115]

Radiolabelling strategies have also been used to determine kinetic parameters and substrate selectivity for transport proteins. Human glucose transporters GLUT-1, GLUT-2, and GLUT-3 were characterized by this strategy using the non-metabolizable glucose analogues 3-*O*-methyl-D-glucose (3-*O*-MG) and 2-deoxy-D-glucose (deGlc) (Figure 1.28).[116] These permeases, along with other GLUT glucose transporters, move glucose down its concentration gradient by facilitated diffusion and are found throughout the body. Many GLUT isoforms exist and their affinity for glucose, as well as their spatiotemporal expression, varies significantly in order to regulate the distribution of this essential nutrient.[117] In order to study the kinetics of GLUTs 1 – 3, an equilibrium-exchange assay was employed in which labelled and unlabelled 3-*O*-MG were allowed to exchange across the membrane of frog oocytes.[116] The oocyte system was chosen due to their particularly low levels of endogenous glucose transport activity.[116] Oocytes, each expressing one of the GLUT isoforms, were incubated with a range of unlabelled 3-*O*-MG concentrations, and time was allowed for the extra- and intracellular concentrations to equilibrate. Labelled 3-*O*-MG was then added to the equilibrated oocytes, and exchange was allowed to occur. At multiple timepoints the reaction was stopped by aspirating the solution and washing the oocytes thrice with ice-cold buffer

containing the transport inhibitor phloretin. The oocytes were then transferred to scintillation vials and lysed with sodium docecyl sulphate, and the radioactivity arising from the internalized 3-*O*-MG was measured. A first-order plot was used to obtain the observed rate constants (k_{obs}) for each unlabelled 3-*O*-MG concentration, and a Lineweaver-Burk plot was used to determine the K_M for GLUTs 1 – 3 by plotting $1/k_{obs}$·[3-*O*-MG] against 1/[3-*O*-MG]. The substrate specificity of GLUTs 1 – 3 was also examined (Figure 1.28), but deGlc was used since its rapid intracellular phosphorylation prevents efflux, which allowed study of the outward-facing binding site of the GLUT transporters in isolation.[116] Oocytes were first incubated with an unlabelled sugar (one of either the D- or L-isomer of glucose, mannose, arabinose, or xylose, or the D-isomer of fructose), and then an aliquot of labelled deGlc was added. After time was allowed for transport to occur, the reaction was stopped and the radioactivity was measured for each oocyte as before. Interestingly, differential selectivity between the three isoforms was observed. While D-glucose and D-mannose competed with deGlc for transport in all three isoforms, GLUT-2 was the only isoform for which D-fructose showed any significant competition, and only GLUT-3 showed competition by D-xylose. None of the isoforms showed competition by the L-isomers, which is unsurprising since these isomers are not native to animal biochemistry.

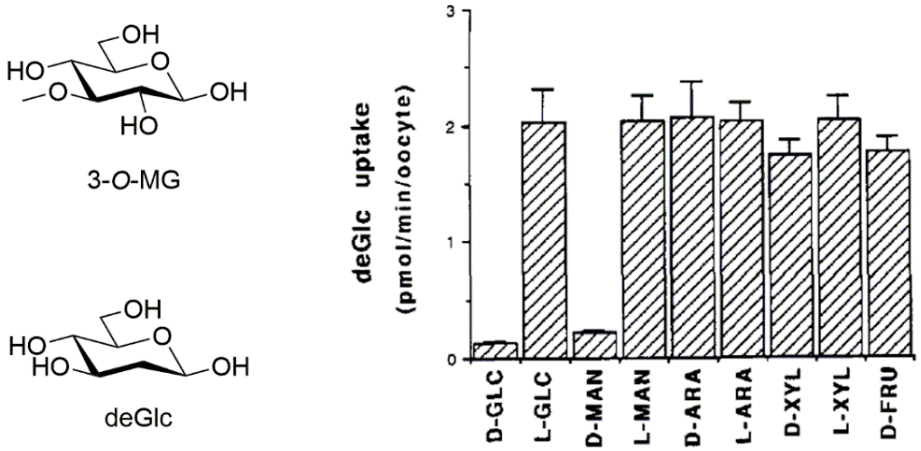


Figure 1.28. Non-metabolizable glucose analogues used in equilibrium-exchange transport assays to study GLUT transporters. A competition assay for GLUT-1 showed that only D-glucose and D-mannose are substrates.
Adapted, with permission, from Gould *et al.* 1991.[116]

Intrinsically fluorescent substrates have also been developed to assay glucose transporters, especially as a method to determine cell viability. Earlier cell viability assays used the activity of intracellular enzymes, such as esterases, as a marker to differentiate between living and dead cells.[118] However, since glucose uptake is the phenomenon most closely related to metabolic activity, it is perhaps a more promising indicator of cell viability. While previous work had shown that glucose modified at the C-6 position was internalized by mammalian cells,[119] no uptake was observed in *E. coli*.[120] This is not unsurprising, as the main transport system used by *E. coli* under conditions of glucose excess is the phosphoenolpyruvate:carbohydrate phosphotransferase system (PTS), which internalizes glucose and concomitantly phosphorylates the C-6 position, requiring this position to be unsubstituted.[121] Accordingly, a C-2 substituted fluorescent glucose analogue, 2-(N-(7-nitrobenz-2-oxa-1,3-diazol-4-yl)amino)-2-deoxyglucose (NBD-Glc), was synthesized (Figure 1.29).[70] Initial rates of transport were measured by incubating *E. coli* cells with varying concentrations of NBD-Glc, centrifuging and washing the cells, and quantifying the fluorescence using excitation and emission wavelengths of 475 and 550 nm, respectively. From a Lineweaver-Burk plot, the K_M for NBD-Glc transport mediated by the glucose transporter was determined (Figure 1.29). The presence of D-glucose during the assay inhibited the uptake of NBD-Glc, while neither L-glucose nor sucrose had any impact on NBD-Glc internalization. Fluorescence microscopy was also performed to determine whether NBD-Glc could be used as a marker for cell viability. Native *E. coli* cells that had been incubated with NBD-Glc exhibited fluorescence in the green region of the visible spectrum, while those which had been killed by 10 min of pre-treatment with ethanol showed none. Further, a 1:1 mixture of native and ethanol-treated *E. coli* were incubated with NBD-Glc and propidium iodide, a red fluorescent indicator for dead cells. Upon examination, approximately half of the cells were coloured green while the remainder exhibited red fluorescence. Due in large part to its synthetic simplicity and ease-of-use, NBD-Glc is still widely used to determine the viability of bacterial cells, and has also been used to measure glucose transport in mammalian cells.[122]

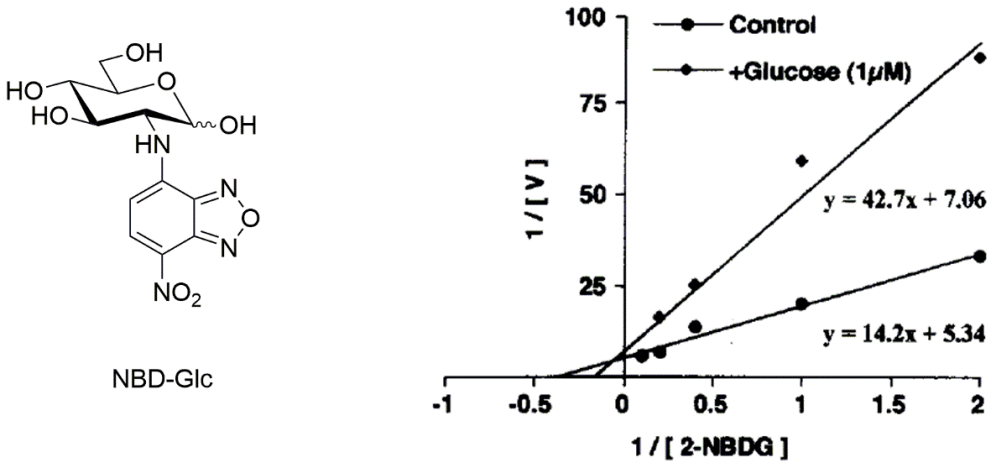


Figure 1.29. Fluorescent glucose analogue NBD-Glc (2-NBDG), and a Lineweaver-Burk plot showing its uptake by *E. coli* cells and inhibition by glucose.
Adapted, with permission, from Yoshioka *et al.* 1996.[70]

Despite the fact that most transporters do not catalyze the chemical conversion of a substrate, real-time assays of these transporters have been developed by monitoring non-rate-determining reactions that are subsequent to transport. Such an assay exists for the *Lactobacillus pentosus* xyloside transporter XylP, a member of the galactoside-pentoside-hexuronide (GPH) transporter family.[123] Xylosides are the building blocks of the structural hemicelluloses of plants, and it is likely that bacteria harbour xyloside transporters such as XylP in order to use these saccharides as a source of energy. The assay indirectly measures the XylP-mediated transport of the disaccharide isoprimeverose by detecting rapid subsequent oxidation by soluble quinoprotein glucose dehydrogenase (sGDH) – a process that converts the artificial electron acceptor 2,6-dichlorophenol-indophenol (Cl_2Ind) from a blue to colourless state (Figure 1.30).[124] Histidine-tagged XylP was expressed in *Lactococcus lactis*, purified by affinity chromatography using a nickel column, and reconstituted into proteoliposomes. Next, Ca^{2+}-dependent sGDH was expressed from *E. coli*, reconstituted with its redox cofactor pyrroloquinoline quinone (PQQ),[125] and enclosed within the prepared proteoliposomes. While preloading with sGDH was essential, Cl_2Ind is highly membrane permeable and so its preloading was not necessary, as it easily permeates the proteoliposomes during the assay. To minimize background signal, excess sGDH associated with the exterior of the proteoliposomes was deactivated using EDTA to complex extravesicular Ca^{2+}. This

process did not affect the sGDH enclosed within the proteoliposomes, since EDTA is membrane impermeable. After washing and resuspension of the sGDH-containing proteoliposomes in buffer containing EDTA, the suspensions were used to determine the kinetics of isoprimeverose transport. To do so, the proteoliposomes were added to solutions containing Cl_2Ind and varying concentrations of isoprimeverose, and the decrease in absorbance at 600 nm resulting from reduction of Cl_2Ind was monitored over time to obtain initial rates of isoprimeverose transport. Any residual activity from external sGDH was taken into account by carrying out a control experiment with the disaccharide maltose instead of isoprimeverose, which is a substrate for sGDH but not for XylP. Differences in the K_M of isoprimeverose and maltose for sGDH were also taken into account by appropriately adjusting the concentrations of the substrates used in the assay. The corrected initial rates were plotted against the isoprimeverose concentrations, and these data were fitted to the Michaelis-Menten equation to obtain K_M and V_{max} values for isoprimeverose transport by XylP. To ensure that the rate of isoprimeverose oxidation by sGDH was not rate-limiting, the kinetics were carried out using proteoliposomes containing varying sGDH concentrations, and a saturating sGDH concentration was used for the K_M and V_{max} determinations. Interestingly, preloading the proteoliposomes with the non-oxidizable methyl-isoprimeverose resulted in a two-fold stimulation of the oxidation rate, indicating that exchange transport occurs faster than does unidirectional uptake. Further, while other members of the GPH transporter family transport monosaccharides, XylP was unable to transport xylose, and instead seems to require a glycosidic linkage at the anomeric position. While this assay provides a real-time approach to the study of XylP, the labour required to prepare the proteoliposomes for the assay is substantial. Perhaps more convenient would be the use of a quenching approach, in which a membrane-impermeable extracellular compound quenches the fluorescence of a substrate until it is transported into a cell, as has been done to observe the transport of lipids into adipocytes.[126]

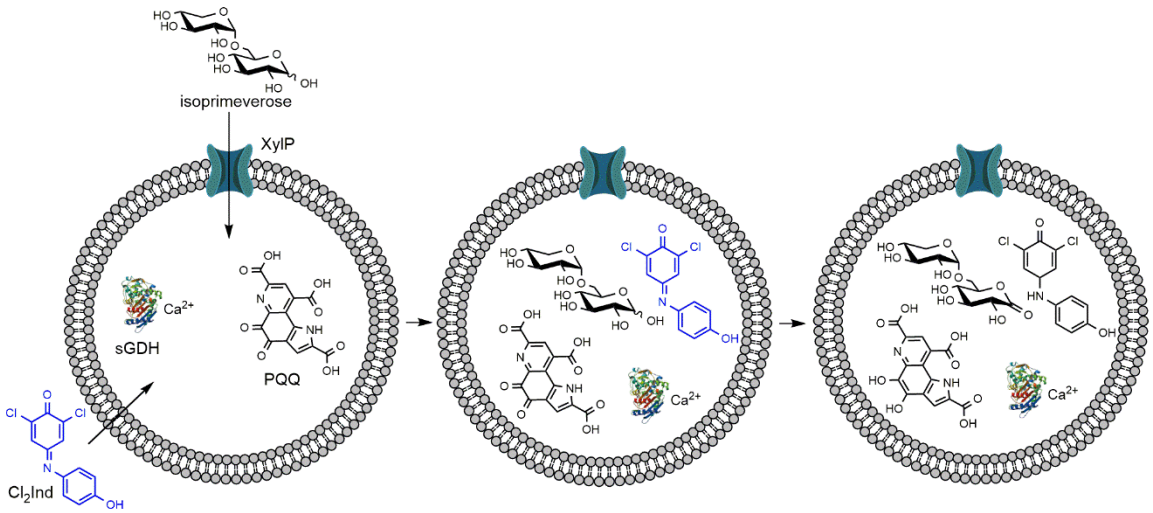


Figure 1.30. A representation of the XylP transport assay.
External sGDH is deactivated by EDTA, which cannot pass the proteoliposome membrane. Incubation with isoprimeverose and Cl_2Ind results in internalization by XylP and passive diffusion, respectively. Upon entry into the proteoliposome isoprimeverose is rapidly oxidized by sGDH, with concomitant reduction of the cofactor PQQ to $PQQH_2$. In this process, the electron acceptor Cl_2Ind is converted from a blue to colourless state.

1.4. Aims of Thesis

This thesis details the development of tools and methods to study glycan processing proteins in the relevant environment of a living system. In particular, studies will focus on the Gram-negative inner membrane permease AmpG and the human β-hexosaminidase *O*-GlcNAcase (hOGA). While AmpG plays an important role in bacterial cell wall recycling and in β-lactam antibiotic resistance, hOGA is involved in a number of processes from neurodegeneration and cancer to stabilization of nascent proteins. As alluded to, development of live-cell and *in vivo* techniques to study protein function are not trivial, and so this thesis reports extensions of previous works that describe the design of quenched and inherently fluorescent substrates and their use in assays of glycoside hydrolases and sugar transporters. The techniques described here will not only help improve the understanding of AmpG and hOGA, but may also be expanded upon to enable the study of other glycoside hydrolases and transport proteins.

1.4.1. Development of a Transport Assay for AmpG Permeases

AmpG is a Gram-negative inner membrane permease that translocates peptidoglycan (PG) fragments (GlcNAc-1,6-anhydroMurNAc-peptides) from the periplasm to the cytoplasm for recycling. Despite its involvement in maintaining the structural integrity of the bacterial cell wall and in affording antibiotic resistance, no convenient nor quantitative assays of AmpG transport exist. This thesis details the development of such an AmpG transport assay using live spheroplasts as substrate capture vessels for a fluorescently labelled analogue of the natural AmpG substrate. This project has been a collaborative effort involving chemical synthesis of the fluorescent AmpG probe by Anuj Yadav; assay development, transport quantitation, kinetic measurements, and fluorescence microscopy by me; and genetic manipulations and β-lactamase assays by Judith Winogrodzki. Together we have shown that uptake of the probe is AmpG-mediated, and that the transport assay can be used to examine the kinetics of AmpG-mediated transport and to identify AmpG inhibitors. Further, we applied this assay to study AmpG homologues from the clinically relevant pathogen *Pseudomonas aeruginosa* and found that only one, *Pa*-AmpG, is responsible for the transport of PG fragments that induce antibiotic resistance.

1.4.2. Design of Quenched Substrates for the β-hexosaminidase OGA

Attachment of *N*-acetylglucosamine to the serine and threonine residues of nucleocytoplasmic proteins is known as the *O*-GlcNAc modification. Despite the fact that hundreds of proteins are modified with *O*-GlcNAc, only two proteins are involved in its installation and hydrolysis – *O*-GlcNAc transferase (OGT) and *O*-GlcNAcase (OGA), respectively. This modification is involved in a host of processes from neurodegeneration to the cell cycle, but to date no methods exist to study the responsible proteins in real-time within the relevant context of a living system. This thesis details the development of quenched fluorescent substrates that will enable live-cell or *in vivo* study of human OGA (hOGA). This work has been a collaborative effort between myself, Sami Cecioni, and Nevena Cekic. However only my work, which has focused on installation of a quencher at the C-6 position, will be discussed. I have prepared three hOGA substrates with thioamides installed at the C-6 position to act as fluorescence quenchers of the intact

substrate. I have shown that these substrates are processed by hOGA, and that quenching of up to 90% is observed in comparison to the widely used commercial substrate 4MUGlcNAc. Further I have demonstrated that the quenching efficiency can be tuned through the installation of a linker between the thioamide quencher and the sugar.

Chapter 2.

Design of a Fluorescent Substrate and Transport Assay to Probe AmpG Membrane Permeases and Their Roles in Antibiotic Resistance

2.1. Contributions

This work has been a multidisciplinary undertaking requiring the collaborative efforts of multiple researchers. The substrate GlcNAc-anhMurNAc-AF350 was synthesized by Anuj Yadav, and the genetic manipulations, immunoblots, and β-lactamase assays were carried out by Judith Winogrodzki. My part in this work was focused around development and implementation of the transport assay, although I also carried out corroborative immunoblots when expressing *Pseudomonas aeruginosa* AmpG and AmpP under spheroplasting conditions, as well as single-step chemical reactions and purification to obtain compounds for inhibition experiments. My major contributions include: developing and optimizing the spheroplasting conditions for BW25113 and MG1655 strains; developing the AmpG transport assay and obtaining quantitative measurements of transport for *Escherichia coli* AmpG, and *P. aeruginosa* AmpG and AmpP; fluorescence microscopy; adaptation of the transport assay to obtain Michaelis-Menten and inhibition data; and preparation of the manuscript for publication.

2.2. Abstract

Inducible AmpC β-lactamases deactivate a broad-spectrum of β-lactam antibiotics and are a clinically formidable mechanism of antibiotic resistance within Gram-negative bacteria. Increased peptidoglycan (PG) fragmentation caused by β-lactams leads to accumulation of GlcNAc-1,6-anhydroMurNAc-peptides, that are then

transported by AmpG into the cytoplasm where they induce AmpC β-lactamase expression. Loss of AmpG restores susceptibility toward β-lactams, highlighting AmpG as a potential target for resistance-attenuating therapeutics. We describe the synthesis of a GlcNAc-1,6-anhydroMurNAc-fluorophore conjugate, and its use in convenient quantitative assays to probe AmpG transport activity. Using live *Escherichia coli* spheroplasts, we show that transport of the probe is AmpG-dependent and follows Michaelis-Menten kinetics. We further show that this assay can be used to quantitatively assess inhibition of AmpG. Finally, we illustrate the generality and utility of this probe by assessing the function of two AmpG homologues from the clinically relevant pathogen *Pseudomonas aeruginosa (Pa)*. We show that *Pa*-AmpG, but not *Pa*-AmpP, acts as a transporter of GlcNAc-1,6-anhydroMurNAc PG fragments. We support this finding by monitoring AmpC induction in *ampG*$^{-}$ *E. coli* complemented with either *Pa*-AmpG or *Pa*-AmpP. We anticipate that this AmpG transport assay will enable improved understanding of the PG recycling pathway and sugar transporters from the major facilitator superfamily, as well as identification of AmpG antagonists that could be useful in combating bacteria with AmpC-mediated antibiotic resistance.

2.3. Introduction

Antibiotic drug resistance is a growing problem of major societal concern.[127] Resistance to β-lactams, which constitute more than 50% of the antibiotics in clinical use, is particularly disconcerting.[16] β-lactams block biosynthesis of peptidoglycan (PG), a critical structural component of the bacterial cell wall that confers cellular rigidity and prevents osmotic lysis.[128] Due in large part to overreliance on and misuse of β-lactams, clinically relevant bacteria are acquiring a variety of β-lactamases – enzymes that hydrolytically deactivate β-lactams.[129] The chromosomally-encoded group 1, class C, β-lactamase AmpC,[130,131] which confers resistance against penicillins and cephalosporins,[132] is particularly problematic and has been identified in clinically relevant strains of *Pseudomonas aeruginosa* and *Enterobacteriaceae*.[131,133] Indeed, AmpC-bearing *P. aeruginosa* causes life-threatening nosocomial infections in the immunocompromised, as well as chronic respiratory infections in patients with cystic fibrosis.[134,135] Although β-lactamase inhibitors exist, many only poorly inhibit AmpC.[131] To

block AmpC-mediated resistance to β-lactams, efforts have been made to develop AmpC-targeting inhibitors.[136,137] Unfortunately, β-lactamases such as AmpC can develop mutational resistance toward these inhibitors,[132,138,139] and so new approaches to combatting β-lactam antibiotic resistance are needed. One promising strategy for targeting AmpC involves preventing its induction – a process that is intricately linked to the Gram-negative PG recycling pathway and is stimulated by β-lactams in most AmpC-producing bacteria (Figure 2.1).

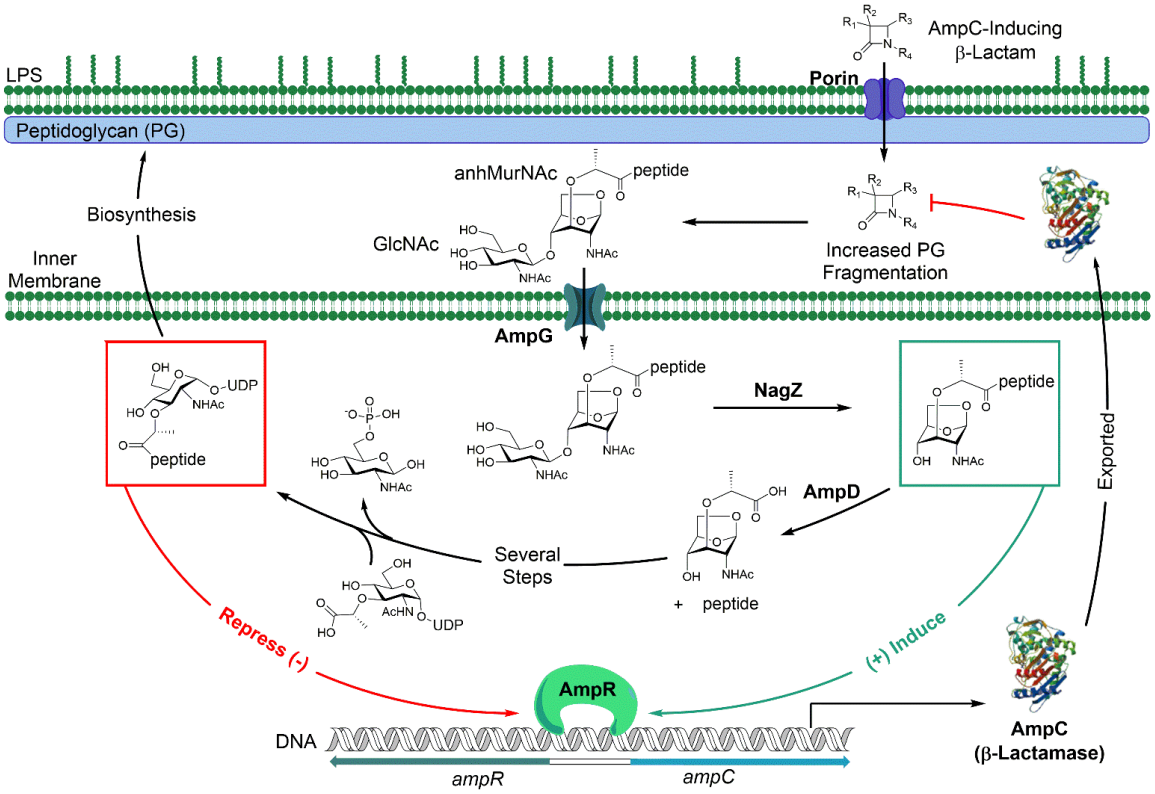


Figure 2.1. Induction of AmpC β-lactamase is stimulated by the impact of β-lactam antibiotics on the Gram-negative PG recycling pathway. In the presence of β-lactam antibiotics PG fragmentation is increased, leading to a buildup of AmpC-inducer molecules (boxed in green). These bind to transcriptional regulator AmpR and convert it from a repressed state, bound to AmpC-repressor molecules (boxed in red), into an activator of *ampC* transcription.

The PG layer is a mesh-like heteropolymer of glycan strands composed of alternating β-1,4-linked *N*-acetylglucosamine (GlcNAc) and *N*-acetylmuramic acid (MurNAc) residues that are cross-linked by short-stem peptides.[140] PG recycling enables cell growth and division while maintaining the structural integrity of the cell wall.[141]

Beginning this process, auto-lysins[142] act on the PG layer to generate a pool of GlcNAc-1,6-anhydroMurNAc disaccharides bearing stem peptides of three to five amino acids that are linked to the lactyl group of MurNAc (GlcNAc-anhMurNAc-peptides, Figure 2.1).[141] These species are then transported from the periplasm into the cytoplasm by the inner membrane permease AmpG,[143] where the β-glucosaminidase NagZ removes GlcNAc to produce anhMurNAc-peptides.[144,145] The anhydrosugar-peptides are metabolized further to yield UDP-MurNAc-pentapeptide, a precursor of PG biosynthesis.[141] Binding of PG recycling intermediates to the LysR-type transcriptional regulator AmpR is thought to control AmpC expression, with the anhMurNAc-peptides[146,147] acting as AmpC inducers while UDP-MurNAc-pentapeptide acts as a repressor.[146] In the absence of β-lactams, UDP-MurNAc-pentapeptide binds AmpR, repressing *ampC* transcription. However, increased PG fragmentation caused by β-lactams results in accumulation of anhMurNAc-peptides in the cytoplasm, one of which binds AmpR and converts it into an activator of *ampC* transcription.[146,147] The resulting AmpC protein deactivates β-lactams to restore PG homeostasis, which in turn brings AmpC expression back to basal levels.[146]

Deletion of the gene encoding NagZ attenuates AmpC induction in *P. aeruginosa*.[148] Moreover, inhibitors targeting NagZ have indicated that a small-molecule therapeutic approach to combatting AmpC-mediated resistance is tractable.[149,150] Compared to NagZ, the highly conserved inner membrane permease AmpG – a member of the Major Facilitator Superfamily of secondary transporters[51] – has received little attention even through it plays a critical role in AmpC induction.[151,152] Indeed, genetic inactivation of AmpG in *P. aeruginosa* reverses AmpC-mediated resistance, even in mutated strains that exhibit constitutive hyperproduction of AmpC.[153] Given its ubiquitous presence in Gram-negative bacteria,[151,154-156] its central role in AmpC induction,[151-154] and its unusual sugar-conjugate transport activity,[143] an improved understanding of AmpG would provide new insights into PG metabolism and could lead to the discovery of therapeutics that attenuate β-lactam antibiotic resistance. However, a convenient method to examine the kinetics of AmpG-mediated transport does not exist, making the characterization of AmpG transporters and the identification of AmpG inhibitors problematic.

One difficulty in studying AmpG transporters is that recombinant transmembrane proteins can be difficult to express due to, among other problems, poor folding efficiency, toxicity at elevated expression levels, and a propensity for mislocalization in cells.[157] Secondly, designing assays for transport proteins, which simply translocate substrates, is more challenging than for enzymes, which form and break chemical bonds. Assays of transport proteins require substrate capture vessels that allow transported substrate to be retained for quantification while extravesicular substrate can be washed away.[113,158] Such assays generally use isotopically labeled substrates to monitor transport. Potentially more convenient than using radiolabelled substrates is to use fluorescently labeled substrates. Problematic, however, is that the structural perturbations associated with fluorophore installation may not be tolerated by the target transporter. Indeed, the difficulty of studying transport proteins using fluorescently labeled substrates is evidenced by the fact that only one fluorescence-based assay of a sugar transporter has been described.[70] However the probe, 2-(N-(7-nitrobenz-2-oxa-1,3-diazol-4-yl)amino)-2-deoxyglucose (NBD-Glc), has proven invaluable for monitoring glucose transport in bacteria[70] and in eukaryotic cells.[122]

When using labelled substrates, post-uptake metabolism of the substrate can lead to errors in measuring substrate transport, especially when using whole cells as substrate capture vesicles. However, this can be overcome by using non-metabolizable analogues of the substrate of interest.[116,159] Errors can also arise after the uptake phase of the assay when excess substrate is being washed away. Typically, these erors arise from internalized substrate being lost from the vesicles during the protracted and somewhat harsh conditions required for washing. Employing transport inhibitors when stopping the assay, using cold wash solutions to decrease the activity of exporters, and minimizing the time required to wash the vesicles, can all reduce the amount of substrate lost during the washing procedure.[113,158] In order to avoid the complications associated with using membrane capture vesicles, and to allow the use of unlabelled substrates, some assays instead use purified membrane transport proteins. In such assays, the extent of molecular binding between unlabelled substrates and the purified protein can be measured using surface plasmon resonance or scintillation proximity assays.[158] However these assays introduce a new difficulty, as membrane proteins are

notoriously difficult to purify due to their lipophilic character and the fact that their stability is heavily reliant on being embedded in the phospholipid bilayer.[160]

Due to the challenges associated with assaying membrane transport proteins, only qualitative information of AmpG substrate preference has been established using a crude assay involving radioactively labelled PG fragments which are complicated to prepare and handle.[143] Here we describe a convenient and quantitative assay of AmpG, enabled by the synthesis of a fluorescent analogue of the natural AmpG substrate and the use of *E. coli* spheroplasts as substrate capture vessels (Figure 2.2). We show that the assay can be used to measure the kinetics of AmpG-mediated transport and to identify compounds that inhibit this process. Moreover, we demonstrate the utility of the assay by examining the transport capabilities of two *P. aeruginosa* (*Pa*) proteins, *Pa*-AmpG and *Pa*-AmpP; and establish that *Pa*-AmpG, but not *Pa*-AmpP, transports PG fragments containing GlcNAc-anhMurNAc. We corroborate this finding by AmpC induction assays in *E. coli*.

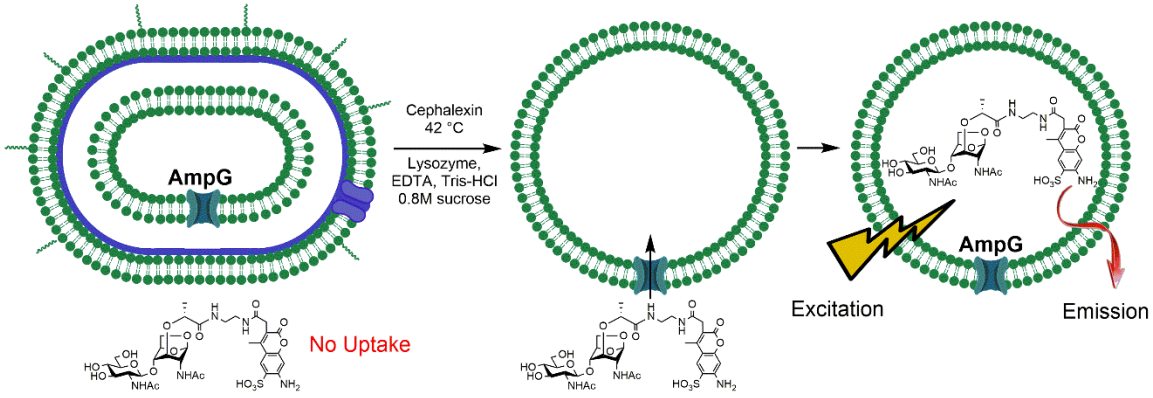


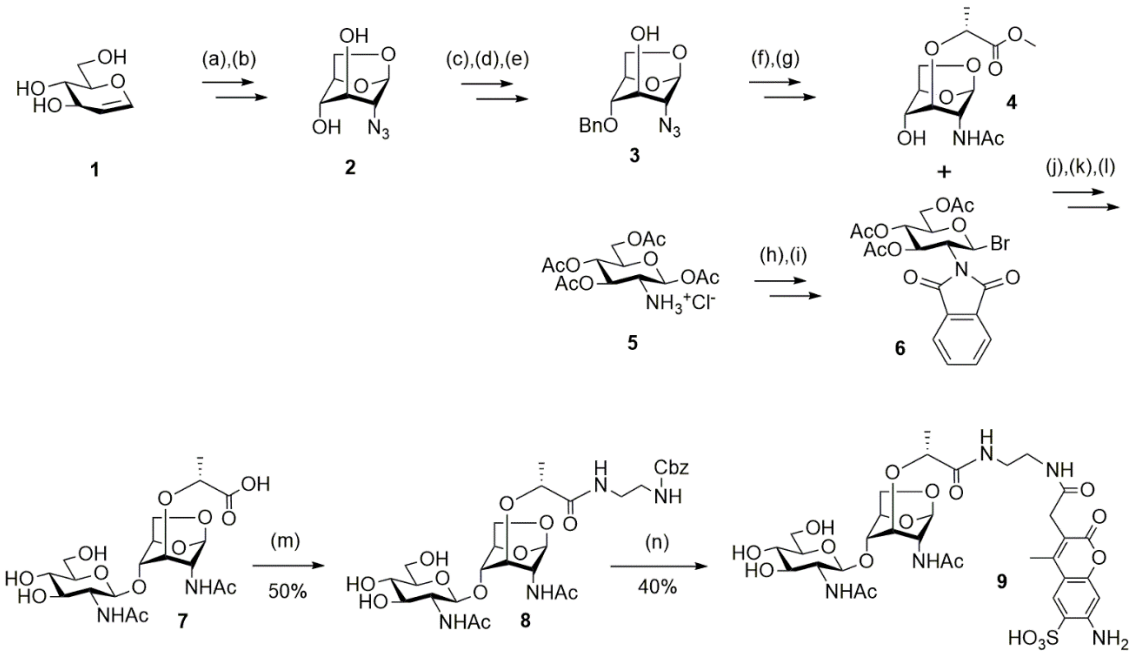
Figure 2.2. General schematic of the AmpG transport assay.
Generation of bacterial spheroplasts exposes AmpG to the extracellular environment, allowing the probe to be internalized.

2.4. Results and Discussion

2.4.1. Design and Synthesis of AmpG Probe 9.

Inspired by previous work on glucose transporters,[70] we envisioned that a fluorescently labeled AmpG substrate could be used to examine the kinetics of AmpG-

mediated transport. To design such a substrate, we considered qualitative studies showing that radiolabeled PG fragments, obtained through partial purification of bacterial cell wall, do not require the pendant peptide in order to be transported by AmpG.[143] This is not unexpected, as metabolism of the PG layer naturally produces structural variation in the peptide portion of endogenous AmpG substrates (GlcNAc-anhMurNAc-peptides). In contrast it was found that modification of the disaccharide structure, whether through hydration of the anhMurNAc sugar or cleavage of the interglycosidic linkage, prevented AmpG-mediated transport.[143] We therefore reasoned that replacing the stem peptide with a chromophore, while retaining the core disaccharide structure, would yield a fluorescent analogue of the GlcNAc-anhMurNAc-peptides without markedly compromising recognition and transport by AmpG permeases. To limit the steric demands of this change, we opted to use the low molecular weight fluorophore Alexa Fluor 350 (AF350). Moreover, AF350 exhibits high photostability, quantum yield, and water solubility, as well as pH independent excitation and emission spectra, qualities that make it suitable for cell-based assays. With these considerations in mind, we carried out the synthesis of probe **9** (GlcNAc-anhMurNAc-AF350, Scheme 2.1).



Scheme 2.1. Synthesis of GlcNAc-anhMurNAc-AF350 (9).
(a) $(Bu_3Sn)_2O$, 3 Å MS, MeCN, reflux, 3 hr; then I_2, RT, 1.5 hr, 84%; (b) NaN_3, DMF:H_2O (9:1), 120 °C, 3 hr, 81%; (c) TBSCl, imidazole, DMF, RT, 82%; (d) BnBr, BaO, $Ba(OH)_2{\cdot}8H2O$, DMF, RT, 2.5 hr, 90%; (e) TBAF, THF, 90%; (f) (S)-MeCH(OTf)CO_2Me, NaH, DCM, 70%; (g) 1) 10% Pd/C, H_2; 2) Ac_2O, Py; 3) 10% Pd/C, H_2, pressure, 80%; (h) 1) Phthalic anhydride, TEA; 2) Ac_2O, Py, 68%; (i) 33% HBr in CH_3COOH; 85% (j) AgOTf, 4 Å MS, DCM, -40 to 0 °C, 85%; (k) 1) N_2H_4, EtOH, reflux; 2) Ac_2O, Py; 3) K_2CO_3, MeOH; 75%; (l) LiOH·H_2O, THF:Dioxane:H_2O (4:2:1), 99%; (m) $NH_2(CH_2)_2$NHCbz, HBTU, DIPEA, DMF, 50%; (n) 1) 10% Pd/C, H_2, MeOH; 2) AF350, DIPEA, DMF, 40%.

Using an established route with minor modifications, we synthesized the disaccharide GlcNAc-anhMurNAc through a convergent approach beginning with commercially available D-glucal **1** and 1,3,4,6-*tetra*-*O*-acetyl-2-amino-2-deoxy-β-D-glucopyranose hydrochloride **5**. D-glucal 1 was first converted to the 1,6-anhydrosugar, and an azide was installed at the 2-position to obtain azidosugar **2**.[161] Protecting group manipulation allowed selective benzylation in three steps to generate azidosugar **3**.[162] Subsequent installation of a methyl lactate substituent at the 3-position via direct displacement to effect stereochemical inversion, followed by debenzylation, acetylation of the resultant alcohol, and conversion of the azide to an acetamide, furnished glycosyl acceptor **4**.[163] The glycosyl bromide donor **6** was obtained in two steps by first protecting the amine of 1,3,4,6-*tetra*-*O*-acetyl-2-amino-2-deoxy-β-D-glucopyranose hydrochloride **5** as a phthalimide and then treating the product with 33% hydrobromic acid in acetic

acid.[164] Coupling of glycosyl acceptor **4** and glycosyl donor **6** using Koenigs-Knorr conditions provided the protected disaccharide which, upon removal of the phthalimide and subsequent acetylation of the resultant amine, followed by hydrolysis of the methyl ester, furnished the disaccharide GlcNAc-anhMurNAc **7**.[163] To facilitate appending a suitable fluorophore and to ensure the fluorophore did not abut the disaccharide, we installed a spacer moiety on disaccharide **7** using benzyl 2-aminoethylcarbamate to yield the Cbz-protected aminoethyl amide **8**. Finally, we removed the carboxybenzyl group using catalytic Pd/C under hydrogen and directly coupled the resultant crude amine with the fluorophore (AF350) to yield the desired GlcNAc-anhMurNAc-AF350 conjugate (probe **9**).

2.4.2. Transport Assay Development.

With probe **9** in hand, we developed a convenient assay to quantify AmpG-mediated transport. Transport assays using whole cells have been developed for different classes of compounds, including but not limited to sugars,[70] nucleotides,[165] and amino acids.[166] In these cases, the natural substrate is scavenged from the extracellular environment. For AmpG, however, the substrates are generated within the periplasmic space by action of endogenous auto-lysins,[142] making it unlikely that probe **9** can be transported directly from the media into intact bacterial cells. Indeed, we observed that *E. coli* cells were unable to take up probe **9** even after several hours (Figure 2.3). In previous work, a freeze-thaw protocol had been used to expose *E. coli* AmpG to biosynthetically-derived radiolabeled substrates. This approach enabled qualitative insights into AmpG specificity, but gave low signal-to-noise and high variability.[143] To develop a more convenient and quantitative method of studying AmpG, we turned to an approach used in electrophysiology studies of ion channels,[167] which is to generate spheroplasts – bacteria that have had their outer membrane and cell wall largely removed but which maintain an intact inner membrane. Notably, we found that the production of bacterial spheroplasts is scalable and the results obtained from independent batches are highly reproducible.

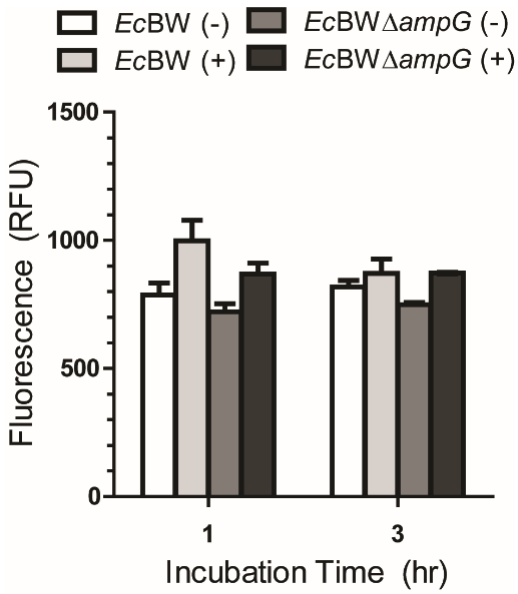


Figure 2.3. Neither whole cell *Ec*BW nor *Ec*BWΔ*ampG* are capable of internalizing probe 9, regardless of incubation time.
Fluorescence observed for native *Ec*BW and *Ec*BWΔ*ampG* bacteria after incubation in the presence (+) or absence (-) of probe **9**. Error bars represent the standard error of the mean.

We prepared spheroplasts from wild-type *E. coli* BW25113 (*Ec*BW),[168] and an isogenic strain lacking AmpG (*ampG*$^{-}$, *Ec*BWΔ*ampG*). This preparation involved elongation of the cells by treatment with cephalexin, followed by treatment with EDTA and lysozyme to digest the cell wall and cause the elongated cells to balloon outward. Elongation of the *E. coli* cells enables formation of larger spheroplasts, making it easier to monitor their formation by light microscopy. However we noticed that obtaining larger spheroplasts came at the expense of the overall spheroplast population, as the elongated *E. coli* had a higher propensity to rupture on addition of the spheroplasting reagents. We found that the published cephalexin, lysozyme, and EDTA concentrations caused our strain to lyse during the elongation or spheroplasting steps,[167] and so the concentrations of these reagents were adjusted. To verify that the spheroplasts generated from the *Ec*BW and *Ec*BWΔ*ampG* strains were comparably viable we used NBD-Glc.[70] This experiment ensured that an inability of *Ec*BWΔ*ampG* spheroplasts to transport probe **9** could be unambiguously assigned to the lack of AmpG and not to a failure of spheroplast viability. After incubation, excess NBD-Glc was removed by repeated cycles of dilution, gentle centrifugation, and careful aspiration of the

supernatant. Three cycles of washing was sufficient to eliminate greater than 99.5% of extravesicular NBD-Glc, putting the supernatant fluorescence from the last wash on par with the background fluorescence of the diluting solution (Figure 2.4).

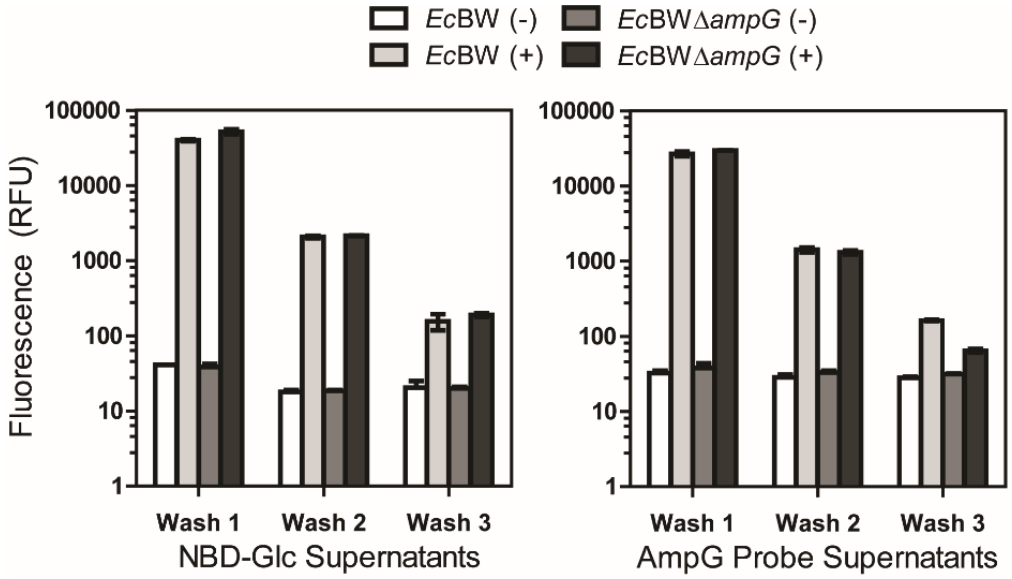


Figure 2.4. Three washes were sufficient to remove the vast majority of excess NBD-Glc or probe 9 from EcBW or EcBWΔampG spheroplasts. Supernatant fluorescence after one to three washes of spheroplasts which were incubated with (+) or without (-) NBD-Glc or probe **9**. Error bars represent the standard error of the mean. Data plotted on a logarithmic y-axis.

Quantitation of spheroplast fluorescence indicated a signal ten-fold higher than background for both the *Ec*BW and *Ec*BWΔ*ampG* strains (Figure 2.5). Using fluorescence microscopy, we observed that for both strains NBD-Glc was localized within the spheroplasts, excluding the possibility that fluorescence originated from debris- or membrane-associated NBD-Glc (Figure 2.5). We noted that some of the elongated *E. coli*, which make up a portion the post-spheroplasting population, also exhibited significant fluorescence. This is not unexpected since NBD-Glc can pass through the cell wall and outer membrane.[70] A fraction of the spheroplasts and the elongated *E. coli* exhibited no fluorescence. Given that that NBD-Glc is used as a measure of cell viability, these cells likely ruptured during the spheroplasting process. Cumulatively, these fluorescence data demonstrate the comparable viability of spheroplasts prepared from the *Ec*BW and *Ec*BWΔ*ampG* strains, and support their use as substrate capture vessels to study transport reactions.

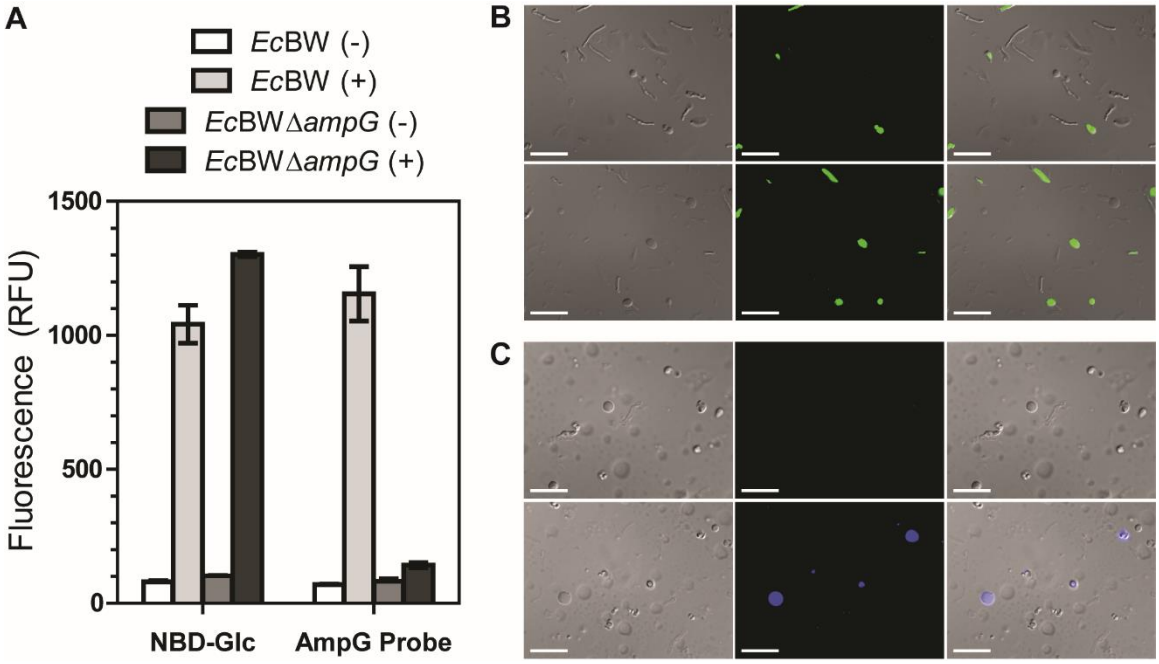


Figure 2.5. Wild-type (*Ec*BW) spheroplasts, but not isogenic *ampG*⁻ (*Ec*BWΔ*ampG*) spheroplasts, are able to internalize probe 9.
A: Fluorescence of *Ec*BW and *Ec*BWΔ*ampG* spheroplasts after incubation in the presence (+) or absence (-) of either NBD-Glc or probe **9**. The spheroplasts uptake NBD-Glc, indicating they are comparably viable. Error bars reflect the standard error of the mean. **B**: DIC (left), fluorescence (middle), and overlay (right) images of *Ec*BWΔ*ampG* (top) and *Ec*BW (bottom) spheroplasts after incubation with NBD-Glc; or probe **9** (**C**). Scale bars indicate 20 μm.

We next examined AmpG-mediated transport of probe **9** using *Ec*BW and *Ec*BWΔ*ampG* spheroplasts. As with NBD-Glc, we found that three washes was sufficient to remove the vast majority of extravesicular probe **9** (Figure 2.4). We saw that only *Ec*BW spheroplasts internalized probe **9**; no significant internalization was seen for *Ec*BWΔ*ampG* spheroplasts (Figure 2.5). Parallel examination of NBD-Glc uptake confirmed comparable viability of the spheroplast preparations. Examination of the post-assay spheroplasts by fluorescence microscopy revealed probe **9** localized within the *Ec*BW spheroplasts, whereas no fluorescence was observed within the *Ec*BWΔ*ampG* spheroplasts (Figure 2.5). Notably, image acquisition was facile for probe **9** in comparison to NBD-Glc, as the high photostability of the AF350 fluorophore prevented photobleaching during imaging, underscoring the utility of probe **9** in quantitative fluorescence microscopy. Interestingly, some of the elongated *Ec*BW cells were also capable of internalizing probe **9**, suggesting that during spheroplasting their cell walls

were made permeable while still retaining sufficient structural integrity to maintain the native rod-shape of the bacterium. Many of the spheroplasts and elongated *E. coli* exhibited no fluorescence whatsoever after incubation with probe **9**. It is likely that these cells suffered damage during elongation or spheroplasting, leading to a destruction of the proton gradient required for the function of AmpG.[143,169] Together, these quantitative fluorescence measurements and fluorescence images indicate that internalization of probe **9** is AmpG-mediated.

2.4.3. Kinetic Measurements of Transport.

We next assessed the kinetics governing the AmpG-mediated transport of probe **9**. Since Michaelis-Menten kinetics have been successfully used to describe membrane transport proteins,[124,170] we applied this model to AmpG-mediated transport. Given the high concentration of probe **9** outside the spheroplasts, we made the assumption that loss of probe **9** from the spheroplasts during the uptake phase of the assay was negligible. To minimize potential export of probe **9** during washing steps (though no mechanisms for active export of PG recycling fragments are known), manipulations were performed at 4 °C and samples were kept on ice.[113] Given the high polarity of probe **9**, and particularly the negatively charged sulfonic acid of the fluorophore moiety, we expected that passive diffusion of probe **9** out of the spheroplasts would be negligible. To establish the Michaelis constant (K_M) for AmpG-mediated transport of probe **9**, we also needed to ensure that the assay does not proceed beyond steady-state conditions, thus allowing initial rates of internalization to be determined. We therefore evaluated the time-dependent uptake of probe **9** using 50 μM of the probe and *Ec*BW spheroplasts. We found that the rate of internalization remained relatively constant over the first 90 minutes of the assay (Figure 2.6). We opted to use a 1-hour incubation for the kinetic experiments, as this time point is within the linear region for uptake of probe **9** while being long enough to minimize errors associated with small time variations in sample handling. As is required for accurate Michaelis-Menten kinetic analysis, we verified that the substrate concentration does not change appreciably during the course of the assay by comparing the fluorescence of the spheroplasts after one hour of uptake (~1200 RFU, Figure 2.5) to that of the first supernatant (~27000 RFU, Figure 2.4). In this way we estimated that only 4% of probe **9** was internalized. Moreover, we noted that the amount

of probe **9** within the supernatant does not change appreciably with time, and thus the amount internalized is negligible when compared to the total concentration in the assay (Figure 2.6). Using these conditions we measured initial rates of probe **9** internalization by *Ec*BW spheroplasts, and fit the data to the Michaelis-Menten equation to obtain an apparent K_M ($K_{M,app}$) of 340 ± 50 μM (Figure 2.7). We also determined the $K_{M,app}$ from a Lineweaver- Burk plot, and found a value matching this one within error (Figure 2.7).

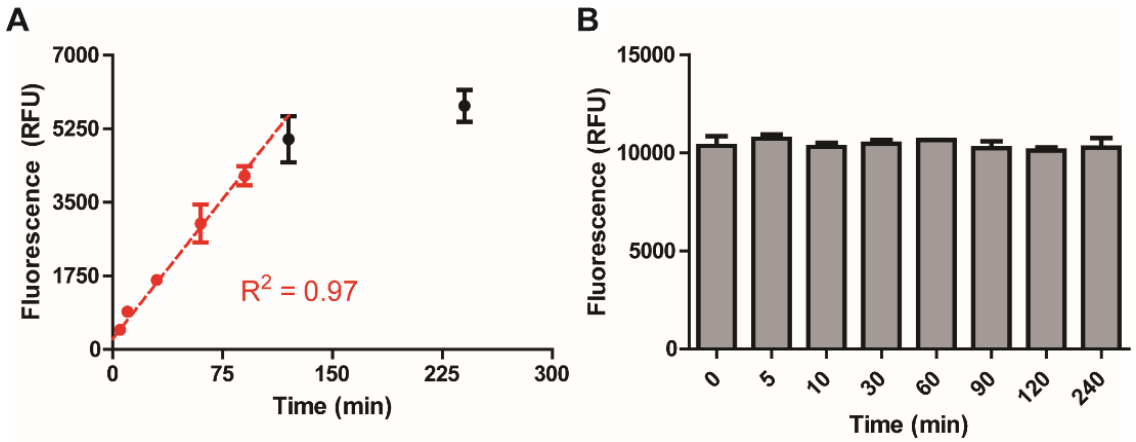


Figure 2.6. ***Ec*BW spheroplasts exhibit linearity in the initial uptake of probe 9, and the amount of extravesicular probe remains constant.**
A: Time course for the uptake of probe **9** into *Ec*BW spheroplasts. A line of best fit is indicated for the first five time points, up to and including 90 min, to indicate the approximately linear rate of probe internalization over this period. **B**: Supernatant fluorescence from time course remains constant, indicating that only a small amount of probe **9** is internalized. Error bars represent the standard error of the mean.

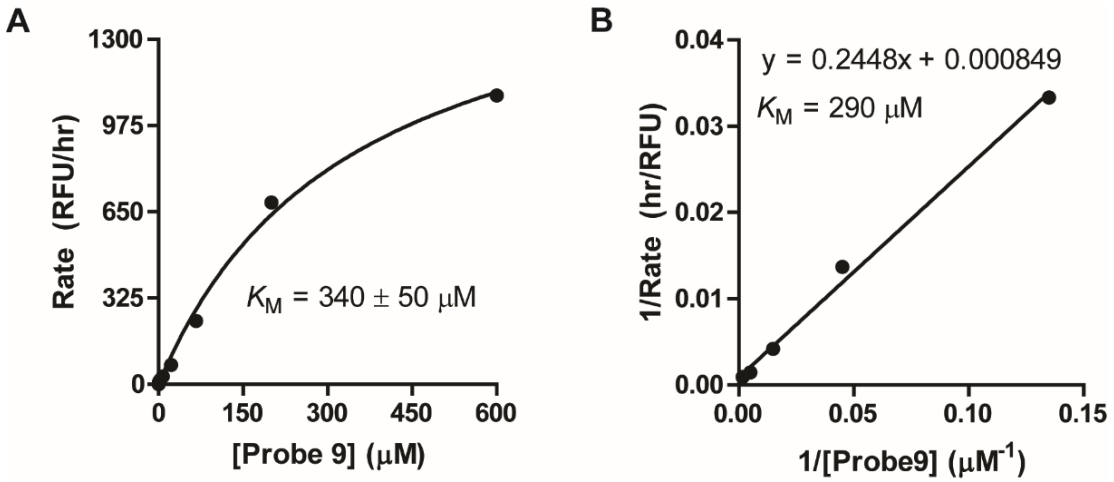


Figure 2.7. Initial rates of probe 9 uptake by *Ec*BW spheroplasts can be fit to the Michaelis-Menten equation.
A: Michaelis-Menten fit of the initial rates of uptake for varying probe **9** concentrations, from which an apparent K_M of 340 μM ± 50 μM was calculated. **B**: Lineweaver-Burk plot of the five highest concentrations from (**A**). K_M from the negative reciprocal of the x-intercept conforms to the value obtained from the Michaelis-Menten equation.

Having established that the uptake assay can quantify the kinetics of AmpG-mediated transport of probe **9**, we next tested whether it could be applied to assess the binding of competitive ligands. Previous work suggested that the disaccharide GlcNAc-anhMurNAc is transported by AmpG whereas the monosaccharide anhMurNAc is not.[143] Thus, we reasoned that GlcNAc-anhMurNAc should interfere with the transport of probe **9** by acting as a competitive substrate for AmpG. Conversely, the monosaccharides GlcNAc and anhMurNAc should have no impact on the ability of AmpG to transport probe **9**. To test this idea, we incubated *Ec*BW spheroplasts with 50 μM of probe **9** for one hour in the presence of varying concentrations of GlcNAc-anhMurNAc, GlcNAc, or anhMurNAc, and quantified the uptake of probe **9**. A dose-response curve showed a decreasing initial rate of probe **9** internalization with increasing GlcNAc-anhMurNAc concentration, indicating that GlcNAc-anhMurNAc does indeed compete with probe **9** for AmpG-mediated transport (Figure 2.8). As expected, the rate of probe **9** internalization was unaffected by the presence of either GlcNAc or anhMurNAc (Figure 2.8). From the dose-response curve, we obtained an IC_{50} for GlcNAc-anhMurNAc of 78 ± 1 μM. It has been shown that a competitive binding assay, in which a labelled and unlabelled substrate compete for a single binding site, can be used to calculate the dissociation

constant for the unlabelled substrate-enzyme complex ($K_{i,app}$).[171] Doing so uses the Cheng-Prusoff equation, which takes into account the concentration of unlabelled substrate causing 50% inhibition of the reaction between the enzyme and the labelled substrate (IC_{50}), the concentration of the labelled substrate (S), and the Michaelis constant for the labelled substrate-enzyme complex (K_M).[171,172] This method to establish a K_i value is convenient as it requires neither parallel determination of V_{max}, nor simultaneous variation of inhibitor and substrate concentrations. With the obtained IC_{50} value along with the concentration of probe **9** used in the assay (50 μM) and the $K_{M,app}$ for probe **9** with respect to AmpG, we calculated a $K_{i,app}$ of 70 ± 10 μM using the Cheng-Prusoff equation.[172] Since the binding of GlcNAc-anhMurNAc-peptides to AmpG has been proposed to depend primarily on the disaccharide moiety, one would expect that both probe **9** and GlcNAc-anhMurNAc should have a similar affinities for AmpG. Indeed, this appears to be the case as the apparent Michaelis constant for the probe **9**-AmpG complex ($K_{M,app}$ = 340 ± 50 μM) is similar to the apparent inhibition constant of the disaccharide-AmpG complex ($K_{i,app}$ = 70 ± 10 μM). The 5-fold difference may be attributable to the pendant fluorophore of probe **9**.

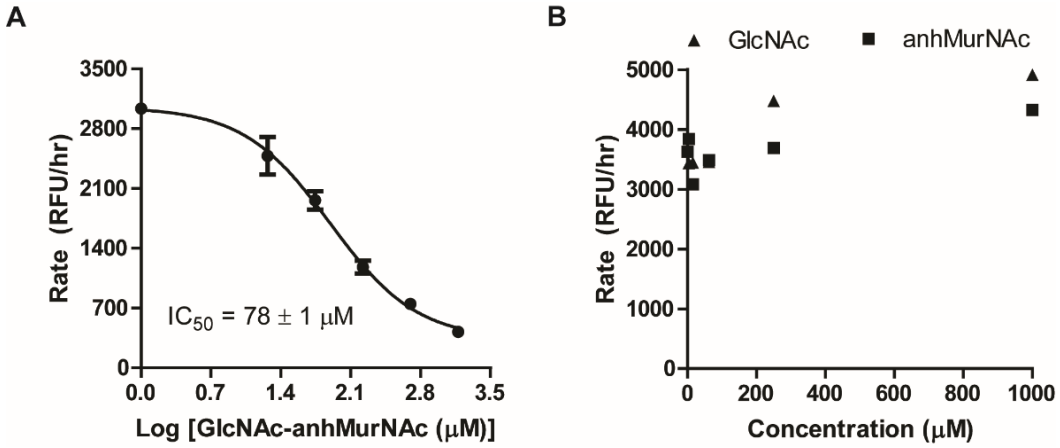


Figure 2.8. Uptake of probe 9 by *Ec*BW spheroplasts is inhibited by the competing substrate GlcNAc-anhMurNAc.
A: Dose-response curve indicating inhibition of AmpG-mediated uptake of probe **9** by the competing substrate GlcNAc-anhMurNAc, from which an IC_{50} of 78 ± 1 μM was determined. **B**: Monosaccharides GlcNAc and anhMurNAc have no impact on the AmpG-mediated transport of probe **9**.

2.4.4. Transport by *P. aeruginosa* AmpG Homologues.

Two *P. aeruginosa* (*Pa*) proteins, *Pa*4393 (*Pa*-AmpG), and *Pa*4218 (*Pa*-AmpP) have recently been shown to have 45% and 41% sequence identity, respectively, to *E. coli* AmpG.[151] Moreover, analysis of their topology shows 14 and 10 transmembrane helices for *Pa*-AmpG and *Pa*-AmpP respectively, lending support to their putative roles as membrane localized sugar transporters.[173] Despite their similarity to *E. coli* AmpG and their topological resemblance to transport proteins, *Pa*-AmpG and *Pa*-AmpP share only 19% sequence identity, suggesting they may not have redundant functions.[151] Previous studies report that *Pa*-AmpG contributes to maintaining intrinsic β-lactam resistance in *P. aeruginosa*. Specifically, *Pa*-AmpG was found to complement *ampG* knockouts in both *P. aeruginosa* and *E. coli*,[151] and to allow maximum expression of β-lactamase AmpC.[173] However, the role of *Pa*-AmpP is less clear, with contradictory reports in the literature regarding its involvement in AmpC expression.[151,153,173] To define the roles of *Pa*-AmpG and *Pa*-AmpP in mediating transport of GlcNAc-anhMurNAc peptidoglycan fragments and in AmpC-mediated resistance to β-lactam antibiotics, we performed AmpG transport assays using probe **9** and spheroplasts transformed with these *P. aeruginosa* transporters.

We cloned *Pa*-AmpG and *Pa*-AmpP fused to V5 and His_6 tags into the pBAD vector (pBAD-*Pa*-AmpG-V5His-Km and pBAD-*Pa*-AmpP-V5His-Km) to enable inducible expression in *E. coli*. For reasons unknown, *Ec*BW transformed with either of these two plasmids would not induce AmpC when co-transformed with a plasmid-borne *ampR-ampC* operon. However, *E. coli* strain MG1655 (*Ec*MG) accommodates this co-expression system and allowed us to evaluate the function of *Pa*-AmpG and *Pa*-AmpP using both probe **9** and AmpC induction assays within a single *E. coli* strain (*see below*). As a host for the uptake assay with probe **9** we used an *ampG*$^-$ strain (*Ec*MGΔ*ampG*). Both *Pa*-AmpG and *Pa*-AmpP expressed well using the pBAD system in *Ec*MGΔ*ampG*, and the amount of protein expressed was dependent on the inducer concentration (Figure 2.9).

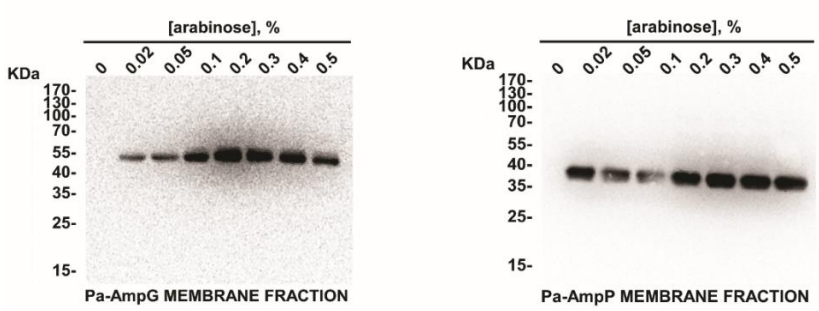


Figure 2.9. Expression of V5His tagged *Pa*-AmpG and *Pa*-AmpP in native *Ec*MGΔ*ampG* is dependent on the dose of inducer.

Spheroplasting of *Ec*MGΔ*ampG* was carried out similarly to the *Ec*BW strains (see Experimental Section). Interestingly, the presence of kanamycin as a selection antibiotic prevented cephalexin from having an effect on the cells, even when the cephalexin concentration was increased from 6 to 300 μg/mL. As a result of the impaired function of cephalexin in the presence of kanamycin, the *E. coli* reached a maximum elongation of 5 μm instead of the typical 50 μm, although cell division was not impeded as the OD600 steadily increased. This result was unexpected as kanamycin is an aminoglycoside antibiotic which inhibits protein synthesis, and so it is unknown how it imparts apparent resistance to cephalexin (a β-lactam) which was lethal at 50 μg/mL in the absence of kanamycin. Nevertheless, we were able to monitor spheroplast formation from these 5 μm cells, which were more stable to the spheroplasting conditions than the 50 μm cells as significantly less lysis was observed. Cellular fractionation of the spheroplasts followed by immunoblot revealed that the inner membrane fraction was enriched in *Pa*-AmpG and *Pa*-AmpP, indicating proper targeting of the proteins under spheroplasting conditions (Figure 2.10). As is typical for membrane proteins, both *Pa*-AmpG and *Pa*-AmpP migrated slightly further during electrophoresis than was predicted by their molecular weight.[174] (Figure 2.9, Figure 2.10).

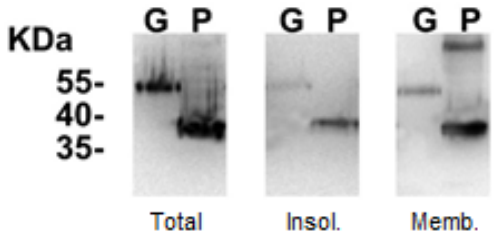


Figure 2.10. ***Ec*MGΔ*ampG* spheroplasts expressing V5His tagged *Pa*-AmpG (G) or *Pa*-AmpP (P) have the proteins properly targeted to the membrane.**

Knowing the *P. aeruginosa* transporters are properly expressed in spheroplasts, we next carried out the AmpG transport assay as was done for the *Ec*BW strains. We made a slight modification to our previous procedure in that we doubled the spheroplast concentration, as well as duration of incubation with probe **9**. This was in an attempt to compensate for the modest expression levels we observed for *Pa*-AmpG. As with *Ec*BW, analysis of the washing procedure indicated that a negligible amount of either NBD-Glc or probe **9** was found in the supernatant after three washes (Figure 2.11).

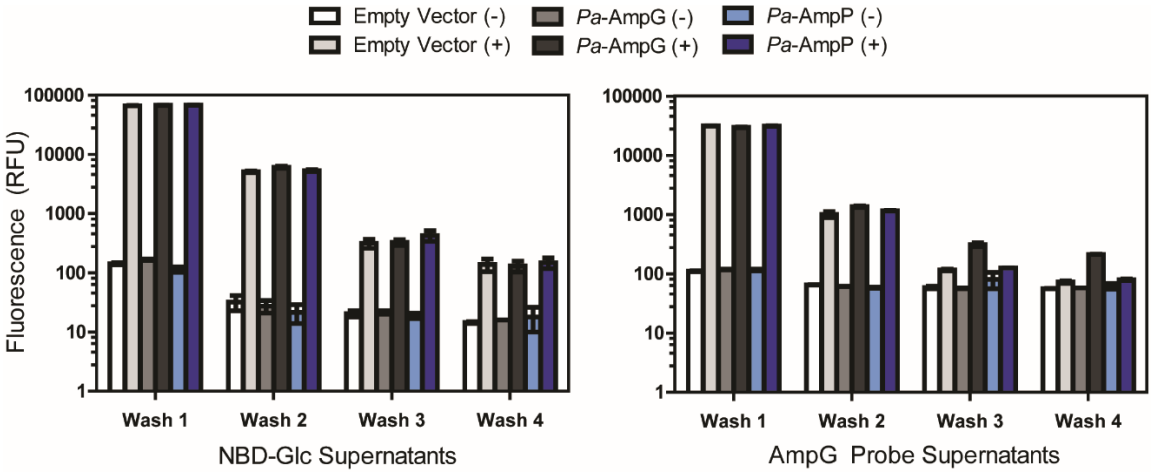


Figure 2.11. **Three washes were sufficient to remove the vast majority of excess NBD-Glc or probe 9 after incubation with *Ec*MGΔ*ampG* spheroplasts bearing the empty vector, or V5His tagged *Pa*-AmpG or *Pa*-AmpP.** Supernatant fluorescence after one to four washes of *Ec*MGΔ*ampG* spheroplasts bearing the empty vector or tagged *Pa*-AmpG or *Pa*-AmpP which were incubated with (+) or without (-) NBD-Glc or probe **9**. Error bars represent the standard error of the mean. Data plotted on a logarithmic y-axis.

Regardless of whether the *Ec*MGΔ*ampG* spheroplasts were transformed with an empty vector (pBAD322-V5His-Km), or one expressing *Pa*-AmpG or *Pa*-AmpP, NBD-Glc

was internalized to the same extent indicating comparable viability (Figure 2.12). However, only the spheroplasts expressing *Pa*-AmpG were able to internalize probe **9**; negligible probe **9** internalization was observed for spheroplasts expressing *Pa*-AmpP or those transformed with empty vector (Figure 2.12). The inability of *Pa*-AmpP to transport probe **9** does not stem from poor expression, since it is more highly expressed than is *Pa*-AmpG (Figure 2.10). These results were corroborated by fluorescence microscopy, which shows that *Ec*MGΔ*ampG* spheroplasts bearing any of the three vectors became fluorescent after incubation with NBD-Glc (Figure 2.12), whereas only the spheroplasts expressing *Pa*-AmpG were fluorescent after incubation with probe **9** (Figure 2.12). These results indicate that only *Pa*-AmpG, but not *Pa*-AmpP, can transport probe **9**, suggesting that *Pa*-AmpG, but not *Pa*-AmpP, codes for a functional transporter of the disaccharide GlcNAc-anhMurNAc in *P. aeruginosa*.

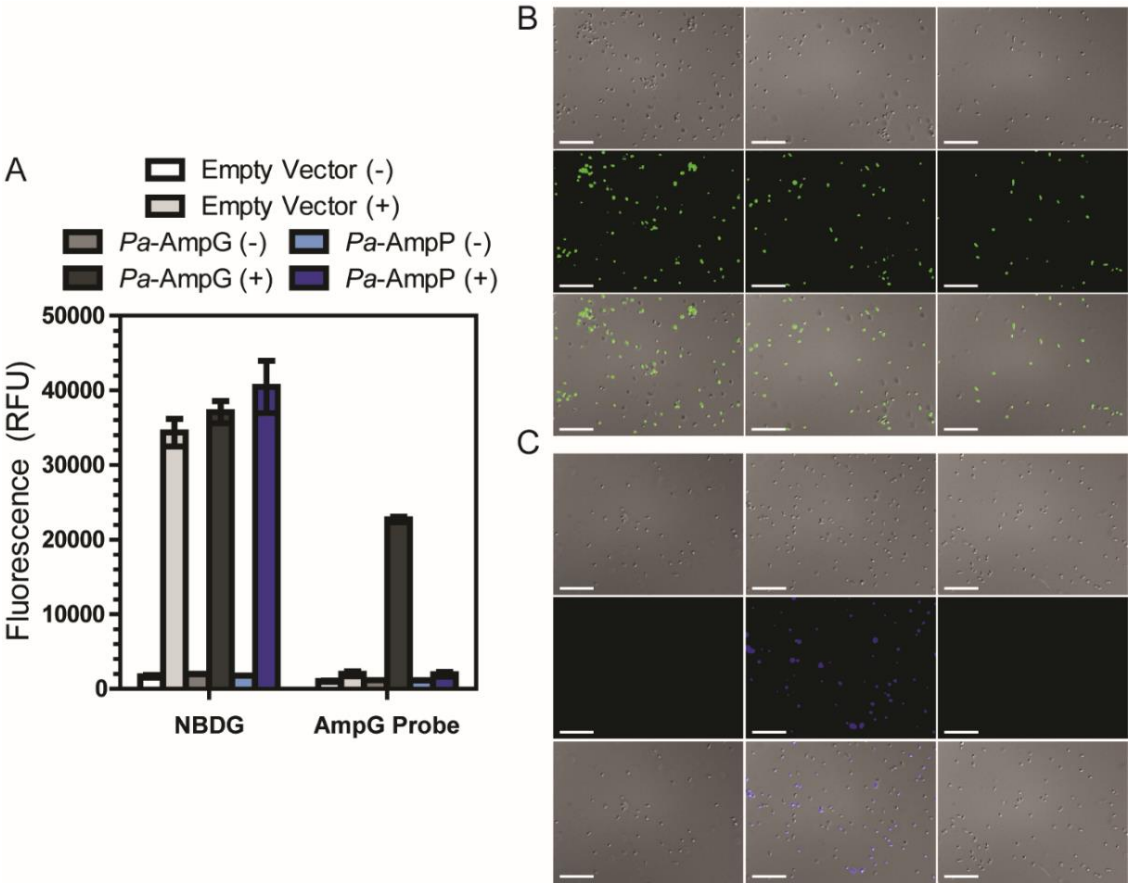


Figure 2.12. Spheroplasts of *ampG*⁻ *E. coli* (*Ec*MGΔampG) transformed with *Pa*-AmpG, but not with *Pa*-AmpP or an empty vector, are able to internalize probe 9.

A: Fluorescence of *Ec*MGΔampG spheroplasts carrying an empty vector, or expressing V5His tagged *Pa*-AmpG or *Pa*-AmpP, after incubation in the presence (+) or absence (-) of either NBD-Glc or probe **9**. The spheroplasts uptake NBD-Glc, indicating they are comparably viable. Error bars reflect the standard error of the mean. **B**: DIC (top), fluorescence (middle), and overlay (bottom) images of *Ec*MGΔampG spheroplasts transformed with the empty vector (right), or plasmids allowing the expression of *Pa*-AmpG (centre), or *Pa*-AmpP (left), after incubation with NBD-Glc; or probe **9** (**C**). Scale bars indicate 20 μm.

We corroborated the results of the uptake assay using an AmpC induction experiment that examined the ability of untagged *Pa*-AmpG (pBAD-*Pa*-AmpG-Km) and *Pa*-AmpP (pBAD-*Pa*-AmpG-Km) to complement an *ampG*⁻ strain (*Ec*MGΔ*ampG*). *Ec*MG and *Ec*MGΔ*ampG* were transformed with a plasmid harboring the *ampR-ampC* operon from *C. freundii* to enable monitoring of AmpC induction.[149] As controls, *Ec*MG and *Ec*MGΔ*ampG* were co-transformed with empty vector (pBAD322-Km), and after

exposure to the AmpC-inducer cefoxitin *Ec*MG exhibited dramatically higher AmpC activity than did *Ec*MGΔ*ampG* (Figure 15). This observation confirmed that *E. coli* AmpG is required for AmpC induction. Co-transformation of *Ec*MGΔ*ampG* with the plasmid expressing *Pa*-AmpG resulted in induction of AmpC in a manner dependent on the amount of *Pa*-AmpG expressed – heightened expression of *Pa*-AmpG from the pBAD promotor lead to heightened levels of AmpC activity (Figure 15). Conversely, co-transformation of *Ec*MGΔ*ampG* with the plasmid expressing *Pa*-AmpP resulted in no AmpC induction, regardless of the amount of *Pa*-AmpP expressed (Figure 15). These data indicate that *Pa*-AmpP is not involved in AmpC induction. Coupled with the results of the transport assay (Figure 2.12), we suggest that this is due to *Pa*-AmpP being unable to transport peptidoglycan fragments akin to probe **9** that are precursors to AmpC-inducer molecules.

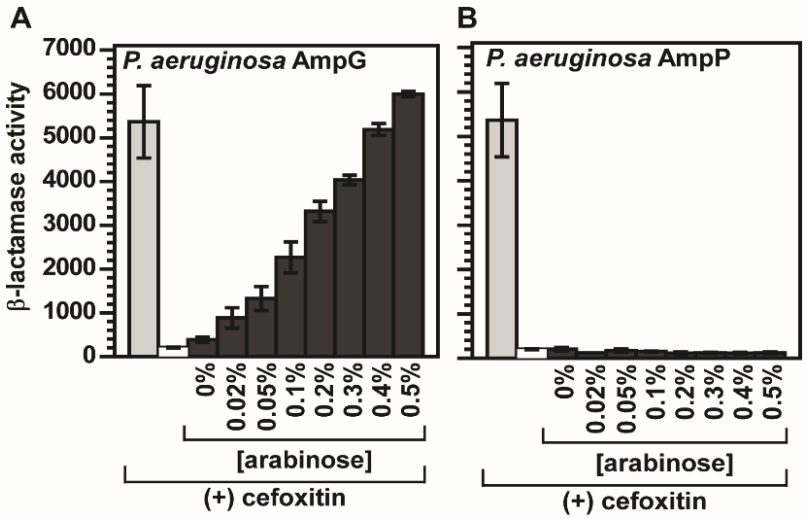


Figure 2.13. ***P. aeruginosa Pa*-AmpG but not *Pa*-AmpP is able to complement *ampG*⁻ *E. coli* (*Ec*MGΔ*ampG*).**
All *E. coli* strains were transformed with a plasmid bearing the *C. freundii ampR-ampC* operon. When transformed with empty vector, wild-type *E. coli* (*Ec*MG, gray), but not an isogenic *ampG*⁻ strain (*Ec*MGΔ*ampG*, white), is capable of expressing AmpC when treated with cefoxitin. **A**: Transformation of *Ec*MGΔ*ampG* with untagged *Pa*-AmpG (black) rescues AmpC expression in a manner dependent upon the amount of *Pa*-AmpG expressed. **B**: Transformation of *Ec*MGΔ*ampG* with untagged *Pa*-AmpP (black) is unable to rescue AmpC expression regardless of the amount of *Pa*-AmpP expressed. AmpC activity given in nanomoles of nitrocefin hydrolyzed per minute per mg protein. Error bars represent standard deviations.

2.5. Conclusions and Future Work

Chemical synthesis of a GlcNAc-1,6-anhydroMurNAc-fluorophore conjugate (probe **9**), in conjunction with adaptation of methods used in electrophysiology to generate *E. coli* spheroplasts,[167] enabled us to develop a quantitative fluorescence-based transport assay of AmpG proteins. AmpG transporters are ubiquitous among Gram-negative bacteria[151,152,154-156] and play a critical role in PG metabolism and in providing resistance against β-lactam antibiotics.[152,175] Surprisingly, this assay, to our knowledge, is only the second fluorescent transport assay developed using a carbohydrate-based substrate,[70] and the first for a periplasmic transporter. Importantly the assay is convenient, allowing facile kinetic analysis of AmpG without the prohibitively complex preparation or radiation safety concerns present in quantitative transport assays of other permeases.[124,170]

We expect this assay can be used generally, in conjunction with engineered bacterial strains, to assess the activity of AmpG homologues from other bacteria. We demonstrated this possibility by concluding that *P. aeruginosa* AmpG (*Pa*-AmpG), but not *Pa*-AmpP, transports GlcNAc-anhMurNAc-peptides that lead to induction of AmpC β-lactamase, a finding supported by complementation experiments using *ampG*$^{-}$ *E. coli* harboring an *ampR-AmpC* operon. These data clarify previous conflicting observations surrounding the activity of *Pa*-AmpP, which were based solely on bacteriological studies.[151,173]

We further expect that adaptation of spheroplasting procedures, as we have done for *E. coli* strains, will enable examination of AmpG transport function directly in other bacterial species. This will afford a more accurate understanding of AmpG function and the PG recycling pathway, as endogenous molecular effectors will be in place within the live bacterial spheroplasts. Additionally, this spheroplast-based assay may be adapted to study other periplasmic transporters, so long as a fluorescent analogue of the natural substrate is both tolerated and synthetically accessible. Of particular interest would be examining PG recycling within Gram-positive bacteria, which appear to lack AmpG homologues but for which PG recycling plays a crucial role in the bacterial life cycle and has also been recently linked to β-lactam antibiotic resistance.[175-178]

This spheroplast-based assay also enabled us to observe inhibition of AmpG-mediated transport of probe **9** by GlcNAc-anhMurNAc. Given this observation we anticipate that, with minor adjustments such as reducing the sample handling steps by employing a rapid-filtration method,[170] the assay could prove useful in high throughput screening of libraries to identify compounds that could serve as AmpG inhibitors. Moreover, improved knowledge of AmpG-substrate interactions could be accessed through this assay, and would aid in the rational design of such inhibitors. Indeed, analogues of GlcNAc-anhMurNAc may themselves prove to be potent inhibitors of AmpG transporters. Accordingly, this spheroplast-based transport assay may not only prove useful in understanding this important class of inner membrane transporters, but could also aid in the identification and characterization of compounds to combat AmpC-mediated β-lactam antibiotic resistance.

2.6. Experimental Section

2.6.1. Synthetic Procedures for Key Chemical Compounds.

Compound 8. To a mixture of **7** (40 mg, 0.0836 mmol) and benzyl 2-aminoethylcarbamate (19.5 mg, 0.100 mmol) in DMF (2 mL) was added HBTU (47.6 mg, 0.126 mmol) and DIPEA (43.7 µL, 0.251 mmol). This mixture was stirred overnight at room temperature, concentrated by rotary evaporation, and purified by silica-gel flash column chromatography (6:3:1 EtOAc:MeOH:H_2O) to give **8** (27 mg, 0.0418 mmol, 50%).

^{1}H NMR (600 MHz, CD_3OD) δ: 7.33 – 7.25 (m, 5H), 5.24 (s, 1H), 5.04 (q, J = 12.5, 2H), 4.54 (d, J = 5.4, 1H), 4.45 (d, J = 8.4, 1H), 4.14 (d, J = 7.5, 1H), 4.05 (q, J = 6.8, 1H), 3.97 (s, 1H), 3.87 (dd, J = 11.9, 1.9, 1H), 3.79 (s, 1H), 3.76 – 3.70 (m, 1H), 3.68 – 3.63 (m, 2H), 3.47 (s, 1H), 3.41 (dd, J = 10.5, 8.2, 1H), 3.37 – 3.33 (m, 1H), 3.32 (s, 1H), 3.31–3.30 (m, 1H), 3.25 (dd, J = 9.7, 3.8, 3H), 2.02 (s, 3H), 2.00 (s, 3H), 1.32 (d, J = 6.8, 3H); ^{13}C NMR (151 MHz, CD_3OD) δ: 176.05, 174.90, 173.52, 158.98, 138.35, 129.53, 129.04, 128.96, 102.05, 101.88, 79.61, 78.26, 77.58, 75.41, 75.32, 74.65, 72.07, 67.53, 65.84, 62.69, 57.14, 49.05, 41.24, 39.98, 23.50, 22.68, 18.83; HRMS: m/z calcd for $C_{29}H_{43}N_4O_{13}^+$ $[M+H]^+$: 655.2821, found 655.2823.

GlcNAc-anhMurNAc-AF350 (**9**). To a solution of **8** (12 mg, 0.018 mmol) in MeOH (5 mL) was added 10% Pd/C (10 mol%), and H_2 was bubbled into the reaction at room temperature. After 4 hr, the reaction mixture was filtered and the filtrate was concentrated by rotary evaporation. To this residue (10 mg, 0.019 mmol) in anhydrous DMF (2 mL) at 0 °C was added AF350 (5 mg, 0.012 mmol) and DIPEA (3.2 µL, 0.018 mmol). The solution was allowed to slowly warm to room temperature and was stirred overnight, after which concentration by rotary evaporation and purification by silica-gel flash column chromatography (7.4:1.6:1 EtOAc:MeOH:H2O) gave **9** (5.9 mg, 0.00732 mmol, 40%).

^{1}H NMR (600 MHz, CD_3OD) δ: 8.03 (s, 1H), 6.55 (s, 1H), 5.25 (s, 1H), 4.41 – 4.39 (m, 2H), 4.01 (q, J = 6.7, 1H), 3.94 (d, J = 7.5, 1H), 3.87 – 3.81 (m, 2H), 3.66 (ddd, J = 17.4, 15.6, 7.1, 3H), 3.57 – 3.51 (m, 4H), 3.44 (dd, J = 11.7, 6.8, 2H), 3.38 – 3.31 (m, 2H), 3.29 (d, J = 0.6, 4H), 2.37 (s, 3H), 1.98 (s, 4H), 1.96 (s, 3H), 1.30 (d, J = 6.7, 4H); ^{13}C NMR (151 MHz, CD_3OD) δ: 176.04, 174.97, 173.40, 173.30, 164.22, 156.43, 152.58, 150.34, 126.85, 126.70, 115.68, 110.83, 102.20, 102.07, 101.94, 79.21, 78.12, 77.26, 75.85, 75.35, 74.78, 72.06, 65.83, 62.68, 57.21, 49.05, 40.12, 39.87, 35.41, 23.49, 22.67, 18.65, 15.54; HRMS: m/z calcd for $C_{33}H_{46}N_5O_{17}S^+$ $[M+H]^+$: 816.2604, found 816.2601.

2.6.2. Preparation of Bacterial Strains and Spheroplasts

Bacterial strains and culture conditions. For routine cloning and plasmid propagation, DH5α (F^- Φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17(r_k^-, m_k^+) *pho*A *sup*E44 *thi*-1 *gyr*A96 *rel*A1 λ^-) or TOP10 (F- *mcrA* Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ *lac*X74 *rec*A1 *ara*D139 Δ(*araleu*)7697 *gal*U *gal*K *rps*L (StrR) *end*A1 *nup*G*)* were used as hosts. *E. coli* BW25113 (CGSC No. 7636) and MG1655 (CGSC No. 6300) were utilized for gene modification and was obtained from the Coli Genetic Stock Centre (Yale University, New Haven, CT, USA). *Pseudomonas aeruginosa* PAO1 was used for chromosomal DNA isolation and was generously provided by Dr. A. Oliver (Instituto Universitario de Investigación en Ciencias de la Salud (IUNICS), Palma de Mallorca, Spain). If not otherwise stated, cells were grown aerobically in LB medium (10 g/L tryptone, 5 g/L yeast extract, 5 g/L sodium chloride) with the following additions: 150

μg/mL ampicillin, 35 μg/mL kanamycin and/or 35 μg/mL chloramphenicol and 0.2% (w/v) arabinose. For β-lactamase assays, the cells were grown in M9CA supplemented with 1 μg/mL thiamine, 50 μg/mL uracil, 35 μg/mL chloramphenicol, 70 μg/mL kanamycin, 0.5% (v/v) glycerol and arabinose at the indicated concentration.

Deletion of *E. coli* AmpG from the chromosome of BW25113 and MG1655. Wild-type *E. coli* BW25113 (CGSC No. 7636) and its isogenic Δ*ampG790::kan* derivative (CGSC No. 8587, Keio Collection No. JW0423-1) were obtained from the Coli Genetic Stock Center. The kanamycin resistance gene of the mutant was eliminated via recombination between the flanking FRT sites using FLP recombinase supplied by the chloramphenicol and ampicillin resistant, temperature-sensitive plasmid, pCP20. This plasmid was subsequently cured by growth at 42 °C and the strain was tested for loss of all antibiotic resistance determinants. Chromosomal deletion of the *ampG* gene in *E. coli* MG1655 (CGSC No. 6300) was carried out using the strategy as described.[179] Briefly, PCR products were amplified using primers ECΔAMPGF: 5'-GATTACTACAGCTAAATAATATTTACAGATTACGTCAGATGCGTTTTTCGattccggggat ccgt-cgacc-3' and ECΔAMPGR: 5'-TTAATTTCTGCCCTCTGGCCCGGTGCAAGCCG-GGCCTGTAGACGCCCATGtgtaggctggagctgcttcg-3' and Phusion High Fidelity DNA polymerase as specified by the manufacturer. These primers were designed to have 20 nucleotide C-terminal ends for priming the FRT sites flanking the kanamycin resistance cassette in pKD13 (lower case letters) and 50 nucleotide ends homologous to the regions upstream and downstream of the targeted gene deletion (upper case letters). Chromosomal integrands were selected as kanamycin resistant colonies after electroporation of the PCR product into MG1655 harbouring the ampicillin resistant λRed expression plasmid, pKD46. The kanamycin resistance gene was then eliminated using pCP20 as described above. The presence of the null deletion in both strains was verified by PCR amplification using the primers ECAMPGF: 5'-GACAACCATCCCAAACTTTTAT-TGTATTTCCTG-3' and ECAMPGR 5'-CGTCACCGTAAGCATTAATGGTGC-3'.

Construction of pBAD-*Pa*-AmpG-Km and pBAD-*Pa*-AmpP-Km. Phusion High Fidelity DNA polymerase was used to PCR amplify the *ampG* (*Pa*4393) and *ampP* (*Pa*4218) ORFs utilising chromosomal DNA isolated from *Pseudomonas aeruginosa*

PAO1 as described.[180] Primers used for amplification were PA4393F: 5'-GATATAgaattcaCCATGACTCAGCAATCCTGGCGAGAG-3' (an introduced EcoR1 site is in lower-case letters) and PA4393R: 5'-CTATATaagcttTCAGTGC-TGCTCGGCGTTCTG-3' (an introduced HindIII site is in lower case letters), and PA4218F: 5'-GATATAgaattcaCCA-TGCTTGAGCTGTAC-3' (an introduced EcoR1 site is in lower-case letters) and PA4218R: 5'-CTATATaagcttTCAGGCCTCTTCCGCCCG-3' (an introduced HindIII site is in lower-case letters), respectively. The purified PCR products were then directionally cloned into pBAD322-Km under the arabinose-inducible, glucose-attenuable P_{BAD} promoter, as EcoR1/ HindIII fragments and the fidelity of the inserts was confirmed by DNA sequencing at the Toronto Hospital for SickKids, Centre for Applied Genomics, DNA Sequencing/Synthesis Facility.

Construction of pBAD-*Pa*-AmpG-V5His-Km and pBAD-*Pa*-AmpP-V5His-Km. Initially, a fragment containing the C-terminal V5 epitope and six histidine tag, together with the downstream *rrn*B T_1 and T_2 transcriptional terminators was amplified by PCR utilizing pBAD-TOPO/lacZ/V5-His plasmid DNA as a template. Primers V5HISF: 5'-GATATAaagcttGGTAA-GCCTATCCCTAAC-3' (an introduced HindIII site is in lower-case letters) and V5HISR: 5'-CTATATtcatgaGAGTTTGTAGAAACGCAAAAAG-3' (an introduced BspH1 site is in lower-case letters) were used, as well as Phusion High Fidelity DNA polymerase under the reaction conditions as suggested by the manufacturer. The resulting purified PCR product was then cloned into pBAD322-Km[5] expression vector as a HindIII /BspH1 fragment yielding the pBAD322-V5His-Km construct. pBAD-Pa-AmpG-V5His-Km and pBAD-Pa-AmpP-V5His-Km were generated essentially as described above, but the reverse primers used for PCR amplification (AMPGV5HISR: 5'-CTATATaagcttGTGCTGCTCGGCGTTCTG-3' and AMPPV5HISR: 5'-CTATATaagcttGG-CCTCTTCCGCCCG-3') were designed to eliminate the native stop codon of both ORFs to preserve reading through the C-terminal V5 epitope and six histidine tags provided by the pBAD322-V5His-Km vector. DNA sequencing at the Toronto Hospital for SickKids, Centre for Applied Genomics, DNA Sequencing/Synthesis Facility, confirmed the fidelity of all inserts.

Preparation of spheroplasts. The protocol was adapted from a previously published procedure.[167] For the BW25113 strains, cultures of *Ec*BW or *Ec*BWΔ*ampG*

were prepared in lysogeny broth (LB) and grown aerobically overnight at 37 °C. The next day, the cells were inoculated, 1/100, into 400 mL of LB and allowed to grow at 37 °C, until the OD600 reached ~ 0.5 - 0.6. To 1 L of LB, pre-heated to 42 °C, was added: 120 mL of OD600 0.5 - 0.6 culture, and 7.2 mL of 2 mg/mL (*Ec*BW) or 1 mg/mL (*Ec*BWΔ*ampG*) cephalexin. The cells were further incubated at 42 °C until they had elongated to 25 – 50 μm as determined by differential interference contrast (DIC) microscopy using an Axioskop 2 Plus microscope. The cultures were then centrifuged at 3000 rcf at 4 °C for 10 min and the resulting pellets resuspended in 100 mL of 0.8 M sucrose. For spheroplasting, the following reagents were added in the indicated order: 6 mL of 0.2 M Tris pH 7.2, 4.8 mL of 1 mg/mL lysozyme, 2 mL deionized water and 6 mL of 0.025 M EDTA. Once spheroplast formation was complete, as monitored by DIC microscopy, 40 mL of ice-cold stop solution (87.5% 0.8 M sucrose, 9.5% deionized water, 2% 1 M $MgCl_2$, 1% 1 M Tris pH 7.2) was added, followed by 280 mL of dilution solution (98% 0.8 M sucrose, 1% $MgCl_2$, 1% 1 M Tris pH 7.2). The spheroplast suspensions were then centrifuged at 2000 rcf at 4 °C for 25 min, washed once with 5 mL of dilution solution, and then resuspended to homogeneity in dilution solution to a concentration of 0.4 - 0.5 g/mL. In the case of MG1655, overnight cultures of *Ec*MGΔ*ampG* transformed with either pBAD322-V5His-Km, pBAD322-*Pa*-AmpG-V5His-Km, or pBAD322-*Pa*-AmpP-V5His-Km, were carried out in the presence of 35 μg/mL kanamycin. The next day, the cells were inoculated, 1/50, into 500 mL LB and allowed to grow at 28 °C until the OD600 reached ~ 0.6 - 0.7. To induce expression, the cells were grown an additional 4 hr at 28 °C in the presence of 0.2% L-arabinose. During the last hour of induction 6 μg/mL of cephalexin was added, after which the cells harvested by centrifugation and spheroplasts were prepared essentially as described above. The spheroplast stocks were stored at 4 °C and used within 48 hours.

2.6.3. Assays and Kinetics

Transport of probe 9 by native *Ec*BW and *Ec*BWΔ*ampG E. coli.* Wild-type or isogenic *ampG*$^-$ BW25113 *E. coli* (*Ec*BW or *Ec*BWΔ*ampG*) were grown until the OD600 reached ~ 0.5 - 0.6, as described in the 'Preparation of Spheroplasts' portion of the Experimental Section. To 1 L of LB, preheated to 37 °C, was added 120 mL of OD600 0.5 - 0.6 culture. The cells were grown aerobically at 37 °C for 1 hr, centrifuged at 3000

rcf at 4 °C for 10 min, and resuspended to homogeneity in dilution solution to a concentration of 0.25 g/mL. The resuspended bacteria was stored at 4 °C. The next day 200 µL of 0.25 g/mL *Ec*BW or *Ec*BWΔ*ampG* was added to microcentrifuge tubes, and the solution was brought up to 250 µL with dilution solution. The experiment was initiated by the addition of 100 µL of 175 µM probe **9**. Blanks were prepared in which dilution solution was added in place of probe **9**. Tubes were mixed by inversion several times, then incubated for 1 or 3 hr at 25 °C in the dark using a digital heatblock. Final reaction conditions were: 350 µL; 50 µM probe **9**, and 0.15 g/mL *E. coli*. Post-experiment washing and fluorescence quantitation was performed as described below.

Uptake assay for probe 9 and NBD-Glc. Experiments were carried out in duplicate using *Ec*BW and *Ec*BWΔ*ampG* spheroplasts, or with spheroplasts derived from *Ec*MGΔ*ampG* transformed with the indicated expression plasmid. Briefly, 50 µL of 0.4 g/mL (*Ec*BW) or 100 µL of 0.4 g/mL (*Ec*MG) spheroplasts were added to microcentrifuge tubes and the volume was brought up to 250 µL with dilution solution. The assay was then initiated by the addition of 100 µL of 175 µM probe **9**, or of 1.5 mM NBD-Glc. Blanks were prepared in which dilution solution was added in place of the fluorescent compound. The tubes were mixed by inversion several times, then incubated for 1 hr (*Ec*BW) or 2 hr (*Ec*MG) at 25 °C in the dark using a digital heatblock. Final reaction conditions were: 350 µL; 50 µM probe **9** or 430 µM NBD-Glc, and 0.06 g/mL (*Ec*BW) or 0.12 g/mL (*Ec*MG) spheroplasts. Post-assay spheroplast washing and fluorescence quantitation was performed as described below.

Time course of probe 9 uptake by *Ec*BW spheroplasts. The experiment was carried out in duplicate using *Ec*BW spheroplasts. In brief, 50 µL of 0.25 g/mL spheroplasts were added to microcentrifuge tubes, followed by the addition of 50 µL of 100 µM probe **9** in dilution solution, beginning with those tubes corresponding to the longest time points. Blanks contained dilution solution in place of probe **9**. The tubes were inverted several times to mix, and were incubated for 0, 5, 10, 30, 60, 90, 120, and 240 minutes at 25 °C in the dark using a digital heatblock. Final reaction conditions were: 100 µL; 50 µM probe **9**, and 0.125 g/mL spheroplasts. Post-assay spheroplast washing and fluorescence quantitation was performed as described below.

Michaelis-Menten kinetic evaluation of the uptake of probe 9. The experiment was carried out in singlicate using spheroplasts derived from *Ec*BW. In microcentrifuge tubes, a series of probe **9** concentrations was obtained via three-fold serial dilution, from a 900 μM stock, into a final volume of 100 μL of dilution solution. Then the assay was initiated by the addition of 50 μL of 0.5 g/mL spheroplasts. The tubes were inverted several times to mix and were incubated for 1 hr at 25 °C in the dark using a digital heatblock. Final reaction conditions were: 150 μL; 600 μM to 0 μM probe **9**, and 0.17 g/mL spheroplasts. Post-assay spheroplast washing and fluorescence quantitation was performed as described below.

Competitive inhibition of probe 9 uptake by GlcNAc-anhMurNAc. The experiment used *Ec*BW spheroplasts, and was carried out in singlicate with GlcNAc or anhMurNAc, or in triplicate with GlcNAc-anhMurNAc. In microcentrifuge tubes, a range of concentrations was initially prepared by four-fold dilution of GlcNAc or anhMurNAc, or three-fold dilution of GlcNAc-anhMurNAc, in 100 μL of dilution solution. Then, 50 μL of 200 μM probe **9** in dilution solution was added and the competitive assay was initiated by the addition of 50 μL of 0.52 g/mL *Ec*BW spheroplasts. Blanks contained dilution solution in place of probe **9**. The tubes were then mixed by inversion and were incubated for 1 hr at 25 °C in the dark using a digital heatblock. Final reaction conditions were: 200 μL; 1 mM to 0 mM GlcNAc or anhMurNAc, or 1.5 mM to 0 mM GlcNAc-anhMurNAc; 50 μM probe **9**; 0.13 g/mL of spheroplasts. Post-assay spheroplast washing and fluorescence quantitation was performed as described below.

Spheroplast washing and fluorescence quantitation. After incubation of spheroplasts with labeled substrate, each tube was placed on ice and 1 mL of ice-cold dilution solution was added in the same order in which the experiment in question was initiated. The tubes were then centrifuged at 2000 rcf at 4 °C for 20 min, and the supernatants were withdrawn using a syringe equipped with an 18 gauged needle and stored on ice for later fluorescence testing. The syringe was subsequently washed twice with water after finishing with one set of tubes (i.e. blank, NBD-Glc, or probe **9**). Dilution, centrifugation, and supernatant withdrawal was repeated two additional times to wash away remaining extracellular fluorogenic substrate for *Ec*BW, or three additional times for *Ec*MG, and the final washed pellets were vortexed to homogeneity in 200 μL of

dilution solution. To a black 96-microwell plate was added 150 µL of each spheroplast suspensions and the supernatant. Fluorescence was quantified at 25 °C with excitation/emission wavelengths of 350/450 nm for probe **9**, and 475/535 nm for NBD-Glc. Data was acquired using a Synergy 4 microplate reader, reported in relative fluorescence units (RFU), and was analyzed using Gen5 and Graphpad software.

β-lactamase assays. *Ec*MGΔ*ampG* bearing the plasmid-borne *Citrobacter freundii ampR-ampC* operon[149] transformed with either pBAD-*Pa*-AmpG-Km or pBAD-*Pa*-AmpP-Km were grown at 37 °C with vigorous aeration in M9CA containing 1 µg/mL thiamine, 50 µg/mL uracil, 35 µg/mL chloramphenicol, 70 µg/mL kanamycin and 0.5% (v/v) glycerol until the OD600 reached approximately 0.4. L-arabinose was then added at the specified concentration and the cells were grown at 23 °C for 1.5 hr. To induce *ampR-ampC* expression, the cultures were supplemented with 16 µg/mL cefoxitin and were grown for an additional 1.5 hr at 23 °C. Crude sonic lysates were then prepared as described previously[149,181] and β-lactamase specific activity was determined using a continuous assay procedure that follows the linear rate of liberation of 2,4-dinitrophenolate from nitrocefin,[182] as determined by measuring the absorption at 486 nm. The experimental mixture contained 100 µL of PBS and 1 µL of an appropriately diluted lysate. The reaction was then initiated by the addition of 100 µM nitrocefin and monitored at room temperature for 5 min. The β-lactamase specific activity is expressed in nanomoles of nitrocefin hydrolyzed per min per mg protein. Wild-type *Ec*MG and *Ec*MGΔ*ampG* transformed with "empty" pBAD322-Km[183] served as positive and negative controls. Protein concentrations of the supernatants were quantified by the Bradford assay using bovine gamma globulin as the standard.

2.6.4. Other Techniques.

Dose-dependent expression of V5His tagged *Pa*-AmpG and *Pa*-AmpP. Overnight cultures of *Ec*MGΔ*ampG* transformed with the pBAD-*Pa*-AmpG-V5His-Km or pBAD-*Pa*-AmpP-V5His-Km were inoculated, one in fifty, into 3.0 mL of the LB containing 35 µg/mL kanamycin, and grown at 28 °C until the optical density at 600 nm reached approximately 0.4. To induce expression, L-arabinose was added at the indicated concentration and the cultures were grown at 28 °C for an additional three hours. The

cells were then pelleted by centrifugation at 6,000 rcf at 4 °C for 5 min, washed once with PBS, resuspended in lysis buffer containing 50 mM Tris-HCl (pH 8.0), 150 mM NaCl, 1% Triton-X-100, 100 µg/mL lysozyme, 50 µg/mL DNase, 0.1 mM PMSF and 0.5 mM TCEP, then mixed by end-over-end rotation for 1 hr at 4 °C. The cellular debris was subsequently removed by centrifugation at 12,000 rcf at 4 °C for 30 min, and the supernatants, which contained solubilized membranes, were analysed by Western blot analysis as described in the Experimental Section. Protein concentrations were quantified by the Bradford assay using bovine gamma globulin as the standard.

Spheroplast fractionation. Approximately 60 mg of spheroplasts were collected by centrifugation at 2,000 rcf at 4 °C for 5 min, washed once in dilution solution containing 0.5 mM TCEP and 0.1 mM PMSF, resuspended in the same buffer, supplemented with 1% Triton-X-100 and 0.1% SDS, then mixed by end-over-end rotation for 1 hr at room temperature. The suspensions were then pelleted by two consecutive spins at 16,000 rcf at 4 °C for 20 min and the resulting supernatants containing solubilized membrane protein were retained. The final pellets, representing the insoluble fraction, were resuspended in 300 µL of dilution solution, sonicated three times at 50% duty cycle, six second bursts, with 10 seconds of incubation on ice between each burst. Western blot analysis was performed as described in the Experimental Section. The Protein concentration of all fractions was quantified by the Bradford assay using bovine gamma globulin as the standard.

Western blot analysis. 50 µg of total protein in a final volume of 20 µL was mixed with 20 µL of standard 2X loading buffer and heated at 37 °C for 15 min. Proteins were resolved on 12% polyacrylamide gels and transferred onto Immobilin-P PVDF Western blotting membrane using the Bio-Rad Mini-Trans-Blot Electrophoretic Wet Transfer Cell according to the manufacturer's instructions. The membranes were blocked overnight at 4 °C in 5% skim milk powder dissolved in 1X TBS (2.4 g/L Tris, 8.8 g/L NaCl, pH 7.4), then washed twice in 1X TBS for 5 min, prior to incubation in the presence of mouse α-V5-HRP conjugate at a dilution of 1:5000 in 1% skim milk powder dissolved in 1X TBS for 1 hr at room temperature. The blots were then rinsed four times in 1X TBS containing 0.05% Tween-20 for 10 min at room temperature and immunoreactive proteins were detected using Luminata Forte Western HRP

Chemiluminescent substrate, as recommended by the supplier. Fluorchem 8900 was used to obtain the resulting signal.

Microscopy. Images of post-assay spheroplasts were captured using an Axioskop 2 Plus light microscope equipped with an Axiocam camera, and were processed using Axiovision software. Fluorescence images were obtained using a Colibri LED light source. A 2 sec exposure was used in all fluorescence images for comparability. For the NBD-Glc images, brightness was adjusted post-capture, as fluorescence intensity rapidly decayed on excitation. Identical brightness settings were used for all NBD-Glc fluorescence images for comparability. Excitation wavelengths of 365 nm and 470 nm were used for probe **9** and NBD-Glc, respectively. Both DIC and fluorescence images were captured at 1000 times magnification.

Chapter 3.

Development of Quenched Fluorescent Substrates for Monitoring *O*-GlcNAcase Activity in Living Systems

3.1. Abstract

Despite being invaluable for assessing the kinetics of enzyme substrates and inhibitors and for screening compound libraries for modulators of enzyme activity, *in vitro* assays are nevertheless limited in the information they supply. This is a result of *in vitro* assays providing neither an accurate measure of native spatiotemporal enzyme activity, nor insight into the molecular effectors and dynamic post-translational modifications occurring in the complex milieu of a living cell. Much more informative in these regards are live-cell and *in vivo* assays, in which factors contributing to the true complexity and dynamism of a living system are retained. However, the complex environments provided by cells, tissues, or whole organisms inherently complicate the design of such assays, and issues such as substrate selectivity, sensitivity, reactivity, and cell permeability must be addressed. These problems can be overcome by employing quenched fluorescent probes – substrates whose background fluorescence is negligible, but which form highly fluorescent products on enzymatic processing. One enzyme which would greatly benefit from live-cell and *in vivo* study is human *O*-GlcNAcase (hOGA). This nucleocytoplasmic protein is one of the two enzymes involved in regulating the *O*-GlcNAc modification, and is responsible for removing this modification from the serine and threonine residues of hundreds of different proteins. *O*-GlcNAc has been implicated in a diversity of disease states ranging from cancer metastasis to Alzheimer's disease, as well as in fundamental cellular processes including cell-cycle progression and stabilization of nascent proteins. However, due to the difficulties associated with developing substrates and assays for live-cell and *in vivo* use, very few such probes exist for glycoside hydrolases and none

are yet known for hOGA. In this work, we describe the synthesis of new hOGA substrates, based on the widely used 4-methylumbelliferyl 2-acetamido-2-deoxy-β-D-glucopyranoside (4MUGlcNAc), whose fluorescence is quenched by thioamides appended to the C-6 position. We demonstrate that the degree of quenching can be controlled by both the linker length and installation of moieties that promote close contact between the fluorophore and quencher. We also evaluate the processing of these quenched substrates by hOGA, and show that structural elaboration of C-6 is tolerated. Together, these results lay out a blueprint for the design of quenched hOGA substrates with potential value for live-cell imaging of hOGA activity, and should also provide insights into in the design of such probes for other glycoside hydrolases.

3.2. Introduction

The development of new chemical tools and methodologies for studying enzymes not only afford us greater insight into the function of these cellular machines, they also facilitate the detection, monitoring, and treatment of diseases associated with deficiencies in enzyme function or spatiotemporal regulation. Traditionally, *in vitro* assays have been employed to study enzyme-catalyzed reactions because they allow experimental conditions to be carefully controlled. Indeed, *in vitro* techniques are ideal for screening compound libraries to identify substrates or inhibitors,[184,185] for determining kinetic parameters such as Michaelis (K_M) or inhibition (K_i) constants,[186,187] and for elucidating enzyme specificity and mechanism of action.[44,188] However, the highly optimized conditions of *in vitro* assays do not accurately reflect the complex cellular environment in which enzymes function. Information gleaned using *in vitro* assays is therefore inherently limited. To gain a fuller understanding of enzyme function, live-cell or *in vivo* assays can be used. Doing so ensures that factors regulating enzyme localization and activity are unaffected, allowing the enzyme of interest to be influenced by the full repertoire of activators, inhibitors, small-molecule or protein co-factors, and post-translational modifications present within cells.[90]

Proteases have received the most attention from researchers aiming to better understand enzyme function through the development of live-cell and *in vivo* assays.[189-192] Few examples exist for glycoside hydrolases (GHs),[108,193,194] likely due to the

sterically constrained pocket-shaped architectures of *exo*-acting GH active sites as compared to the more accommodating cleft-like active sites of proteases. Cleft-shaped active sites enable installation of fluorescent reporter groups at either end of a peptide, whereas for GHs the locations at which such groups can be appended are highly limited. A GH of particular interest for live-cell and *in vivo* study is human O-GlcNAcase (hOGA) – a GH family 84 β-hexosaminidase which acts reciprocally to *O*-GlcNAc transferase (OGT)[195] to catalyze the removal of 2-acetamido-2-deoxy-β-D-glucopyranose (GlcNAc) from the serine and threonine residues of nucleocytoplasmic proteins (*O*-GlcNAc).[196-198] This post-translational modification is found on hundreds of proteins, and its importance is highlighted by the fact that loss of hOGA or OGT in mammals is lethal.[199,200] *O*-GlcNAc has been found to play a role in many cellular processes including cell cycle progression,[201] stress response,[202,203] and stabilization of proteins and nascent polypeptide chains.[204] Further, *O*-GlcNAc has been implicated in disease states including neurodegeneration and Alzheimer's disease,[205,206] and cancer progression and metastasis.[207,208] Despite the many biological roles in which *O*-GlcNAc is involved, little progress has been made toward studying the activities of the associated enzymes in living systems, and no methods yet exist for live-cell or *in vivo* monitoring of hOGA activity.

Live-cell and *in vivo* experiments pose a unique set of challenges in terms of substrate design that are not encountered *in vitro*, suggesting why such assays have yet to be developed for hOGA. Foremost, reaction of the substrate with the target enzyme must result in a signal that is distinguishable from the background. This requirement can be addressed in multiple ways. Firstly, the substrate can be tailored to improve the rate of its processing by the target enzyme, resulting in a higher turnover and thus a greater signal. Additionally, installation of a quencher to reduce the background fluorescence of the intact substrate, while also ensuring that the product of catalysis is comparatively bright, can afford an increase in the signal-to-noise ratio.[108,192] Also important in such assays is the issue of selectivity – the designed substrate must preferentially target the enzyme of interest to prevent off-target activities from obscuring the results.[209,210] An additional issue, in the case of fluorescent substrates, is ensuring that the emission wavelength of the product is shifted to the near infrared (NIR) region of the electromagnetic spectrum, as cell autofluorescence is minimal in this spectral region and

maximum tissue penetration can be achieved.[93] Finally, the substrate must be able to reach the enzyme of interest by penetrating the cell membrane. The permeability of a substrate depends on properties such as its molecular weight, overall polarity, and presence of hydrogen bonding or ionizable functional groups. These physicochemical properties can be tuned by optimizing the structure of the substrate and through the installation of specific moieties.[97,99]

Low background fluorescent substrates quenched by Förster resonance energy transfer (FRET) are the most widely used tools for studying enzyme activity in living systems.[81,83,97,108,211] These substrates rely on overlap between the emission profile of the fluorophore and the excitation profile of the quencher to enable fluorescence quenching by non-radiative energy transfer between the two chromophores.[69] However, because of the dependence on spectral overlap, most FRET quenchers are bulky aromatic moieties. Therefore, the design of FRET quenched substrates that can be accommodated within an enzyme active site, and whose physical properties are amenable to live-cell or *in vivo* assays, is challenging. Several approaches have been taken to mitigate such steric problems including installation of the quencher at a position on the substrate that does not form extensive contacts with the enzyme, appending the quencher to the substrate by a linker so that it can orient in a way that minimizes unfavourable interactions, or by a combination of these techniques.[108,192] An alternative approach that could circumvent these issues is to employ a particularly small quencher that minimizes these concerns.

Surprisingly poorly explored is the use of alternative quenching modes that do not rely, as FRET does, on spectral overlap between the fluorophore and quencher. Photoinduced electron transfer (PET) is one such mechanism, which instead depends on the redox potentials of the fluorophore and quencher and the formation of a charge transfer complex.[69] Notably, simple thioamides have been found to be efficient PET quenchers for many fluorophores,[86] and their use in the design of quenched substrates has enabled facile study of protease function as well as protein dynamics and stability.[212,213] Thioamides have many beneficial properties in comparison to traditional quenchers that recommend their use in the design of quenched hOGA substrates for live-cell or *in vivo* assays. Firstly, unlike bulky aromatic quenchers,[214] thioamides are

small and so should not markedly congest the hOGA active site or hinder catalysis when tethered to a substrate. Further, their small size should result in a minimal impact on the solubility and membrane permeability of the substrate. Finally, thioamides themselves are not fluorescent, unlike most aromatic quenchers,[73] and so should allow for more sensitive measurements by making no contribution to the background whatsoever.

To test the concept of thioamide-quenched hOGA substrates, we synthesized three derivatives of the widely used substrate 4-methylumbelliferyl 2-acetamido-2-deoxy-β-D-glucopyranoside (4MUGlcNAc) appended with thioamide substituents at the C-6 position (Figure 3.1). We show that these compounds are hOGA substrates, and we compare their hOGA-catalyzed processing relative to 4MUGlcNAc. These data provide insight into the active site tolerance of hOGA for specific substrate modifications and installation of the thioamide quencher. We also find that the fluorescence of these thioamide-containing compounds is quenched relative to 4MUGlcNAc, and that the efficiency of quenching can be tuned by adjusting the structure of the probe.

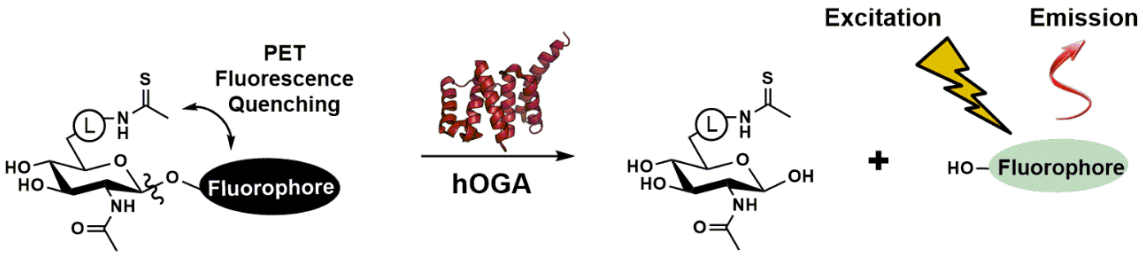


Figure 3.1. Concept for thioamide PET quenched hOGA substrates. The fluorescence of the intact substrate is quenched by PET between the C-6 thioamide and the anomeric fluorophore. Quenching can be tuned by modification of the linker "L". Processing of the substrate by hOGA liberates the fluorophore which diffuses away, bringing an end to quenching.

3.3. Results and Discussion

3.3.1. Substrate Design and Synthesis

Although the fluorophore 4-methylumbelliferone (4MU) is not ideally suited for assays in living systems because of its relatively blue-shifted fluorescence, it is convenient for use in the synthesis of probe molecules to test the concept of PET-quenched substrates for hOGA. In addition, 4MU is inexpensive in comparison to other

fluorophores and is stable to a range of reactive conditions. Moreover, by employing 4MU the commercially available and widely used hOGA substrate 4MUGlcNAc can be used as a reference compound when measuring kinetics and quantifying quenching. Since PET quenching is dependent on the relative redox potentials of the fluorophore and quencher,[69] not all fluorophores are quenched to the same extent in the presence of a thioamide.[86] Thus, before undertaking synthesis we examined whether the fluorescence of 4MU could be quenched by thioacetamide in solution. We performed four-fold serial dilutions of thioacetamide in water to obtain eight concentrations (0 - 300 mM), to which we added either 4MUGlcNAc (2.5 mM) or 4MU (1 μM), or resorufin (1 μM) as a positive control that is known to be quenched by thioacetamide.[86] We observed that both 4MUGlcNAc and free 4MU were quenched in a concentration dependent manner by thioacetamide, as was resorufin (Figure 3.2). Since PET quenching requires formation of a charge-transfer complex and is therefore highly dependent on the distance between the fluorophore and quencher,[69] this concentration dependence was expected. At higher thioacetamide concentrations, the average distance between fluorophore and thioacetamide molecules is smaller, leading to heightened quenching. The converse is true at low thioacetamide concentrations, resulting in comparatively lessened quenching. While upwards of 75% quenching of 4MU was observed at high thioacetamide concentrations, quenching at low thioacetamide concentrations was negligible. Given that tethering two moieties can yield high effective concentrations,[215] we anticipated that a substrate bearing both 4MU and a suitably positioned thioamide should exhibit effective PET quenching while intact. However, given that low mM or high μM substrate concentrations are generally used in enzyme assays, quenching of the liberated 4MU by the remaining thioamide-sugar should be negligible. Knowing that 4MU is efficiently quenched by thioacetamide, we proceeded with this fluorophore in these proof-of-concept experiments.

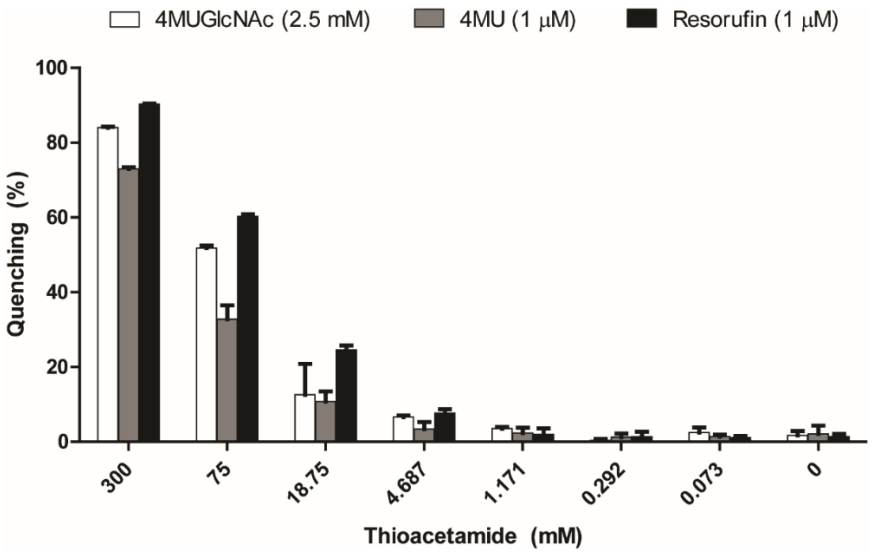
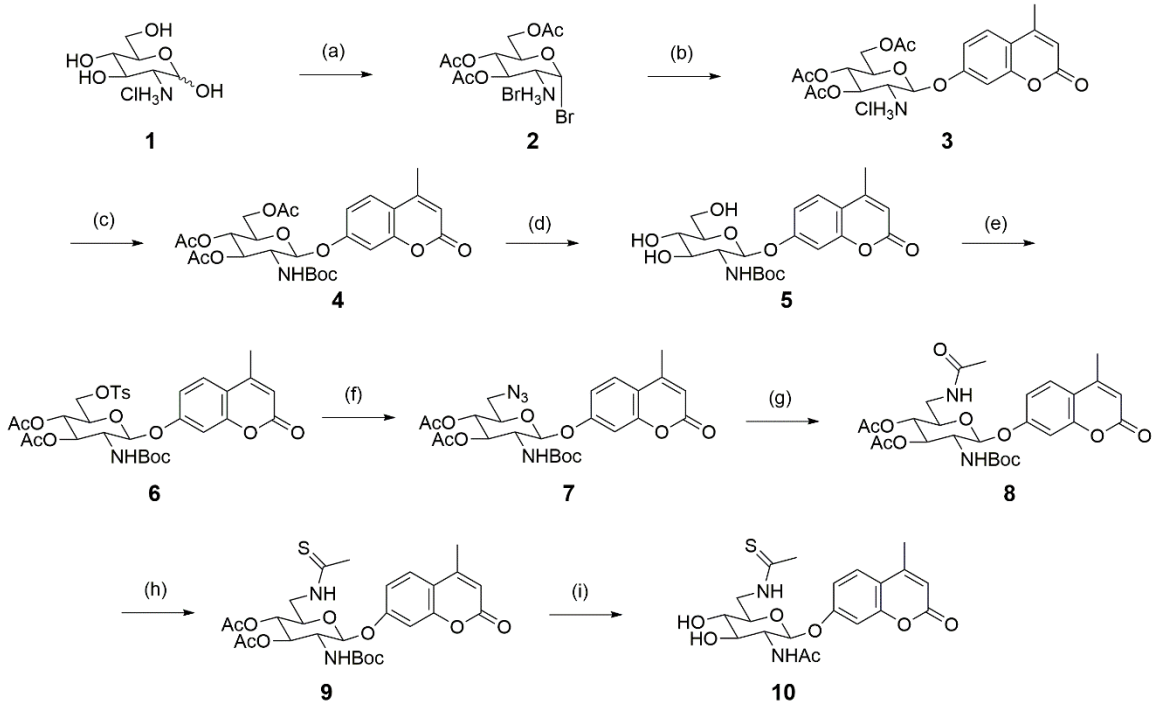


Figure 3.2. 4MUGlcNAc, 4MU, and resorufin are all quenched by thioacetamide in aqueous solution.
The dependence of quenching on thioacetamide concentration highlights the distance dependence of PET.

A co-crystal structure of the hOGA inhibitor NAG-thiazoline with an hOGA homologue from *Bacteroides thetaiotaomicron* shows C-6 to be oriented away from the enzyme active site.[216] Based on this observation, we expected that hOGA may tolerate substituents at C-6. Unfortunately, simple conversion of the C-2 *N*-acyl substituent of 4MUGlcNAc to a thioamide was not an option. Such glycosides are known to be unstable, undergoing cyclization to form thiazolines.[188] Moreover, such substrates are not ideal for studying enzyme kinetics, as enzymatic processing has been shown to produce thiazolines that inhibit enzymes which use the same catalytic mechanism as hOGA.[188] Thus, we set out to synthesize a 4MUGlcNAc analogue in which the C-6 hydroxyl was simply substituted with thioacetamide to give (6NSAc)4MUGlcNAc **10** (Scheme 3.1). We hoped that this simple substitution would minimally perturb the activity of hOGA towards the substrate, while giving an indication of whether our proposed strategy results in significant quenching. Beginning with D-glucosamine hydrochloride **1**, we prepared 4-methylumbelliferyl 3,4,6-tri-*O*-acetyl-2-amino-2-deoxy-β-D-glucopyranoside hydrochloride **3**,[217] from intermediate glycosyl bromide **2**,[218] as described previously. Subsequent neutralization and reaction with di-*tert*-butyl-dicarbonate in the presence of trimethylamine (TEA) gave carbamate **4**. Protection of the C-2 amine as a carbamate, as opposed to immediate *N*-acylation to give GlcNAc, was

required to allow later use of Lawesson's reagent to selectively install the thioamide. Deacetylation of **4** using NaOMe in MeOH afforded **5**, after which the C-6 hydroxyl was selectively tosylated using *p*-TsCl in Py followed by acetylation of the remaining alcohols on addition of Ac_2O to give **6**. Displacement of the tosyl substituent using NaN_3 produced azidosugar **7**, after which hydrogenation using activated Pd/C under H_2 provided an intermediate amine that was directly acetylated using Ac_2O in Py to afford amide **8**. Reflux with Lawesson's reagent in toluene cleanly converted amide **8** to thioamide **9**. Finally, the carbamate group of **9** was removed using trifluoroacetic acid (TFA) to give a crude amine salt that was neutralized and directly acetylated with Ac_2O and Py, followed by concentration and de-*O*-acetylation of the crude product using NaOMe in MeOH to furnish (6NSAc)4MUGlcNAc **10**, which was purified by HPLC.



Scheme 3.1. Synthesis of (6NSAc)4MUGlcNAc (10).
a) AcBr, Ar, 48 hr, 83%; (b) 4MU, $4MU^-Na^+$, acetone, Ar, 24 hr, 41%; (c) Boc_2O, TEA, DCM, Ar, 24 hr, 88%; (d) NaOMe, MeOH, 24 hr, 99%; (e) 1) 3 eq. p-TsCl, Py, DCM, Ar, 1 hr; then Ac_2O, 2 hr, 92%; (f) NaN_3, DMF, Ar, 5 d, 88%; (g) 1) EtOAc, EtOH, Pd/C, H_2, 24 hr; 2) Ac_2O, Py, 4 hr, 69%; (h) Lawesson's reagent, PhMe, Ar, reflux, 30 min, 77%; (i) 1) TFA, DCM, Ar, 3 hr; 2) Ac_2O, Py, 2 hr; 3) NaOMe, MeOH, Ar, 90 min, 63%.

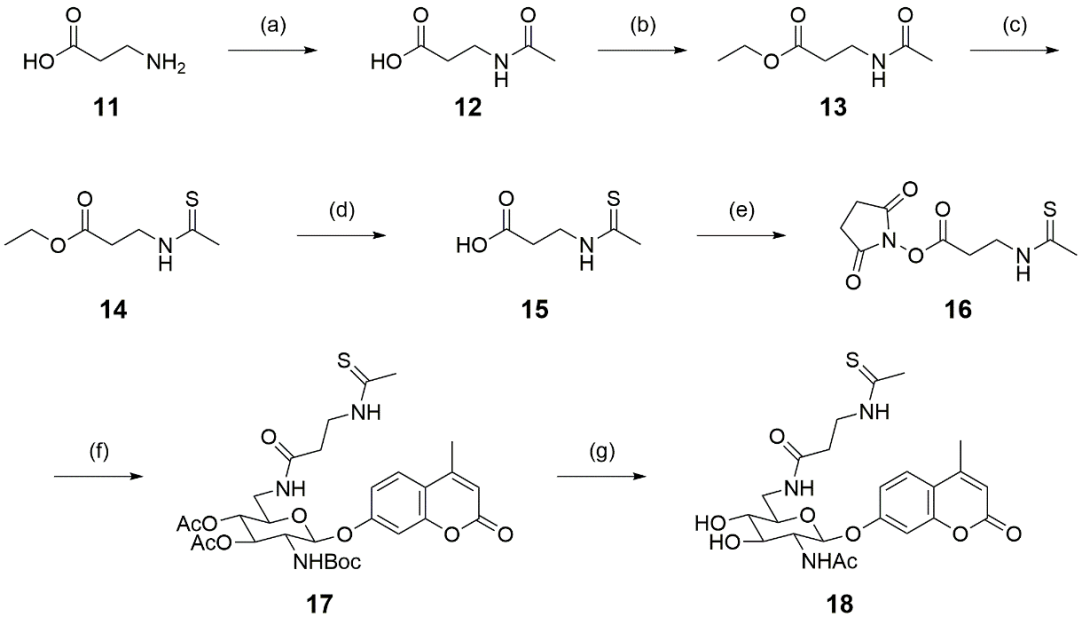
The efficiency of PET, a collisional quenching process, is dependent on the proximity of the fluorophore and quencher, as described by equation (3.1).[69]

$$E_c = \frac{1}{1 + \frac{1}{\tau A e^{-\beta(r-r_c)}}} \quad (3.1)$$

In equation (3.1), the collisional quenching efficiency (E_c) depends on: r, the distance between the centres of the fluorophore and quencher; r_c, their distance at molecular contact; A and β, constants that relate to the distance and time dependence of orbital interactions, respectively; and τ, the lifetime of the fluorescent state.[69] E_c decreases exponentially as the distance between the fluorophore and quencher increases. On this basis, we reasoned that quenching would be dependent on the exact location and orientation of the thioamide. We therefore set out to prepare additional substrates in which the position of the thioamide is altered by inserting a linker between the sugar and the thioamide. We use two linkers having the same length but different structures, one based on β-alanine and coupled to the C-6 position via an amide ((6NAcNSAc)4MUGlcNAc **18**, Scheme 3.2), the other based on propargylamine and bound to C-6 through a 1,2,3-triazole ((6tAzNSAc)4MUGlcNAc **23**, Scheme 3.3). We chose these linkers because in both cases four atoms span the distance between C-6 and the thioamide, and so differences in quenching efficiencies between the two should stem from differences in orientation and intramolecular interactions and not from variation in linker length. Further, because amides, thioamides, and 1,2,3-triazoles are generally accepted as isosteres,[219,220] we anticipated that observed differences in processing of **10**, **18**, and **23** by hOGA would be more easily rationalized in terms of specific structural changes.

The linker-containing compounds were prepared using a convergent approach making use of intermediate azidosugar **7** from the synthesis of (6NSAc)4MUGlcNAc **10**. The linker of (6NAcNSAc)4MUGlcNAc **18** was prepared from β-alanine **11**, which was acetylated to give amide **12**,[221] and then esterified to mask the carboxylic acid as ethyl ester **13** (Scheme 3.2).[222] We found that this protection step was required because thioamide **15**, produced from the reaction of amide **12** with Lawesson's reagent, was inseparable from polar reaction by-products. Thus, amide **13** was selectively converted to thioamide **14** using Lawesson's reagent. Acid-catalyzed ester hydrolysis of **14** at reflux produced carboxylic acid **15**, which was then activated by reaction with *N*-

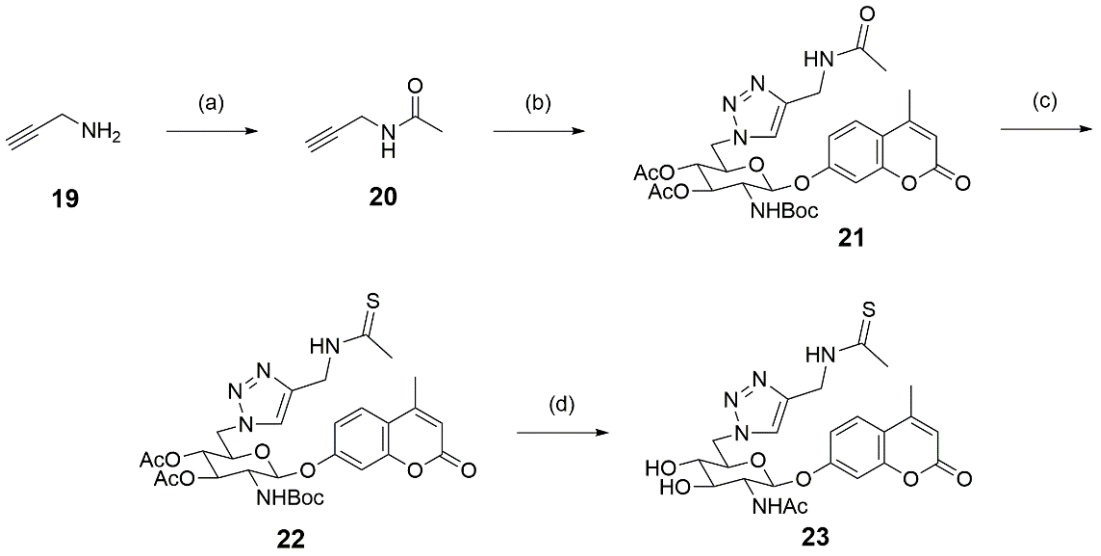
hydroxysuccinimide (HOSu) and 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC·HCl) to yield *N*-hydroxysuccinimidyl ester **16**. Next, the linker and sugar were coupled using the Staudinger-Vilarrasa[223] reaction in which azide **7** was reduced to an amine using PBu_3, and reacted *in situ* with *N*-hydroxysuccinimidyl ester **16** to yield amide **17**. Subsequent removal of the *tert*-butoxycarbonyl group using TFA, acetylation of the resulting amine using Ac_2O and TEA, and de-*O*-acetylation of the crude product using K_2CO_3 in MeOH gave (6NAcNSAc)4MUGlcNAc **18**, which was purified by HPLC.



Scheme 3.2. Synthesis of (6NAcNSAc)4MUGlcNAc (18).
(a) Ac_2O, MeOH, Ar, reflux, 90 min 77%; (b) PTSA, EtOH, reflux, 2 hr, 85%; (c) Lawesson's reagent, dioxane, Ar, 1 hr, 82%; (d) TFA, H_2O, reflux, 24 hr, 85%; (e) HOSu, EDC·HCl, dioxane, Ar, 24 hr, 43%; (f) 7, PBu_3, Ar, 2 hr, 70%; (g) 1) TFA, DCM, Ar, 1 hr; 2) Ac_2O, TEA, 1 hr; 3) K_2CO_3, MeOH, Ar, 2 hr, 88%.

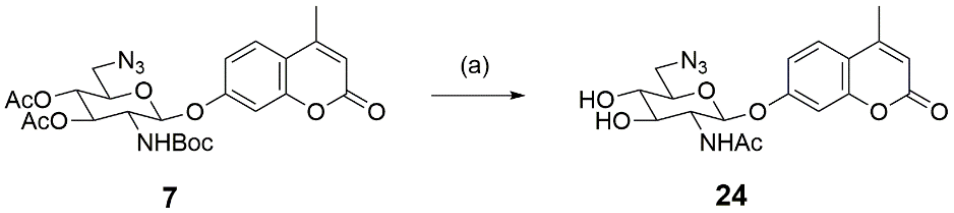
The linker of (6tAzNSAc)4MUGlcNAc **23** was prepared by reacting propargylamine **19** with acetyl chloride in the presence of TEA to give propargyl amide **20** (Scheme 3.3).[224] Multiple attempts failed to convert amide **20** to the corresponding thioamide, and so coupling of the linker to azidosugar **7** was performed before thioamide installation using Cu-assisted azide-alkyne cycloaddition in a mixture of DMF and water to give triazole **21**. Subsequent reaction with Lawesson's reagent in 1,4-dioxane at reflux cleanly afforded thioamide **22**. Removal of the *tert*-butoxycarbonyl group, *N*-acetylation

of the resultant amine, and de-*O*-acetylation furnished (6tAzNSAc)4MUGlcNAc **23**, which was purified by HPLC.



Scheme 3.3. Synthesis of (6tAzNSAc)4MUGlcNAc (23)
(a) AcCl, TEA, DCM, Ar, 4 hr, 91%; (b) 7, $CuSO_4$, $Na^+Ascorbate^-$, DMF, H_2O, 4 hr, 85%; (c) Lawesson's reagent, dioxane, Ar, reflux, 2 hr, 92%; (d) 1) TFA, DCM, Ar, 2 hr; 2) Ac_2O, Py, 2 hr; 3) K_2CO_3, MeOH, Ar, 2 hr, 62%.

To provide insight into the importance of the C-6 hydroxyl group in catalysis by hOGA, we prepared a simple derivative of 4MUGlcNAc bearing an azide in place of the C-6 hydroxyl group. The azido substrate ($6N_3$)4MUGlcNAc **24** was obtained readily from azidosugar **7** by reaction with TFA followed by acetylation of the resultant crude amine and de-*O*-acetylation (Scheme 3.4).



Scheme 3.4. Synthesis of Azidosugar (24).
(a) 1) TFA, DCM, Ar, 1 hr; 2) Ac_2O, TEA, DCM, 1 hr; 3) K_2CO_3, MeOH, Ar, 16 hr, 99%.

3.3.2. Kinetics of Substrate Processing by hOGA

With the substrates in hand, *in vitro* kinetic experiments were performed in parallel alongside commercial 4MUGlcNAc to quantify the effects of structural alterations at C-6. Recombinant hOGA was expressed in *E. coli* and purified by metal chelate affinity chromatography as previously described.[225] Initial rates of hydrolysis of all four substrates by hOGA were measured as a function of substrate concentration using a continuous assay in combination with a standard curve for 4MU. These data were then plotted and fit to the Michaelis-Menten equation (Figure 3.3). The Michaelis constant (K_M, 220 ± 10 μM), turnover number (k_{cat}, 1.94 ± 0.04 min^{-1}, 0.203 ± 0.005 $\mu mol \cdot min^{-1} mg^{-1}$), and second-order rate constant (k_{cat}/K_M, 8.8 ± 0.06 $min^{-1} mM^{-1}$, 0.92 ± 0.06 $\mu mol \cdot min^{-1} mg^{-1} mM^{-1}$) calculated for 4MUGlcNAc were comparable to those previously reported (Table 3.1).[43] Interestingly, the K_M value (72 ± 9 μM) of (6NSAc)4MUGlcNAc **10** was slightly lower than that of 4MUGlcNAc, which suggests that hOGA can accommodate this change with respect to binding, though the k_{cat} value (0.025 ± 0.001 min^{-1}) was significantly lower than that of 4MUGlcNAc, resulting in a k_{cat}/K_M value (0.35 ± 0.06 $min^{-1} mM^{-1}$) 25 times lower than that measured for 4MUGlcNAc (Table 3.1). ($6N_3$)4MUGlcNAc **24** exhibited a slightly lower K_M value (78 ± 8 μM) and a slightly higher k_{cat} value (6.3 ± 0.2 min^{-1}) in comparison to 4MUGlcNAc, resulting in a 9-fold higher k_{cat}/K_M value (80 ± 10 $min^{-1} mM^{-1}$) (Table 3.1). These results give a good indication that the C-6 hydroxyl substituent is not required for catalysis by hOGA. The fact that (6NSAc)4MUGlcNAc **10** shows a decrease in k_{cat} compared to 4MUGlcNAc suggests that that steric bulk at C-6 can negatively impact catalysis by hOGA.

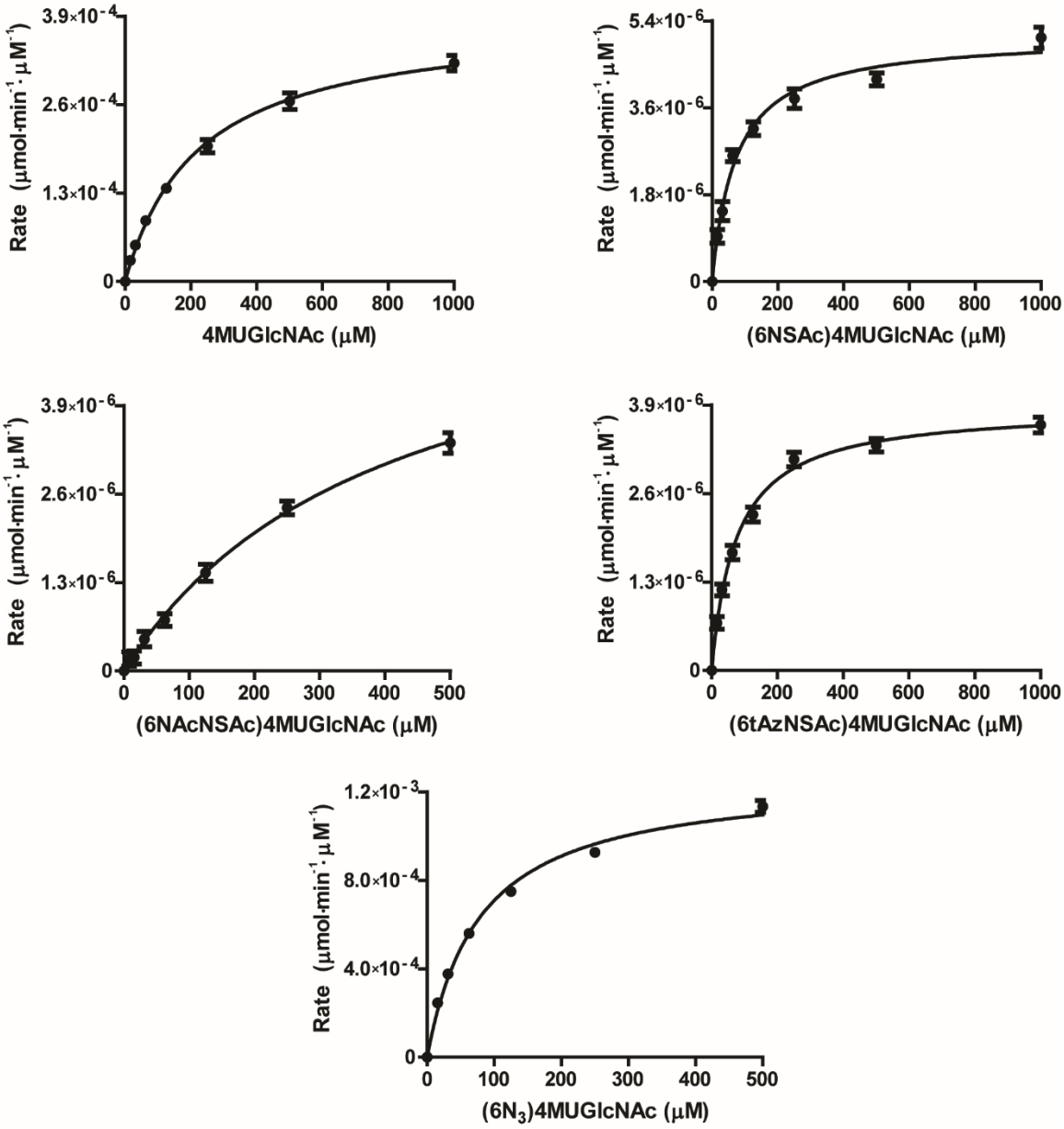


Figure 3.3. Michaelis-Menten fits of the hOGA-catalyzed hydrolysis data for 4MUGlcNAc, and the C-6 derivatives (6NSAc) 10, (6NAcNSAc) 18, (6tAzNSAc) 23, and ($6N_3$) 24.
Rate expressed in units of µmol of 4MU liberated per minute per µM enzyme.

The hOGA-catalyzed processing of (6NAcNSAc)4MUGlcNAc **18** and (6tAzNSAc)4MUGlcNAc **23** were examined in the same fashion and fit to the Michaelis-Menten equation (Figure 3.3). For β-alanine-based **18** the K_M value (400 ± 30 µM) was found to be higher and the k_{cat} value (0.030 ± 0.001 min^{-1}) to be lower than those of

4MUGlcNAc, resulting in the k_{cat}/K_M value (0.075 ± 0.008 $min^{-1}mM^{-1}$) being 117-fold lower (Table 3.1). On the other hand the K_M (76 ± 5 μM) and k_{cat} (0.019 ± 0.001 min^{-1}) values of propargylamine-based **23** both decreased relative to 4MUGlcNAc, resulting in a k_{cat}/K_M value (0.25 ± 0.02 $min^{-1}mM^{-1}$) only 35-fold lower than that of 4MUGlcNAc and on par with that of (6NSAc)4MUGlcNAc **10** (Table 3.1). Due to solubility issues, the highest concentration of β-alanine-based **18** used in the enzyme assay was 500 μM, and as a result saturation was not observed during the kinetic studies (Figure 3.3). To confirm the accuracy of the second-order rate constant obtained for β-alanine-based **18**, we examined processing by hOGA at substrate concentrations at least 10-fold below the determined K_M. At these low substrate concentrations, and since $V_{max} = k_{cat}\cdot[E]$ as described in section 1.3.1, the Michaelis-Menten equation can be simplified to a linear relationship in which the slope represents k_{cat}/K_M (3.2).

$$v = \frac{V_{max}[S]}{K_M + [S]} \quad \rightarrow \quad v = \frac{k_{cat}}{K_M} \cdot [E][S] \qquad (3.2)$$

In this way a k_{cat}/K_M value of 0.075 ± 0.002 $min^{-1}\mu M^{-1}$ was obtained for **18**, a value which is the same as from the Michaelis-Menten fit. Interestingly, despite the differences in the size of the C-6 substituent between (6NSAc)4MUGlcNAc **10** and the linker-containing compounds **18** and **23**, the k_{cat} values are remarkably similar (Table 3.1). This lends strength to the argument that size of the C-6 substituent does not greatly affect catalysis. More significant changes are seen, however, in the K_M values for these substrates. While (6NSAc)4MUGlcNAc **10** and (6tAzNSAc)4MUGlcNAc **23** have similar K_M values, a higher K_M value is observed for (6NAcNSAc)4MUGlcNAc **18** (Table 3.1). This difference suggests that properties other than linker size influence binding by hOGA. The lower K_M value for (6tAzNSAc)4MUGlcNAc **23** as compared to that of (6NAcNSAc)4MUGlcNAc **18** may stem from the rigidity of the triazole, which could limit the number of possible conformations that can be adopted within the hOGA active site.

Table 3.1. Michaelis constants (K_M), turnover numbers (k_{cat}) and second-order rate constants (k_{cat}/K_M) for 4MUGlcNAc and the synthesized substrates with respect to hOGA.

Compound	K_M (μM)	k_{cat} (min^{-1})	k_{cat}/K_M ($min^{-1}mM^{-1}$)	$\frac{(k_{cat}/K_M)_{4MUGlcNAc}}{(k_{cat}/K_M)_{\#}}$
4MUGlcNAc	220 ± 10	1.94 ± 0.04	8.8 ± 0.6	---
(6NSAc) **10**	72 ± 9	0.025 ± 0.001	0.35 ± 0.06	25
(6NAcNSAc) **18**	400 ± 30	0.030 ± 0.001	0.075 ± 0.008	117
(6tAzNSAc) **23**	76 ± 5	0.019 ± 0.001	0.25 ± 0.02	35
($6N_3$) **24**	78 ± 8	6.3 ± 0.2	80 ± 10	0.11

3.3.3. Quenching Measurements

Knowing that the three thioamide-containing substrates **10**, **18**, and **23** are processed by hOGA, we next evaluated and compared their quenching efficiencies. To do so, we first purified commercial 4MUGlcNAc by HPLC to remove any free 4MU that would prevent its use as a standard. Excitation and emission scans were performed with a fluorescence spectrophotometer using quartz cuvettes and substrate concentrations of 10, 50, and 500 μM (Figure 3.4). At each concentration the excitation and emission profiles were scanned iteratively, until the observed excitation maximum was used to scan the emission profile and vice versa. The same cuvette was used for all measurements, and rigorous cleaning with methanol and water, followed by drying with compressed nitrogen, was carried out when changing between concentrations or compounds. Quenching efficiency was determined as the percentage decrease in emission of the thioamide-containing compounds compared to that of 4MUGlcNAc at a given wavelength and concentration (Figure 3.4). We found the quenching efficiency to be independent of the emission wavelength examined, even though the shape of the emission profiles changed at 500 μM relative to lower concentrations, and so at each concentration we obtained an average quenching efficiency from the quenching efficiencies at 350, 380, and 440 nm emission. By this method, we calculated the quenching efficiency for (6NSAc)4MUGlcNAc **10** to be 64%, 63%, and 59% at 10, 50, and 500 μM, respectively, resulting in an average quenching efficiency of 62 ± 3% (Table 3.2). These results contrast with the experiment using thioacetamide in solution, in which a strong dependence of the quenching efficiency on thioamide concentration

was observed. Given that quenching of 4MUGlcNAc by thioacetamide was negligible at concentrations below 1 mM (Figure 3.2), we posit that the observed quenching for **10** is due to intramolecular rather than intermolecular PET.

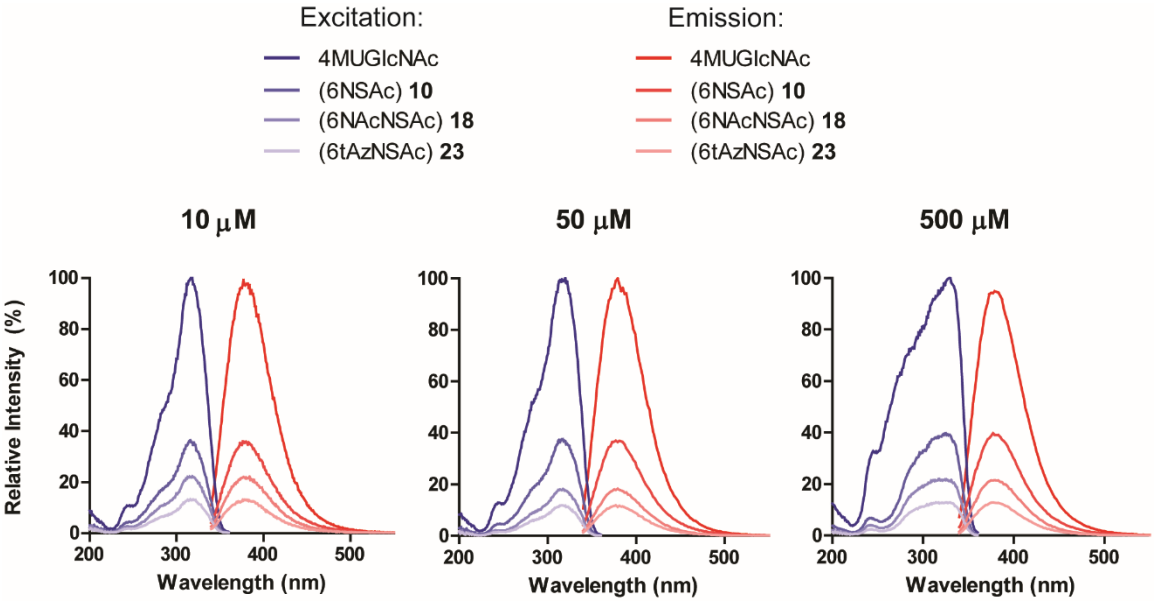


Figure 3.4. Excitation and emission scans for 4MUGlcNAc, and the thioamide-containing C-6 derivatives (6NSAc) 10, (6NAcNSAc) 18, (6tAzNSAc) 23, each at 10, 50, and 500 μM.

In the same way as for (6NSAc)4MUGlcNAc **10**, we determined the quenching efficiencies for (6NAcNSAc)4MUGlcNAc **18** and (6tAzNSAc)4MUGlcNAc **23** with respect to 4MUGlcNAc. For **18** we calculated the quenching efficiencies to be 77%, 82%, and 77% at 10, 50, and 500 μM, respectively, resulting in an average quenching efficiency of 79 ± 3% (Table 3.2). For **23** we calculated the quenching efficiencies to be 87%, 88%, and 86% at the same respective concentrations, resulting in an average quenching efficiency of 87 ± 1% (Table 3.2). The improvement in quenching efficiency for **18** and **23** over (6NSAc)4MUGlcNAc **10** likely stems from the installation of the linker between the sugar and thioamide, which facilitates closer contact between these moieties. Interestingly, the average quenching efficiency of (6tAzNSAc)4MUGlcNAc **23** was greater than that of (6NAcNSAc)4MUGlcNAc **18**. It is unlikely that the difference in quenching efficiencies between these two compounds is a result of linker length. In actuality, based on a purely length-based argument it would be expected that (6tAzNSAc)4MUGlcNAc **23** would have poorer quenching efficiency, since this

unsaturated linker is expected to be slightly shorter than that of (6NAcNSAc)4MUGlcNAc **18**. Therefore, we reason that the greater quenching efficiency observed for (6tAzNSAc)4MUGlcNAc **23** is a result of improved fluorophore-quencher proximity brought about by π-stacking between the triazole of the linker and the 4MU group.

Table 3.2. Quenching efficiencies relative to 4MUGlcNAc of thioamide-containing compounds (6NSAc) 10, (6NAcNSAc) 18, and (6tAzNSAc) 23, at 10, 50, and 500 μM.

Compound	Quenching at 10 μM (%)	Quenching at 50 μM (%)	Quenching at 500 μM (%)	Average Quenching (%)
(6NSAc) **10**	64	63	59	62 ± 3
(6NAcNSAc) **18**	77	82	77	79 ± 3
(6tAzNSAc) **23**	87	88	86	87 ± 1

Finally, we wanted to confirm that the observed quenching efficiencies are not due to a change of shape or shift in position of the excitation or emission profiles for the thioamide-containing compounds relative to that of 4MUGlcNAc. To do so we prepared an overlay of normalized excitation and emission profiles of 50 μM 4MUGlcNAc, (6NSAc)4MUGlcNAc **10**, (6NAcNSAc)4MUGlcNAc **18**, and (6tAzNSAc)4MUGlcNAc **23** (Figure 3.5). The normalized excitation and emission profiles for 1 μM 4MU were included for comparison. From the overlay, it was clear that the excitation and emission profiles of 4MUGlcNAc overlay almost exactly with those of the thioamide-containing compounds. This indicates that the C-6 substituents of the thioamide-containing compounds do not affect the excitation and emission profiles, and so the lower fluorescence emission observed for **10**, **18**, and **23** must be a result of quenching. Moreover, the fact that the excitation and emission profiles of **10**, **18**, and **23** overlap so exactly indicates that the differences in quenching efficiencies between them is also accurate. The small discrepancy in overlap between 4MUGlcNAc and the thioamide-containing compounds in the UV-region of the excitation spectrum does not likely affect the observed quenching, since the excitation maximum was used to scan the emission profiles and here the overlap is precise.

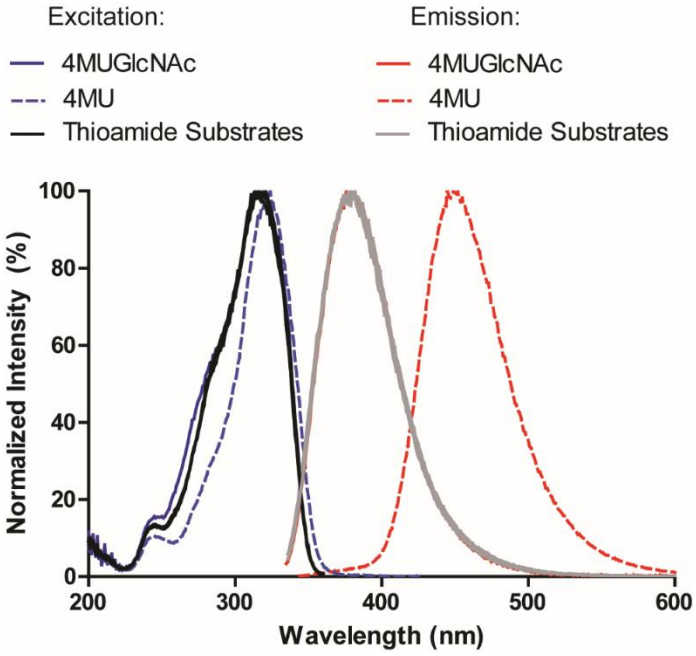


Figure 3.5. Overlay of the normalized excitation and emission scans for 50 μM 4MUGlcNAc, (6NSAc) 10, (6NAcNSAc) 18, (6tAzNSAc) 23, and 1 μM 4MU.

3.4. Conclusions and Future Work

Through chemical synthesis we have prepared four substrates, each bearing a 4MU fluorophore, for the human β-hexosaminidase hOGA. Three of these have thioamide moieties appended to the C-6 position of the sugar. We demonstrate that the fluorescence of the thioamide-containing substrates is quenched relative to 4MUGlcNAc with efficiencies of up to 87%. Further, we show that the quenching efficiency can be tuned through variations in the structure of the linker, which can presumably bring the aromatic fluorophore and thioamide quencher in close contact, a necessity for efficient PET quenching.[69] Further improvement in PET quenching efficiency should be accessible through fine tuning of the length and other physical properties of the linker. Moreover, quenching efficiency may be synergistically increased by the presence of multiple thioamides, as this would result in an increase in effective thioamide concentration. Design of a synergistically quenched substrate should be facile, simply requiring installation of a short thioamide-containing peptide at C-6, such as one based on triglycine.

Examination of the *in vitro* kinetics of these substrates indicated that replacement of the C-6 hydroxyl does not abolish catalysis by hOGA. Indeed, we showed that a C-6 azidosugar has an improved second-order rate constant relative to the standard commercial substrate 4MUGlcNAc. The kinetic results suggest that the C-6 hydroxyl is not essential for catalysis by hOGA, and that large substituents at C-6 are relatively well tolerated. Moreover, these results lend confidence to the use of thioamides as quenchers with minimal steric impact for other enzymes with pocket-shaped active sites.

Despite having blue-shifted fluorescence, substrates based on 4-methylumbelliferone,[226] as well as on other blue-shifted fluorophores,[105] have been used previously in live-cell or even *in vivo* assays. In these cases, the substrates in question were fluorogenic but did not bear quencher moieties to further reduce background. Since substrates **10**, **18**, and **23** are fluorogenic but also have particularly low background fluorescence due to the appended thioamides, they may also find use in live-cell or *in vivo* assays. Nevertheless, we believe that the chief value of these substrates is to serve as a proof-of-concept demonstrating that small thioamides can be used to design PET quenched substrates for glycoside hydrolases, and for hOGA in particular. Furthermore, because the red-shifted flurorophore resorufin is known to be quenched by thioamides[86] and has been successfully used in living systems previously,[193,227] development of PET quenched hOGA substrates using this fluorophore, or similar red-shifted fluorophores, may be a promising next step toward live-cell and *in vivo* monitoring of hOGA.

To create thioamide-containing PET quenched substrates better suited to use in living systems, the issues of selectivity and cell permeability will need to be adressed. It is known that inhibitors can be made selective for hOGA over the other human β-hexosaminidases (HexA and HexB) by modification of the C-2 position.[43,228] Thus, making analogous modifications to the C-2 postion of a series of quenched substrates should also provide an increase in selectivity for hOGA and facilitate their use in live-cell or *in vivo* assays. Given the small size of these thioamide-containing probes we anticipate that they will be able to penetrate cell membranes to interrogate OGA activity in live cells. If, however, these probes are not cell-penetrating, alternative approaches to increasing permeability are possible. Cell-penetrating peptides have been used in the design of live-cell and *in vivo* susbtrates to increase their cell permeability and allow the

substrates to reach the target enzyme.[97,98] Installation of a cell-penetrating peptide on these thioamide-quenched susbtrates should accordingly circumvent such problems. Moreover, if multiple thioamides indeed synergistically improve quenching efficiency as we suspect, then installation a C-6 cell-penetrating peptide whose peptide bonds have been converted to thioamides may act to simultaneously afford hightened quenching efficiency and improved cell permeability, thus providing a valuable substrate for live-cell or *in vivo* monitoring of hOGA.

3.5. Experimental Section

3.5.1. General Procedures for Synthesis

Reagents and solvents were purchased directly from Sigma-Aldrich, or from other reputable chemical vendors through the distributor Cedarlane. Unless otherwise noted, all reagents were used without further purification. Amberlite IR120 H^+ resin was regenerated by rinsing several times each with 1 M HCl, water, and methanol, and was allowed to dry under vacuum prior to use. In order to dry solvents, glassware containing molecular sieves was flame dried under vacuum and flushed with Ar while still hot, after which the solvent was added to the cooled apparatus by syringe and allowed to dry for one hour. All reactions were mounted on magnetic stirring plates and were carried out with stirring using a Teflon-coated stir bars. Reactions were monitored using aluminum backed 0.2 mm silica gel 60 UV_{254} TLC plates. TLC plates were stained with 50 mM $KMnO_4$, 360 mM K_2CO_3, 16 mM NaOH in water ($KMnO_4$ stain); or 85 mM ninhydrin, 500 mM acetic acid in n-butanol (ninhydrin stain); or 40 mM $(NH_4)_6Mo_7O_{24}$, 6 mM $Ce(SO_4)_2$, 1.9 M H_2SO_4, in water (CAM stain). After staining, TLC plates were developed with a heat gun until compound spots appeared. Column chromatography was performed under positive pressure, and using 230-400 mesh, grade 60 silica gel as stationary phase and the specified solvents as mobile phase. HPLC purification was carried out using a Semiprep Eclipse XDB-C18 5 µm particle size column (Agilent) fitted with a 2 mL loop, and all solvents used were HPLC grade. NMR spectroscopy was carried out on either a Bruker Avance III 400 MHz or 500 MHz spectrometer equipped with a 5mm TXI Inverse $^1H/^{13}C/^{19}F$ probe, or on a Bruker Avance II 600 MHz spectrometer equipped with

a 5mm QNP $^{13}C/^{15}N/^{31}P$ cryoprobe. Compounds were dissolved in $CDCl_3$, CD_3OD, DMSO-d6, or acetone-d6 where appropriate, and the chemical shifts were reported with respect to these solvents. LRMS and HRMS of compounds was performed at the Simon Fraser University MS Facility.

3.5.2. Synthesis of Common Intermediate (7)

3,4,6-tri-*O*-acetyl-2-amino-2-deoxy-α-D-glucopyranosyl bromide hydrobromide (2). Prepared as previously described from commercially available D-glucosamine hydrochloride **1** and acetyl bromide,[218] and used without further purification. The ^1H NMR spectrum conformed to the literature description.

4-methylumbelliferyl 3,4,6-tri-*O*-acetyl-2-amino-2-deoxy-β-D-glucopyranoside hydrochloride (3). Prepared as previously described from bromosugar **2** and a mixture of 4-methylumbelliferone and 4-methylumbelliferyl sodium.[217] The compound was purified by recrystallization at -20 °C from Et_2O and MeOH. Recrystallization was repeated twice on remnants of the mother liquor. The ^1H NMR spectrum conformed to the literature description.

4-methylumbelliferyl 3,4,6-tri-*O*-acetyl-2-amino-2-*N*-*tert*-butoxycarbonyl-2-deoxy-β-D-glucopyranoside (4). To an oven dried flask under Ar containing hydrochloride salt **3** (362 mg, 0.72 mmol, 1.0 eq) was added dried DCM (10 mL). To the mixture was added TEA (200 µL, 1.44 mmol, 2.0 eq), which caused the starting material to dissolve, after which Boc_2O was added (629 mg, 2.88 mmol, 4.0 eq) and the reaction mixture was protected from light and stirred under Ar at room temperature. After 16 hr the reaction was judged complete by TLC analysis and the solvent was removed *in vacuo*. The material was purified by silica-gel flash column chromatography (9:1; DCM:EtOAc, 0. % TEA) to yield the desired compound as a white solid (360 mg, 0.64 mmol, 88%).

^1H NMR (400 MHz, acetone-d6) δ: 7.72 (d, J = 8.6 Hz, 1H, AR), 7.03 (m, J = 8.6, 2.4 Hz, 2H, AR), 6.36 (d, J = 9.3 Hz, 1H, NH), 6.19 (d, J = 1.2 Hz, 1H, 4MU H-3), 5.60 (d, J = 8.4 Hz, 1H, H-1), 5.39 (dd, J = 10.4, 9.4 Hz, 1H, H-3), 5.08 (t, J = 9.6 Hz, 1H, H-4), 4.33 – 4.25 (m, 1H, H-6), 4.21 (m, 2H, H-6, H-5), 3.95 (dt, J = 10.4, 8.6 Hz, 1H, H-2),

2.45 (d, J = 1.2 Hz, 3H, 4MU CH_3), 2.04 (s, 3H, OAc CH_3), 2.02 (s, 3H, OAc CH_3), 2.00 (s, 3H, OAc CH_3), 1.39 (s, 9H, Boc $(CH_3)_3$); ^{13}C NMR (126 MHz, acetone-d6) δ: 169.80, 169.53, 169.20, 159.95, 159.78, 155.01, 152.64, 126.27, 114.99, 113.53, 112.41, 103.44, 98.48, 78.53, 72.65, 71.95, 68.75, 61.92, 55.35, 27.60 (3C, Boc $(CH_3)_3$), 19.76 (3C, OAc CH_3), 17.64; HRMS: m/z calcd for $C_{27}H_{33}NO_{12}Na^+$ $[M+Na]^+$: 586.1901, found: 586.1893.

4-methylumbelliferyl 2-amino-2-*N*-*tert*-butoxycarbonyl-2-deoxy-β-D-glucopy-ranoside (5). To an oven dried flask under Ar containing acetylated sugar **4** (250 mg, 0.44 mmol, 1.0 eq) was added dried MeOH (5 mL). Upon dissolution of the sugar NaOMe was added (24 mg, 0.44 mmol, 1.0 eq), and the reaction mixture was protected from light and stirred under Ar at room temperature. After 16 hr the reaction was judged complete by TLC analysis (9:1; EtOAc:MeOH, 0.5% TEA). Regenerated Amberlite IR120 H^+ resin was added until the reaction mixture became slightly acidic (pH ~ 5) as determined using pH paper. The resin was then removed by filtration and washed several times with MeOH, after which the combined filtrates were evaporated *in vacuo* to yield the pure product as a white solid (192 mg, 0.44 mmol, 99%).

^1H NMR (500 MHz, CD_3OD) δ: 7.70 (d, J = 8.8 Hz, 1H, AR), 7.04 (dd, J = 8.8, 2.4 Hz, 1H, AR), 7.00 (d, J = 2.3 Hz, 1H, AR), 6.20 (d, J = 0.9 Hz, 1H, 4MU H-3), 5.09 (d, J = 8.3 Hz, 1H, H-1), 3.93 (dd, J = 12.1, 2.0 Hz, 1H, H-6), 3.72 (dd, J = 12.1, 5.8 Hz, 1H, H-6), 3.66 – 3.60 (dd, J = 10.0 Hz, 8.5 Hz, 1H, H-2), 3.54 – 3.46 (m, 2H, H-3, H-5), 3.42 (t, J = 9.21 Hz, 1H, H-4), 2.45 (d, J = 0.9 Hz, 3H, 4MU CH_3), 1.45 (s, 9H, Boc $(CH_3)_3$); ^{13}C NMR (126 MHz, CD_3OD) δ: 161.86, 160.69, 157.09, 154.66, 154.04, 125.93, 114.71, 113.63, 111.51, 103.43, 99.74, 77.00, 74.25, 70.48, 61.13, 27.36 (3C, Boc $(CH_3)_3$), 17.23; HRMS: m/z calcd for $C_{21}H_{37}NO_9Na^+$ $[M+Na]^+$: 460.1578, found: 460.1576.

4-methylumbelliferyl 3,4-di-*O*-acetyl-2-amino-2-*N*-*tert*-butoxycarbonyl-2-deo-xy-6-*O*-(p-toluenesulfonyl)-β-D-glucopyranoside (6). To a flame-dried flask under Ar containing deacetylated sugar **5** (1.63 g, 3.73 mmol, 1.0 eq) was added dried DCM (16 mL) and commercial anhydrous Py (7 mL), after which the solution was cooled to 0 °C in an ice water bath. Once cooled, freshly recrystallized (from Et_2O at -20 °C, dried under vacuum) *p*-TsCl (2.49 g, 13.06 mmol, 3.5 eq) was added and the reaction

mixture was protected from light and stirred under Ar at 0 °C for one hour, and then allowed to warm to room temperature. After 1 hr at room temperature the reaction was judged complete by TLC analysis (94:6; EtOAc:MeOH). Ac_2O (9 mL) was added and the reaction mixture was protected from light and stirred under Ar at room temperature. After 16 hr the reaction was judged complete by TLC analysis. Solvent was removed *in vacuo*, and the resulting crude product was dissolved in DCM and washed successively with water, twice with saturated $NaHCO_3$ solution, a second time with water, twice with 1 M HCl, and twice with saturated NaCl solution. The organic extracts were dried over anhydrous Na_2SO_4, filtered, and concentrated *in vacuo*. The material was purified by silica-gel flash column chromatography (6:4; hexanes:EtOAc) to yield the desired compound as a white solid (2.33 g, 3.45 mmol, 92%).

^{1}H NMR (500 MHz, acetone-d6) δ: 7.75 (d, J = 8.3 Hz, 2H, OTs *o*-H), 7.71 (d, J = 8.8 Hz, 1H, 4MU AR), 7.32 (d, J = 8.1 Hz, 2H, OTs *m*-H), 6.97 (dd, J = 8.7, 2.2 Hz, 1H, 4MU AR), 6.92 (d, J = 2.0 Hz, 1H, 4MU AR), 6.32 (d, J = 9.2 Hz, 1H, NH), 6.21 (d, J = 1.0 Hz, 1H, 4MU H-3), 5.54 (d, J = 8.4 Hz, 1H, H-1), 5.34 (t, J = 9.9 Hz, 1H, H-3), 5.05 (t, J = 9.6 Hz, 1H, H-4), 4.36 – 4.30 (m, 1H, H-6), 4.28 – 4.20 (m, 2H, H-5, H-6), 3.89 (dt, J = 8.1, 9.9 Hz, 1H, H-2), 2.47 (d, J = 1.0 Hz, 3H, 4MU CH_3), 2.34 (s, 3H, OTs CH_3), 1.98 (s, 3H, OAc CH_3), 1.97 (s, 3H, OAc CH_3), 1.37 (s, 9H, Boc $(CH_3)_3$); ^{13}C NMR (126 MHz, acetone-d6) δ: 169.56, 169.00, 159.82, 155.31, 154.95, 152.66, 145.05, 132.79, 129.81 (2C, OTs *o*-C), 127.92 (2C, OTs *m*-C), 126.27, 115.04, 113.45, 112.45, 103.36, 98.27, 78.57, 72.49, 72.46, 71.36, 68.31, 67.78, 55.11, 27.59 (3C, Boc $(CH_3)_3$), 20.58, 19.76, 19.73, 17.68; HRMS: m/z calcd for $C_{32}H_{38}NO_{13}S^+$ $[M+H]^+$: 676.2064, found: 676.2057.

4-methylumbelliferyl 3,4-di-*O*-acetyl-2-amino-6-azido-2-*N*-*tert*-butoxycarbon-yl-2,6-di-deoxy-β-D-glucopyranoside (7). To an oven dried flask under Ar containing tosylated sugar **6** (1.30 g, 1.92 mmol, 1.0 eq) was added dried DMF (20 mL) and NaN_3 (880 mg, 13.47 mmol, 7.0 eq). The reaction mixture was protected from light and stirred under Ar at 60 °C. Over 5 days increasing amounts of precipitate formed, and the reaction mixture blackened. No information could be gained by TLC analysis as the R_f of the starting material and product were identical in a variety of solvent systems. Solvent was removed *in vacuo* by co-evaporation with toluene, and the resulting crude product was dissolved in EtOAc and washed five times with water and

once with saturated NaCl solution. Extensive washing with water was necessary to remove all traces of NaN_3, which can form explosive $CH_2(N_3)_2$ in the presence of DCM.[229] The organic extracts were dried over anhydrous Na_2SO_4, filtered, and concentrated *in vacuo*. The material was loaded onto a silica-gel column using DCM, and purified by flash chromatography (6:4; hexanes:EtOAc) to yield the desired compound as a pale yellow solid (930 mg, 1.70 mmol, 88%).

^{1}H NMR (400 MHz, acetone-d6) δ: 7.74 – 7.71 (m, 1H, AR), 7.06 – 7.02 (m, 2H, AR), 6.36 (d, J = 9.1 Hz, 1H, NH), 6.19 (d, J = 1.2 Hz, 1H, 4MU H-3), 5.63 (d, J = 8.4 Hz, 1H, H-1), 5.37 (dd, J = 10.5, 9.4 Hz, 1H, H-3), 5.06 (t, J = 9.7 Hz, 1H, H-4), 4.18 (ddd, J = 9.1, 6.3, 2.9 Hz, 1H, H-5), 3.96 (dt, J = 10.1, 8.5 Hz, 1H, H-2), 3.58 (dd, J = 13.5, 2.9 Hz, 1H, H-6), 3.51 (dd, J = 13.5, 6.4 Hz, 1H, H-6), 2.45 (d, J = 1.2 Hz, 3H, 4MU CH_3), 2.03 (s, 3H, OAc CH_3), 2.00 (s, 3H, OAc CH_3), 1.39 (s, 9H, Boc $(CH_3)_3$); ^{13}C NMR (126 MHz, acetone-d6) δ: 170.44, 170.08, 160.78, 160.70, 156.24, 155.89, 153.54, 127.21, 115.96, 114.39, 113.34, 104.38, 99.32, 79.45, 74.05, 73.38, 70.45, 56.16, 51.81, 28.49 (3C, Boc $(CH_3)_3$), 20.66 (2C, OAc CH_3) 18.54; HRMS: m/z calcd for $C_{25}H_{30}N_4O_{10}Na^+$ $[M+Na]^+$: 569.1860, found: 569.1866.

3.5.3. Synthesis of (6NSAc)4MUGlcNAc (10)

4-methylumbelliferyl 6-acetamido-3,4-di-*O*-acetyl-2-amino-2-*N*-*tert*-butoxycarbonyl-2,6-di-deoxy-β-D-glucopyranoside (8). To a flask containing azidosugar **7** (100 mg, 0.18 mmol, 5.0 eq) was added EtOH (5 mL), and EtOAc (1 mL). The reaction mixture was sealed and Ar was bubbled into the solution 5 times for 30 second intervals, with intermittent introduction of a vacuum to evacuate the flask. Subsequently, 10% Pd/C (4 mg, 0.036 mmol, 1.0 eq) was added and the bubbling/evacuation procedure was repeated with Ar, then with H_2. The reaction mixture was protected from light and stirred under H_2 at room temperature. After 16 hr the reaction was judged complete by TLC analysis (8:2; DCM:EtOAc, 0.5% TEA). The reaction mixture was filtered over Celite, which was washed several times with EtOAc, and the filtrate was concentrated *in vacuo*. The resulting crude amine was dissolved in Ac_2O (3 mL) and Py (1.5 mL), protected from light, and stirred at room temperature. After 4 hr the reaction was judged complete by TLC analysis, and the solvent was removed *in vacuo*. The material was purified by silica-

gel flash column chromatography (8:2; DCM:Acetone) to yield the desired compound as a white solid (71 mg, 0.126 mmol, 69%).

^{1}H NMR (500 MHz, acetone-d6) δ: 7.72 (d, J = 8.5 Hz, 1H, AR), 7.29 (t, J = 5.1 Hz, 1H, NHAc), 7.00 (m, 2H, AR), 6.30 (d, J = 9.4 Hz, 1H, Boc NH), 6.19 (d, J = 0.8 Hz, 1H, 4MU H-3), 5.49 (d, J = 8.5 Hz, 1H, H-1), 5.31 (t, J = 9.9 Hz, 1H, H-3), 4.96 (t, J = 9.7 Hz, 1H, H-4), 4.08 – 3.99 (ddt, J = 9.6, 4.7 Hz, 1H, H-5), 3.94 (dt, J = 10.2, 8.6 Hz, 1H, H-2), 3.47 (m, 2H, H-6), 2.45 (d, J = 0.8 Hz, 3H, 4MU CH_3), 2.02 (s, 3H, O/NAc CH_3), 1.98 (s, 3H, O/NAc CH_3), 1.93 (s, 3H, O/NAc CH_3), 1.39 (s, 9H, Boc $(CH_3)_3$); ^{13}C NMR (126 MHz, acetone-d6) δ: 170.44, 170.29, 170.07, 160.65, 155.91, 153.52, 127.16, 115.82, 114.32, 113.29, 104.31, 99.53, 73.72, 73.60, 70.49, 56.19, 40.17, 28.49 (3C, Boc $(CH_3)_3$), 22.86, 20.75, 20.67, 18.53; HRMS: m/z calcd for $C_{27}H_{35}N_2O_{11}^+$ $[M+H]^+$: 563.2241, found: 563.2235.

4-methylumbelliferyl 3,4-di-*O*-acetyl-2-amino-2-*N*-*tert*-butoxycarbonyl-2,6-di-deoxy-6-thioacetamido-β-D-glucopyranoside (9). To an oven dried flask containing amide **8** (61 mg, 0.11 mmol, 1.5 eq) under Ar was added commercial anhydrous toluene (3 mL) and Lawesson's reagent (29 mg, 0.072 mmol, 1.0 eq). The reaction mixture was protected from light and refluxed under Ar. After 30 min the reaction was judged complete by TLC analysis. The reaction mixture was concentrated *in vacuo,* and the material was purified by silica-gel flash column chromatography (9:1; DCM:Acetone). Three rounds of chromatography were required to remove all the impurities that tended to co-elute and yield the desired compound as a pale yellow solid (48 mg, 0.083 mmol, 77%).

^{1}H NMR (500 MHz, acetone-d6) δ: 9.39 (s, 1H,NHSAc), 7.72 (d, J = 8.7 Hz, 1H, AR), 7.03 – 6.93 (m, 2H, AR), 6.30 (d, J = 9.4 Hz, 1H, Boc NH), 6.19 (d, J = 0.9 Hz, 1H, 4MU H-3), 5.52 (d, J = 8.4 Hz, 1H, H-1), 5.34 (t, J = 9.6 Hz, 1H, H-3), 5.01 (t, J = 9.7 Hz, 1H, H-4), 4.30 (ddd, J = 10.1, 7.5, 2.5 Hz, 1H, H-5), 4.11 – 4.03 (m, 1H, H-6), 3.97 (dt, J = 10.2, 8.5 Hz, 1H, H-2), 3.84 – 3.75 (m, 1H, H-6), 2.55 (s, 3H, NSAc CH_3), 2.45 (d, J = 0.9 Hz, 3H 4MU CH_3), 2.04 (s, 3H, OAc CH_3), 1.99 (s, 3H, OAc CH_3), 1.39 (s, 9H, Boc $(CH_3)_3$); ^{13}C NMR (126 MHz, acetone-d6) δ 202.74, 170.42, 170.17, 160.77, 160.65, 156.27, 155.88, 153.52, 127.18, 115.85, 114.27, 113.31, 104.30, 99.41, 79.44, 73.47,

72.30, 70.74, 56.10, 47.00, 33.60, 33.52, 28.48 (3C Boc $(CH_3)_3$), 20.77, 20.65, 18.53; HRMS: m/z calcd for $C_{27}H_{35}N_2O_{10}S^+$ $[M+H]^+$: 579.2007, found: 579.2002.

4-methylumbelliferyl 2-acetamido-2,6-di-deoxy-6-thioacetamido-β-D-glucopyranoside ((6NSAc)4MUGlcNAc 10). To a flame dried flask containing thioamide **9** (30 mg, 0.052 mmol, 1.0 eq) under Ar was added dried DCM (2 mL), and the reaction mixture was cooled to 0 °C. TFA (0.3 mL) was added by syringe, the reaction mixture was protected from light and stirred under Ar while warming the mixture to room temperature. After 3 hr the reaction was judged complete by TLC analysis (85:15; DCM:Acetone, 0.5% TEA), and the solvent was removed *in vacuo*. To the resulting crude amine was added Ac_2O (1 mL) and Py (1 mL), and the reaction mixture was protected from light and stirred at room temperature. After 2 hr the reaction was judged complete by TLC analysis (75:25; DCM:Acetone), and the solvent was removed *in vacuo*. The resulting crude amide was dissolved in EtOAc and washed successively twice with water, twice with 1 M HCl, once with water, twice with saturated $NaHCO_3$ solution, and twice with saturated NaCl solution. The organic extracts were dried over anhydrous Na_2SO_4, filtered, and concentrated *in vacuo*. The crude amide was allowed to dry under high vacuum, after which the flask was flushed with Ar and commercial anhydrous MeOH (2 mL) was added. After dissolution, NaOMe (2.8 mg, 0.052 mmol, 1.0 eq) was added and the reaction mixture was protected from light and stirred under Ar at room temperature. After 2 hr the reaction was judged complete by TLC analysis. Regenerated Amberlite IR120 H^+ resin was added until the reaction mixture became slightly acidic (pH ~ 5) as determined using pH paper. The resin was then removed by filtration and washed several times with MeOH, after which the combined filtrates were evaporated *in vacuo*. For silica-gel column chromatography, the compound was loaded onto the column by dissolving the crude material in MeOH, concentrating *in vacuo* in the presence of silica, and loading the resulting silica slurry with minimum portions of DCM. Purification by silica-gel flash column chromatography (9:1; EtOAc:MeOH) gave the desired compound as a white solid (14 mg, 0.032 mmol, 63%). Afterwards, the compound was further purified by HPLC using the following conditions: 100 µL injections, 30 mg/mL in DMSO, 2 mL/min flow, 10 to 80%B over 30 min, A = H_2O + 5% MeOH, B = MeOH.

^{1}H NMR (600 MHz, CD_3OD) δ: 7.69 (d, J = 8.8 Hz, 1H, AR), 6.99 (dd, J = 8.8, 2.4 Hz, 1H, AR), 6.96 (d, J = 2.2 Hz, 1H, AR), 6.21 (d, J = 1.0 Hz, 1H, 4MU H-3), 5.15 (d, J = 8.4 Hz, 1H, H-1), 4.21 (dd, J = 13.9, 2.6 Hz, 1H, H-6), 3.97 (dd, J = 10.2, 8.6 Hz, 1H, H-2), 3.90 – 3.85 (ddd, J = 10.0, 7.5, 2.8 Hz, 1H, H-5), 3.79 (dd, J = 13.8, 7.5 Hz, 1H, H-6), 3.61 (dd, J = 10.3, 8.9 Hz, 1H, H-3), 3.33 (m, J = 6.0 Hz, 1H, H-4), 2.53 (s, 3H, NSAc CH_3), 2.45 (d, J = 1.1 Hz, 3H, 4MU CH_3), 1.99 (s, 3H, NAc CH_3); ^{13}C NMR (151 MHz, CD_3OD) δ: 203.29, 173.93, 163.07, 161.64, 156.08, 155.33, 127.31, 116.17, 115.07, 113.04, 104.60, 100.04, 75.14, 74.85, 73.56, 57.11, 48.08, 33.22, 22.95, 18.62; HRMS: m/z calcd for $C_{20}H_{25}N_2O_7S^+$ $[M+H]^+$: 437.1377, found: 437.1371.

3.5.4. Synthesis of (NAcNSAc)4MUGlcNAc (18)

***N*-acetyl-3-aminopropanoic acid (12)**. Prepared as previously described from commercially available β-alanine **11** and acetic anhydride and used without further purification.[221] The ^{1}H NMR spectrum conformed to the literature description.

Ethyl *N*-acetyl-3-aminopropanoate (13). Prepared as previously described from carboxylic acid **12** and EtOH with catalytic PTSA.[222] The product was isolated by extraction from water using DCM, and was used without further purification. The ^{1}H NMR spectrum conformed to the literature description.

Ethyl *N*-thioacetyl-3-aminopropanoate (14). To an oven dried flask containing amide **13** (406 mg 2.55 mmol, 1.5 eq) under Ar was added dried dioxane (25 mL) and Lawesson's reagent (687 mg, 0.170 mmol, 1.0 eq). The reaction mixture was stirred under Ar at room temperature. After 1 hr the reaction was judged complete by TLC analysis. The mixture was concentrated *in vacuo*, and the material was purified by silica-gel flash column chromatography using a gradient solvent system (30:70 – 55:45; Et_2O:hexanes). Three rounds of column chromatography were required to remove all of the co-eluting impurities and yield the desired compound as a pale yellow solid (368 mg, 2.10 mmol, 82%).

^{1}H NMR (500 MHz, CDCl3) δ: 8.11 (s, 1H, NH), 4.09 (q, J = 7.2 Hz, 2H, Et CH_2), 3.88 (dt, J = 6.3, 5.6 Hz, 2H, N CH_2), 2.64 (t, J = 6.0 Hz, 2H, CO CH_2), 2.47 (s, 3H, NSAc CH_3), 1.21 (t, J = 7.2 Hz, 3H, Et CH_3); ^{13}C NMR (126 MHz, CDCl3) δ 201.09,

172.60, 60.97, 41.26, 34.11, 32.12, 14.11; HRMS: m/z calcd for $C_7H_{14}N_1O_2S^+$ $[M+H]^+$: 176.0740, found: 176.0740.

***N*-thioacetyl-3-aminopropanoic acid (15)**. To a flask containing ester **14** (238 mg, 1.36 mmol, 1.0 eq) was added water (3 mL) and TFA (0.2 mL, 2.72 mmol, 2.0 eq), and the reaction mixture was stirred at 90 °C. After 16 hr the reaction was judged complete by TLC analysis (93:7; DCM:MeOH, 0.5% TEA). The mixture was concentrated by co-evaporation with toluene, and the resulting crude product was dissolved in a minimum amount of EtOAc, precipitated with hexanes, filtered, and dried *in vacuo* to give the product with 85% purity as a white solid (170 mg, 1.16 mmol, 85%). Amide 17 was produced as a by-product (15%), however, the impure title compound was used without further purification.

^{1}H NMR (500 MHz, CD_3OD) δ: 3.82 (t, J = 6.8 Hz, 2H, N CH_2), 2.70 (t, J = 6.8 Hz, 2H, O CH_2), 2.47 (s, 3H, CH_3); HRMS: m/z calcd for $C_5H_{10}N_1O_2S^+$ $[M+H]^+$: 148.0427, found: 148.0427.

***N*-hydroxysuccinimidyl *N*-thioacetyl-3-aminopropanoate (16)**. To a flame dried flask containing 85% pure carboxylic acid **15** (178 mg, 1.21 mmol, 1.0 eq) under Ar was added dried dioxane (12 mL), HOSu (167 mg, 1.45 mmol, 1.2 eq), and EDC·HCl (278 mg, 1.45 mmol, 1.2 eq). The reaction mixture was stirred under Ar, and after 5 hr the reaction was judged complete by TLC analysis. The reaction mixture was filtered to remove the urea precipitate, which was washed with dioxane. The combined filtrates were concentrated *in vacuo* and purified by silica-gel flash column chromatography (7:3; DCM:EtOAc) to give the product as a pale beige solid (127 mg, 0.52 mmol, 43%).

^{1}H NMR (500 MHz, $CDCl_3$) δ: 8.02 (s, 1H, NH), 4.12 (dt, J = 6.4, 5.7 Hz, 2H, N CH_2), 3.01 (dd, J = 6.6, 5.2 Hz, 2H, CO CH_2), 2.88 (s, 4H, OSu $2CH_2$), 2.55 (s, 3H NSAc CH_3); HRMS: m/z calcd for $C_9H_{13}N_2O_4S^+$ $[M+H]^+$: 245.0591, found: 245.0594.

4-methylumbelliferyl 3,4-di-*O*-acetyl-2-amino-2-*N*-*tert*-butoxycarbonyl-2,6-di-deoxy-6-(3-(thioacetamido)propanamido)-β-D-glucopyranoside (17). To a flame dried flask under Ar containing azide **7** (53 mg, 0.096 mmol, 1.0 eq) and ester **16** (70, 0.29 mmol, 3.0 eq) was added dioxane (4 mL) and PBu_3 (58 mg, 0.29 mmol, 3.0 eq).

The reaction mixture was protected from light and stirred under Ar at room temperature. After 90 min the reaction was judged complete by TLC analysis, and was concentrated *in vacuo*. Water was added to the resulting crude product to precipitate PBu_3O, and the mixture was washed twice with DCM. The combined organic extracts were washed successively twice with saturated $NaHCO_3$ solution, once with water, and once with saturated NaCl solution. The organic extracts were then dried over anhydrous Na_2SO_4, filtered, and concentrated *in vacuo*. The material was purified by silica-gel flash column chromatography (7:3 EtOAc:DCM) to yield the desired compound as a white solid (44 mg, 0.068 mmol, 70%).

^{1}H NMR (500 MHz, acetone-d6) δ: 9.03 (s, 1H, NHSAc), 7.75 (d, J = 9.4 Hz, 1H, AR), 7.51 (t, J = 5.0 Hz, 1H, C-6 NH), 7.03 – 6.98 (m, 2H, AR), 6.31 (d, J = 9.4 Hz, 1H, Boc NH), 6.22 (d, J = 1.0 Hz, 1H, 4MU H-3), 5.51 (d, J = 8.4 Hz, 1H, H-1), 5.32 (t, J = 9.9 Hz, 1H, H-3), 4.95 (t, J = 9.7 Hz, 1H, H-4), 4.08 (ddd, J = 9.8, 6.1, 2.4 Hz, 1H, H-5), 3.98 – 3.82 (m, 3H, NSAc CH_2, H-2), 3.64 – 3.56 (m, 1H, H-6), 3.50 – 3.43 (m, 1H, H-6), 2.68 (dt, J = 14.6, 7.2 Hz, 1H, N(CO) CH_2), 2.59 (dt, J = 15.2, 6.1 Hz, 1H, N(CO) CH_2), 2.47 (s, 3H, NSAc CH_3), 2.47 (d, J = 1.0 Hz, 3H, 4MU CH_3), 2.05 (s, 3H, OAc CH_3), 1.97 (s, 3H, OAc CH_3), 1.39 (s, 9H, Boc $(CH_3)_3$); ^{13}C NMR (126 MHz, acetone-d6) δ 200.98, 171.67, 170.44, 161.09, 160.87, 155.84, 153.91, 127.28, 115.87, 114.44, 113.21, 104.29, 99.49, 79.42, 73.68, 73.54, 70.33, 56.19, 42.63, 42.48, 39.83, 34.07, 33.70, 33.62, 28.49 (3C Boc $(CH_3)_3$), 20.92, 20.67, 18.62; HRMS: m/z calcd for $C_{30}H_{39}N_3O_{11}S^+$ $[M+H]^+$: 672.2203, found: 672.2209.

4-methylumbelliferyl 2-acetamido-2,6-di-deoxy-6-(3-(thioacetamido)propan-amido)-β-D-glucopyranoside ((6NAcNSAc)4MUGlcNAc 18). To a flame dried flask containing **17** (7.5 mg, 0.012 mmol, 1.0 eq) and dried DCM (1.5 mL) under Ar was added TFA (0.5 mL) by syringe, and the reaction mixture was protected from light and stirred under Ar at room temperature. After 45 min the reaction was judged complete by TLC analysis (93:7; DCM:MeOH), and the solvent was removed *in vacuo*. To the resulting crude amine was added Ac_2O (1.0 mL) and TEA (1.0 mL), and the reaction mixture was protected from light and stirred at room temperature. After 1 hr the reaction was judged complete by TLC analysis (93:7; DCM:MeOH), the solvent was removed *in vacuo* and the resulting crude amide was dried under high vacuum. To the crude amide

under Ar was added commercial anhydrous MeOH (2.0 mL) followed by K_2CO_3 (0.8 mg, 0.006 mmol, 0.5 eq). The reaction mixture was protected from light and stirred under Ar at room temperature. After 16 hr the reaction was judged complete by TLC analysis. For silica-gel column chromatography, the compound was loaded onto the column by dissolving the crude material in MeOH, concentrating *in vacuo* in the presence of silica, and loading the resulting silica slurry with minimum portions of DCM. Purification by silica-gel flash column chromatography using a gradient solvent system (88:12 – 85:15; DCM:MeOH) gave the desired compound as a white solid (5.4 mg, 0.011 mmol, 88%). Afterwards, the compound was further purified by HPLC using the following conditions: 100 µL injections, 20 mg/mL in DMSO, 2 mL/min flow, 15 to 80% B over 30 min, A = H_2O + 5% MeOH, B = MeOH.

^{1}H NMR (600 MHz, DMSO) δ: 9.93 (t, J = 4.6 Hz, 1H, NHSAc), 8.03 (t, J = 5.5 Hz, 1H, C-6 NH), 7.83 (d, J = 9.0 Hz, 1H, NHAc), 7.69 (d, J = 8.7 Hz, 1H, AR), 6.96 (d, J = 2.4 Hz, 1H, AR), 6.93 (dd, J = 8.8, 2.4 Hz, 1H, AR), 6.26 (d, J = 1.0 Hz, 1H, 4MU H-3), 5.24 (d, J = 5.2 Hz, 1H, OH-4), 5.15 (d, J = 5.4 Hz, 1H, OH-3), 5.10 (d, J = 8.5 Hz, 1H, H-1), 3.71 (dt, J = 9.4, 8.4 Hz, 1H, H-2), 3.67 – 3.60 (m, 2H, NSAc CH_2), 3.55 – 3.48 (m, 2H, H-5, H-6), 3.44 (ddd, J = 10.2, 8.6, 5.5 Hz, 1H, H-3), 3.19 (ddd, J = 13.1, 8.2, 6.1 Hz, 1H, H-6), 3.09 (ddd, J = 9.8, 8.3, 5.2 Hz, 1H, H-4), 2.49 – 2.45 (m, 2H, N(CO) CH_2), 2.40 (d, J = 1.0 Hz, 3H, 4MU CH_3), 2.33 (s, 3H, NSAc CH_3), 1.80 (s, 3H, NAc CH_3); ^{13}C NMR (151 MHz, DMSO) δ: 199.06, 170.47, 169.28, 160.03, 159.90, 154.37, 153.17, 126.47, 114.33, 113.25, 111.89, 103.25, 98.60, 74.57, 73.42, 71.81, 55.35, 41.73, 40.05, 32.97, 32.74, 23.03, 18.15; HRMS: m/z calcd for $C_{23}H_{30}N_3O_8S^+$ $[M+H]^+$: 508.1748, found: 508.1746.

3.5.5. Synthesis of (tAzNSAc)4MUGlcNAc (23) and Azidosugar (24)

***N*-acetyl propargylamine (20)**. Prepared as previously described from commercially available propargylamine **19** and acetyl chloride,[224] and purified by silica-gel flash column chromatography (8:2; DCM:EtOAc). The ^{1}H NMR spectrum conformed to the literature description.

4-methylumbelliferyl 3,4-di-*O*-acetyl-2-amino-2-*N*-*tert*-butoxycarbonyl-2,6-di-deoxy-6-(4-(acetamidomethyl)-1,2,3-triazol-1-yl)-β-D-glucopyranoside (21) – To a flask containing alkyne **20** (11.8 mg, 0.12 mmol, 3.6 eq) and azidosugar **7** (60 mg, 0.11 mmol, 3.3 eq) was added DMF (2 mL) and the mixture was stirred until it became homogeneous. In a vial, sodium ascorbate (13.1 mg, 0.066 mmol, 2.0 eq) was dissolved in water (250 µL), and the resulting solution was added to the reaction mixture. The vial was washed with water (250 µL), and the wash was also added to the reaction mixture. This procedure was repeated to add $CuSO_4 \cdot 5H_2O$ (8.2 mg, 0.033 mmol, 1.0 eq). The reaction mixture was protected from light and stirred at room temperature. After 16 hr the reaction was judged complete by TLC analysis. The reaction mixture was diluted with water and DCM, and the aqueous phase was washed several times with DCM. The combined organic extracts were dried over anhydrous Na_2SO_4, filtered, and co-evaporated with toluene to remove any residual DMF. The material was purified by silica-gel flash column chromatography (97:3; DCM:MeOH) to yield the desired compound as a white solid (60.5 mg, 0.094 mmol, 85%).

^{1}H NMR (500 MHz, acetone-d6) δ: 7.76 (s, 1H, tAz H-5), 7.68 (d, J = 8.8 Hz, 1H, AR), 7.49 (s, 1H, NHAc), 6.78 (dd, J = 8.8, 2.2 Hz, 1H, AR), 6.63 (d, J = 2.2 Hz, 1H, AR), 6.34 (d, J = 9.6 Hz, 1H, Boc NH), 6.20 (d, J = 1.0 Hz, 1H, 4MU H-3), 5.45 (d, J = 8.4 Hz, 1H, H-1), 5.39 (t, J = 9.9 Hz, 1H, H-3), 5.00 (t, J = 9.7 Hz, 1H, H-4), 4.73 (dd, J = 14.4, 2.6 Hz, 1H, H-6), 4.57 (dd, J = 14.4, 8.8 Hz, 1H, H-6), 4.43 (m, 2H, H-5, [tAz]CH_2[NAc]), 4.34 (dd, J = 15.2, 5.5 Hz, 1H, [tAz]CH_2[NAc]), 3.96 (dt, J = 10.0, 8.4 Hz, 1H, H-2), 2.47 (d, J = 1.1 Hz, 3H, 4MU CH_3), 2.08 (s, 3H, O/NAc CH_3), 2.00 (s, 3H, O/NAc CH_3), 1.84 (s, 3H, O/NAc CH_3), 1.38 (s, 9H, Boc $(CH_3)_3$); ^{13}C NMR (151 MHz, acetone-d6) δ 170.45, 170.30, 169.90, 161.02, 160.63, 156.26, 155.70, 153.81, 146.58, 127.27, 124.08, 116.05, 114.21, 113.31, 104.44, 99.46, 79.48, 73.39, 73.34, 71.07, 56.16, 51.26, 35.48, 28.48 (3C, Boc $(CH_3)_3$), 22.74, 20.76, 20.66, 18.59; HRMS: m/z calcd for $C_{30}H_{38}N_5O_{11}^+$ $[M+H]^+$: 644.2568, found: 644.2578.

4-methylumbelliferyl 3,4-di-*O*-acetyl-2-amino-2-*N*-*tert*-butoxycarbonyl-2,6-di-deoxy-6-(4-(thioacetamidomethyl)-1,2,3-triazol-1-yl)-β-D-glucopyranoside (22). To an oven dried flask containing amide **21** (108 mg 0.18 mmol, 1.0 eq) under Ar was added dried dioxane (4 mL) and Lawesson's reagent (73 mg, 0.18 mmol, 1.0 eq). The

reaction mixture was protected from light and stirred under Ar at reflux. After 2 hr the reaction was judged complete by TLC analysis. The reaction mixture was concentrated *in vacuo,* and the material was purified by silica-gel flash column chromatography using a gradient solvent system (98:2 - 96:4; DCM:MeOH). Three rounds of chromatography were required to remove all of the co-eluting impurities and yield the desired compound as a pale yellow solid (102 mg, 0.16 mmol, 92%).

^{1}H NMR (500 MHz, acetone-d6) δ: 9.49 (s, 1H, NHSAc), 7.92 (s, 1H,tAz H-5), 7.65 (d, J = 8.8 Hz, 1H, AR, 6.79 (dd, J = 8.7, 1.9 Hz, 1H, AR), 6.58 (d, J = 2.0 Hz, 1H, AR), 6.36 (d, J = 9.2 Hz, 1H, Boc NH), 6.20 (d, J = 0.8 Hz, 1H, 4MU H-3), 5.47 (d, J = 8.4 Hz, 1H, H-1), 5.40 (t, J = 9.9 Hz, 1H, H-3), 5.01 (t, J = 9.7 Hz, 1H, H-4), 4.95 (dd, J = 15.2, 5.4 Hz, 1H, CH_2), 4.80 – 4.73 (m, 2H, CH_2, H-6), 4.60 (dd, J = 14.4, 8.9 Hz, 1H, H-6), 4.43 (ddd, J = 10.2, 8.1, 2.0 Hz, 1H, H-5), 3.97 (dt, J = 10.1, 8.4 Hz, 1H, H-2), 2.45 (d, J = 0.8 Hz, 3H, 4MU CH_3), 2.45 (s, 3H, NSAc CH_3), 2.09 (s, 3H OAc CH_3), 2.01 (s, 3H OAc CH_3), 1.38 (s, 9H Boc $(CH_3)_3$); ^{13}C NMR (126 MHz, acetone-d6) δ: 201.38, 170.45, 170.30, 161.13, 160.52, 156.23, 155.67, 153.80, 144.01, 127.13, 124.82, 116.01, 114.27, 113.31, 104.32, 99.26, 79.50, 73.35 (2C), 71.04, 56.13, 51.31, 41.84, 33.37, 28.47 (3C, Boc $(CH_3)_3$), 20.77, 20.66, 18.65; HRMS: m/z calcd for $C_{30}H_{38}N_5O_{10}S^+$ $[M+H]^+$: 660.2334, found: 660.2336.

4-methylumbelliferyl 2-acetamido-2,6-di-deoxy-6-(4-(thioacetamidomethyl)-1,2,3-triazol-1-yl)-β-D-glucopyranoside ((6tAzNSAc)4MUGlcNAc 23). – To a flame dried flask containing **22** (49 mg, 0.075 mmol, 1.0 eq) and dried DCM (4 mL) under Ar was added TFA (1 mL) by syringe, and the reaction mixture was protected from light and stirred under Ar at room temperature. After 3 hr the reaction was judged complete by TLC analysis (98:2; DCM;MeOH, 0.5% TEA), and the solvent was removed *in vacuo*. To the resulting crude amine was added Ac_2O (1.5 mL) and Py (1.5 mL), and the reaction mixture was protected from light and stirred at room temperature. After 16 hr the reaction was judged complete by TLC analysis (98:2; DCM;MeOH), and the solvent was removed *in vacuo*. The resulting crude amide was dissolved in EtOAc and washed successively twice with water, twice with 1 M HCl, once with water, twice with saturated $NaHCO_3$ solution, and twice with saturated NaCl solution. The organic extracts were dried over anhydrous Na_2SO_4, filtered, and concentrated *in vacuo*. The crude amide was allowed to

dry under high vacuum, after which the flask was flushed with Ar and commercial anhydrous MeOH (2 mL) was added. After dissolution, K_2CO_3 (2.1 mg, 0.015 mmol, 0.2 eq) was added and the reaction mixture was protected from light and stirred under Ar at room temperature. After 2 hr the reaction was judged complete by TLC analysis. Regenerated Amberlite IR120 H^+ resin was added until the reaction mixture became slightly acidic (pH ~ 5) as determined by pH paper. The resin was then removed by filtration and washed several times with MeOH, after which the combined filtrates were evaporated *in vacuo*. For silica-gel column chromatography, the compound was loaded onto the column by dissolving the crude material in MeOH, concentrating *in vacuo* in the presence of silica, and loading the resulting silica slurry with minimum portions of DCM. Purification by silica-gel flash column chromatography (9:1; EtOAc;MeOH) gave the desired compound as a white solid (24 mg, 0.047 mmol, 62%). Afterwards, the compound was further purified by HPLC using the following conditions: 100 µL injections, 50 mg/mL in DMSO, 2 mL/min flow, 15 to 80% B over 30 min, A = H_2O + 5% MeOH, B = MeOH.

^{1}H NMR (601 MHz, DMSO) δ: 10.33 (t, J = 5.1 Hz, 1H, NHSAc), 7.86 (d, J = 8.9 Hz, 1H, NHAc), 7.85 (s, 1H, tAz H-5), 7.60 (d, J = 8.8 Hz, 1H, AR), 6.71 (dd, J = 8.8, 2.4 Hz, 1H, AR), 6.66 (d, J = 2.4 Hz, 1H, AR), 6.25 (d, J = 1.2 Hz, 1H, 4MU H-3), 5.63 (d, J = 5.6 Hz, 1H, OH-4), 5.28 (d, J = 5.6 Hz, 1H, OH-3), 5.09 (d, J = 8.5 Hz, 1H, H-1), 4.78 (dd, J = 14.4, 2.2 Hz, 1H, H-6), 4.71 (dd, J = 15.0, 5.5 Hz, 1H, CH_2), 4.65 (dd, J = 15.0, 5.1 Hz, 1H, CH_2), 4.47 (dd, J = 14.4, 8.7 Hz, 1H, H-6), 3.85 (ddd, J = 10.5, 8.5, 2.2 Hz, 1H, H-5), 3.72 (dt, J = 9.7, 7.9 Hz, 1H, H-2), 3.48 (ddd, J = 10.1, 8.8, 5.7 Hz, 1H, H-3), 3.20 – 3.14 (m, 1H, H-4), 2.39 (d, J = 1.0 Hz, 3H, 4MU CH_3), 2.34 (s, 3H, NSAc CH_3), 1.81 (s, 3H, OAc CH_3); ^{13}C NMR (151 MHz, DMSO) δ: 199.31, 169.37, 159.97, 159.44, 154.28, 153.05, 142.30, 126.33, 124.41, 114.46, 113.16, 112.01, 103.19, 98.14, 74.56, 73.43, 71.47, 55.12, 50.67, 40.47, 32.54, 23.01, 18.14. HRMS: m/z calcd for $C_{23}H_{28}N_5O_7S^+$ $[M+H]^+$: 518.1704, found: 518.1708.

4-methylumbelliferyl 2-acetamido-6-azido-2,6-di-deoxy-β-D-glucopyranoside (24). To an oven dried flask under Ar containing azidosugar **7** (12.4 mg, 0.023 mmol, 1.0 eq) was added dried DCM (1.5 mL) and TFA (0.5 mL), and the mixture was stirred under Ar protected from light at room temperature. After 1 hr the

reaction was judged complete by TLC analysis and the solvent was removed by co-evaporation with toluene *in vacuo*. To the resulting crude amine was added DCM (2 mL), TEA (0.095 mL, 0.681 mmol, 30 eq), and Ac_2O (0.063 mL, 0.681 mmol, 30 eq), and the reaction mixture was protected from light and stirred at room temperature. After 1 hr the reaction was judged complete by TLC analysis and the solvent was removed by co-evaporation with toluene. The resulting crude amide was flushed with Ar and commercial anhydrous MeOH (2.0 mL) was added followed by K_2CO_3 (1.7 mg, 0.012 mmol, 0.5 eq). The reaction mixture was protected from light and stirred under Ar at room temperature. After 16 hr the reaction was judged complete by TLC analysis. For silica-gel column chromatography, the compound was loaded onto the column by dissolving the crude material in MeOH, concentrating *in vacuo* in the presence of silica, and loading the resulting silica slurry with minimum portions of DCM. Purification by silica-gel flash column chromatography (9:1; DCM:MeOH) gave the desired compound as a white solid (9.1 mg, 0.023 mmol, 99%).

^{1}H NMR (600 MHz, CD_3OD) δ: 7.71 (d, J = 8.8 Hz, 1H, AR), 7.04 (dd, J = 8.7, 2.4 Hz, 1H, AR), 7.01 (d, J = 2.4 Hz, 1H, AR), 6.21 (d, J = 1.3 Hz, 1H, 4MU H-3), 5.24 (d, J = 8.4 Hz, 1H, H-1), 3.95 (dd, J = 10.4, 8.4 Hz, 1H, H-2), 3.67 (ddd, J = 9.5, 7.0, 2.3 Hz, 1H, H-5), 3.61 (dd, J = 13.9, 1.8 Hz, 1H, H-6), 3.60 (t, J = 9.6 Hz, 1H, H-3), 3.48 (dd, J = 13.3, 7.0 Hz, 1H, H-6), 3.40 (dd, J = 9.7, 8.8 Hz, 1H, H-4), 2.45 (d, J = 1.3 Hz, 3H, 4MU CH_3), 1.99 (d, J = 1.8 Hz, 3H, NAc CH_3); ^{13}C NMR (151 MHz, CD_3OD) δ: 172.03, 161.37, 159.69, 154.17, 153.54, 125.50, 114.39, 113.17, 111.14, 102.90, 98.04, 75.37, 73.48, 70.76, 55.27, 50.95, 21.05, 16.74; HRMS: m/z calcd for $C_{18}H_{21}N_4O_7^+$ $[M+H]^+$: 405.1405, found: 405.1408.

3.5.6. Enzyme Preparation, Kinetics, and Quenching Measurements

Fluorescence quenching by dissolved thioacetamide. Thioacetamide dissolved in water (200 µL, 600 mM) was added to a black opaque 96-microwell plate in duplicate and serial diluted (1:1) into water (100 µL). To the wells was added 100 µL of 4MUGlcNAc (5 mM), or 4MU (2 µM), or resorufin (2 µM), in 5% DMSO in water. A blank containing no fluorescent compound was prepared by instead adding 100 µL of 5% DMSO in water. Final conditions were: 200 µL; 300 to 0 mM thioacetamide; 2.5 mM

4MUGlcNAc, or 1 μM 4MU, or 1 μM resorufin; and 2.5% DMSO. Fluorescence was quantified at 25 °C with excitation/emission wavelengths of 365/445 nm for 4MU/4MUGlcNAc, and 365/590 for resorufin. Data was acquired using a Synergy 4 microplate reader and analyzed using Gen5 and Graphpad software. Blanks were subtracted, and quenching at each concentration of thioacetamide was calculated as the percentage decrease in fluorescence in comparison to the fluorescence at 0 mM thioacetamide.

Expression and purification of human *O*-GlcNAcase. Lysogeny broth (LB) containing kanamycin (50 μg/mL) was inoculated with *E. coli* transformed with a plasmid expressing His6 tagged hOGA, and was incubated aerobically at 37 °C overnight. The next day, 6x 1 L of LB containing kanamycin (50 μg/mL) was inoculated with overnight culture (1:100) and incubated aerobically at 37 °C to an OD600 of ~ 0.6. Once at the correct OD600 1 mL of IPTG (0.5 M, 0.5 mM final) was added, and protein was expressed aerobically at 21 °C for 4 hr. Afterward, the cultures were transferred to 1 L centrifuge bottles and spun at 5000 rcf and 4 °C for 10 min. Each pellet was resuspended in 20 mL binding buffer (20 mM Na_2HPO_4, 500 mM NaCl, 5 mM imidazole, pH 7.4) and stored at -80 °C in 50 mL centrifuge tubes. When pure protein was needed a protease inhibitor tablet and lysozyme (1mg/mL) were added to the required pellets, which were thawed on a rocker at 4 °C. Once thawed, the pellets were sonicated using a sonic dismembrator at 60% amplitude on ice for 2 min (20 sec on, 40 sec off). The sonicated pellets were centrifuged at 20,000 rcf and 4 °C for 1 hr, during which time a HisTrap column for affinity chromatography was prepared by passing through 50 mL Milli-Q water and 50 mL binding buffer using a peristaltic pump at a flow of 8 mL/min. After centrifugation, the supernatant was carefully withdrawn with a pipette and loaded onto the column. Loading, washing, and elution was carried out at 5 mL/min. Once loaded, 50 mL washing buffer (20 mM Na_2HPO_4, 500 mM NaCl, 60 mM imidazole, pH 7.4) was passed through the column, followed by 50 mL of elution buffer (20 mM Na_2HPO_4, 500 mM NaCl, 250 mM imidazole, pH 7.4) that was collected in 6 mL fractions. Bradford reagent was used to check the protein content of the fractions, and the three most concentrated fractions were combined. The combined fractions were dialyzed for 20 hr twice at 4 °C in PBS using a membrane with a molecular weight cut-off of 14,000 Da. After dialysis 0.01% NaN_3 was added and the concentration was

determined using a Nanodrop 2000 UV-Vis spectrophotometer. The protein was stored at 4 °C and used within three weeks.

Kinetics of substrate processing by human O-GlcNAcase. 80 µL solutions in DMSO of 20 mM 4MUGlcNAc, (6NSAc)4MUGlcNAc, or (6tAzNSAc)4MUGlcNAc; or of 10 mM (6NAcNSAc)4MUGlcNAc or ($6N_3$)4MUGlcNAc, was added to a black opaque 96-microwell plate and serial diluted (1:1) into DMSO (40 µL). 10 µL in triplicate of the serial diluted substrates was added to 90 µL PBS. The plate was warmed to 37 °C, at which point 100 µL of enzyme (200, 800, 1200, 1200, or 100 nM, respectively) in PBS was added and the reaction was monitored at 37 °C on a Synergy 4 microplate reader with Gen5 software using excitation/emission wavelengths of 365/445 nm. Final conditions were: 200 µL; 1000 to 0 µM 4MUGlcNAc, (6NSAc)4MUGlcNAc, or (6tAzNSAc)4MUGlcNAc, or 500 to 0 µM (6NAcNSAc)4MUGlcNAc or ($6N_3$)4MUGlcNAc; 100, 400, 600, 600, or 50 nM hOGA, respectively; and 5% DMSO. The initial rates of 4MU product formation were determined for each substrate concentration with the aid of a 4MU standard curve using Graphpad software. A plot of initial rate against substrate concentration was fit to the Michaelis-Menten equation to obtain kinetic parameters.

Excitation/emission scans and quenching measurements. The same 100 µL quartz fluorescence cuvette was used for all measurements. Before the first measurement and between all subsequent measurements, the cuvette was washed twice with methanol and thrice with 10% DMSO in water, and a blank (10% DMSO in water) was recorded using the same excitation/emission ranges used for the substrates. Before adding the fluorescent substrate, the cuvette was washed twice more with methanol and dried with compressed nitrogen. From 10 mM stocks in DMSO, 100 µL of 4MUGlcNAc, (6NSAc)4MUGlcNAc, (6tAzNSAc)4MUGlcNAc, or (6NAcNSAc)4MUGlcNAc at 10, 50, and 500 µM in 10% DMSO in water were prepared. Working solutions were prepared immediately prior to measurement to avoid precipitation. For all the concentrations of each compound, the excitation and emission profiles were scanned iteratively using a fluorescence spectrophotometer until the wavelength corresponding to the excitation maximum was used to scan the emission profile, and vice versa. The same instrument settings were used for all compounds at a given concentration to ensure comparability. Blanks were subtracted and the excitation

and emission profiles were plotted with relative intensity using Graphpad. Quenching was determined as the percentage decrease in the emission of the thioamide-containing compound compared to that of 4MUGlcNAc at a given wavelength and concentration.

References

(1) *Comprehensive Glycoscience: From Chemistry to Systems Biology*; First ed.; Elsevier: Oxford, UK, 2007.

(2) Laine, R. A. *Glycobiology* **1994**, *4*, 759.

(3) Brooks, S. A.; Dwek, M. V.; Schumacher, U. *Functional and Molecular Glycobiology*; BIOS Scientific Publishers Ltd: Oxford, UK, 2002.

(4) *Essentials of Glycobiology*; Second ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 2009.

(5) Linden, S. K.; Sutton, P.; Karlsson, N. G.; Korolik, V.; McGuckin, M. A. *Mucosal Immunol.* **2008**, *1*, 183.

(6) Zachara, N. E.; Hart, G. W. *Biochim. et Biophys. Acta* **2006**, *1761*, 599.

(7) Kleizen, B.; Braakman, I. *Curr. Opin. Cell Biol.* **2004**, *16*, 343.

(8) Berg, J. M.; Tymoczko, J. L.; Stryer, L. *Biochemistry, Fifth Edition*; W.H. Freeman, 2002.

(9) Candy, D. J. *Biological Functions of Carbohydrates*; Blackie and Son Ltd: Glasgow, Scotland, 1980.

(10) Tessier, F. J. *Pathol. Biol.* **2010**, *58*, 214.

(11) Bunn, H. F.; Higgins, P. J. *Science* **1981**, *213*, 222.

(12) Gibbons, B. J.; Roach, P. J.; Hurley, T. D. *J. Mol. Biol.* **2002**, *319*, 463.

(13) Klemm, D.; Heublein, B.; Fink, H. P.; Bohn, A. *Angew. Chem. Int. Ed. Engl.* **2005**, *44*, 3358.

(14) Jung, H.-J. G. *J. Nutr.* **1995**, *125*, 1025.

(15) Demchick, P.; Koch, A. L. *J. Bacteriol.* **1996**, *178*, 768.

(16) Fisher, J. F.; Meroueh, S. O.; Mobashery, S. *Chem. Rev.* **2005**, *105*, 395.

(17) Parodi, A. J. *Annu. Rev. Biochem.* **2000**, *69*, 69.

(18) Molinari, M. *Nat. Chem. Biol.* **2007**, *3*, 313.

(19) Frenkel, Z.; Gregory, W.; Kornfeld, S.; Lederkremer, G. Z. *J. Biol. Chem.* **2003**, *278*, 34119.

(20) Dean, L.; National Center for Biotechnology Information: Bethesda (MD), 2005.

(21) Olsson, M. L.; Clausen, H. *Br. J. Haematol.* **2008**, *140*, 3.

(22) Lairson, L. L.; Henrissat, B.; Davies, G. J.; Withers, S. G. *Annu. Rev. Biochem.* **2008**, *77*, 521.

(23) Kitagawa, H.; Shimakawa, H.; Sugahara, K. *J. Biol. Chem.* **1999**, *274*, 13933.

(24) Lombard, V.; Golaconda Ramulu, H.; Drula, E.; Coutinho, P. M.; Henrissat, B. *Nucleic Acids Res.* **2014**, *42*, 490.

(25) Gloster, T. M. *Curr. Opin. Struct. Biol.* **2014**, *28*, 131.

(26) Lesk, A. M. *Curr. Opin. Struct. Biol.* **1995**, *5*, 775.

(27) Charnock, S. J.; Davies, G. J. *Biochemistry* **1999**, *38*, 6380.

(28) Vrielink, A.; Rueger, W.; Driessen, H. P. C.; Freemont, P. S. *EMBO J.* **1994**, *13*, 3413.

(29) Lizak, C.; Gerber, S.; Numao, S.; Aebi, M.; Locher, K. P. *Nature* **2011**, *474*, 350.

(30) Wolfenden, R.; Lu, X.; Young, G. *J. Am. Chem. Soc.* **1998**, *120*, 6814.

(31) Henrissat, B.; Davies, G. *Curr. Opin. Struct. Biol.* **1997**, *7*, 637.

(32) Davies, G.; Henrissat, B. *Structure (London, England : 1993)* **1995**, *3*, 853.

(33) Dvir, H.; Harel, M.; McCarthy, A. A.; Toker, L.; Silman, I.; Futerman, A. H.; Sussman, J. L. *EMBO Rep.* **2003**, *4*, 704.

(34) Tsuji, H.; Nishimura, S.; Inui, T.; Kado, Y.; Ishikawa, K.; Nakamura, T.; Uegaki, K. *FEBS J.* **2010**, *277*, 2683.

(35) Rouvinen, J.; Bergfors, T.; Teeri, T.; Knowles, J. K. C.; Jones, T. A. *Science* **1990**, *249*, 380.

(36) Michel, G.; Chantalat, L.; Duee, E.; Barbeyron, T.; Henrissat, B.; Kloareg, B.; Dideberg, O. *Structure (London, England : 1993)* **2001**, *9*, 513.

(37) Koshland, D. E., Jr. *Biol. Rev. Cambridge Philos. Soc.* **1953**, *28*, 416.

(38) McCarter, J. D.; Withers, S. G. *Curr. Opin. Struct. Biol.* **1994**, *4*, 885.

(39) Mhlongo, N. N.; Skelton, A. A.; Kruger, G.; Soliman, M. E. S.; Williams, I. H. *Proteins: Struct., Funct., Bioinf.* **2014**, *82*, 1747.

(40) Dall'Acqua, W.; Carter, P. *Protein Sci.* **2000**, *9*, 1.

(41) Vocadlo, D. J.; Withers, S. G. *Biochemistry* **2005**, *44*, 12809.

(42) Gloster, T. M.; Vocadlo, D. J. *Curr. Signal Transduction Ther.* **2010**, *5*, 74.

(43) Macauley, M. S.; Whitworth, G. E.; Debowski, A. W.; Chin, D.; Vocadlo, D. J. *J. Biol. Chem.* **2005**, *280*, 25313.

(44) Watts, A. G.; Damager, I.; Amaya, M. L.; Buschiazzo, A.; Alzari, P.; Frasch, A. C.; Withers, S. G. *J. Am. Chem. Soc.* **2003**, *125*, 7532.

(45) Rajan, S. S.; Yang, X.; Collart, F.; Yip, V. L. Y.; Withers, S. G.; Varrot, A.; Thompson, J.; Davies, G. J.; Anderson, W. F. *Structure (London, England : 1993)* **2004**, *12*, 1619.

(46) Huang, S.; Czech, M. P. *Cell Metab.* **2007**, *5*, 237.

(47) Zierath, J. R.; Houseknecht, K. L.; Kahn, B. B. *Seminars in Cell & Dev. Biol.* **1996**, *7*, 295.

(48) Gadsby, D. C.; Vergani, P.; Csanady, L. *Nature* **2006**, *440*, 477.

(49) Pao, S. S.; Paulsen, I. T.; Saier, M. H., Jr. *Microbiol. Mol. Biol. Rev.* **1998**, *62*, 1.

(50) Henderson, P. J. F.; Maiden, M. C. J. *Philos. Trans. R. Soc. London, B* **1990**, *326*, 391.

(51) Reddy, V. S.; Shlykov, M. A.; Castillo, R.; Sun, E. I.; Saier, M. H., Jr. *FEBS J.* **2012**, *279*, 2022.

(52) Saier, M. H., Jr. *Mol. Microbiol.* **2000**, *35*, 699.

(53) Abramson, J.; Smirnova, I.; Kasho, V.; Verner, G.; Kaback, H. R.; Iwata, S. *Science* **2003**, *301*, 610.

(54) Huang, Y.; Lemieux, M. J.; Song, J.; Auer, M.; Wang, D.-N. *Science* **2003**, *301*, 616.

(55) Kaback, H. R.; Smirnova, I.; Kasho, V.; Nie, Y.; Zhou, Y. *J. Membr. Biol.* **2011**, *239*, 85.

(56) Yan, N. *Trends Biochem. Sci.* **2013**, *38*, 151.

(57) Radestock, S.; Forrest, L. R. *J. Mol. Biol.* **2011**, *407*, 698.

(58) Wuethrich, K. *J. Biol. Chem.* **1990**, *265*, 22059.

(59) Palmer, A. G., III *Acc. Chem. Res.* **2015**, *48*, 457.

(60) Copeland, R. A.; Editor *Evaluation of Enzyme Inhibitors in Drug Discovery: A Guide to Medicinal Chemists and Pharmacologists*; John Wiley & Sons, 2005.

(61) Brown, A. J. *J. Chem. Soc., Trans.* **1902**, *81*, 373.

(62) Michaelis, L.; Menten, M. L. *Biochem. J.* **1913**, *49*, 333.

(63) Briggs, G. E.; Haldane, J. B. *Biochem. J.* **1925**, *19*, 338.

(64) Copeland, R. A. *Enzymes: A Practical Introduction to Structure, Mechanism, and Data Analysis, Second Edition*; Wiley-VCH, 2000.

(65) Bisswanger, H. *Perspect. Sci.* **2014**, *1*, 41.

(66) Nishihira, T.; Miyano, A.; Ohnuma, T.; Gotoh, T.; Takahashi, S.; Narihiro, K.; Yamashita, K.; Fukamizo, T. *J. Appl. Glycosci.* **2014**, *61*, 113.

(67) Hastie, C. J.; McLauchlan, H. J.; Cohen, P. *Nat. Protoc.* **2006**, *1*, 968.

(68) Harder, D.; Fotiadis, D. *Nat. Protoc.* **2012**, *7*, 1569.

(69) Lakowicz, J. R. *Principles of fluorescence spectroscopy, 3rd Edition*; Springer: New York, NY, USA, 2006.

(70) Yoshioka, K.; Takahashi, H.; Homma, T.; Saito, M.; Oh, K.-B.; Nemoto, Y.; Matsuoka, H. *Biochim. Biophys. Acta, Gen. Subj.* **1996**, *1289*, 5.

(71) Lee, J.; Bogyo, M. *ACS Chem. Biol.* **2010**, *5*, 233.

(72) Hsu, Y.-L.; Nandakumar, M.; Lai, H.-Y.; Chou, T.-C.; Chu, C.-Y.; Lin, C.-H.; Lo, L.-C. *J. Org. Chem.* **2015**, *80*, 8458.

(73) Drake, C. R.; Miller, D. C.; Jones, E. F. *Curr. Org. Synth.* **2011**, *8*, 498.

(74) Yee, D. J.; Balsanek, V.; Sames, D. *J. Am. Chem. Soc.* **2004**, *126*, 2282.

(75) de Silva, A. P.; Gunaratne, H. Q. N.; Gunnlaugsson, T.; Huxley, A. J. M.; McCoy, C. P.; Rademacher, J. T.; Rice, T. E. *Chem. Rev.* **1997**, *97*, 1515.

(76) Beija, M.; Afonso, C. A. M.; Martinho, J. M. G. *Chem. Soc. Rev.* **2009**, *38*, 2410.

(77) Engstler, M.; Talhouk, J. W.; Smith, R. E.; Schauer, R. *Anal. Biochem.* **1997**, *250*, 176.

(78) Gee, K. R.; Sun, W.-C.; Bhalgat, M. K.; Upson, R. H.; Klaubert, D. H.; Latham, K. A.; Haugland, R. P. *Anal. Biochem.* **1999**, *273*, 41.

(79) Ho, N.-H.; Weissleder, R.; Tung, C.-H. *ChemBioChem* **2007**, *8*, 560.

(80) Latt, S. A.; Auld, D. S.; Vallee, B. L. *Anal. Biochem.* **1972**, *50*, 56.

(81) Matayoshi, E. D.; Wang, G. T.; Krafft, G. A.; Erickson, J. *Science* **1990**, *247*, 954.

(82) Charitos, C.; Tzougraki, C.; Kokotos, G. *J. Pept. Res.* **2000**, *56*, 373.

(83) Mittoo, S.; Sundstrom, L. E.; Bradley, M. *Anal. Biochem.* **2003**, *319*, 234.

(84) Dorjsuren, D.; Wilson, D. M., III; Beard, W. A.; McDonald, J. P.; Austin, C. P.; Woodgate, R.; Wilson, S. H.; Simeonov, A. *Nucleic Acids Res.* **2009**, *37*, e128/1.

(85) Mansoor, S. E.; Dewitt, M. A.; Farrens, D. L. *Biochemistry* **2010**, *49*, 9722.

(86) Goldberg, J. M.; Batjargal, S.; Chen, B. S.; Petersson, E. J. *J. Am. Chem. Soc.* **2013**, *135*, 18651.

(87) Sanchez, K. M.; Schlamadinger, D. E.; Gable, J. E.; Kim, J. E. *J. Chem. Educ.* **2008**, *85*, 1253.

(88) Carmel, A.; Zur, M.; Yaron, A.; Katchalski, E. *FEBS Lett.* **1973**, *30*, 11.

(89) Sato, E.; Sakashita, M.; Kanaoka, Y.; Kosower, E. M. *Bioorg. Chem.* **1988**, *16*, 298.

(90) Baruch, A.; Jeffery, D. A.; Bogyo, M. *Trends Cell Biol.* **2004**, *14*, 29.

(91) Rosano, G. L.; Ceccarelli, E. A. *Front. Microbiol.* **2014**, *5*, 172.

(92) Burke, H. M.; Gunnlaugsson, T.; Scanlan, E. M. *Chem. Commun.* **2015**, *51*, 10576.

(93) Hilderbrand, S. A.; Weissleder, R. *Curr. Opin. Chem. Biol.* **2010**, *14*, 71.

(94) Richard, J.-A.; Massonneau, M.; Renard, P.-Y.; Romieu, A. *Org. Lett.* **2008**, *10*, 4175.

(95) Rut, W.; Kasperkiewicz, P.; Byzia, A.; Poreba, M.; Groborz, K.; Drag, M. *Biol. Chem.* **2015**, *396*, 329.

(96) Mueller, J.; Kretzschmar, I.; Volkmer, R.; Boisguerin, P. *Bioconjug. Chem.* **2008**, *19*, 2363.

(97) Maxwell, D.; Chang, Q.; Zhang, X.; Barnett, E. M.; Piwnica-Worms, D. *Bioconjug. Chem.* **2009**, *20*, 702.

(98) Qiu, X.; Johnson, J. R.; Wilson, B. S.; Gammon, S. T.; Piwnica-Worms, D.; Barnet, E. M. *PLoS One* **2014**, *9*, e88855/1.

(99) Xing, B.; Khanamiryan, A.; Rao, J. *J. Am. Chem. Soc.* **2005**, *127*, 4158.

(100) Han, J.; Han, M. S.; Tung, C.-H. *Mol. BioSyst.* **2013**, *9*, 3001.

(101) Motabar, O.; Shi, Z.-D.; Goldin, E.; Liu, K.; Southall, N.; Sidransky, E.; Austin, C. P.; Griffiths, G. L.; Zheng, W. *Anal. Biochem.* **2009**, *390*, 79.

(102) *The Metabolic and Molecular Basis of Inherited Disease*; McGraw-Hill: New York City.

(103) Okumiya, T.; Kroos, M. A.; Vliet, L. V.; Takeuchi, H.; Van der Ploeg, A. T.; Reuser, A. J. J. *Mol. Genet. Metab.* **2007**, *90*, 49.

(104) Simeonov, A.; Jadhav, A.; Thomas, C. J.; Wang, Y.; Huang, R.; Southall, N. T.; Shinn, P.; Smith, J.; Austin, C. P.; Auld, D. S.; Inglese, J. *J. Med. Chem.* **2008**, *51*, 2363.

(105) Asanuma, D.; Sakabe, M.; Kamiya, M.; Yamamoto, K.; Hiratake, J.; Ogawa, M.; Kosaka, N.; Choyke, P. L.; Nagano, T.; Kobayashi, H.; Urano, Y. *Nat. Commun.* **2015**, *6*.

(106) Chatterjee, S. K.; Bhattacharya, M.; Barlow, J. J. *Cancer Res.* **1979**, *39*, 1943.

(107) Bosmann, H. B.; Hall, T. C. *Proc. Nat. Acad. Sci. U. S. A.* **1974**, *71*, 1833.

(108) Yadav, A. K.; Shen, D. L.; Shan, X.; He, X.; Kermode, A. R.; Vocadlo, D. J. *J. Am. Chem. Soc.* **2015**, *137*, 1181.

(109) Meikle, P. J.; Hopwood, J. J.; Clague, A. E.; Carey, W. F. *JAMA* **1999**, *281*, 249.

(110) Sidransky, E.; Lopez, G. *Lancet Neurol.* **2012**, *11*, 986.

(111) Witte, M. D.; Kallemeijn, W. W.; Aten, J.; Li, K.-Y.; Strijland, A.; Donker-Koopman, W. E.; van, d. N. A. M. C. H.; Bleijlevens, B.; Kramer, G.; Florea, B. I.; Hooibrink, B.; Hollak, C. E. M.; Ottenhoff, R.; Boot, R. G.; van, d. M. G. A.; Overkleeft, H. S.; Aerts, J. M. F. G. *Nat. Chem. Biol.* **2010**, *6*, 907.

(112) Chen, L.-Q.; Cheung, L. S.; Feng, L.; Tanner, W.; Frommer, W. B. *Annu. Rev. Biochem.* **2015**, *84*, 865.

(113) Jarvis, S. M. In *Membrane Transport: A Practical Approach*; Baldwin, S. A., Ed.; Oxford University Press: Oxford, U.K., 2000, p 1.

(114) Buttin, G.; Cohen, G. N.; Monod, J.; Rickenberg, H. V. *Ann. Inst. Pasteur* **1956**, *91*, 829.

(115) Booth, I. R.; Mitchell, W. J.; Hamilton, W. A. *Biochem. J.* **1979**, *182*, 687.

(116) Gould, G. W.; Thomas, H. M.; Jess, T. J.; Bell, G. I. *Biochemistry* **1991**, *30*, 5139.

(117) Mueckler, M.; Thorens, B. *Mol. Aspects Med.* **2013**, *34*, 121.

(118) Brunius, G. *Curr. Microbiol.* **1980**, *4*, 321.

(119) Speizer, L.; Haugland, R.; Kutchai, H. *Biochim. Biophys. Acta, Biomembr.* **1985**, *815*, 75.

(120) Yoshioka, K.; Saito, M.; Oh, K.-B.; Nemoto, Y.; Matsuoka, H.; Natsume, M.; Abe, H. *Biosci., Biotechnol., Biochem.* **1996**, *60*, 1899.

(121) Steinsiek, S.; Bettenbrock, K. *J. Bacteriol.* **2012**, *194*, 5897.

(122) Yamada, K.; Saito, M.; Matsuoka, H.; Inagaki, N. *Nat. Protoc.* **2007**, *2*, 753.

(123) Poolman, B.; Knol, J.; van der Does, C.; Henderson, P. J. F.; Liang, W.-J.; Leblanc, G.; Pourcher, T.; Mus-Veteau, I. *Mol. Microbiol.* **1996**, *19*, 911.

(124) Heuberger, E. H. M. L.; Smits, E.; Poolman, B. *J. Biol. Chem.* **2001**, *276*, 34465.

(125) Oubrie, A.; Dijkstra, B. W. *Protein Sci.* **2000**, *9*, 1265.

(126) Liao, J.; Sportsman, R.; Harris, J.; Stahl, A. *J. Lipid Res.* **2005**, *46*, 597.

(127) Laxminarayan, R.; Duse, A.; Wattal, C.; Zaidi, A. K. M.; Wertheim, H. F. L.; Sumpradit, N.; Vlieghe, E.; Hara, G. L.; Gould, I. M.; Goossens, H.; Greko, C.; So, A. D.; Bigdeli, M.; Tomson, G.; Woodhouse, W.; Ombaka, E.; Peralta, A. Q.; Qamar, F. N.; Mir, F.; Kariuki, S.; Bhutta, Z. A.; Coates, A.; Bergstrom, R.; Wright, G. D.; Brown, E. D.; Cars, O. *Lancet Infect. Dis.* **2013**, *13*, 1057.

(128) Vollmer, W.; Holtje, J.-V. *Curr. Opin. Microbiol.* **2001**, *4*, 625.

(129) Livermore, D. M. *Clin. Microbiol. Rev.* **1995**, *8*, 557.

(130) Bush, K.; Jacoby, G. A. *Antimicrob. Agents Chemother.* **2010**, *54*, 969.

(131) Jacoby, G. A. *Clin. Microbiol. Rev.* **2009**, *22*, 161.

(132) Drawz, S. M.; Bonomo, R. A. *Clin. Microbiol. Rev.* **2010**, *23*, 160.

(133) Hanson, N. D.; Sanders, C. C. *Curr. Pharm. Des.* **1999**, *5*, 881.

(134) Lister, P. D.; Wolter, D. J.; Hanson, N. D. *Clin. Microbiol. Rev.* **2009**, *22*, 582.

(135) Lyczak, J. B.; Cannon, C. L.; Pier, G. B. *Clin. Microbiol. Rev.* **2002**, *15*, 194.

(136) Livermore, D. M.; Mushtaq, S.; Warner, M. *J. Antimicrob. Chemother.* **2010**, *65*, 2382.

(137) Hendershot, J. M.; Mishra, U. J.; Smart, R. P.; Schroeder, W.; Powers, R. A. *Bioorg. Med. Chem.* **2014**, *22*, 3351.

(138) Livermore, D. M.; Mushtaq, S.; Barker, K.; Hope, R.; Warner, M.; Woodford, N. *J. Antimicrob. Chemother.* **2012**, *67*, 1354.

(139) Lahiri, S. D.; Walkup, G. K.; Whiteaker, J. D.; Palmer, T.; McCormack, K.; Tanudra, M. A.; Nash, T. J.; Thresher, J.; Johnstone, M. R.; Hajec, L.; Livchak, S.; McLaughlin, R. E.; Alm, R. A. *J. Antimicrob. Chemother.* **2015**, *70*, 1650.

(140) Holtje, J.-V. *Microbiol. Mol. Biol. Rev.* **1998**, *62*, 181.

(141) Park, J. T.; Uehara, T. *Microbiol. Mol. Biol. Rev.* **2008**, *72*, 211.

(142) Vollmer, W.; Joris, B.; Charlier, P.; Foster, S. *FEMS Microbiol. Rev.* **2008**, *32*, 259.

(143) Cheng, Q.; Park, J. T. *J. Bacteriol.* **2002**, *184*, 6434.

(144) Cheng, Q.; Li, H.; Merdek, K.; Park, J. T. *J. Bacteriol.* **2000**, *182*, 4836.

(145) Votsch, W.; Templin, M. F. *J. Biol. Chem.* **2000**, *275*, 39032.

(146) Jacobs, C.; Frere, J. M.; Normark, S. *Cell* **1997**, *88*, 823.

(147) Dietz, H.; Pfeifle, D.; Wiedemann, B. *Antimicrob. Agents Chemother.* **1997**, *41*, 2113.

(148) Asgarali, A.; Stubbs, K. A.; Oliver, A.; Vocadlo, D. J.; Mark, B. L. *Antimicrob. Agents Chemother.* **2009**, *53*, 2274.

(149) Stubbs, K. A.; Balcewich, M.; Mark, B. L.; Vocadlo, D. J. *J. Biol. Chem.* **2007**, *282*, 21382.

(150) Mondon, M.; Hur, S.; Vadlamani, G.; Rodrigues, P.; Tsybina, P.; Oliver, A.; Mark, B. L.; Vocadlo, D. J.; Bleriot, Y. *Chem. Commun.* **2013**, *49*, 10983.

(151) Zhang, Y.; Bao, Q.; Gagnon, L. A.; Huletsky, A.; Oliver, A.; Jin, S.; Langaee, T. *Antimicrob. Agents Chemother.* **2010**, *54*, 4772.

(152) Lindquist, S.; Weston-Hafer, K.; Schmidt, H.; Pul, C.; Korfmann, G.; Erickson, J.; Sanders, C.; Martin, H. H.; Normark, S. *Mol. Microbiol.* **1993**, *9*, 703.

(153) Zamorano, L.; Reeve, T. M.; Juan, C.; Moyá, B.; Cabot, G.; Vocadlo, D. J.; Mark, B. L.; Oliver, A. *Antimicrob. Agents Chemother.* **2011**, *55*, 1990.

(154) Korfmann, G.; Sanders, C. C. *Antimicrob. Agents Chemother.* **1989**, *33*, 1946.

(155) Garcia, D. L.; Dillard, J. P. *J. Bacteriol.* **2008**, *190*, 3799.

(156) Adin, D. M.; Engle, J. T.; Goldman, W. E.; McFall-Ngai, M. J.; Stabb, E. V. *J. Bacteriol.* **2009**, *191*, 2012.

(157) Ward, A.; Sanderson, N. M.; O'Reilly, J.; Rutherford, N. G.; Poolman, B.; Henderson, P. J. F. In *Membrane Transport: A Practical Approach*; Baldwin, S. A., Ed.; Oxford University Press: Oxford, U.K., 2000, p 141.

(158) Xie, H. *Acta Biochim. Biophys. Sin.* **2008**, *40*, 269.

(159) Narahara, H. T.; Green, J. D. *Biochim. Biophys. Acta, Biomembr.* **1982**, *688*, 5.

(160) Smith, S. M. *Methods Mol. Biol.* **2011**, *681*, 485.

(161) Tailler, D.; Jacquinet, J.-C.; Noirot, A.-M.; Beau, J.-M. *J. Chem. Soc., Perkin Trans. 1* **1992**, 3163.

(162) Ganguli, A. R. S.; Coward, J. K. *Tetrahedron: Asymmetry* **2005**, *16*, 411.

(163) Paulsen, H.; Himpkamp, P.; Peters, T. *Liebigs Ann. Chem.* **1986**, 664.

(164) Bianchi, A.; Bernardi, A. *J. Org. Chem.* **2006**, *71*, 4565.

(165) Haferkamp, I.; Schmitz-Esser, S.; Wagner, M.; Neigel, N.; Horn, M.; Neuhaus, H. E. *Mol. Microbiol.* **2006**, *60*, 1534.

(166) Russell, J. B.; Strobel, H. J.; Driessen, A. J.; Konings, W. N. *J. Bacteriol.* **1988**, *170*, 3531.

(167) Martinac, B.; Rohde, P.; Cranfield, C.; Nomura, T. In *Bacterial Cell Surfaces*; Delcour, A. H., Ed.; Humana Press: 2013; Vol. 966, p 367.

(168) Baba, T.; Ara, T.; Hasegawa, M.; Takai, Y.; Okumura, Y.; Baba, M.; Datsenko, K. A.; Tomita, M.; Wanner, B. L.; Mori, H. *Mol. Syst. Biol.* **2006**, No pp. given.

(169) Prabhananda, B. S.; Kombrabail, M. H. *Biochim. Biophys. Acta* **1997**, *1323*, 137.

(170) Tabas, L. B.; Dantzig, A. H. *Anal. Biochem.* **2002**, *310*, 61.

(171) Auld, D. S.; Farmen, M. W.; Kahl, S. D.; Kriauciunas, A.; McKnight, K. L.; Montrose, C.; Weidner, J. R. *Receptor Binding Assays for HTS and Drug Discovery* Bethesda MD, 2004.

(172) Cheng, Y.-C.; Prusoff, W. H. *Biochem. Pharmacol.* **1973**, *22*, 3099.

(173) Kong, K.-F.; Aguila, A.; Schneper, L.; Mathee, K. *BMC Microbiol.* **2010**, *10*, 328.

(174) Rath, A.; Glibowicka, M.; Nadeau, V. G.; Chen, G.; Deber, C. M. *Proc. Nat. Acad. Sci. U. S. A.* **2009**, *106*, 1760.

(175) Johnson, J. W.; Fisher, J. F.; Mobashery, S. *Ann. N. Y. Acad. Sci.* **2013**, *1277*, 54.

(176) Reith, J.; Mayer, C. *Appl. Microbiol. Biotechnol.* **2011**, *92*, 1.

(177) Amoroso, A.; Boudet, J.; Berzigotti, S.; Duval, V.; Teller, N.; Mengin-Lecreulx, D.; Luxen, A.; Simorre, J.-P.; Joris, B. *PLoS Pathog.* **2012**, *8*, e1002571.

(178) Litzinger, S.; Duckworth, A.; Nitzsche, K.; Risinger, C.; Wittmann, V.; Mayer, C. *J. Bacteriol.* **2010**, *192*, 3132.

(179) Baba, T.; Ara, T.; Hasegawa, M.; Takai, Y.; Okumura, Y.; Baba, M.; Datsenko, K. A.; Tomita, M.; Wanner, B. L.; Mori, H. *Mol. Syst. Biol.* **2006**, *2*, 1.

(180) Chen, W. P.; Kuo, T. T. *Nucleic Acids Res.* **1993**, *21*, 2260.

(181) Balcewich, M. D.; Reeve, T. M.; Orlikow, E. A.; Donald, L. J.; Vocadlo, D. J.; Mark, B. L. *J. Mol. Biol.* **2010**, *400*, 998.

(182) O'Callaghan, C. H.; Morris, A.; Kirby, S. M.; Shingler, A. H. *Antimicrob. Agents Chemother.* **1972**, *1*, 283.

(183) Cronan, J. E. *Plasmid* **2006**, *55*, 152.

(184) Motabar, O.; Liu, K.; Southall, N.; Marugan, J. J.; Goldin, E.; Sidransky, E.; Zheng, W. *Curr. Chem. Genomics* **2010**, *4*, 67.

(185) Patel, P. R.; Sun, H.; Li, S. Q.; Shen, M.; Khan, J.; Thomas, C. J.; Davis, M. I. *Bioorg. Med. Chem. Lett.* **2013**, *23*, 4398.

(186) Stubbs, K. A.; Bacik, J.-P.; Perley-Robertson, G. E.; Whitworth, G. E.; Gloster, T. M.; Vocadlo, D. J.; Mark, B. L. *ChemBioChem* **2013**, *14*, 1973.

(187) Eidam, O.; Romagnoli, C.; Dalmasso, G.; Barelier, S.; Caselli, E.; Bonnet, R.; Shoichet, B. K.; Prati, F. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 17448.

(188) Knapp, S.; Vocadlo, D.; Gao, Z.; Kirk, B.; Lou, J.; Withers, S. G. *J. Am. Chem. Soc.* **1996**, *118*, 6804.

(189) Edgington, L. E.; Berger, A. B.; Blum, G.; Albrow, V. E.; Paulick, M. G.; Lineberry, N.; Bogyo, M. *Nature Med.* **2009**, *15*, 967.

(190) Kim, J.-H.; Lee, S.; Park, K.; Nam, H. Y.; Jang, S. Y.; Youn, I.; Kim, K.; Jeon, H.; Park, R.-W.; Kim, I.-S.; Choi, K.; Kwon, I. C. *Angew. Chem., Int. Ed.* **2007**, *46*, 5779.

(191) Law, B.; Weissleder, R.; Tung, C.-H. *ChemBioChem* **2005**, *6*, 1361.

(192) Blum, G.; Mullins, S. R.; Keren, K.; Fonovic, M.; Jedeszko, C.; Rice, M. J.; Sloane, B. F.; Bogyo, M. *Nat. Chem. Biol.* **2005**, *1*, 203.

(193) Ibatullin, F. M.; Banasiak, A.; Baumann, M. J.; Greffe, L.; Takahashi, J.; Mellerowicz, E. J.; Brumer, H. *Plant Physiol.* **2009**, *151*, 1741.

(194) Tung, C.-H.; Zeng, Q.; Shah, K.; Kim, D.-E.; Schellingerhout, D.; Weissleder, R. *Cancer Res.* **2004**, *64*, 1579.

(195) Lubas, W. A.; Frank, D. W.; Krause, M.; Hanover, J. A. *J. Biol. Chem.* **1997**, *272*, 9316.

(196) Dong, D. L. Y.; Hart, G. W. *J. Biol. Chem.* **1994**, *269*, 19321.

(197) Gao, Y.; Wells, L.; Comer, F. I.; Parker, G. J.; Hart, G. W. *J. Biol. Chem.* **2001**, *276*, 9838.

(198) Torres, C.-R.; Hart, G. W. *J. Biol. Chem.* **1984**, *259*, 3308.

(199) Yang, Y. R.; Song, M.; Lee, H.; Jeon, Y.; Choi, E.-J.; Jang, H.-J.; Moon, H. Y.; Byun, H.-Y.; Kim, E.-K.; Kim, D. H.; Lee, M. N.; Koh, A.; Ghim, J.; Choi, J. H.; Lee-Kwon, W.; Kim, K. T.; Ryu, S. H.; Suh, P.-G. *Aging Cell* **2012**, *11*, 439.

(200) Shafi, R.; Iyer, S. P. N.; Ellies, L. G.; O'Donnell, N.; Marek, K. W.; Chui, D.; Hart, G. W.; Marth, J. D. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 5735.

(201) Slawson, C.; Zachara, N. E.; Vosseller, K.; Cheung, W. D.; Lane, M. D.; Hart, G. W. *J. Biol. Chem.* **2005**, *280*, 32944.

(202) Zachara, N. E.; O'Donnell, N.; Cheung, W. D.; Mercer, J. J.; Marth, J. D.; Hart, G. W. *J. Biol. Chem.* **2004**, *279*, 30133.

(203) Ho, S.-R.; Wang, K.; Whisenhunt, T. R.; Huang, P.; Zhu, X.; Kudlow, J. E.; Paterson, A. J. *FEBS Lett.* **2010**, *584*, 49.

(204) Zhu, Y.; Liu, T.-W.; Cecioni, S.; Eskandari, R.; Zandberg, W. F.; Vocadlo, D. J. *Nat. Chem. Biol.* **2015**, *11*, 319.

(205) Yuzwa, S. A.; Shan, X.; Macauley, M. S.; Clark, T.; Skorobogatko, Y.; Vosseller, K.; Vocadlo, D. J. *Nat. Chem. Biol.* **2012**, *8*, 393.

(206) Yuzwa, S. A.; Vocadlo, D. J. *Chem. Soc. Rev.* **2014**, *43*, 6839.

(207) Gu, Y.; Mi, W.; Ge, Y.; Liu, H.; Fan, Q.; Han, C.; Yang, J.; Han, F.; Lu, X.; Yu, W. *Cancer Res.* **2010**, *70*, 6344.

(208) Mi, W.; Gu, Y.; Han, C.; Liu, H.; Fan, Q.; Zhang, X.; Cong, Q.; Yu, W. *Biochim. Biophys. Acta, Mol. Basis Dis.* **2011**, *1812*, 514.

(209) Poras, H.; Duquesnoy, S.; Fournie-Zaluski, M.-C.; Ratinaud-Giraud, C.; Roques, B. P.; Ouimet, T. *Anal. Biochem.* **2013**, *441*, 152.

(210) Park, S.; Hyun, S.; Do, S.-i.; Yu, J. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 2441.

(211) Leriche, G.; Budin, G.; Darwich, Z.; Weltin, D.; Mely, Y.; Klymchenko, A. S.; Wagner, A. *Chem. Commun.* **2012**, *48*, 3224.

(212) Petersson, E. J.; Goldberg, J. M.; Wissner, R. F. *Phys. Chem. Chem. Phys.* **2014**, *16*, 6827.

(213) Goldberg, J. M.; Batjargal, S.; Petersson, E. J. *J Am Chem Soc* **2010**, *132*, 14718.

(214) Cottaz, S.; Brasme, B.; Driguez, H. *Eur. J. Biochem.* **2000**, *267*, 5593.

(215) Krishnamurthy, V. M.; Semetey, V.; Bracher, P. J.; Shen, N.; Whitesides, G. M. *J. Am. Chem. Soc.* **2007**, *129*, 1312.

(216) Dennis, R. J.; Taylor, E. J.; Macauley, M. S.; Stubbs, K. A.; Turkenburg, J. P.; Hart, S. J.; Black, G. N.; Vocadlo, D. J.; Davies, G. J. *Nat. Struct. Mol. Biol.* **2006**, *13*, 365.

(217) Roeser, K. R.; Legler, G. *Biochim. Biophys. Acta, Enzymol.* **1981**, *657*, 321.

(218) Billing, J. F.; Nilsson, U. J. *Tetrahedron* **2005**, *61*, 863.

(219) Hou, J.; Liu, X.; Shen, J.; Zhao, G.; Wang, P. G. *Expert Opin. Drug Discovery* **2012**, *7*, 489.

(220) Choudhary, A.; Raines, R. T. *ChemBioChem* **2011**, *12*, 1801.

(221) Naturale, G.; Lamblin, M.; Commandeur, C.; Felpin, F.-X.; Dessolin, J. *Eur. J. Org. Chem.* **2012**, *2012*, 5774.

(222) Zhao, L.; Li, C.-J. *Angew. Chem. Int. Ed. Engl.* **2008**, *47*, 7075.

(223) Burés, J.; Martín, M.; Urpí, F.; Vilarrasa, J. *J. Org. Chem.* **2009**, *74*, 2203.

(224) Rajagopal, B.; Chen, Y.-Y.; Chen, C.-C.; Liu, X.-Y.; Wang, H.-R.; Lin, P.-C. *J. Org. Chem.* **2014**, *79*, 1254.

(225) Macauley, M. S.; Stubbs, K. A.; Vocadlo, D. J. *J. Am. Chem. Soc.* **2005**, *127*, 17202.

(226) Ribbens, J.; Whiteley, G.; Furuya, H.; Southall, N.; Hu, X.; Marugan, J.; Ferrer, M.; Maegawa, G. H. B. *Anal. Biochem.* **2013**, *434*, 15.

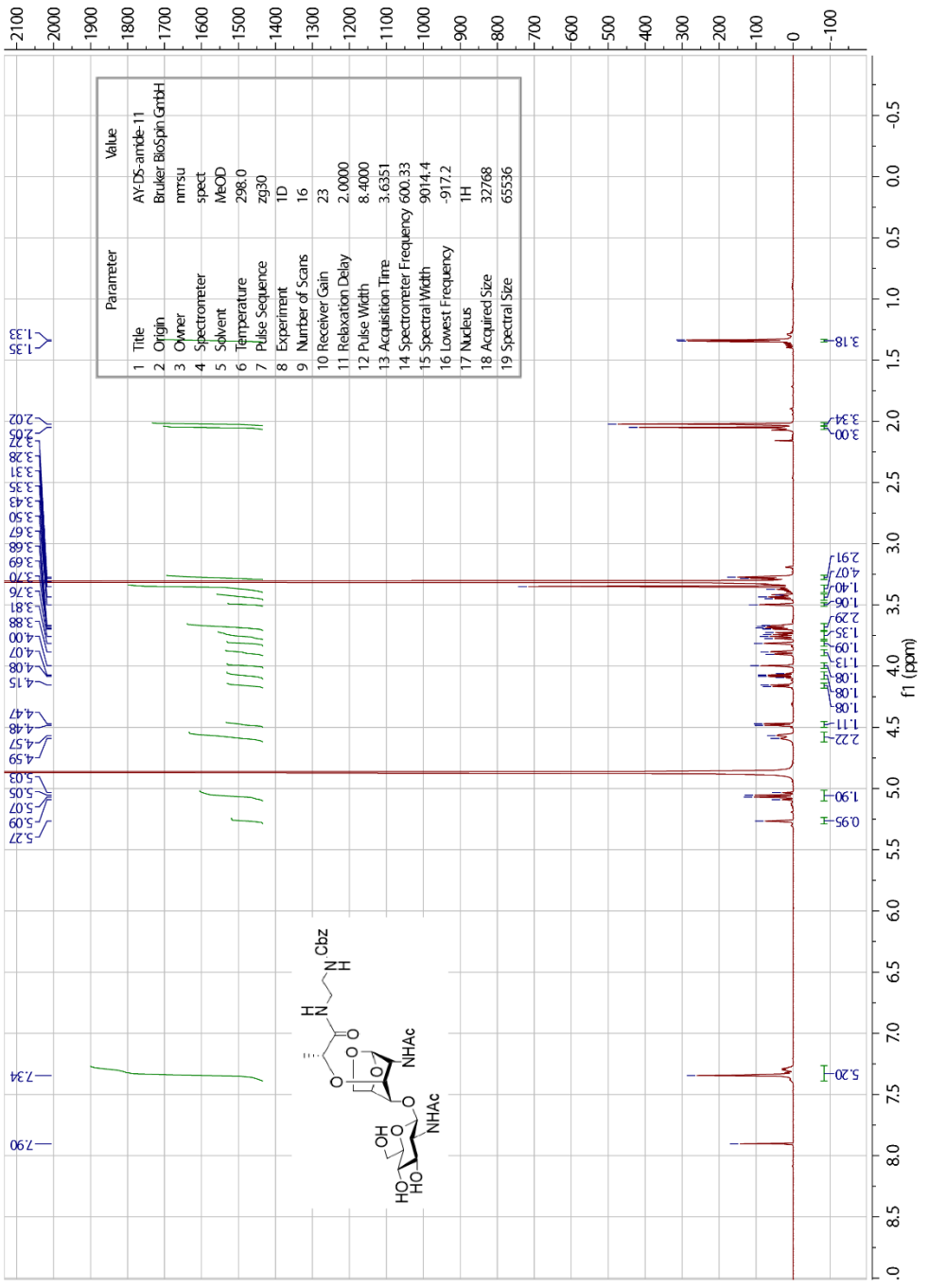
(227) Gao, W.; Xing, B.; Tsien, R. Y.; Rao, J. *J. Am. Chem. Soc.* **2003**, *125*, 11146.

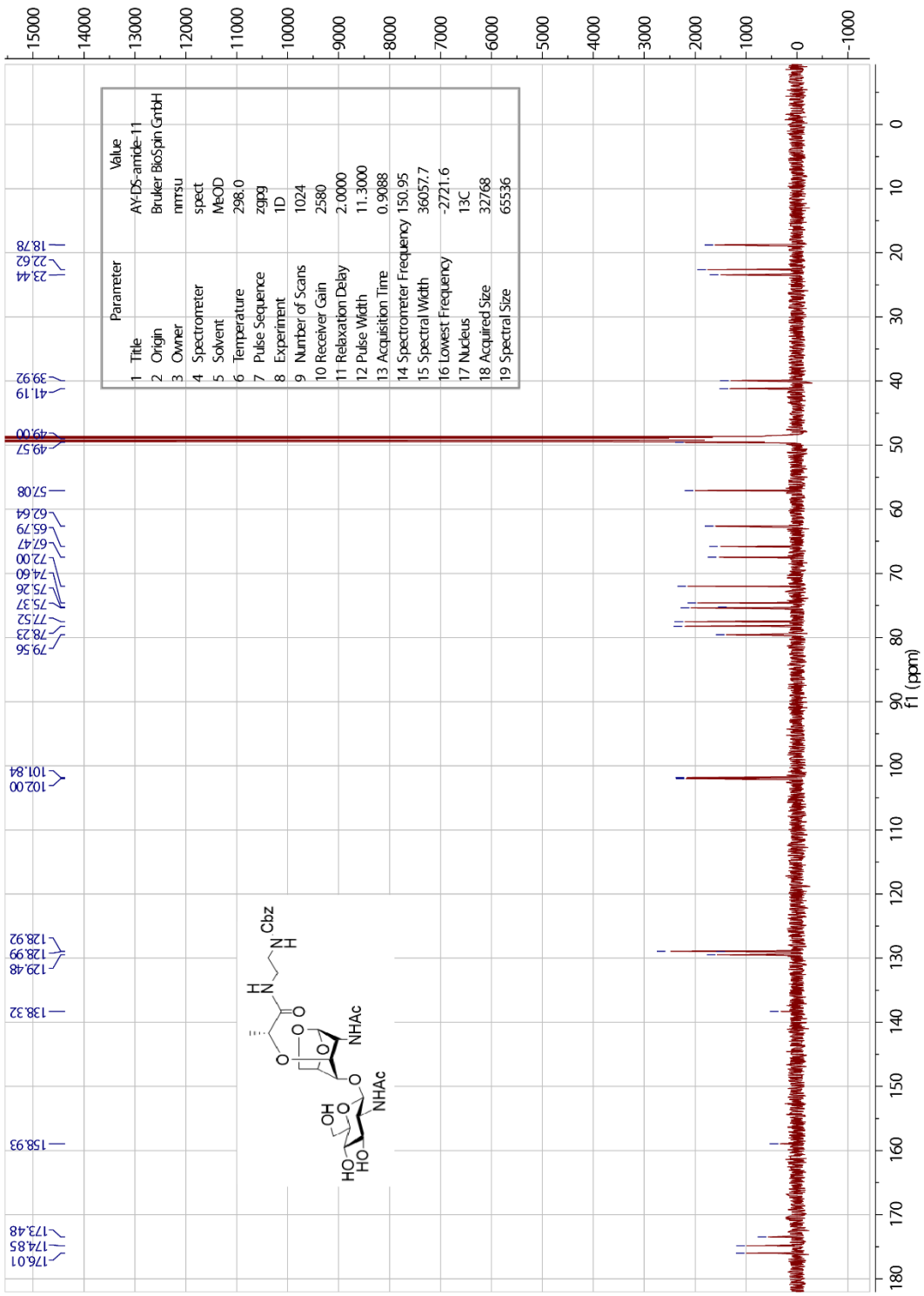
(228) Stubbs, K. A.; Zhang, N.; Vocadlo, D. J. *Org. Biomol. Chem.* **2006**, *4*, 839.

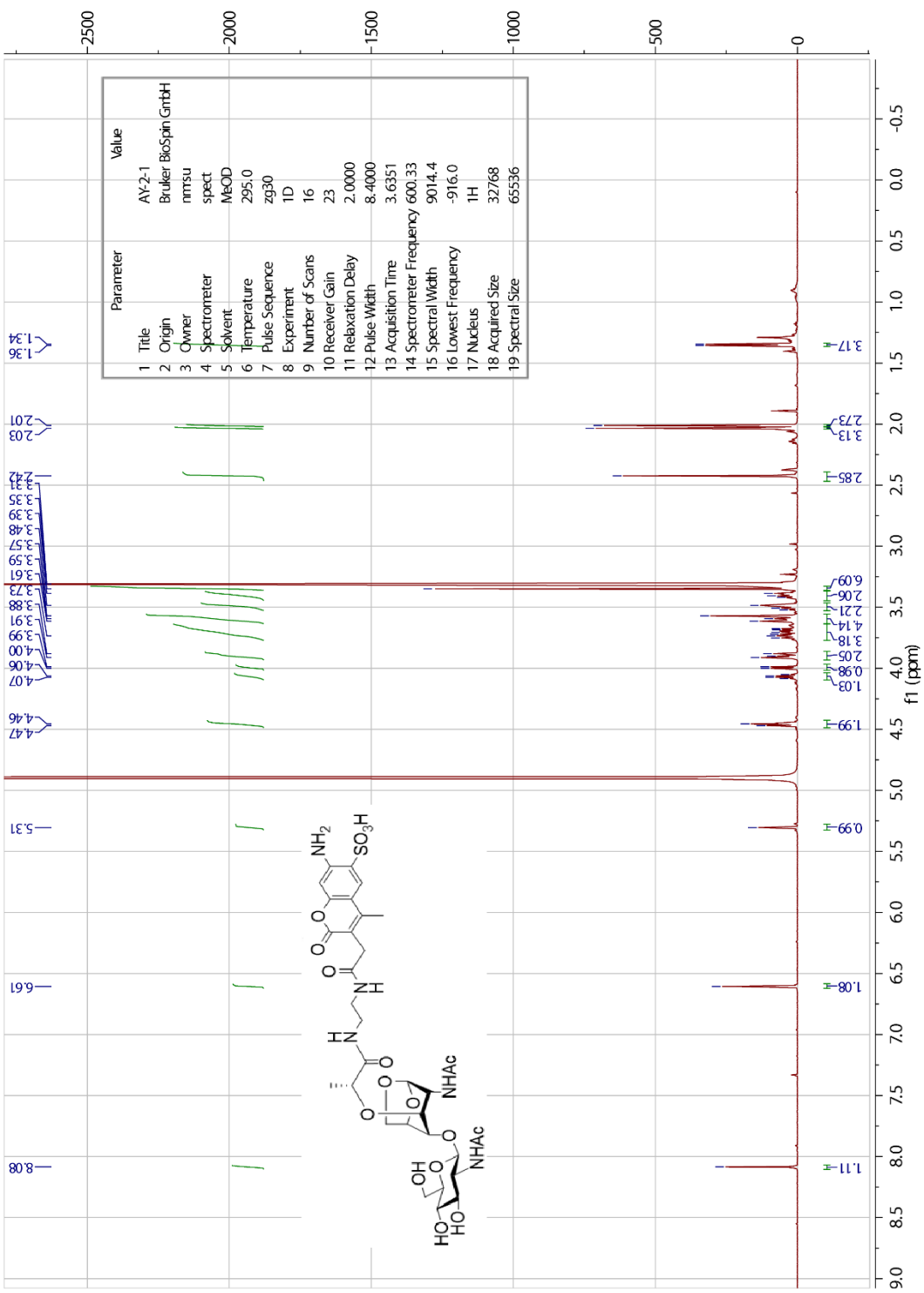
(229) Conrow, R. E.; Dean, W. D. *Org. Process Res. Dev.* **2008**, *12*, 1285.

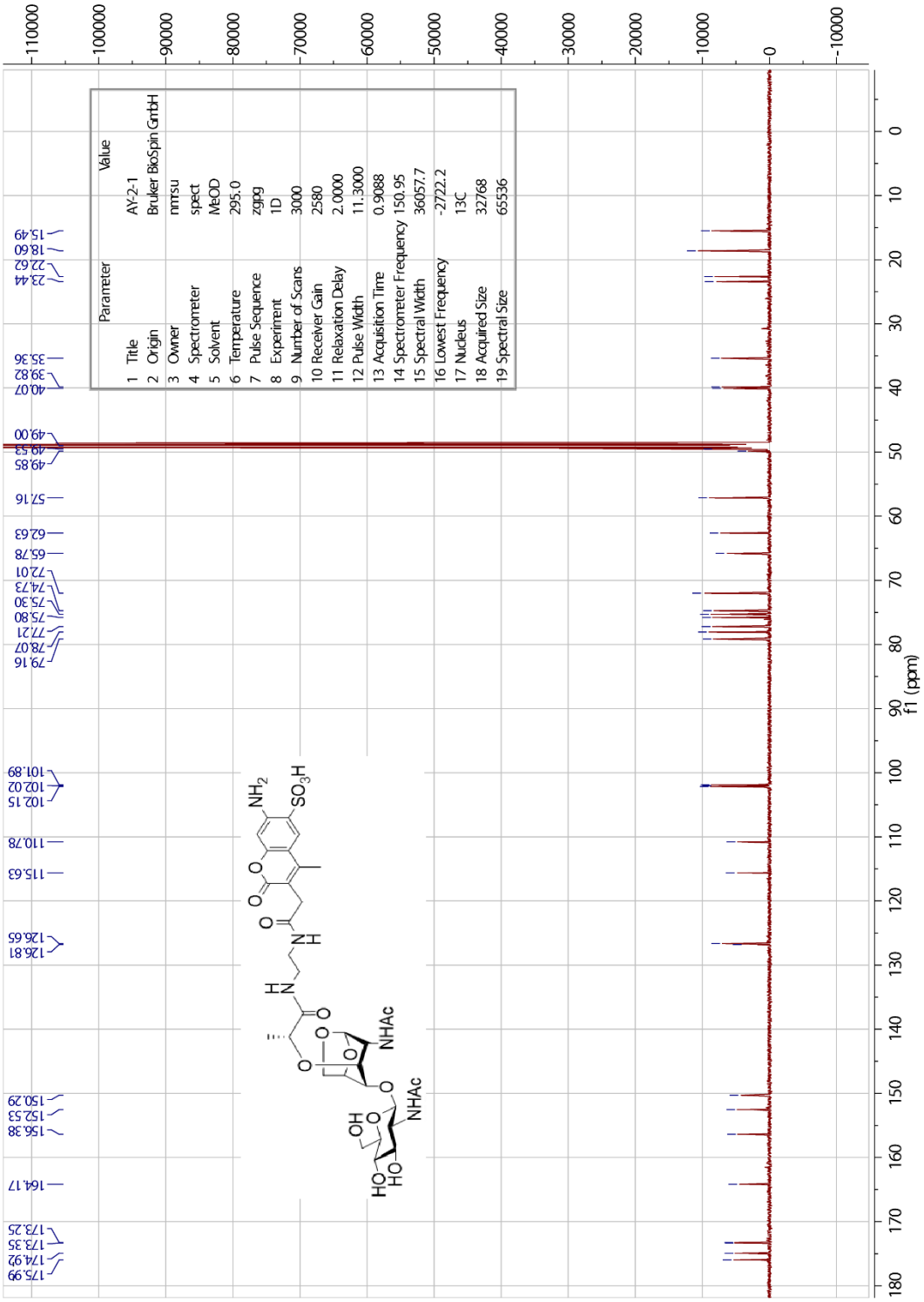
Appendix.

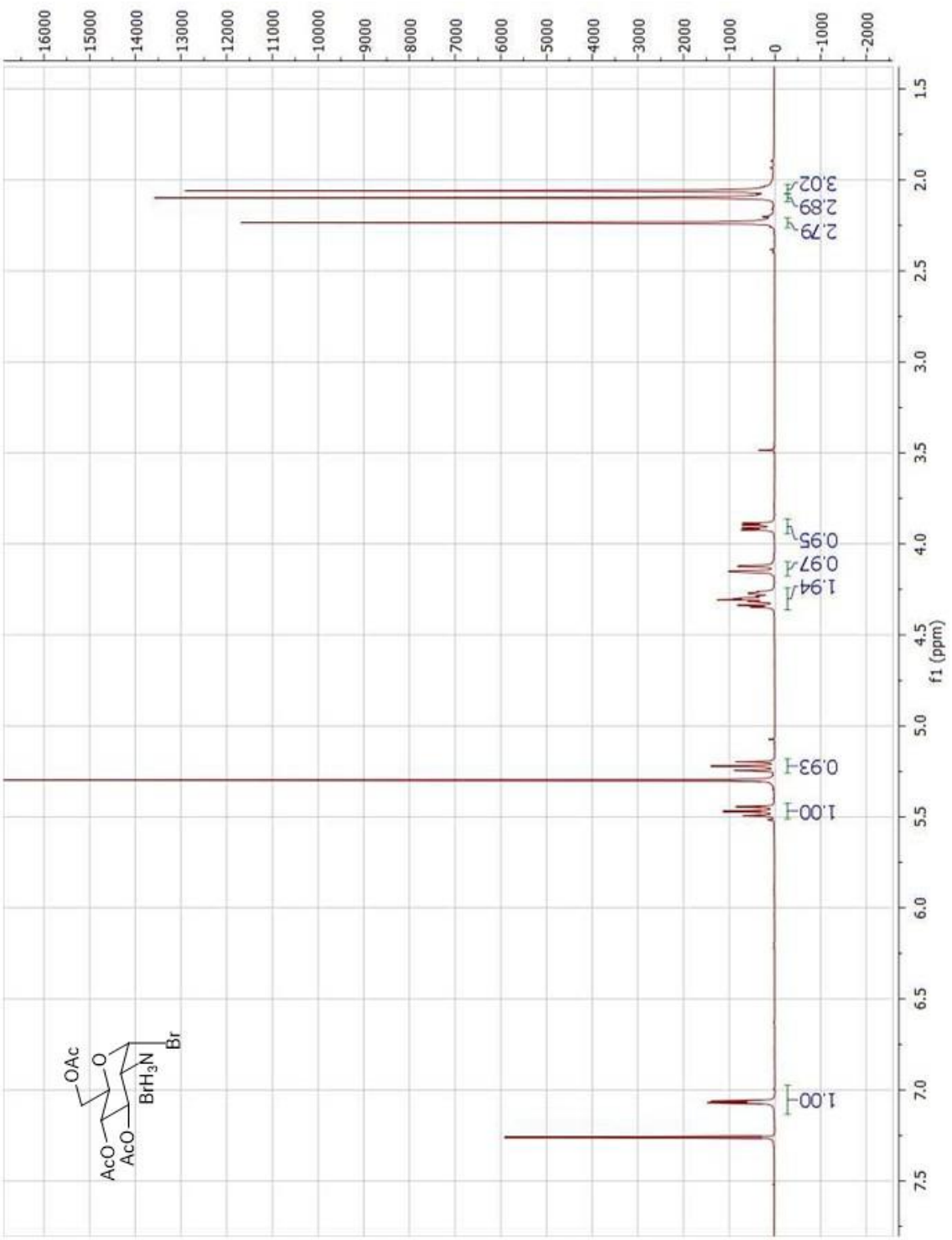
NMR Spectra

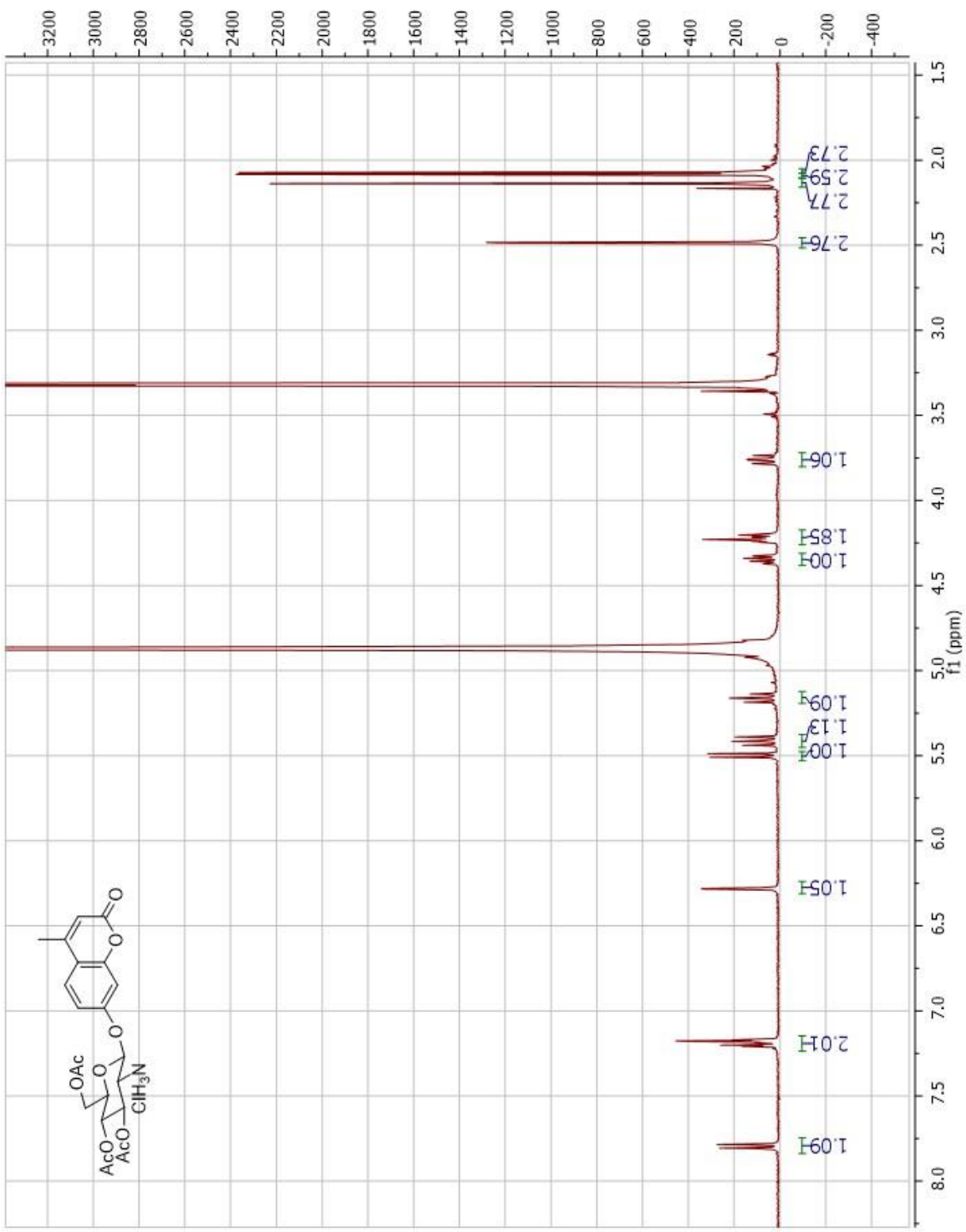


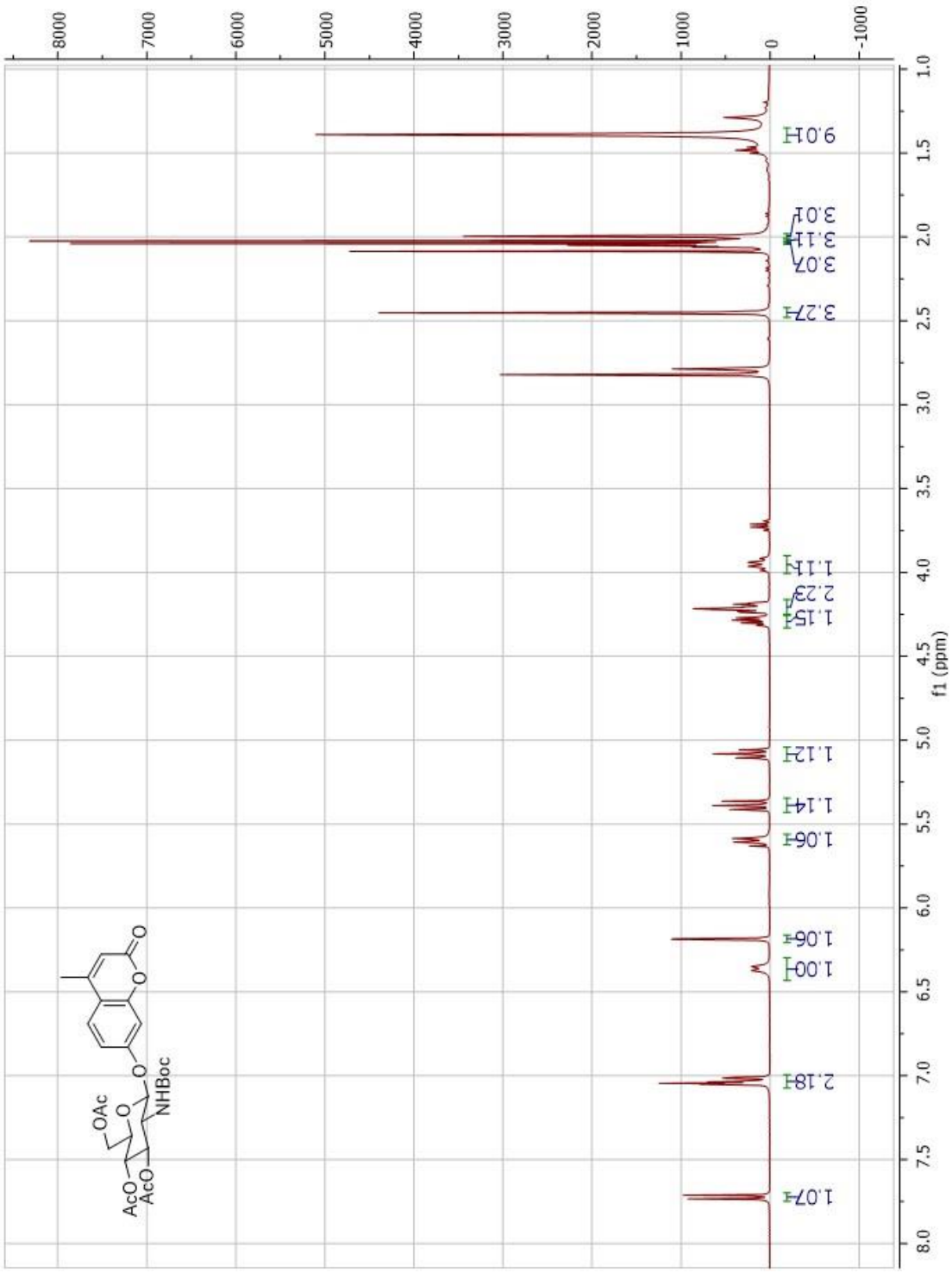

Parameter Value
1 Title AY-DS-amide-11
2 Origin Bruker BioSpin GmbH
3 Owner nmrsu
4 Spectrometer spect
5 Solvent MeOD
6 Temperature 298.0
7 Pulse Sequence zgpg
8 Experiment 1D
9 Number of Scans 1024
10 Receiver Gain 2580
11 Relaxation Delay 2.0000
12 Pulse Width 11.3000
13 Acquisition Time 0.9088
14 Spectrometer Frequency 150.95
15 Spectral Width 36057.7
16 Lowest Frequency -2721.6
17 Nucleus 13C
18 Acquired Size 32768
19 Spectral Size 65536
18.78
22.62
23.44
39.92
41.19
49.57
49.00
57.08
62.64
65.79
67.47
72.00
74.60
75.26
75.37
77.52
78.23
79.56
101.84
102.00
128.92
128.99
129.48
138.32
158.93
173.48
174.85
176.01
Cbz
NHAc
NHAc
f1 (ppm)

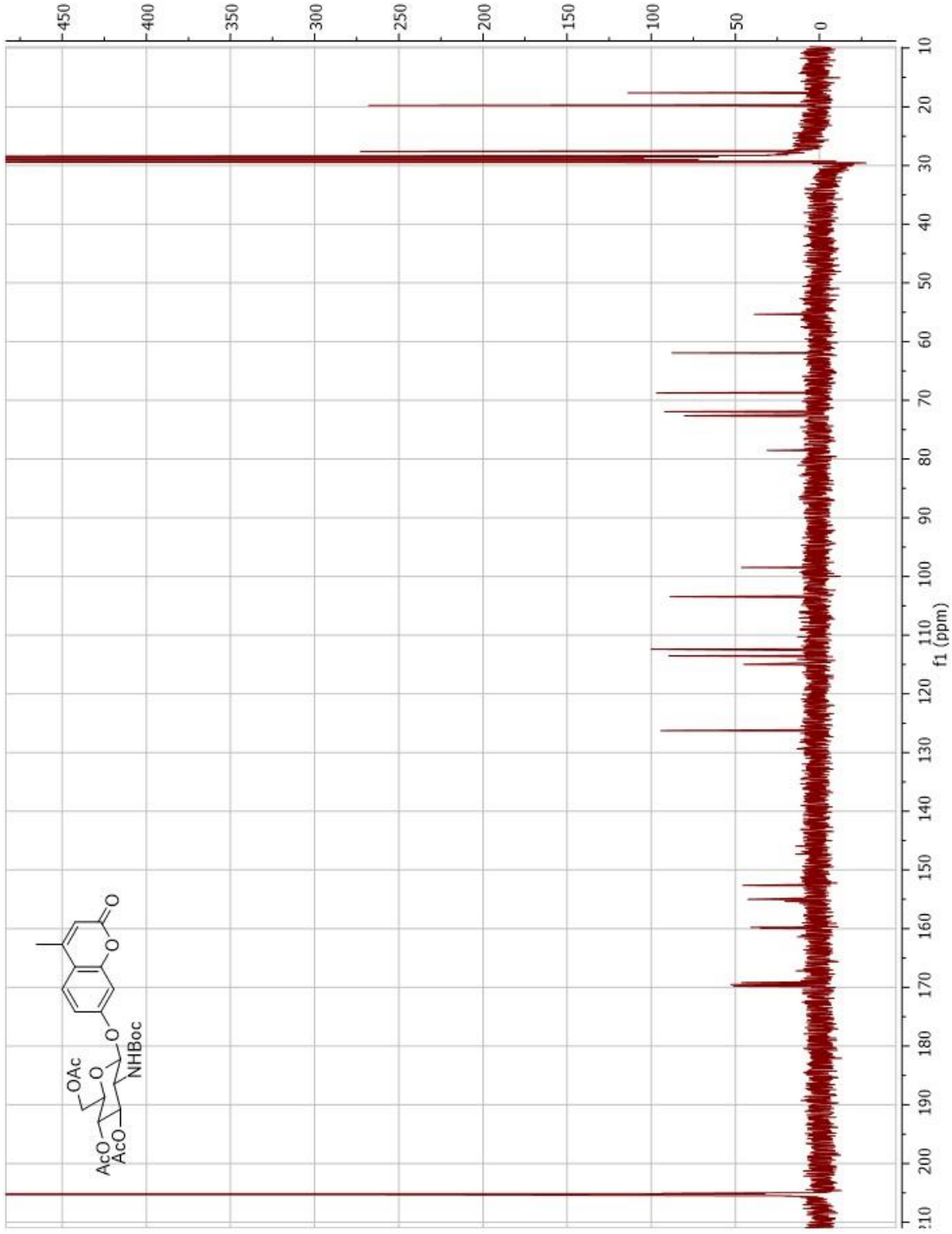

Parameter Value
1 Title AY-2-1
2 Origin Bruker BioSpin GmbH
3 Owner nmrsu
4 Spectrometer spect
5 Solvent MeOD
6 Temperature 295.0
7 Pulse Sequence zg30
8 Experiment 1D
9 Number of Scans 16
10 Receiver Gain 23
11 Relaxation Delay 2.0000
12 Pulse Width 8.4000
13 Acquisition Time 3.6351
14 Spectrometer Frequency 600.33
15 Spectral Width 9014.4
16 Lowest Frequency -916.0
17 Nucleus 1H
18 Acquired Size 32768
19 Spectral Size 65536
f1 (ppm)

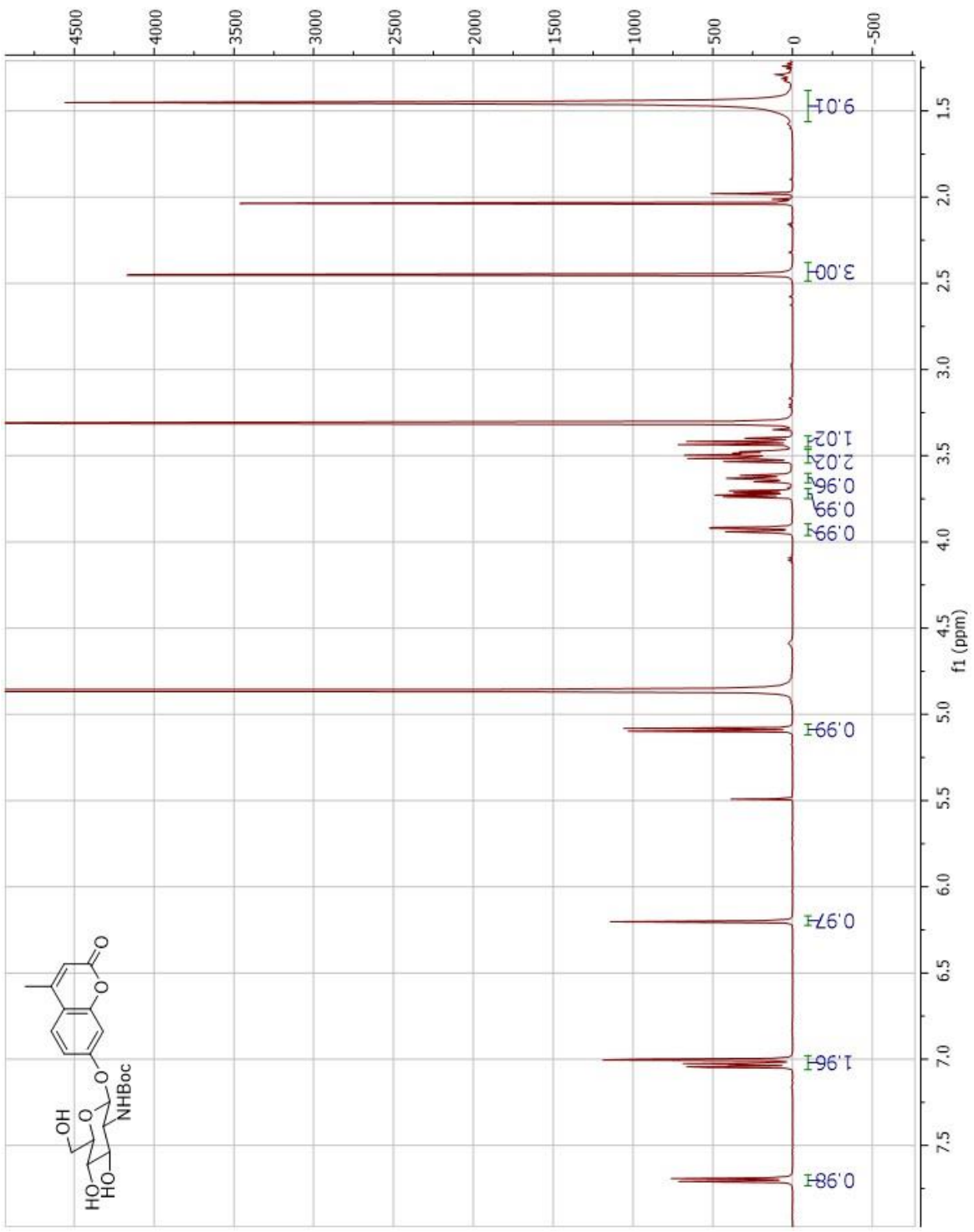

Parameter
Value
1 Title
AY-2-1
2 Origin
Bruker BioSpin GmbH
3 Owner
nmrsu
4 Spectrometer
spect
5 Solvent
MeOD
6 Temperature
295.0
7 Pulse Sequence
zgpg
8 Experiment
1D
9 Number of Scans
3000
10 Receiver Gain
2580
11 Relaxation Delay
2.0000
12 Pulse Width
11.3000
13 Acquisition Time
0.9088
14 Spectrometer Frequency
150.95
15 Spectral Width
36057.7
16 Lowest Frequency
-2722.2
17 Nucleus
13C
18 Acquired Size
32768
19 Spectral Size
65536
15.49
18.60
22.62
23.44
35.36
39.82
40.07
49.00
49.53
49.85
57.16
62.63
65.78
72.01
74.73
75.30
75.80
77.21
78.07
79.16
101.89
102.02
102.15
110.78
115.63
126.65
126.81
150.29
152.53
156.38
164.17
173.25
173.35
174.92
175.99
NH2
SO3H
NHAc
NHAc
f1 (ppm)

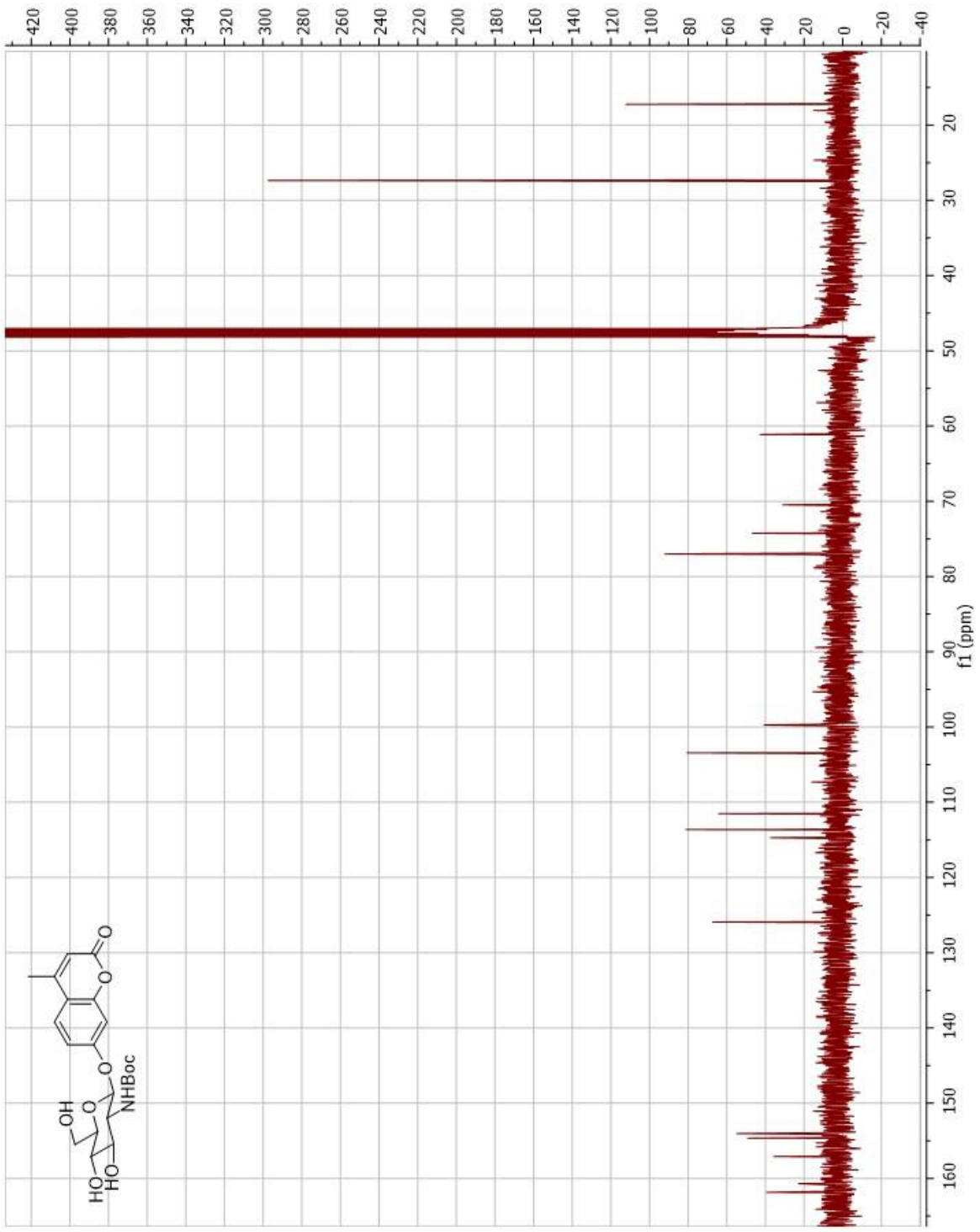

3.02
2.89
2.79
1.94
0.97
0.95
0.93
1.00
1.00
f1 (ppm)
OAc
O
AcO
AcO
BrH3N
Br

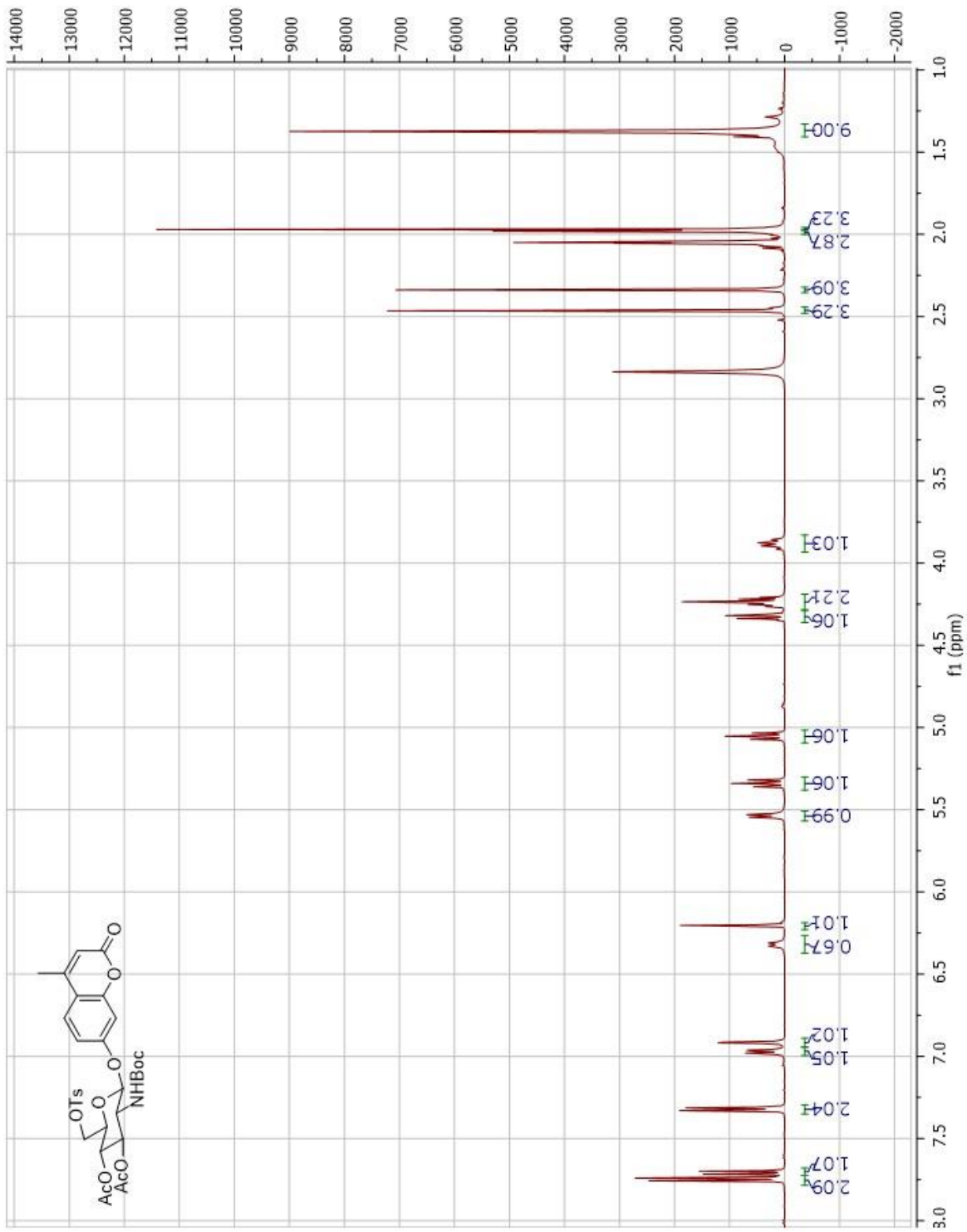

2.73
2.59
2.77
2.76
1.06
1.85
1.00
1.09
1.13
1.00
1.05
2.01
1.09
f1 (ppm)
OAc
AcO
AcO
ClH3N

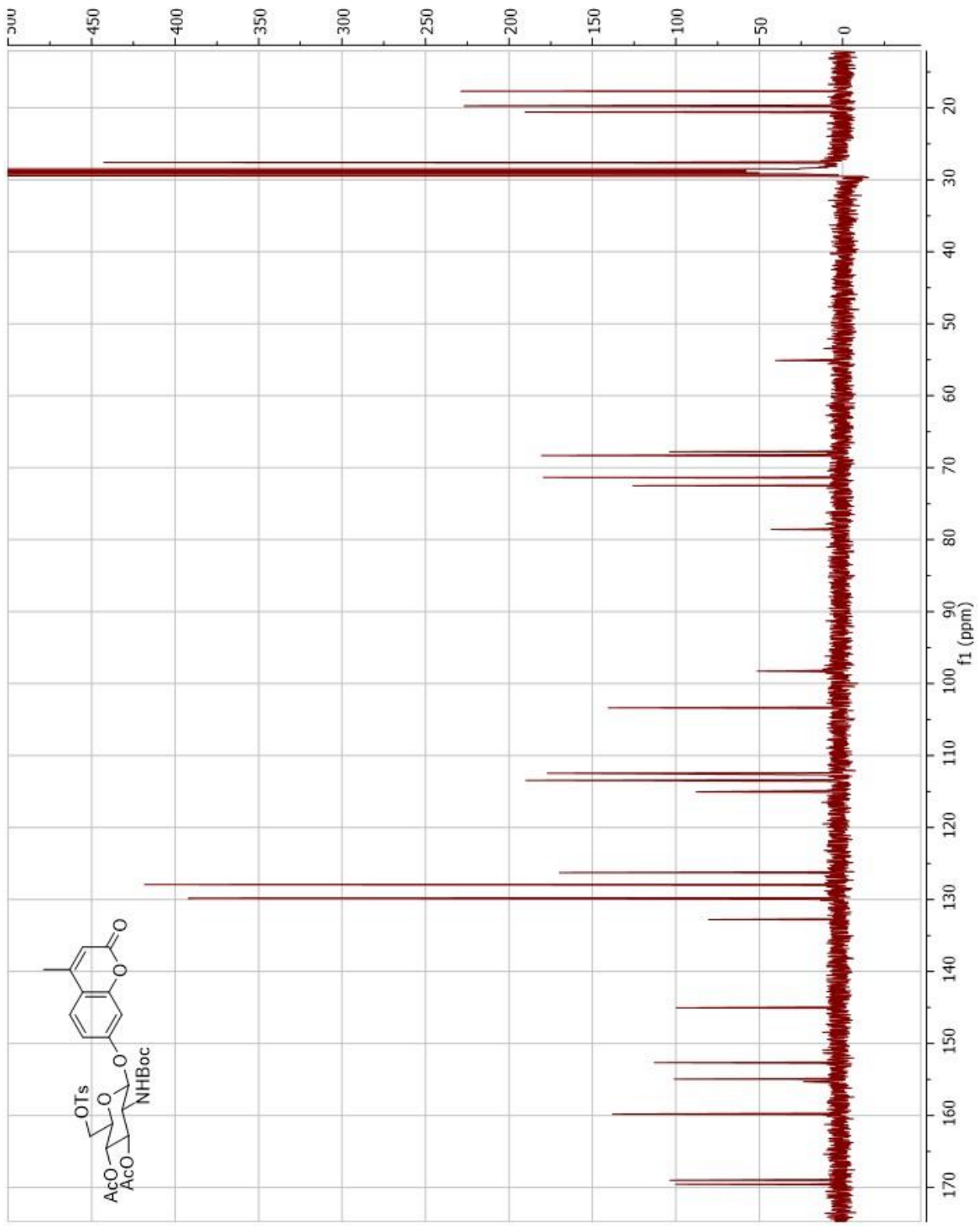

9.01
3.01
3.11
3.07
3.27
1.11
2.23
1.15
1.12
1.14
1.06
1.06
1.00
2.18
1.07
f1 (ppm)
NHBoc
OAc
AcO
AcO

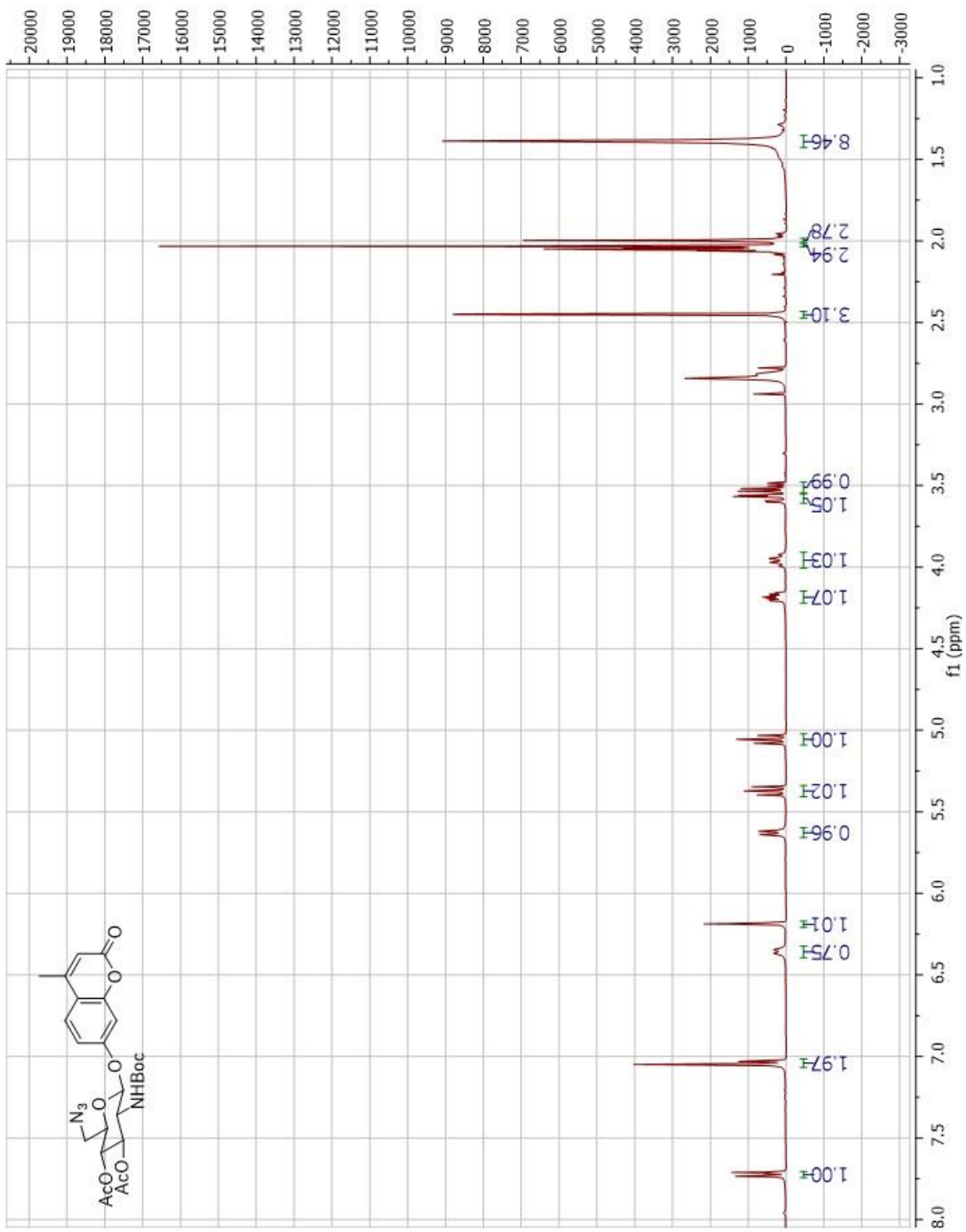

f1 (ppm)
OAc
AcO
AcO
NHBoc

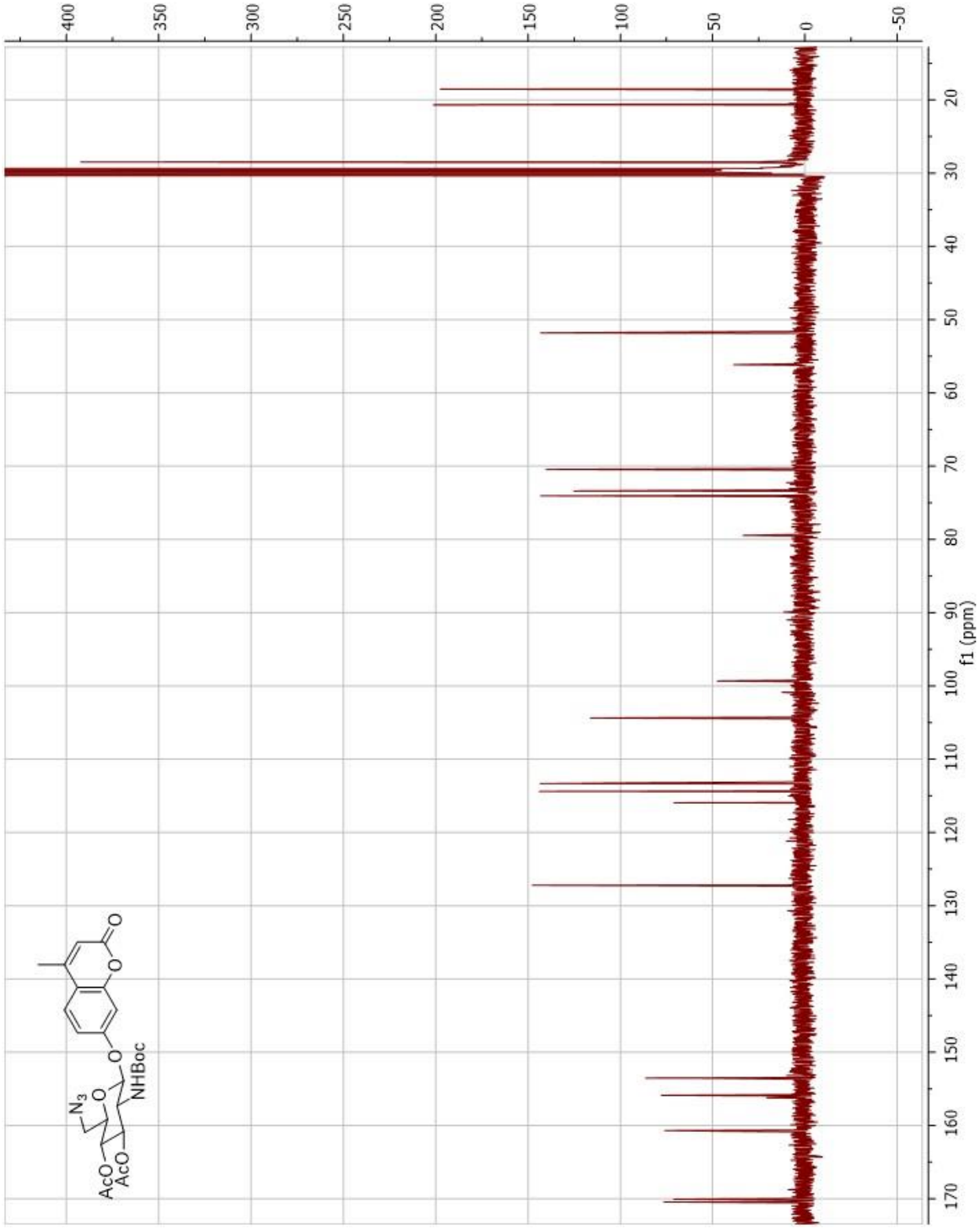

9.01
3.00
1.02
2.02
0.96
0.99
0.99
0.99
0.97
1.96
0.98
f1 (ppm)
NHBoc
OH
HO
HO

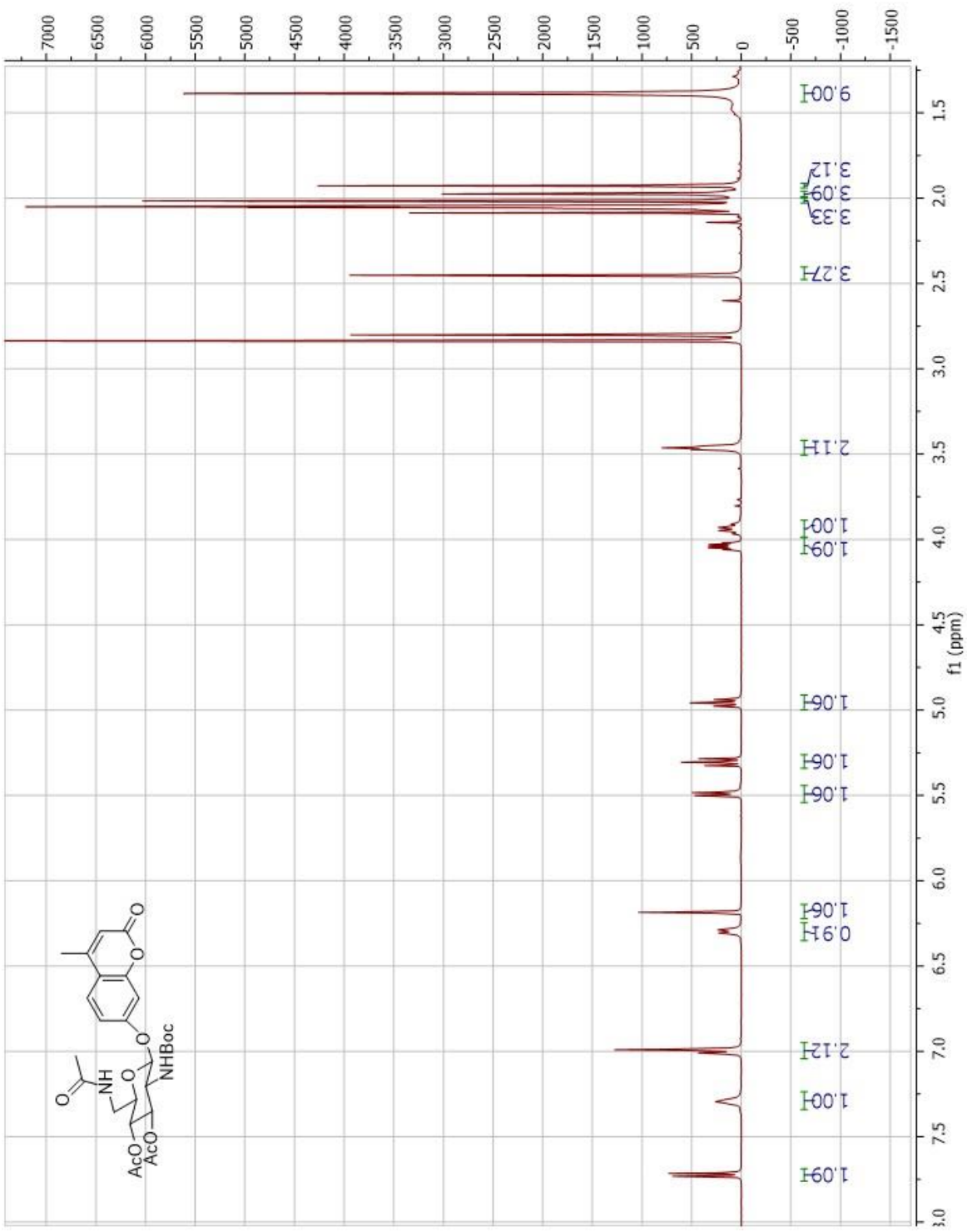

f1 (ppm)
NHBoc
OH
HO
HO

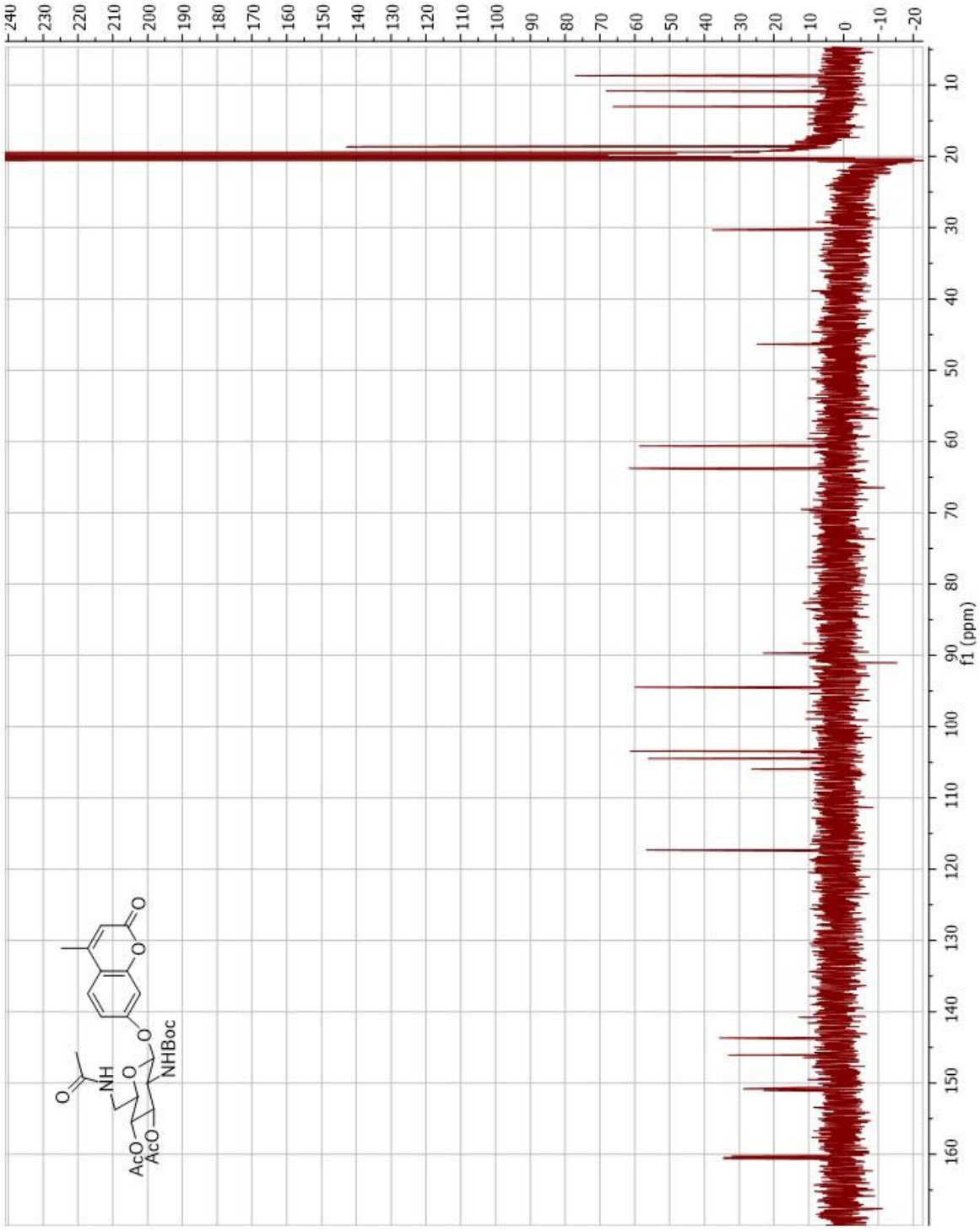

f1 (ppm)
9.00
3.23
2.87
3.09
3.29
1.03
2.21
1.06
1.06
1.06
0.99
0.67
1.01
1.05
1.02
2.04
2.09
1.07
OTs
NHBoc
AcO
AcO

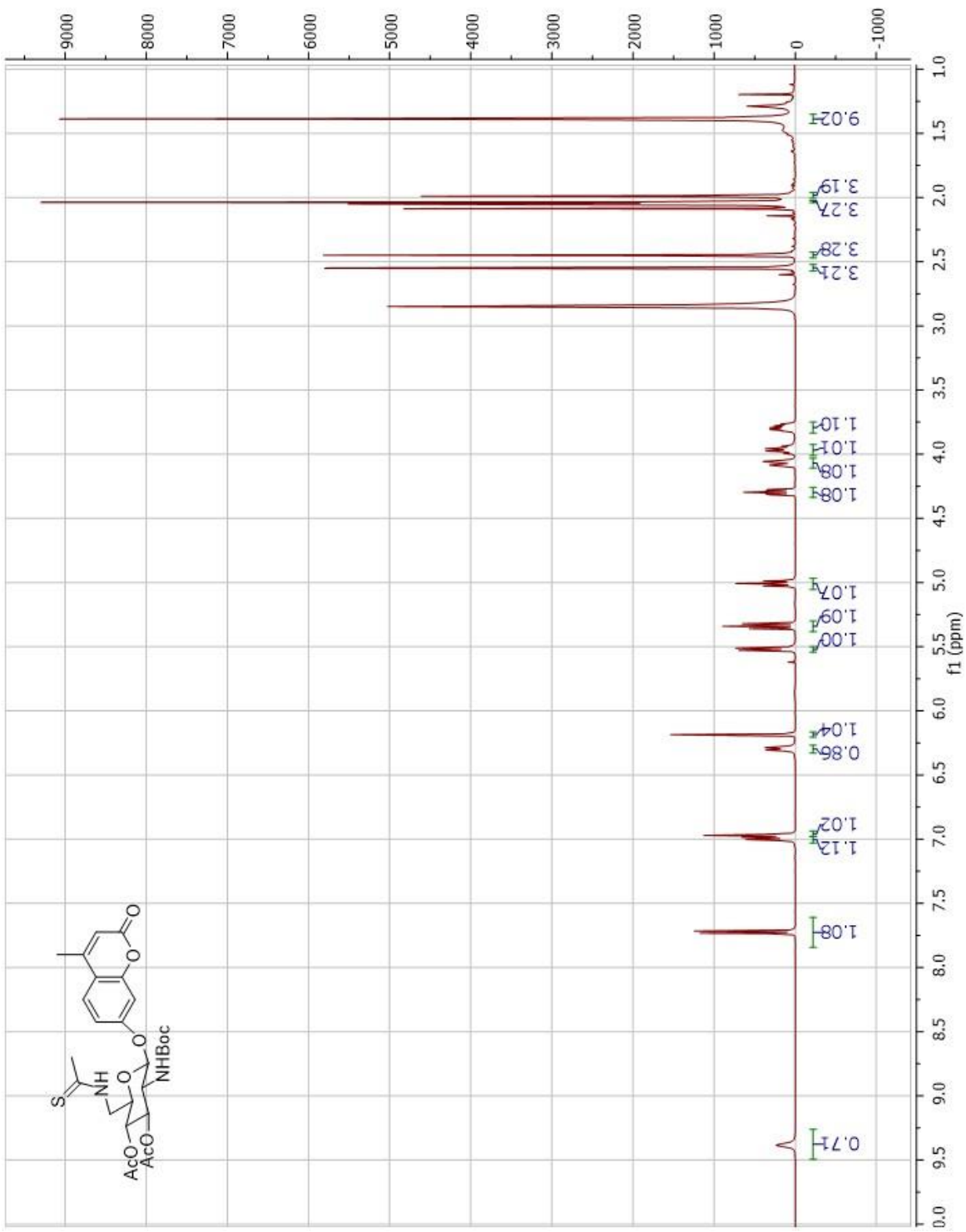

f1 (ppm)
OTs
NHBoc
AcO
AcO

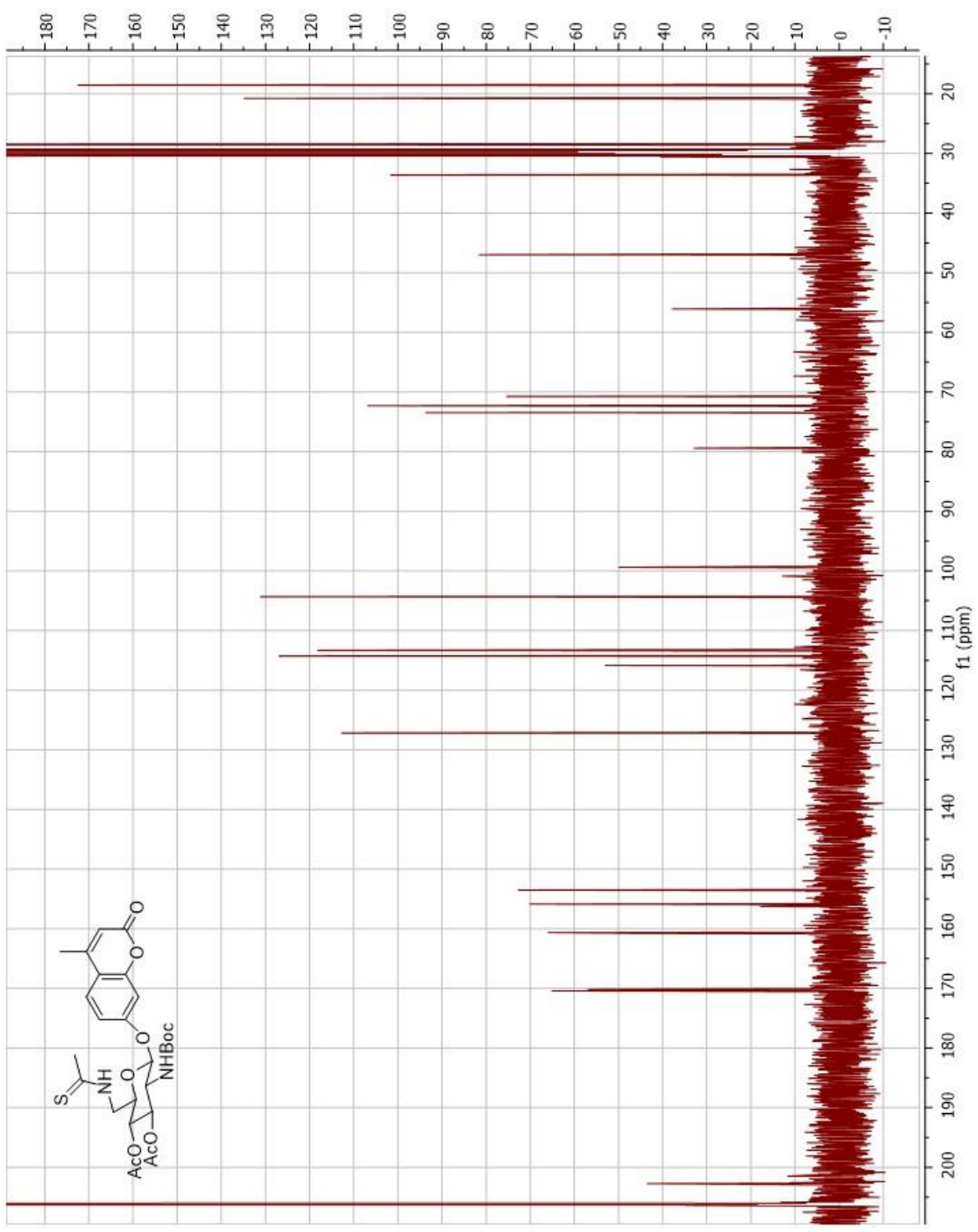

8.46
2.94
2.78
3.10
0.99
1.05
1.03
1.07
1.00
1.02
0.96
1.01
0.75
1.97
1.00
f1 (ppm)
AcO
AcO
N3
NHBoc

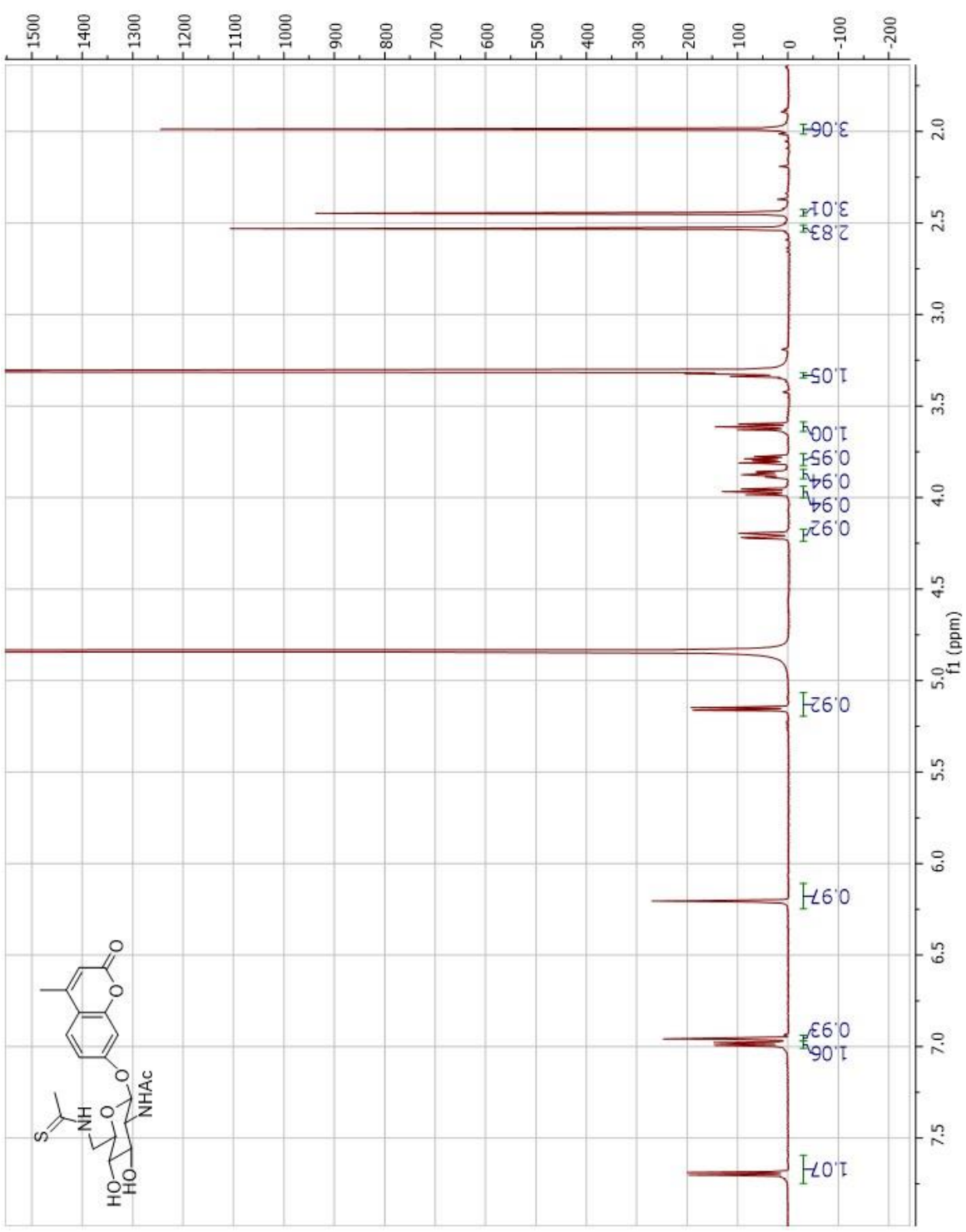

f1 (ppm)
AcO
AcO
N3
NHBoc

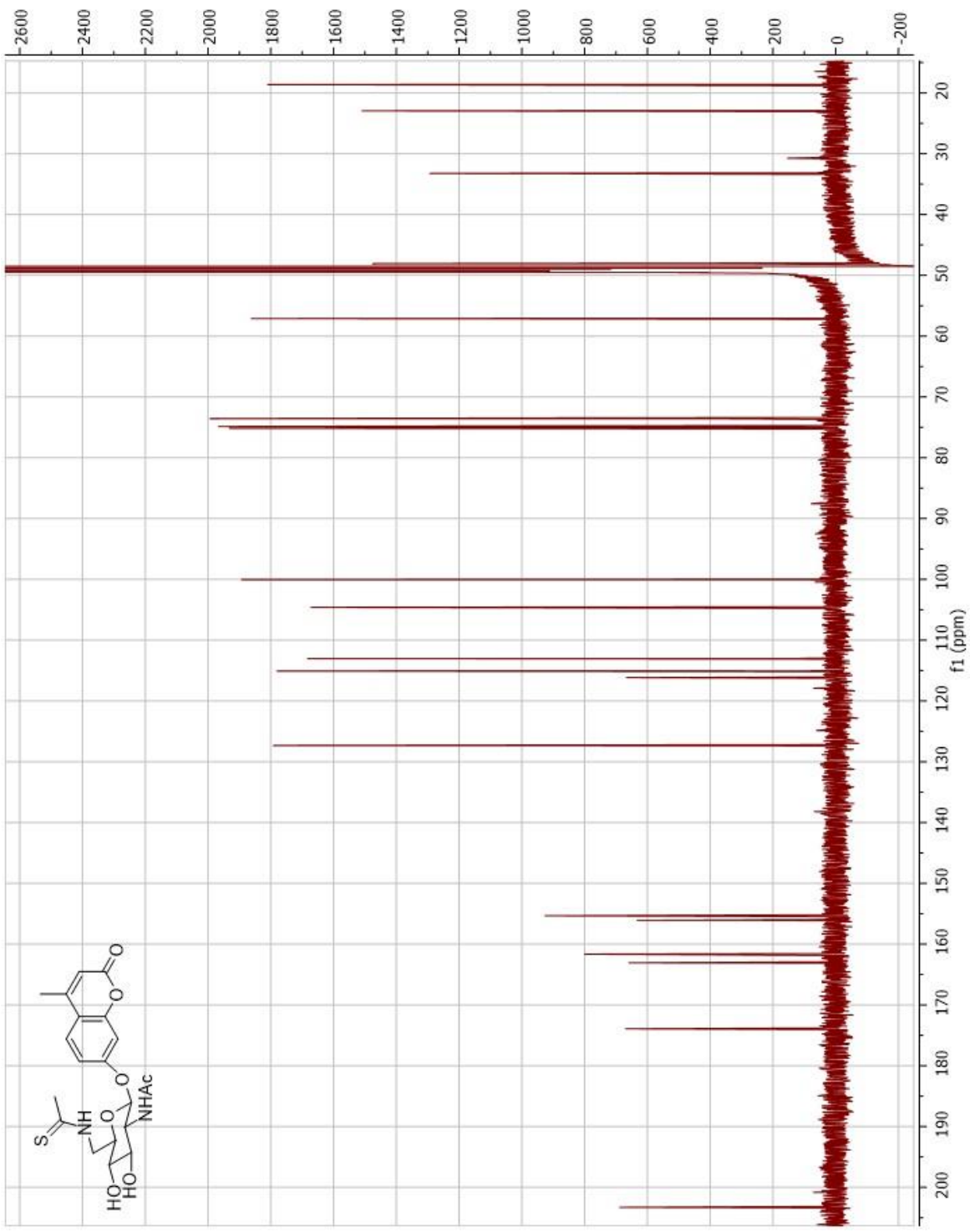

9.00
3.12
3.09
3.33
3.27
2.11
1.00
1.09
1.06
1.06
1.06
1.06
0.91
2.12
1.00
1.09
f1 (ppm)
NHBoc
AcO
AcO
NH

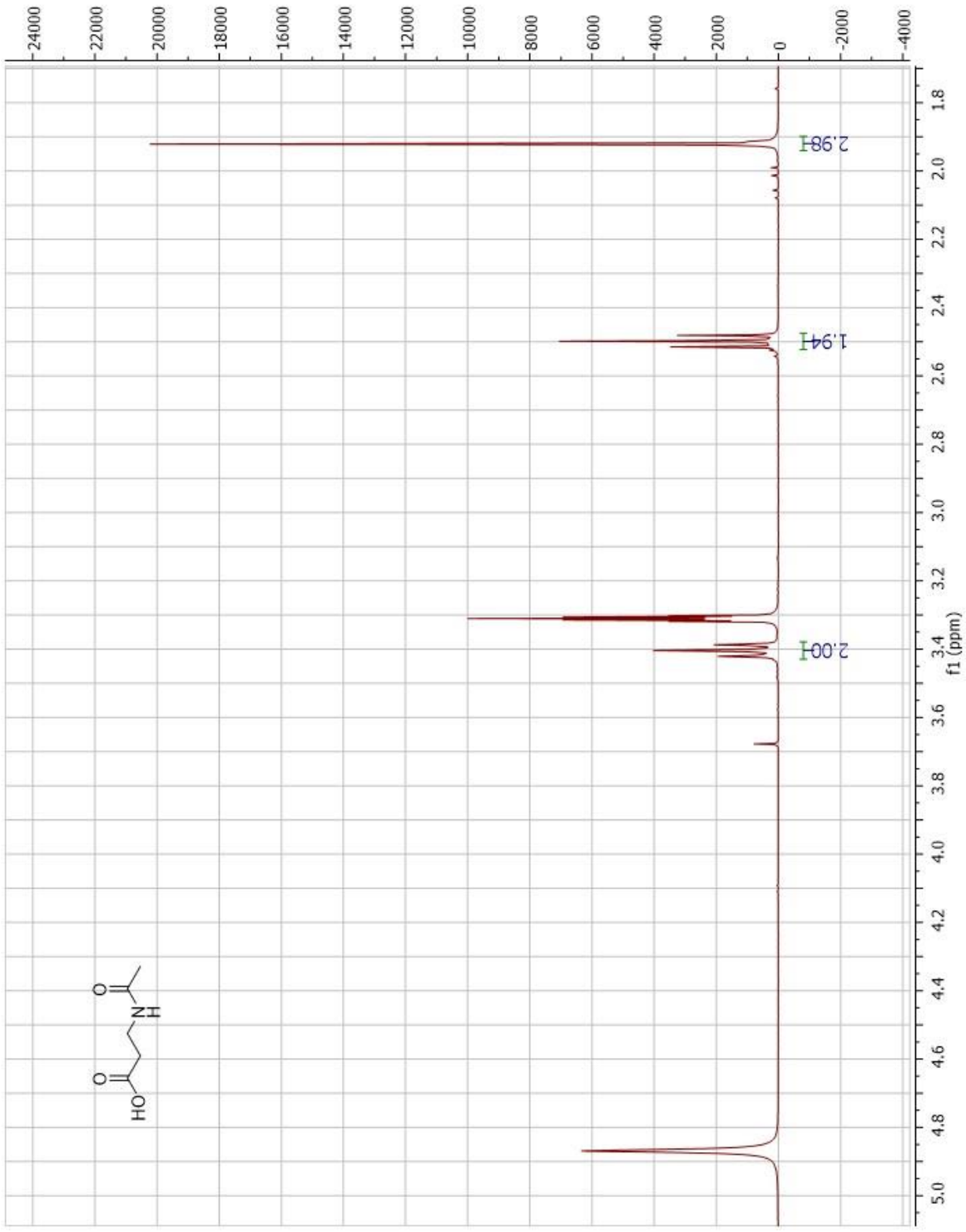

f1 (ppm)
NHBoc
AcO
AcO
NH

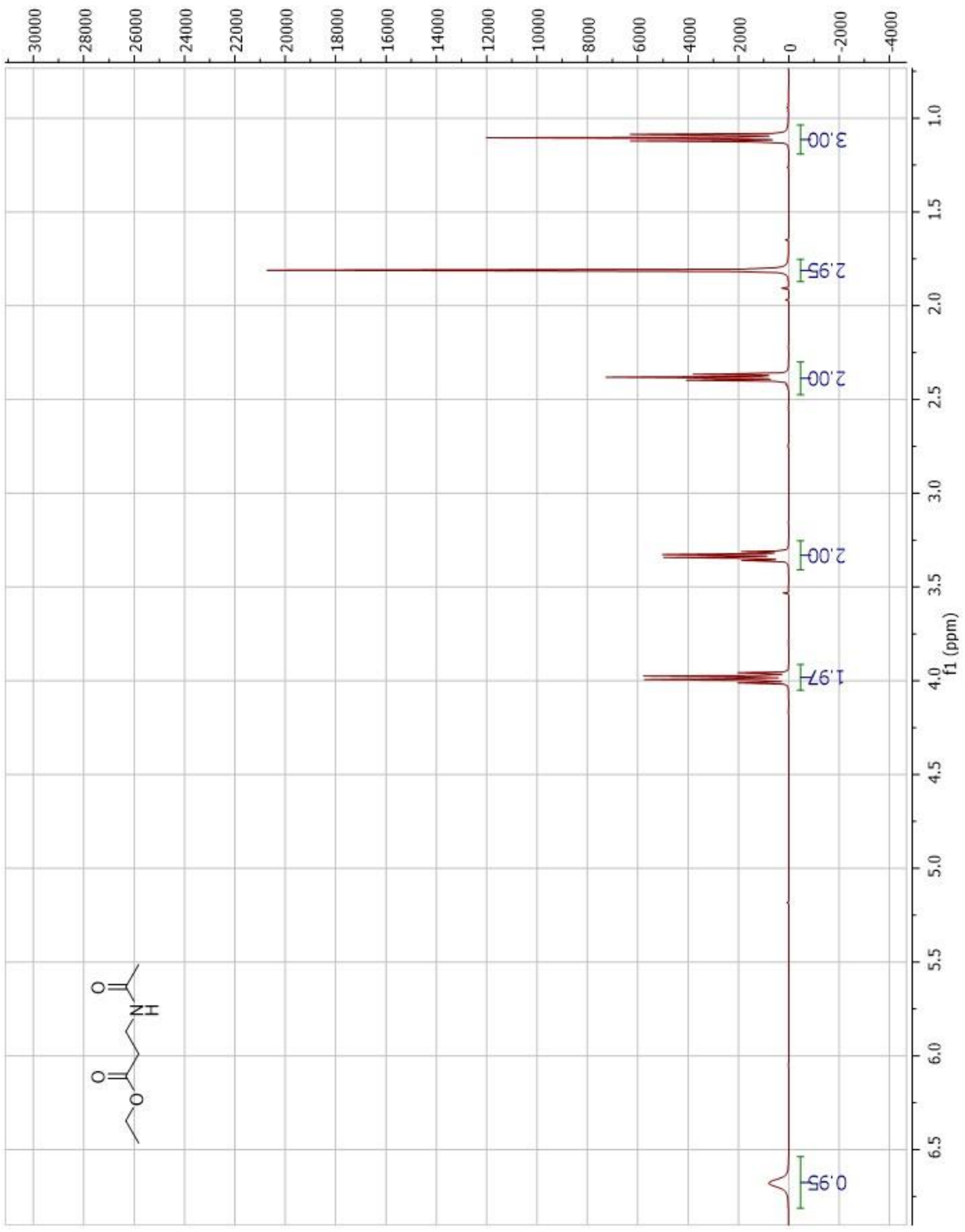

9.02
3.19
3.27
3.28
3.21
1.10
1.01
1.08
1.08
1.07
1.09
1.00
0.86
1.04
1.12
1.02
1.08
0.71
f1 (ppm)
AcO
AcO
NH
NHBoc

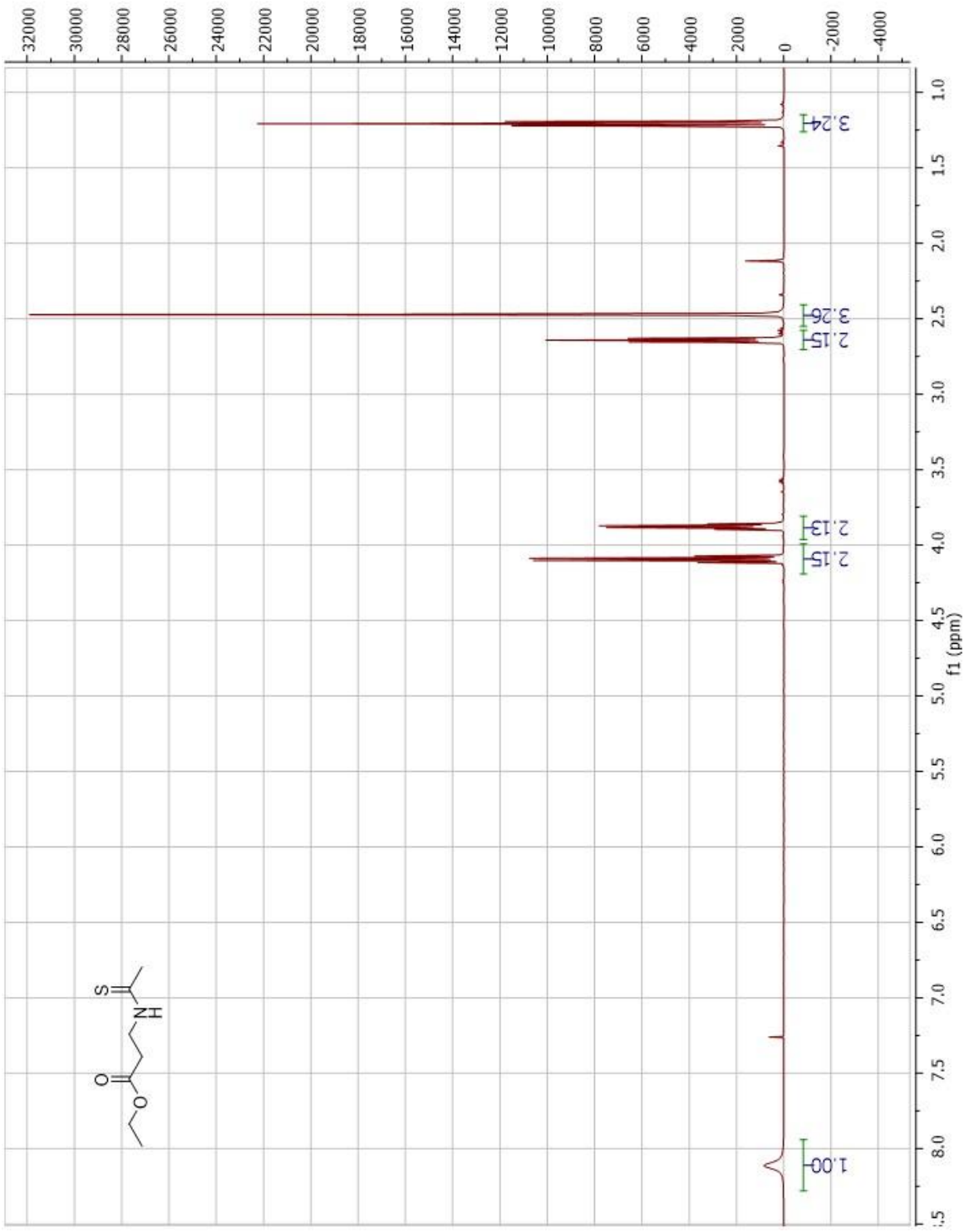

f1 (ppm)
NHBoc
AcO
AcO
NH
S
O

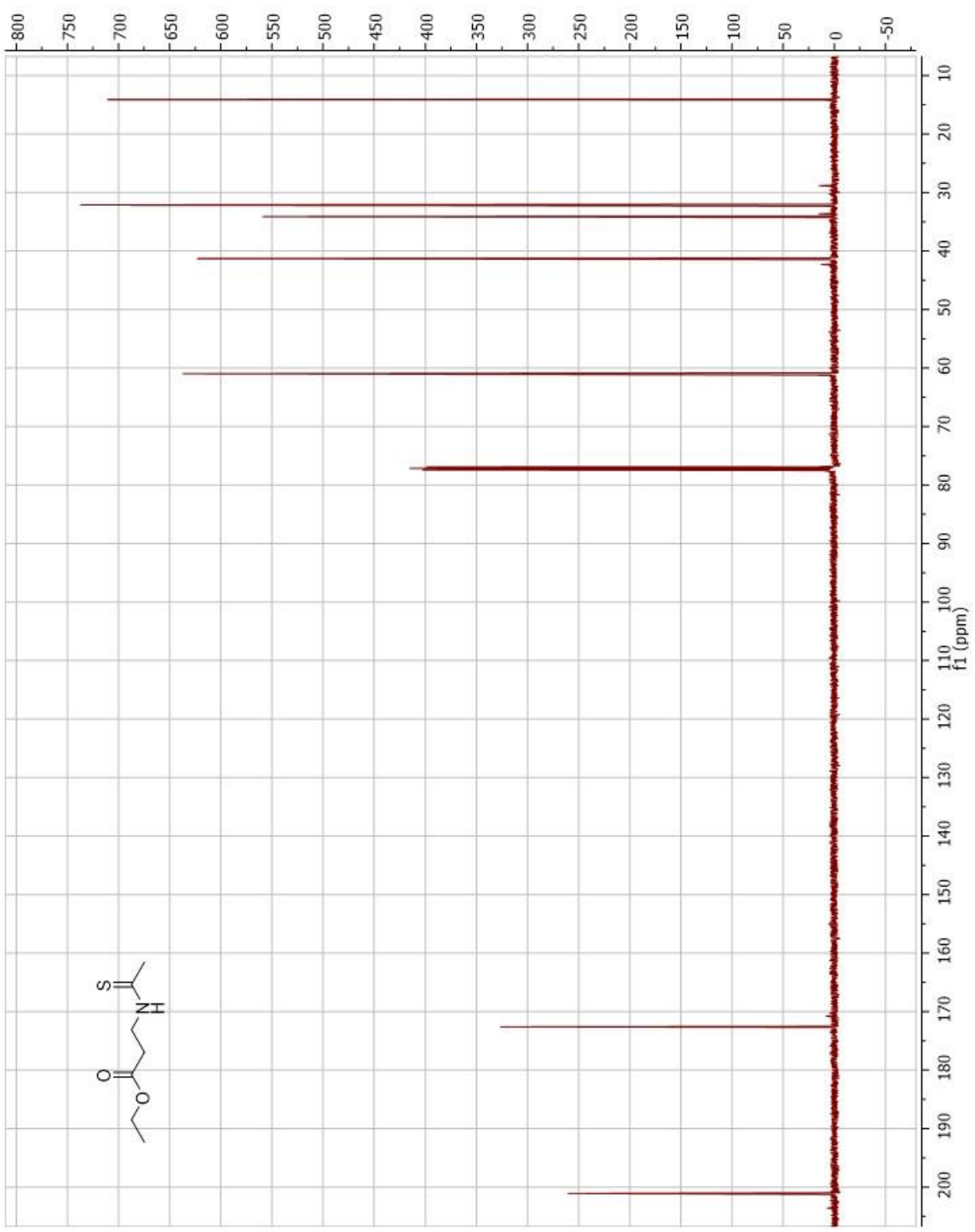

f1 (ppm)
3.06
3.01
2.83
1.05
1.00
0.95
0.94
0.94
0.92
0.92
0.97
1.06
0.93
1.07
NHAc
NH
HO
HO

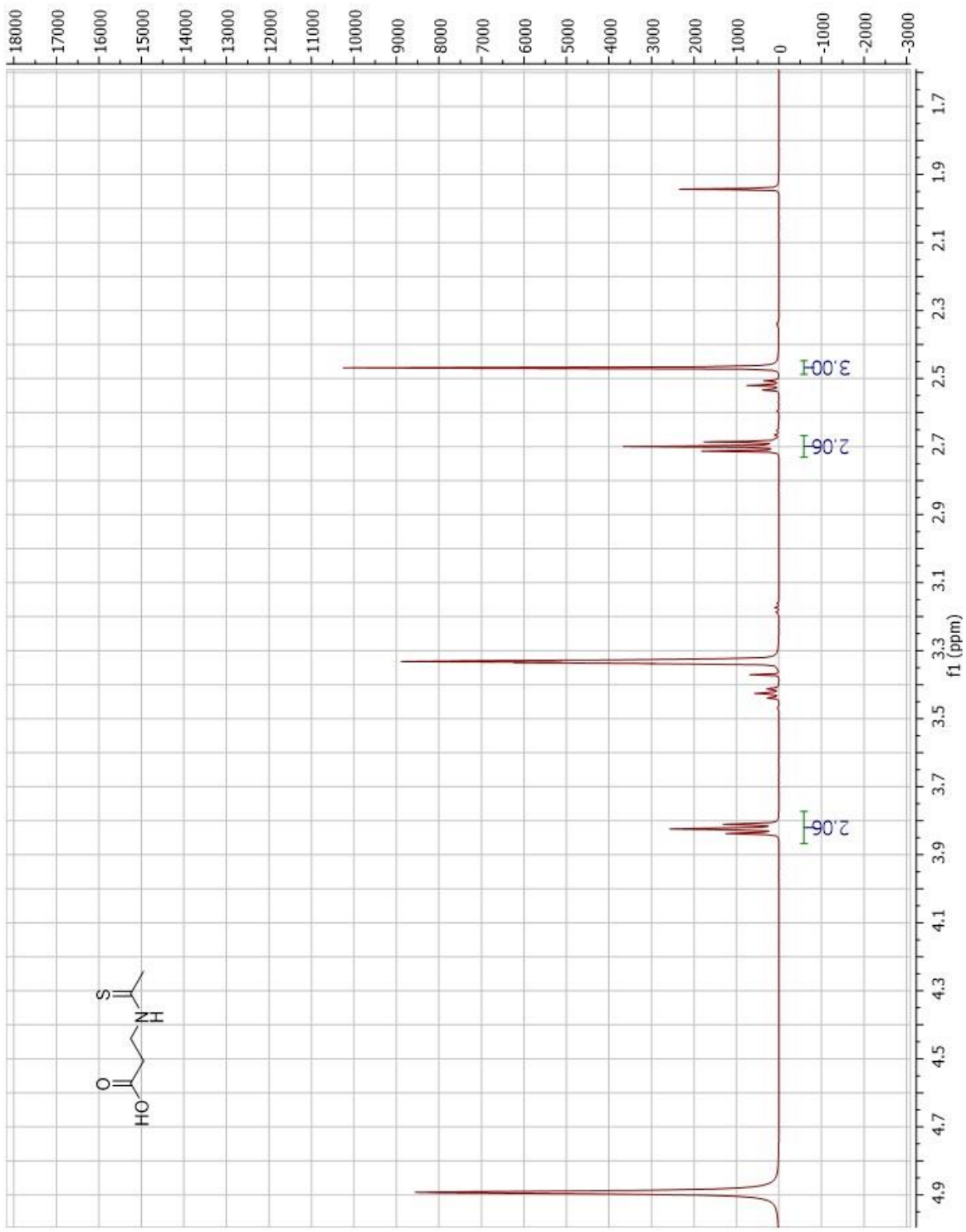

f1 (ppm)
NHAc
HO
HO
NH
S
O

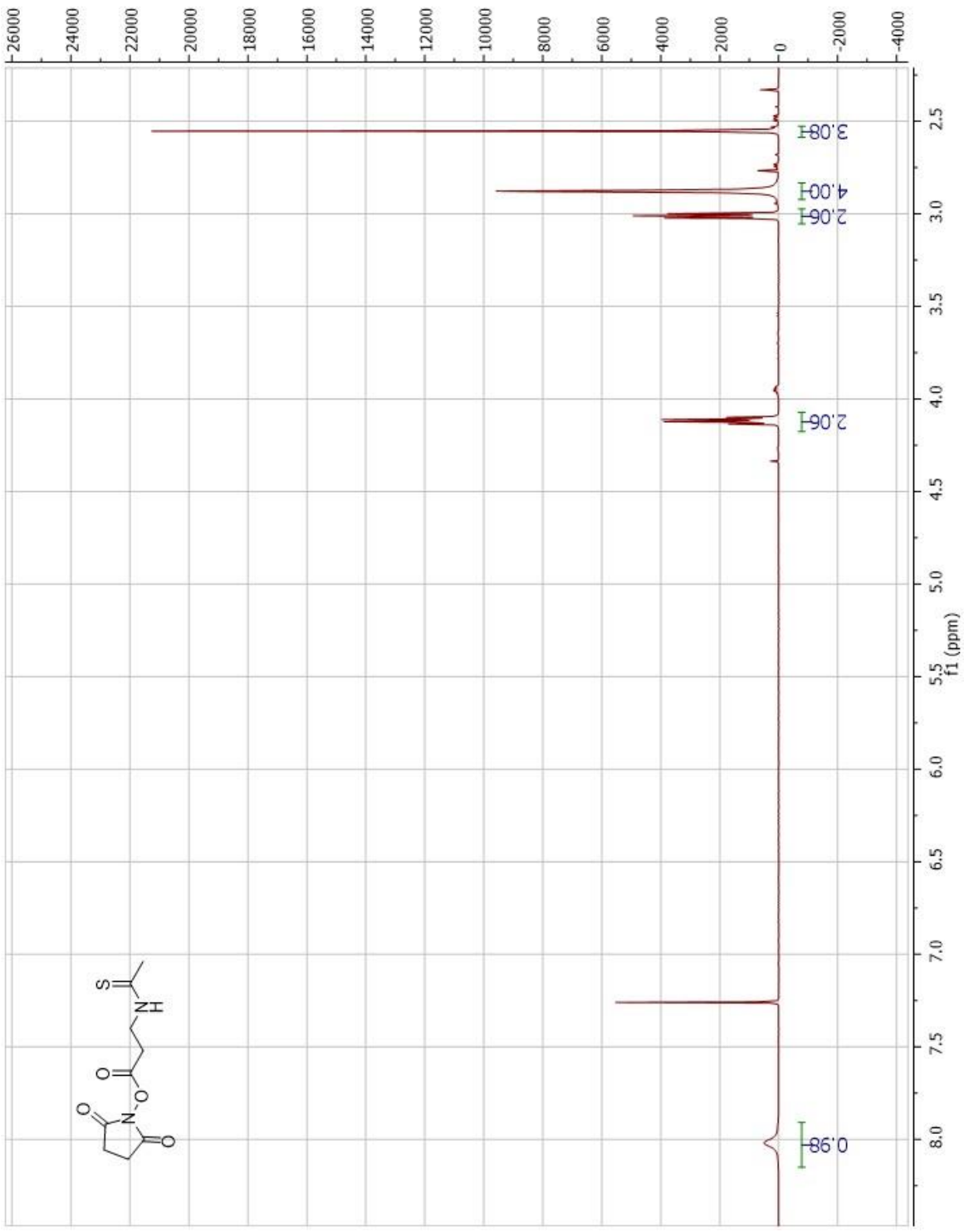

2.98
1.94
2.00
f1 (ppm)

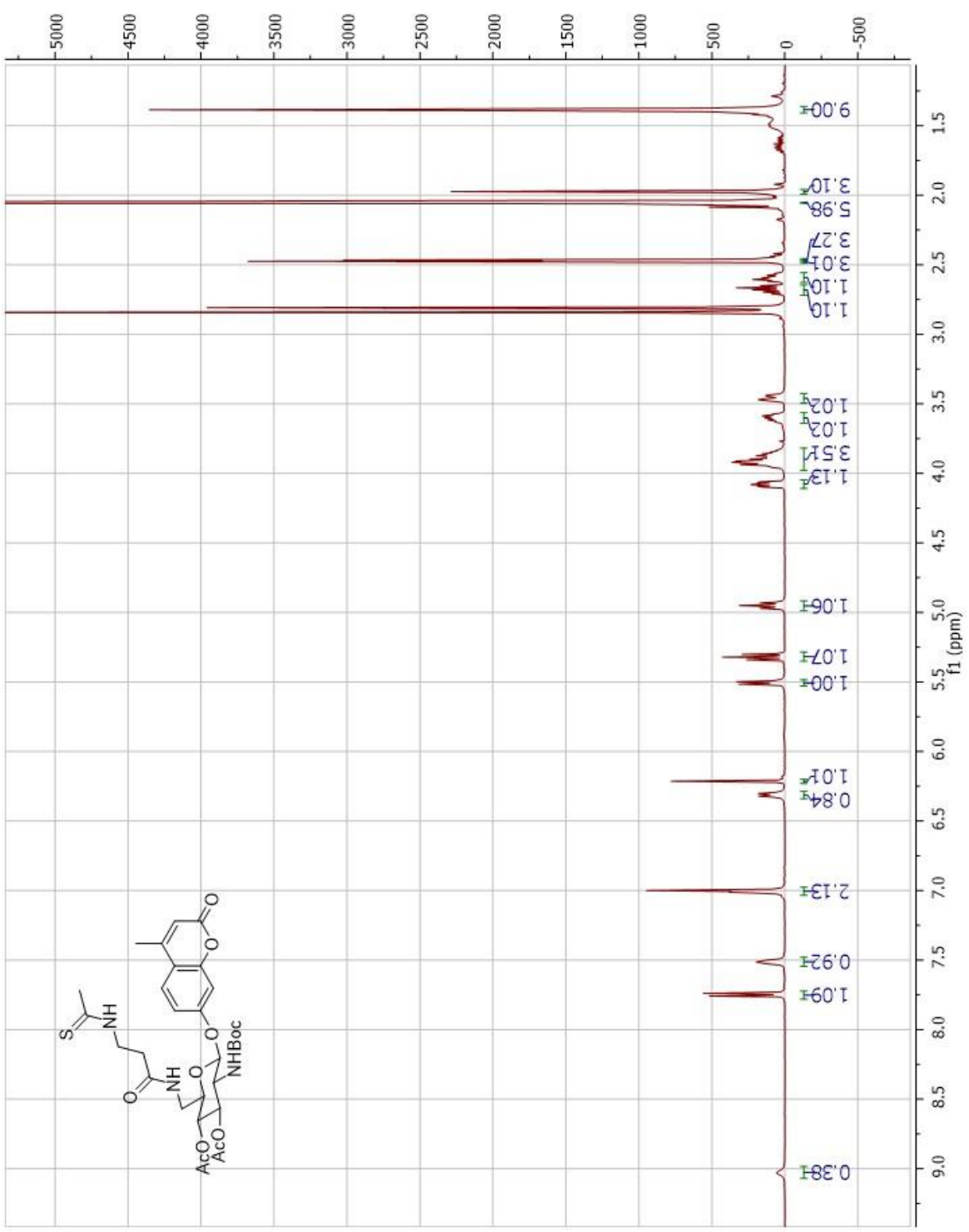

30000
28000
26000
24000
22000
20000
18000
16000
14000
12000
10000
8000
6000
4000
2000
0
-2000
-4000
3.00
2.95
2.00
2.00
1.97
0.95
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
6.5
f1 (ppm)
O
O
O
N
H

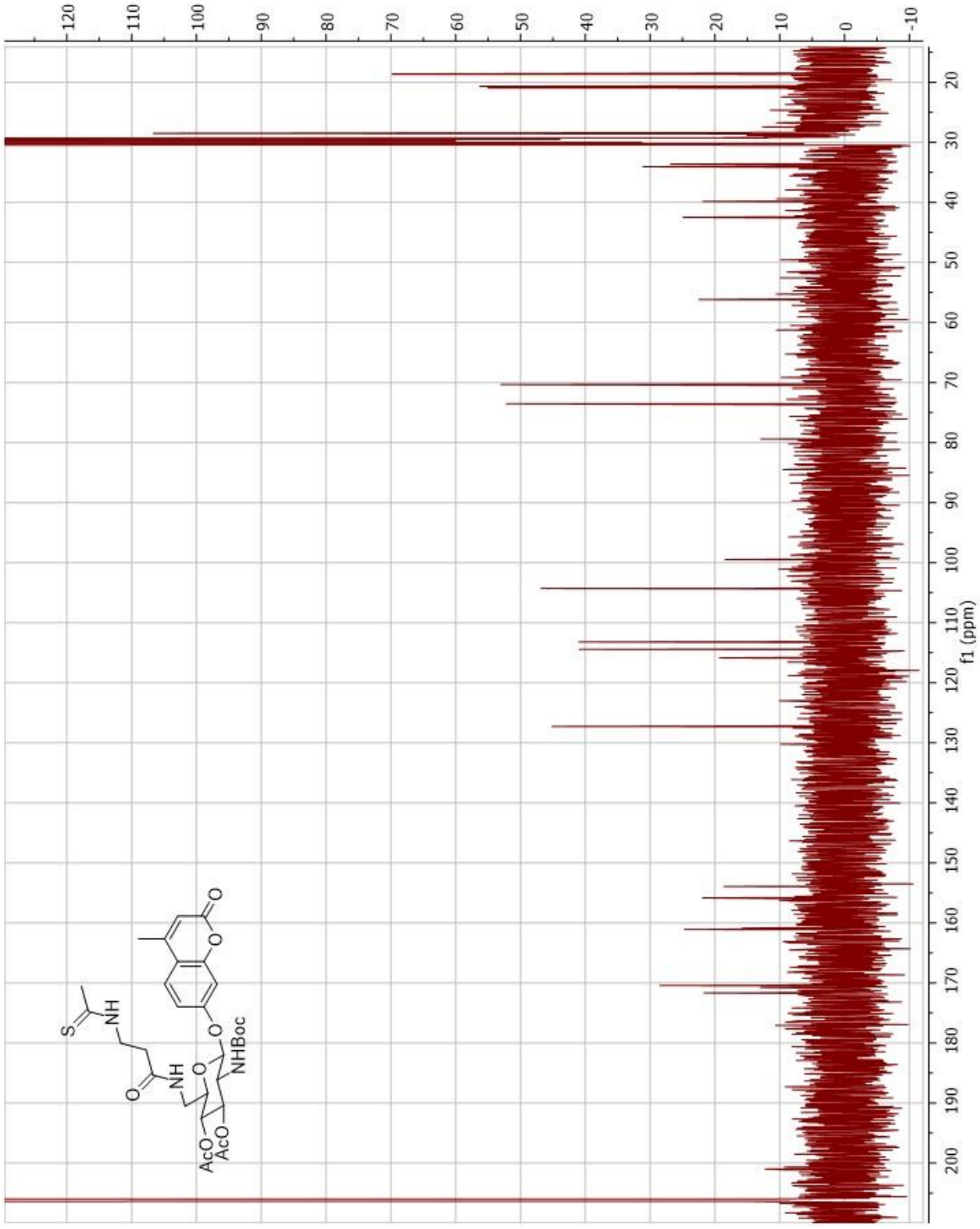

3.24
3.26
2.15
2.13
2.15
1.00
f1 (ppm)

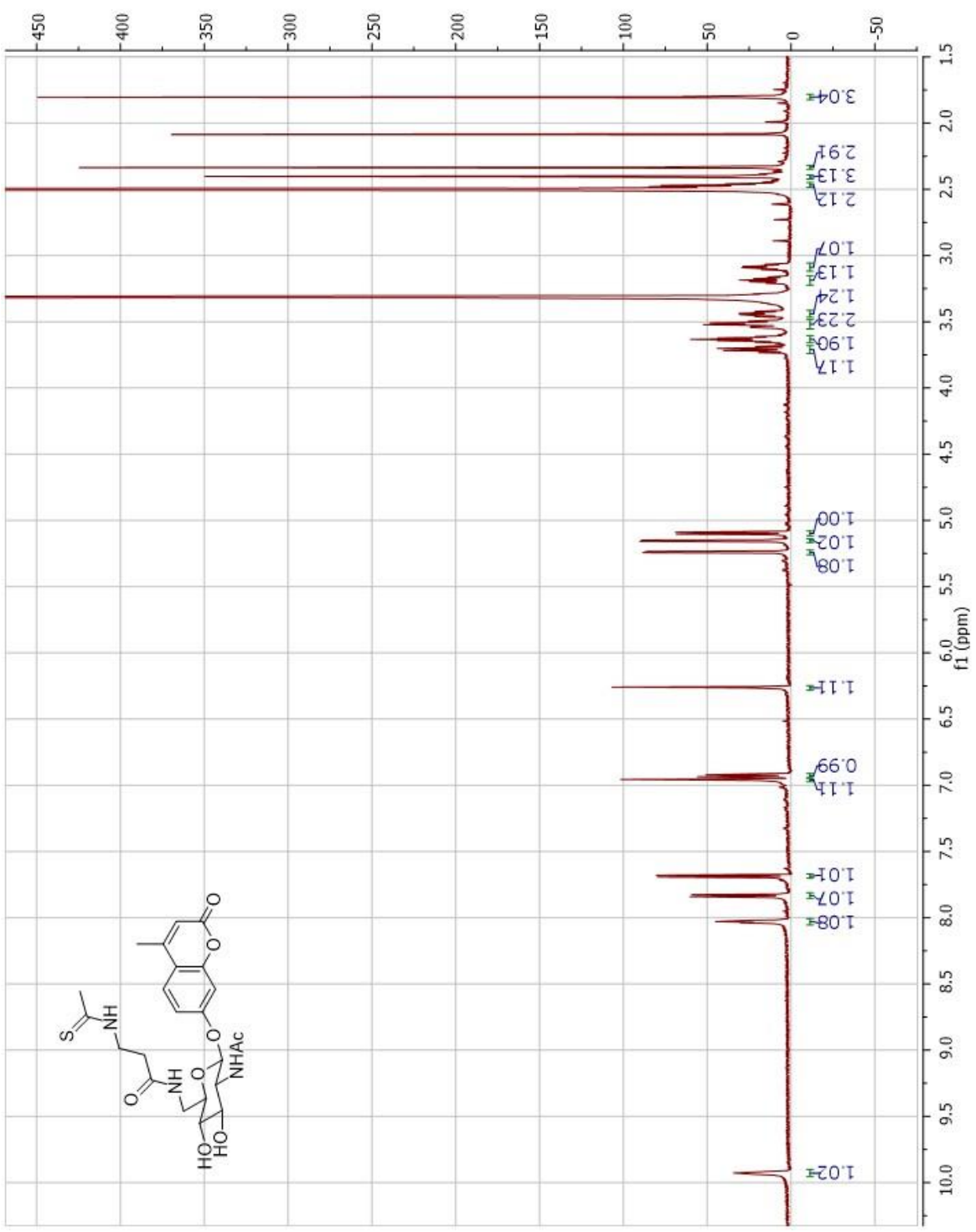

f1 (ppm)

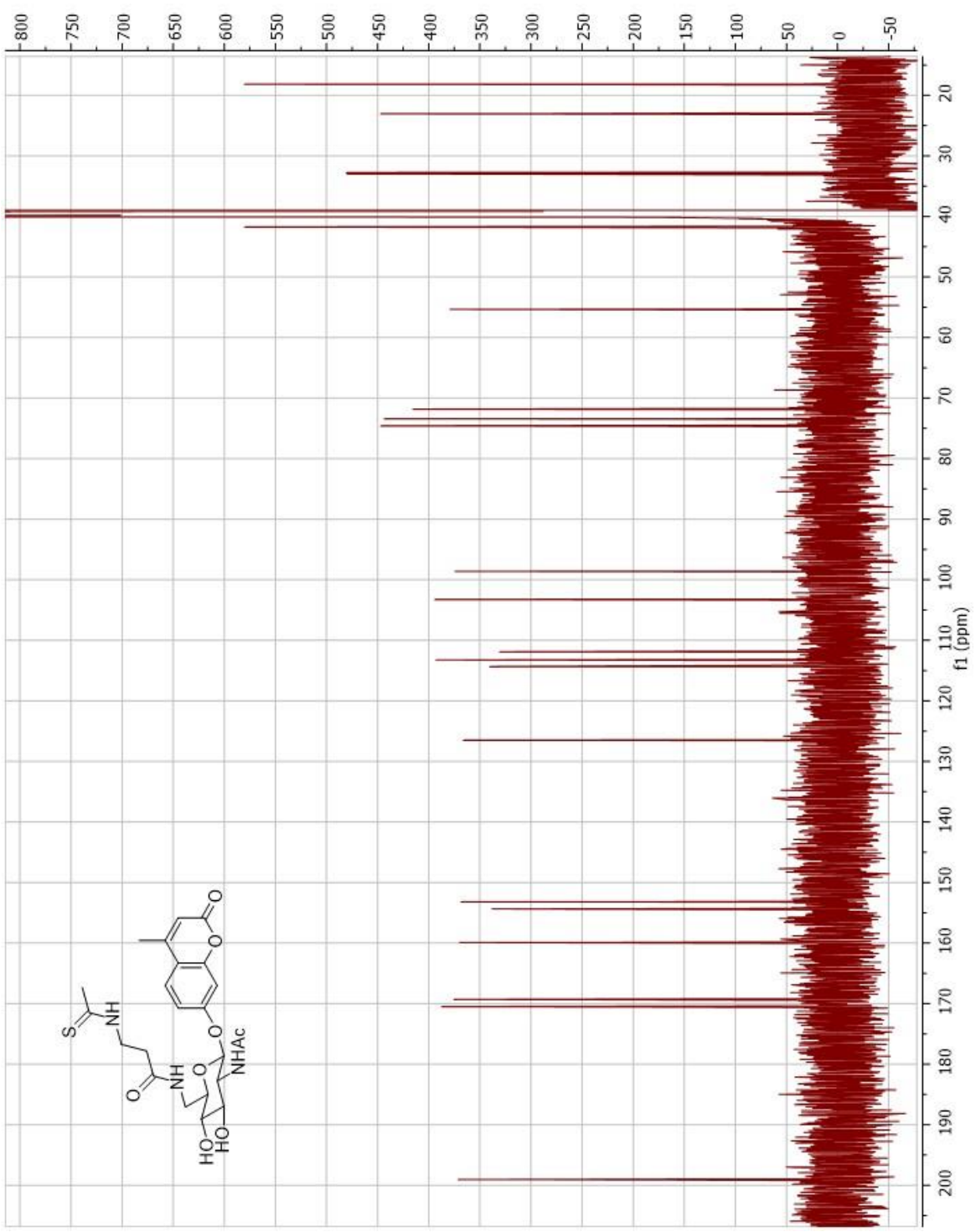

18000
17000
16000
15000
14000
13000
12000
11000
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
-1000
-2000
-3000
1.7
1.9
2.1
2.3
2.5
2.7
2.9
3.1
3.3
3.5
3.7
3.9
4.1
4.3
4.5
4.7
4.9
f1 (ppm)
3.00
2.06
2.06
HO
O
N
H
S

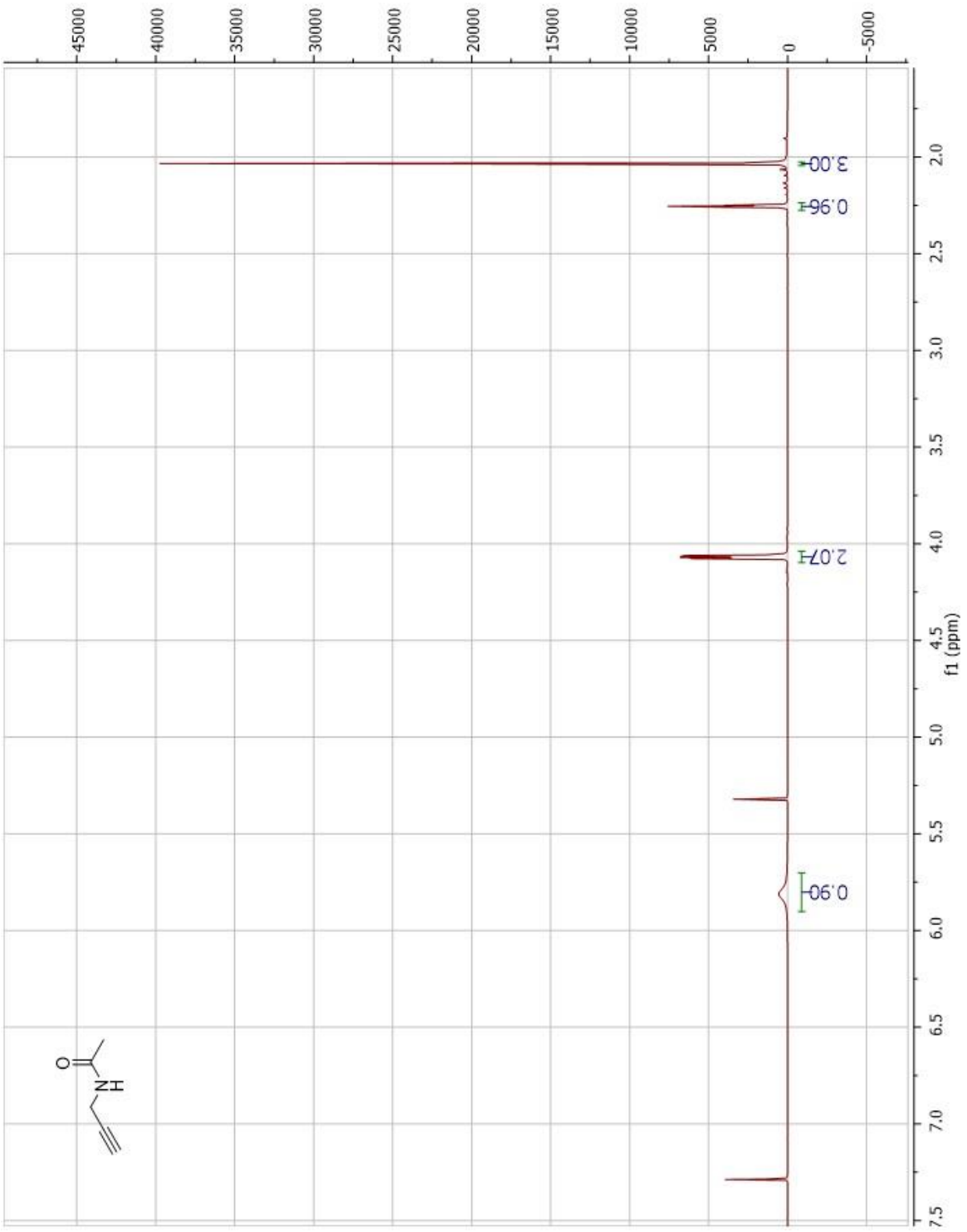

3.08
4.00
2.06
2.06
0.98
f1 (ppm)

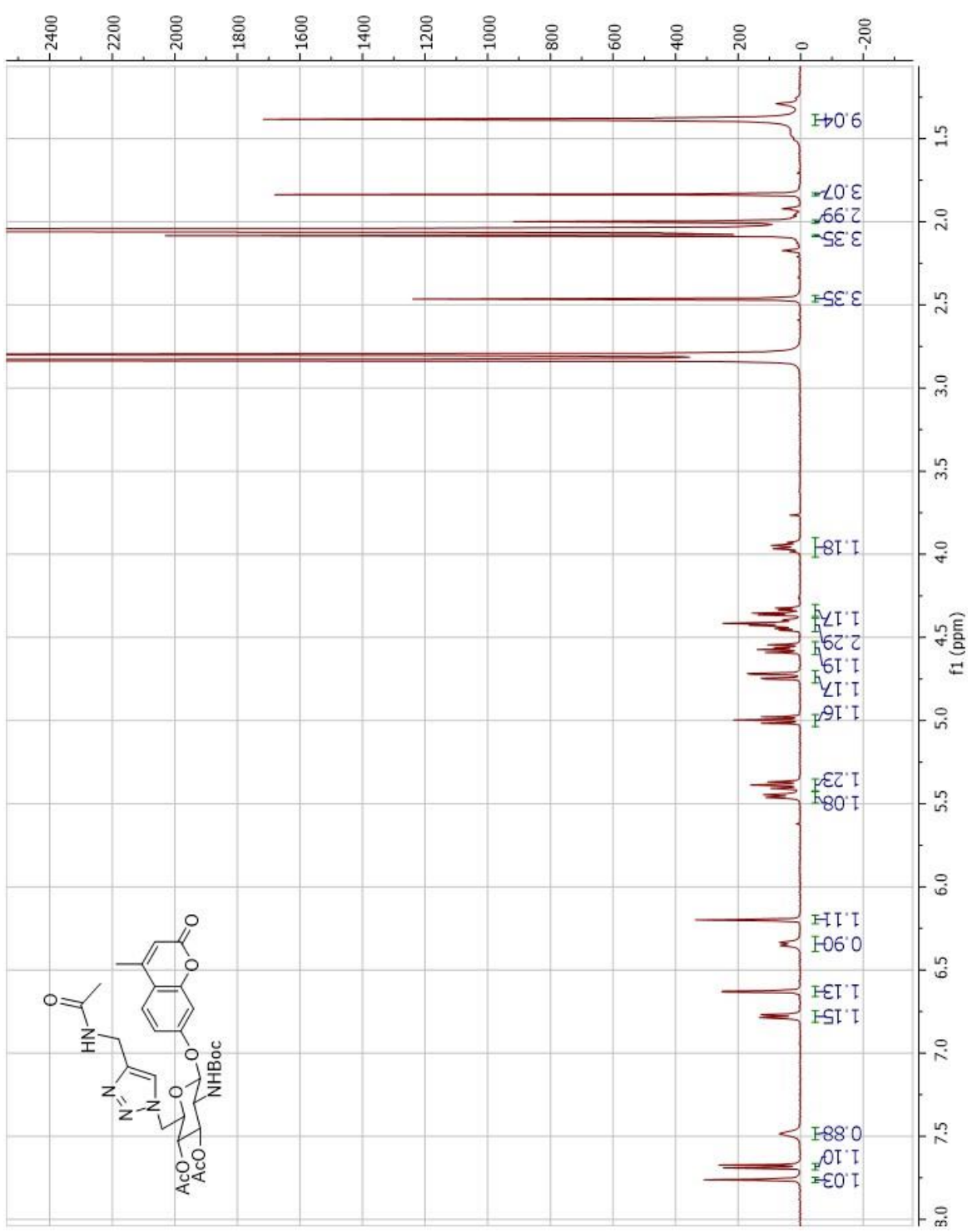

f1 (ppm)
9.00
3.10
5.98
3.27
3.01
1.10
1.10
1.02
1.02
3.51
1.13
1.06
1.00
1.07
0.84
1.01
2.13
0.92
1.09
0.38
NHBoc
AcO
AcO

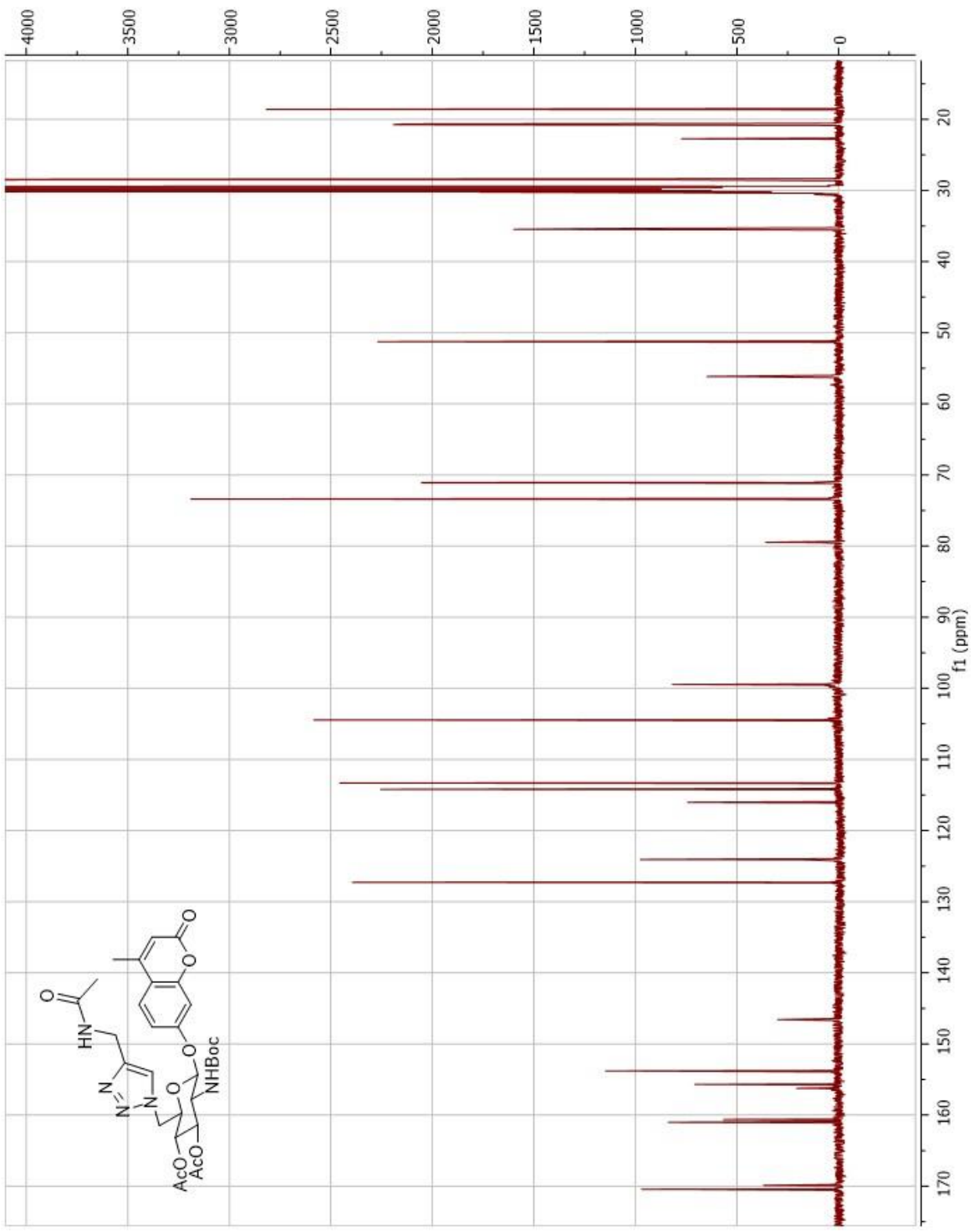

f1 (ppm)
NHBoc
AcO
AcO
NH
NH
S
O

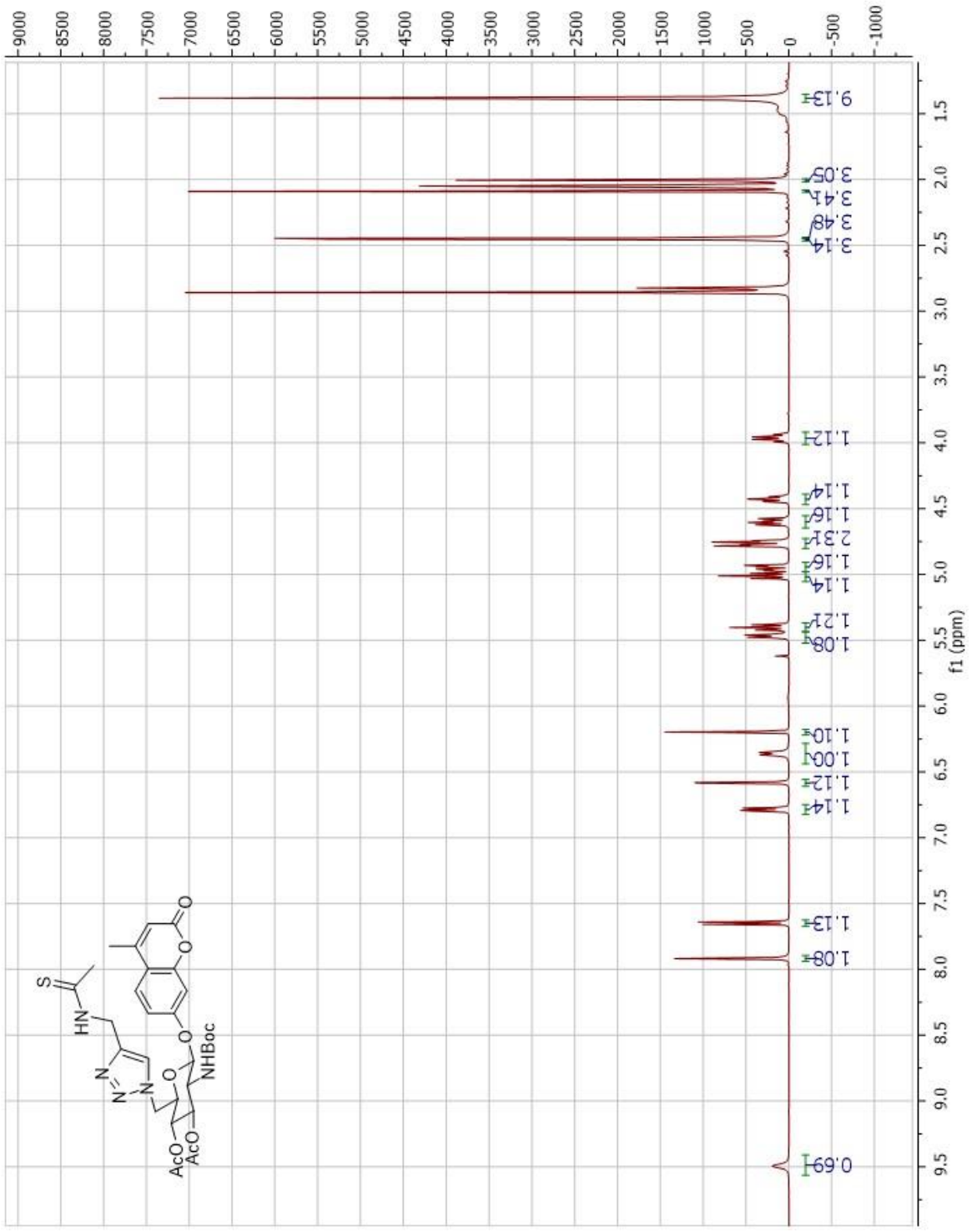

f1 (ppm)
NHAc
HO
HO
NH
NH
S

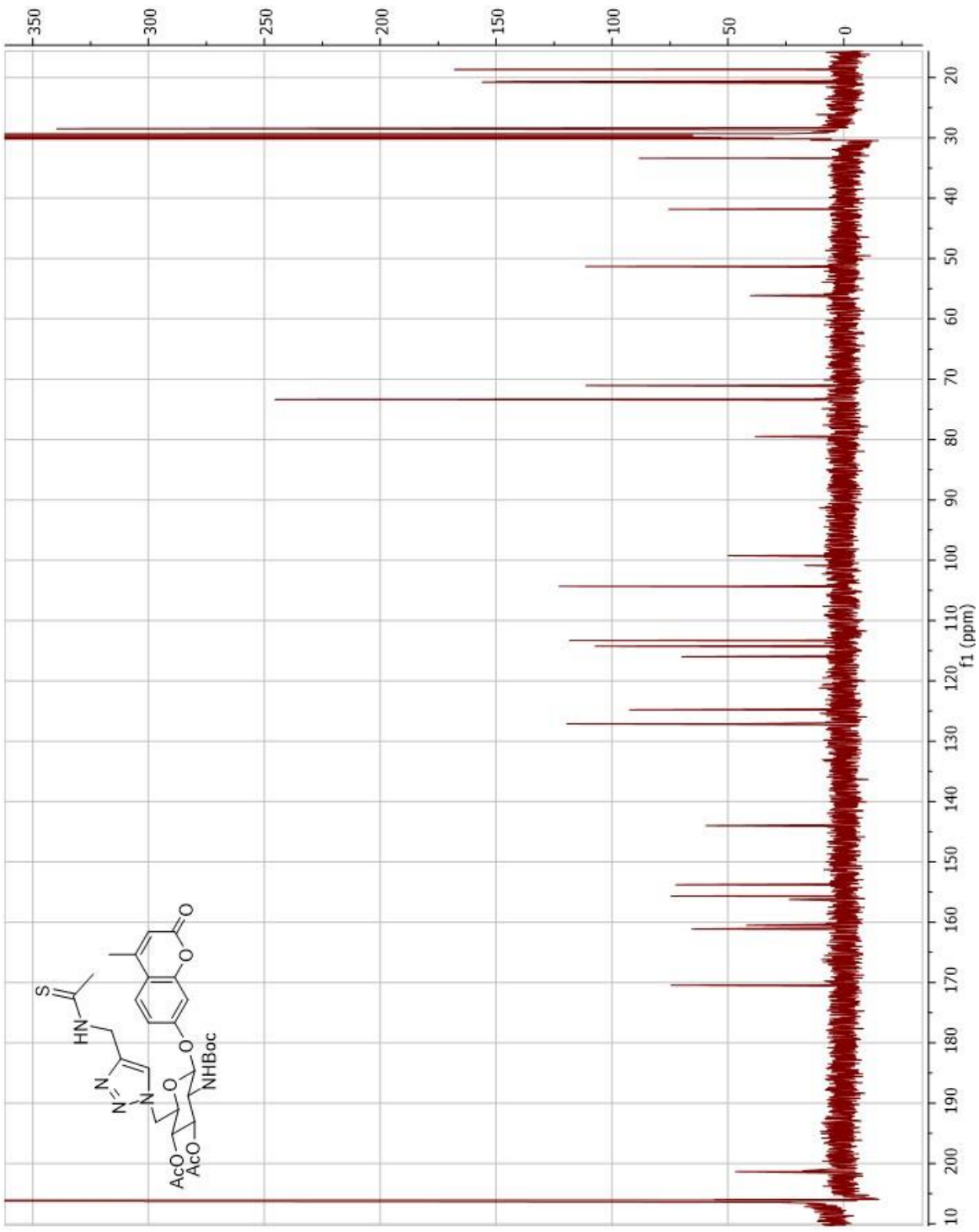

f1 (ppm)

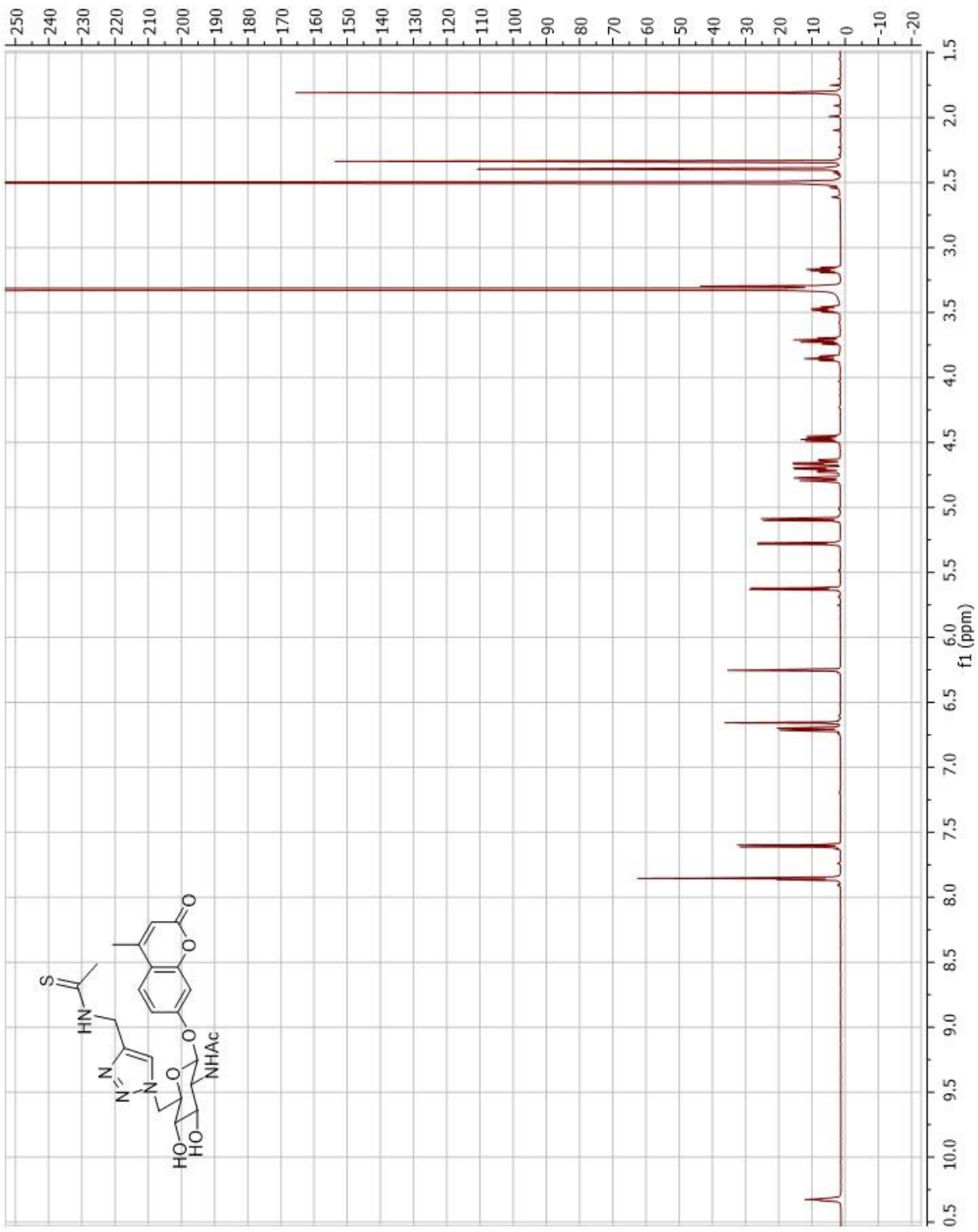

45000
40000
35000
30000
25000
20000
15000
10000
5000
0
-5000
3.00
0.96
2.07
0.90
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
6.5
7.0
7.5
f1 (ppm)
O
N
H

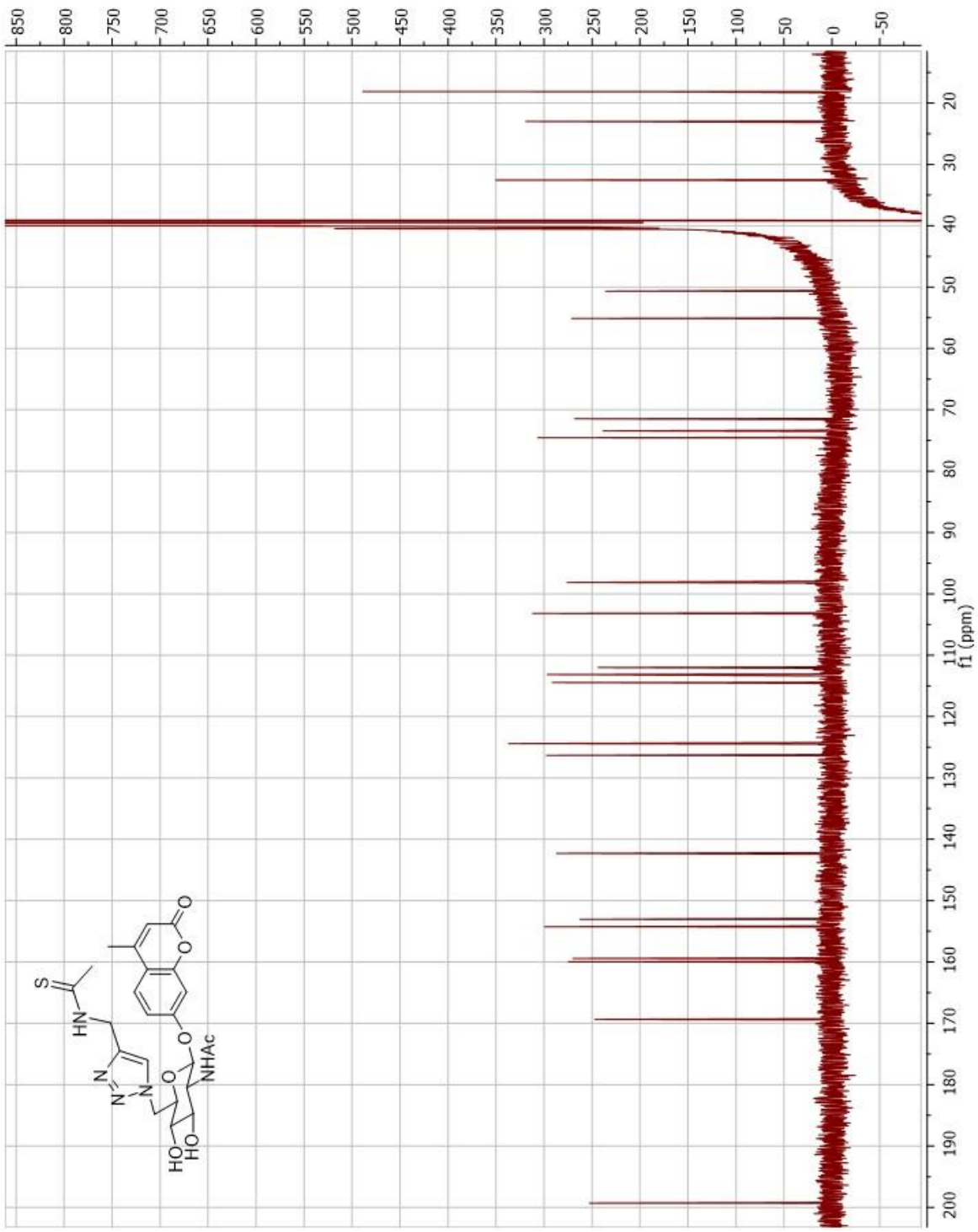

9.04
3.07
2.99
3.35
3.35
1.18
1.17
2.29
1.19
1.17
1.16
1.08
1.23
1.11
0.90
1.13
1.15
0.88
1.03
1.10
f1 (ppm)
AcO
AcO
NHBoc
HN
O
N
N
O
O

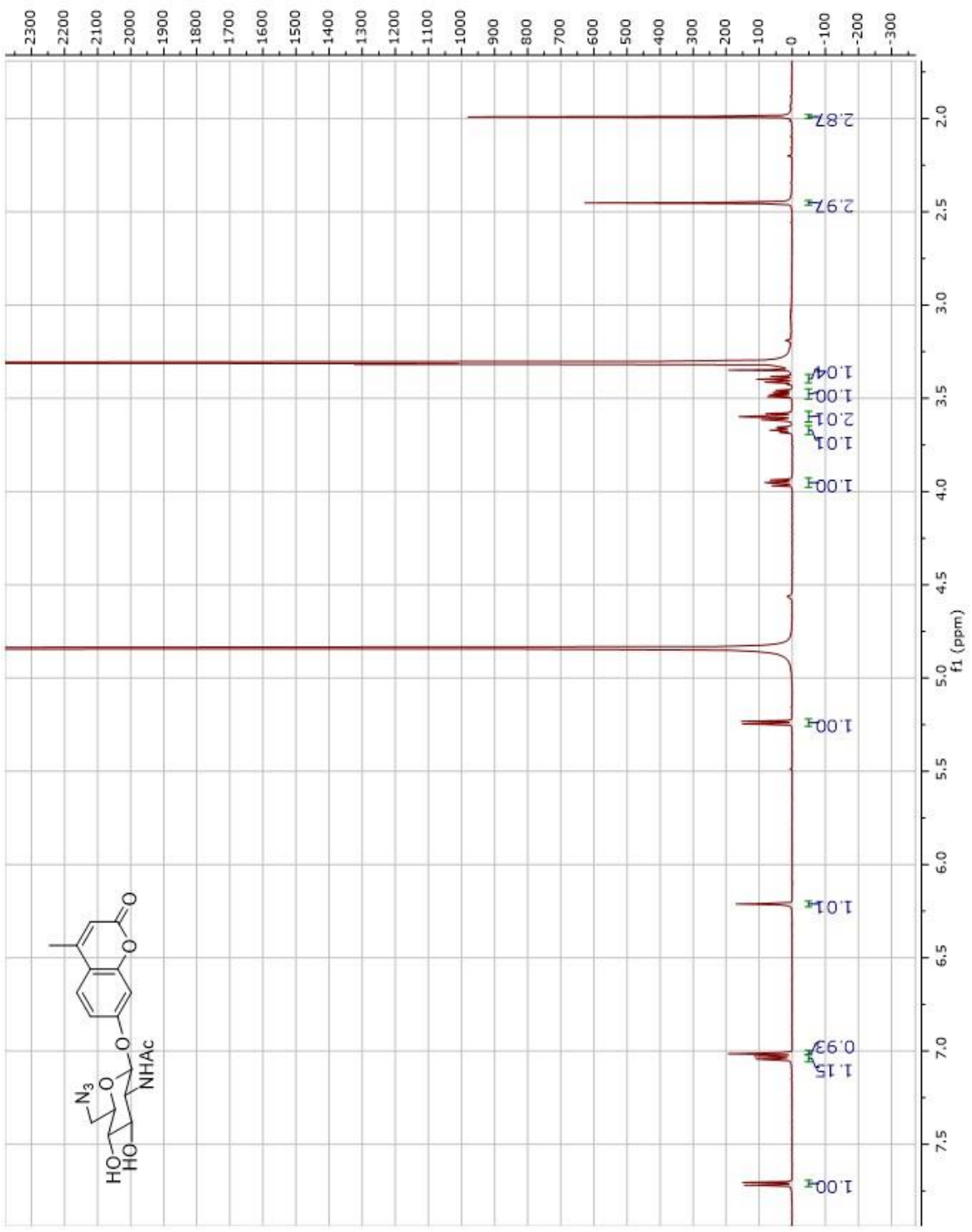

4000
3500
3000
2500
2000
1500
1000
500
0
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
f1 (ppm)
HN
O
N
N
O
O
O
NHBoc
AcO
AcO

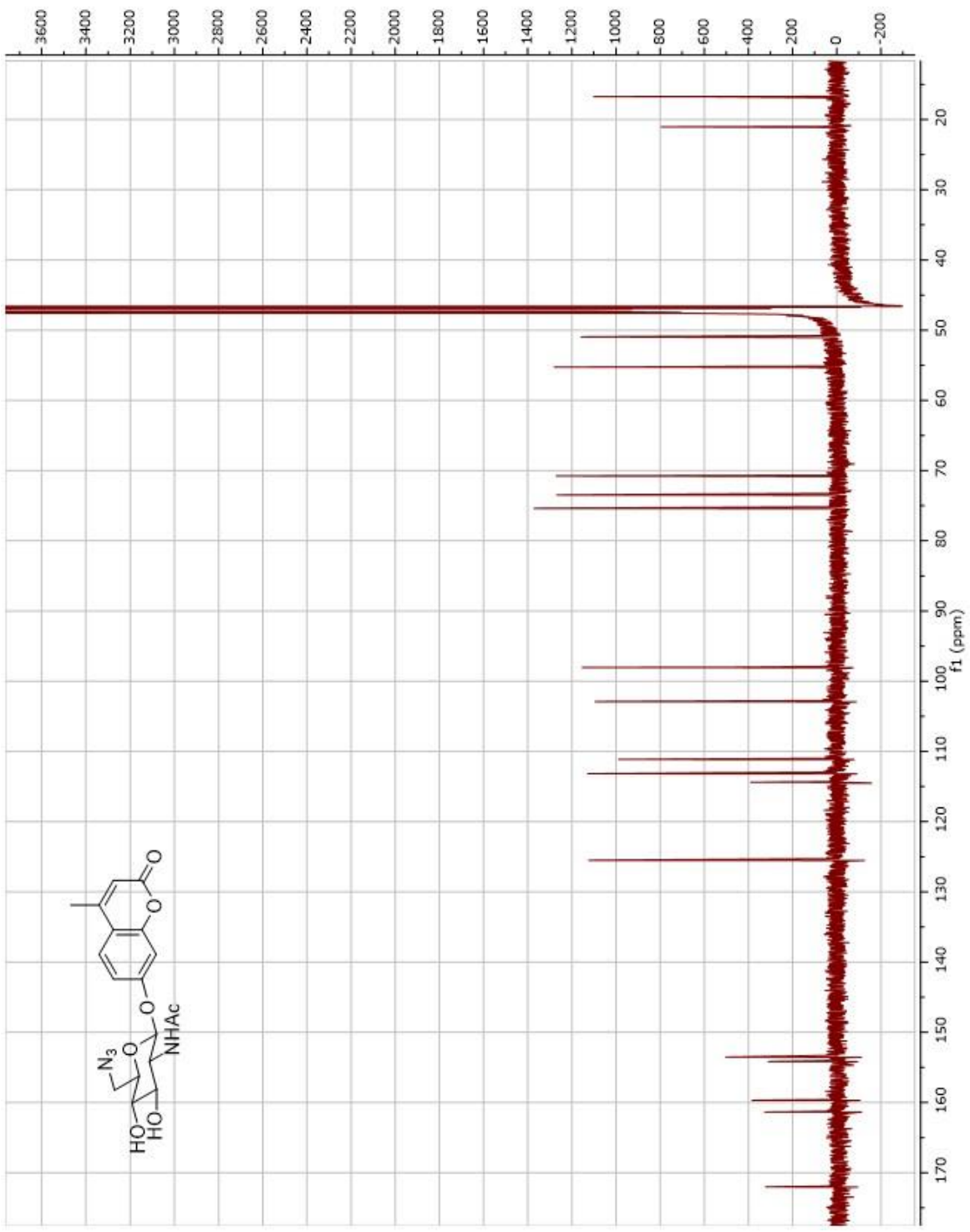

f1 (ppm)
9.13
3.05
3.41
3.48
3.14
1.12
1.14
1.16
2.31
1.16
1.14
1.08
1.21
1.10
1.00
1.12
1.14
1.13
1.08
0.69
NHBoc
AcO
AcO
HN
S


f1 (ppm)
NHBoc
AcO
AcO
HN
S
N
N
O
O


f1 (ppm)
HN
S
N
N
O
O
NHAc
HO
HO
O
O


f1 (ppm)
HN
S
N
N
O
O
NHAc
HO
HO


2.87
2.97
1.04
1.00
2.01
1.01
1.00
1.00
1.01
1.15
0.93
1.00
f1 (ppm)
NHAc
HO
HO
N3


f1 (ppm)
NHAc
N3
HO
HO