A mechanism based approach for screening metagenomic libraries for unusual glycosidases

by

Seyed Amirhossein Nasseri

B.Sc. in Chemistry, Sharif University of Technology, 2015

B.Sc. in Chemical Engineering, Sharif University of Technology, 2015

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(CHEMISTRY)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

March 2018

Abstract

Unsaturated glucuronidases are among the enzymes that a number of pathogens employ to break down glycosaminoglycans, the main structural polysaccharides of our bodies. In order to find out more about these relatively less studied enzymes and the bacteria that produce them and their relative abundance in the microbial populations of our environment, we set out to screen metagenomic libraries made from environmental samples such as the human gut microbiome for unsaturated glucuronidase activity.

In functional metagenomics, proteins from various uncultured organisms are reproduced in labs, by means of transferring the DNA extracted from environmental samples into host bacterial cell lines. This method offers an exciting novel way of enzyme discovery as these expressed proteins can be readily screened for novel enzymatic activity. However, screening for some of these enzymes is not straightforward, as background activity from the host cells and/or other similar unwanted enzymes from metagenomics genes, can mask the desired activity. Unsaturated glucuronidases are one of these masked enzymatic activities.

In this study, a novel strategy has been developed for screening metagenomic libraries for unsaturated glucuronidases. This was achieved based on the differences in the mechanisms of unsaturated glucuronidases and β-glucuronidases, the main source of background activity. These differences make β-glucuronidases inefficient in hydrolyzing thioglycoside substrates, while unsaturated glucuronidases cleave them rapidly. Two fluorogenic thioglycoside substrates with two self-immolative linkers were designed and synthesized. A small metagenomic library was then successfully screened with these new substrates and the usefulness of the selective substrates were

established. We believe that the same strategy is going to be useful when screening metagenomic libraries for some of the other examples of masked activities.

Lay Summary

Enzymes are yet another one of the impressive things that occur in nature; they are the catalysts that nature produces to speed up the reactions that need to be completed rapidly. Although many enzymes have been studied, many more exciting enzymes are still waiting to be discovered.

Functional metagenomics is a technique that enables us to discover unprecedented enzymes from environmental samples. Using this technique, large numbers of enzymes from environmental organisms are reproduced in the laboratory. The resulting metagenomic libraries are then screened for a particular enzymatic activity.

However, screening for a desired enzyme in metagenomic libraries is similar to searching for a needle in a haystack. In this thesis, a new strategy has been developed to screen these libraries for some specific examples of enzymes. In other words, in this thesis, a magnet has been made for finding the needle in the haystack.

Preface

All of the results presented in this thesis are the work of the author, unless mentioned otherwise. Analysis of the data was done in consultation with my supervisor, Professor Withers. The experiments presented in chapters 2.1 and 2.2 were carried out together with Daria Levitskaya, a visiting graduate student from Ludwig-Maximilian University of Munich.

The beaver and human fecal metagenomic libraries described in chapter 2 have been made by Zach Armstrong and Dr. Peter Rahfeld in the Withers lab.

Dr. Hong-Ming Chen provided compound **10**, GA-ClMU.

Table of Contents

List of Tables

List of Figures

List of Schemes

List of Abbreviations

Abg *Agrobacterium* sp. β-glucosidase

Ac Acetyl group

ACN Acetonitrile

BGB β-Glucuronidase from bovine liver

BGE β-Glucuronidase from *Escherichia coli*

BLAST Basic local alignment search tool

BSA Bovine serum albumin

ClMU 6-Chloro-4-methylumbelliferone

δ Chemical shift

ΔGA Unsaturated glucuronic acid

ΔGASC (6-Chloro-4-methylumbelliferone-methyl) 1-thio-4-deoxy-α-L-threo-hex-4-enopyranosiduronic acid

ΔGASP 4-(6-Chloro-4-methylumbelliferone-methyl) phenyl 1-thio 4-deoxy-α-l-threo-hex-4-enopyranosiduronic acid

DABCO 1,4-Diazabicyclo[2.2.2]octane

DBU 1,8-Diazabicyclo[5.4.0]undec-7-ene

DCE 1,2-Dichloroethane

DCM Dichloromethane

DMF *N,N*-Dimethylformamide

DMSO Dimethyl sulfoxide

DTT Dithiothreitol

EA Ethyl acetate

E. coli *Escherichia coli*

GA Glucuronic acid

Hz Hertz

HRMS High resolution mass spectroscopy

GH Glycoside hydrolase

IPTG Isopropyl β-*D*-thiogalactoside

J Coupling constant

LB Lysogeny broth, also known as Luria broth

LRMS Low resolution mass spectroscopy

MU 4-Methylumbelliferone

NMR Nuclear magnetic resonance

ORF Open reading frame

PE Petroleum ether

PMHS Poly(methylhydrosiloxane)

SDS-PAGE Sodium dodecyl sulfate polyacrylamide gel electrophoresis

SMU 4-Methyl-7-thioumbelliferone

THF Tetrahydrofuran

TLC Thin layer chromatography

Tris Tris(hydroxymethyl)aminomethane

TYP Tryptone-yeast-phosphate growth medium

UGC Unsaturated glucuronyl hydrolase from *Clostridium perfringens*

UGL Unsaturated glucuronyl hydrolase

UV Ultraviolet

Acknowledgements

Thinking about writing the acknowledgments, I realized, again, that I truly am very lucky. And that I can never possibly be thankful enough for all the opportunities that I have had. It is also for some time now, that I have given up the hope that I can properly thank all the people that I have to thank. Nevertheless, I will have to at least try.

First of all, I would like to thank my supervisor, Professor Withers for all his support and kindness. I truly feel honoured for the opportunity to work with him. I am also especially grateful to Daria Levitskaya, for her hard work and her enthusiasm, during the period that she was working with me. Next, I would like to thank all members of the Withers group, present and past, for all the help and useful discussions and suggestions and for the great atmosphere of the lab. Specifically, I would like to thank Dr. Hong-Ming Chen for numerous enlightening discussions about synthesis, Ms. Emily Kwan for teaching me about microbiology techniques, Dr. Peter Rahfeld and Zach Armstrong for walking me through the screening procedure, Dr. Feng Liu and Phillip Danby for useful discussions and Jacob Wardman and Grace Ho for proofreading my thesis. My special thanks are reserved for Dr. Leo Betschart, with whom I had the opportunity and pleasure of working closely on the second half of this project and Kyle Robinson, for a history of support, useful discussions, many valuable suggestions and the list goes on.

This would not be even imaginable without all the support of my wonderful family, whom I miss day to day. The exception is my brother, that I don't really miss! I would like to take this opportunity to thank my brother, Soroush, who is the primary reason why I am here, in almost every sense and who has been basically babysitting me for the past two years.

"And our last call, that all the praise is the Lord's…"

To the true motivation of all this, Baba Saeed

And to the most tender of all people I know, Maman Jaaleh

And to my best friends of life, Soroush va Ali

Chapter 1: Introduction

1.1 Glycosidases

Carbohydrate-active enzymes are the catalysts that living organisms produce to manipulate carbohydrates, the most abundant of organic biomolecules. Carbohydrates are omnipresent in nature and play numerous different roles in the process of life. These roles range from being the main class of energy storage and delivery molecules in the living world to their function as the main structural molecules of many living organisms, from cellulose in plants to glycosaminoglycans in vertebrates. Thus, carbohydrate-active enzymes are also ever-present in nature in a huge abundance.

Glycosidases are one of the main classes of carbohydrate-active enzymes. Glycosidase (or glycoside hydrolase) is a general term used for the enzymes that catalyze the hydrolysis of a glycoside, shown in Figure 1.

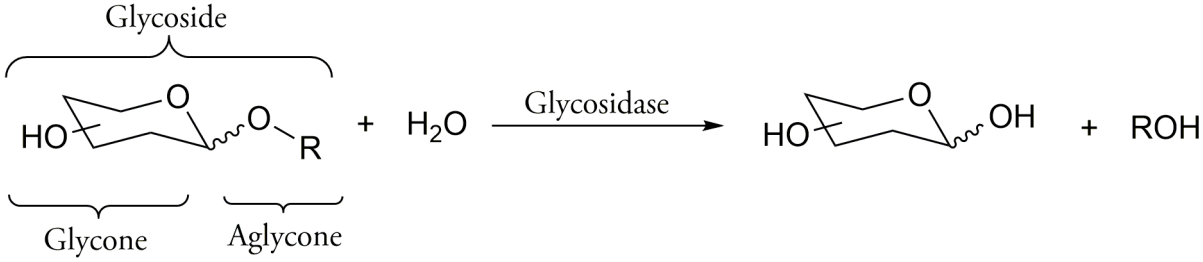


Figure 1: General representation of a glycosidase catalyzed reaction

1.1.1 Classification of glycosidases

Glycosidases are often classified by their substrate and anomeric specificities. For example, a β-glucuronidase is a glycosidase that is most efficient when catalyzing hydrolysis of the substrate shown in Figure 2. The sugar unit in this case is a *D*-glucuronic acid that is linked to the aglycone with a β-linkage.

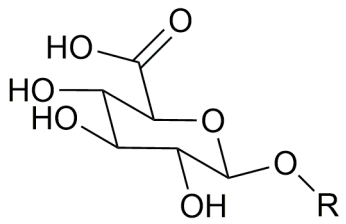


Figure 2: General structure of a substrate for β-glucuronidases

Another classification method that has proven to be useful is based on the observation that the structures of two proteins with similar sequences are related to one another[1]. Accordingly, it was proposed that carbohydrate active enzymes can be classified into "families" based on their sequence homology[2]. The enzymes within such a family generally have a similar folding pattern and more interestingly, they seem to follow the same general mechanism for hydrolysis[3]. The substrate specificity however, can vary within a family.

Using this method, glycosidases have been classified into 152 families up to date. The data for these families of enzymes can be accessed through the Carbohydrate-Active Enzymes (CAZy) database, http://www.cazy.org.

The enzymes that are the primary focus of this thesis are from glycoside hydrolase families GH2, GH88 and GH105.

1.2 β-Glucuronidases from the GH2 family and their mechanism

The GH2 family of glycosidases is a diverse and relatively well-studied family, consisting of nearly 11,000 enzymes. Two enzymes from GH2 family of glycosidases have been used in this study: β-glucuronidase from *E. coli* and β-glucuronidase from bovine liver.

The general mechanism that is followed by the members of this family of enzymes (Figure 3), was first proposed by Koshland[4] and is called 'the retaining mechanism' since the configuration of the glycosidic bond in the product of the enzymes that follow this mechanism is the same as that of the starting material.

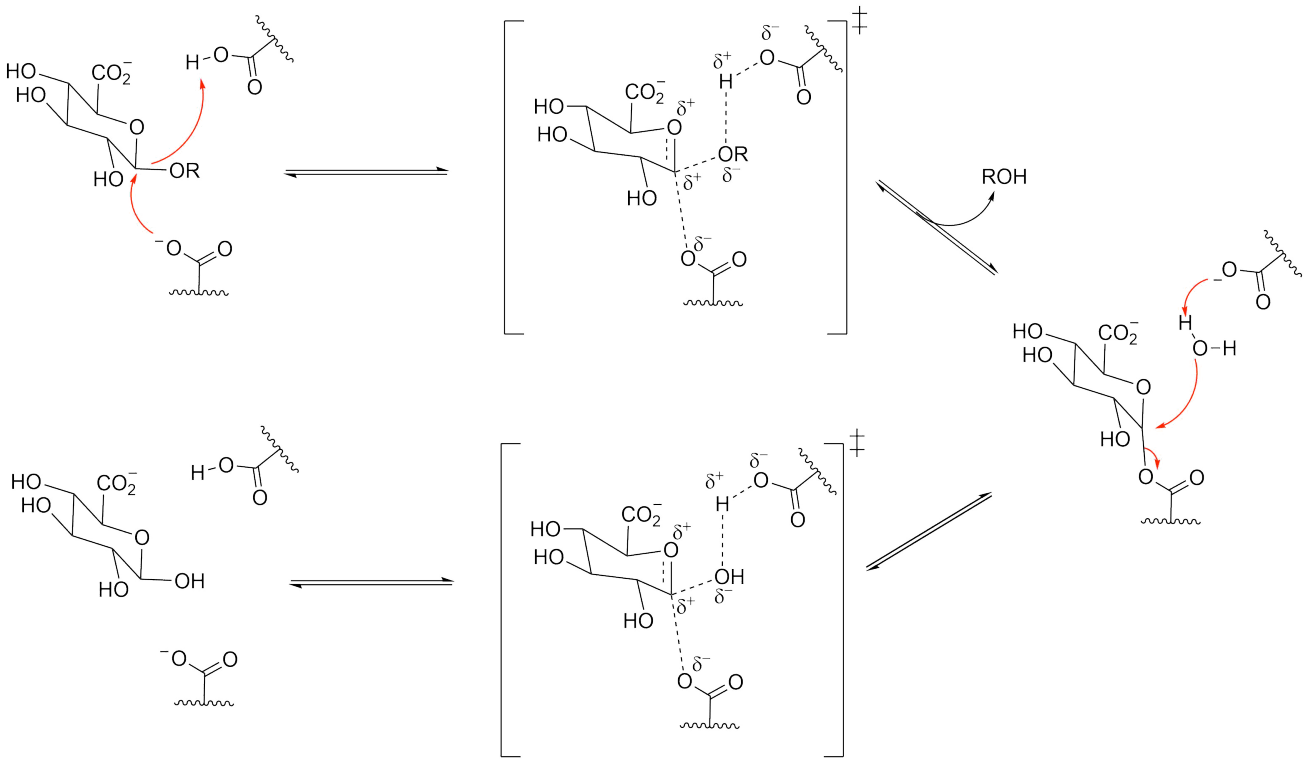


Figure 3: General mechanism for GH2 β-glucuronidases

The first step of this mechanism involves the attack of an enzymatic nucleophile onto the anomeric carbon of the sugar to form a glycosyl-enzyme intermediate with displacement of the aglycone. The departure of the aglycone is assisted through its protonation by an acidic residue. This

acidic residue, now deprotonated, activates a water molecule in the second step to hydrolyze the glycosyl-enzyme intermediate. The nucleophilic and the acid/base residues of the enzymes that follow this mechanism are usually one of the two carboxylic acid-containing amino acids. For the two GH2 enzymes studied in this thesis, they are glutamic acid residues[5].

1.3 Unsaturated glucuronidases

4,5-Unsaturated glucuronidases (UGLs) are the main class of enzymes studied in this thesis. UGLs are relatively uncommon enzymes with an unusual mechanism. These enzymes are classified into two related families of glycosidases, GH88 and GH105. The native substrates of these enzymes, 4,5-unsaturated glucuronic acids, are rare sugars in nature that are only formed in the process of enzymatic degradation of uronic acid-containing polysaccharides.

GH105 enzymes are primarily present in bacterial degradation pathway of polysaccharides that contain galacturonic acid. These polysaccharides are mainly from plant sources. On the other hand, GH88 enzymes are mainly deployed by bacteria to degrade polysaccharides that contain glucuronic acid or iduronic acid. Examples of such polysaccharides are the glycosaminoglycans.

1.3.1.1 Glycosaminoglycans and the bacterial pathway for their degradation

Glycosaminoglycans are long polysaccharides consisting of repeating units of a disaccharide core. This disaccharide motif usually includes an amino-sugar and one of the two uronic acids, either glucuronic acid or iduronic acid[6]. The structure of a simple and common glycosaminoglycan, hyaluronic acid is shown in Figure 4.

Glycosaminoglycans are among the most important structural polysaccharides in mammals. Since they are commonly found in tissues and if hydrolyzed will provide a source of food for bacteria, some of the pathogenic bacteria have evolved enzymes to degrade them. UGLs are one of the enzymes

responsible for the bacterial degradation of glycosaminoglycans. Since UGLs are important for pathogen growth and since they are not expressed by mammalian cells, development of inhibitors for UGLs may provide a novel therapeutic strategy against these bacterial pathogens.

The first enzyme that is involved in the bacterial pathway for degradation of glycosaminoglycans (Figure 4), polysaccharide lyase, cleaves the bond between the glucuronic acid and the *N*-acetylglucosamine and produces a 4,5-unsaturated glucuronic acid in the process. This unsaturated glucuronic acid is then cleaved from the *N*-acetylglucosamine by UGL.

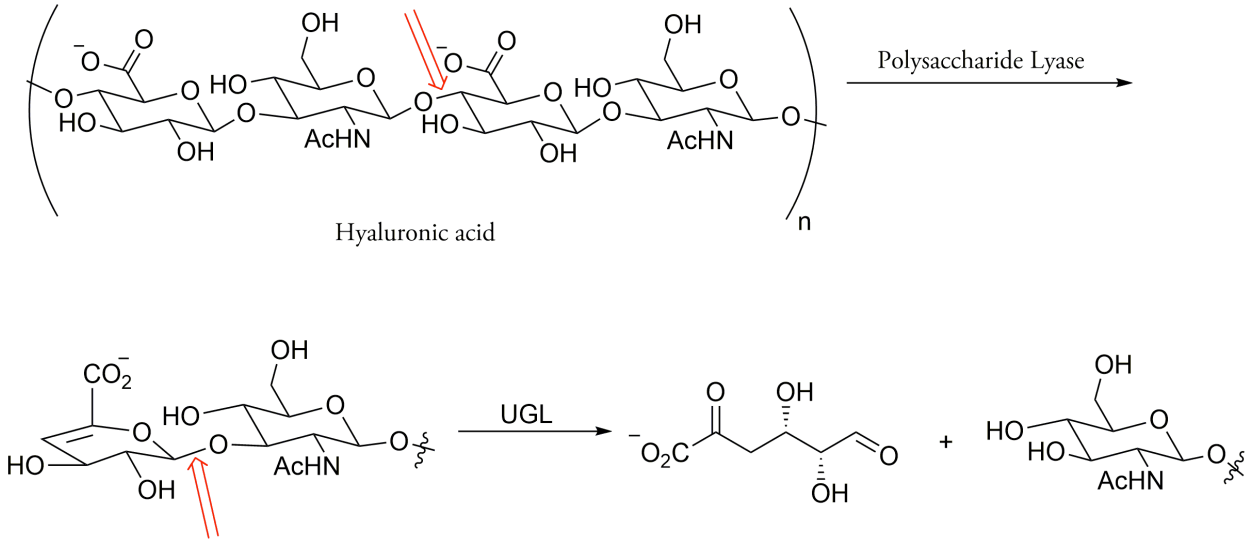


Figure 4: An example of a glycosaminoglycan degradation pathway, the degradation pathway of hyaluronic acid. The bonds that are broken in each step are shown by red arrows

1.3.2 Mechanism of UGLs

UGLs from GH88 and GH105 family of enzymes follow a relatively unconventional mechanism. Despite extensive study in our group[7–9], the exact mechanism is not established yet. Figure 5 shows a basic postulated mechanism of UGLs. The first part of this mechanism is hydration of the C_4 double bond, which takes place in two steps; first, addition of a proton to the double bond to form a carbocation intermediate and then enzyme-assisted nucleophilic attack of water to the positively

charged carbon. The second part of the mechanism consist of the two rearrangements of the resulting hemiketal and then hemiacetal to afford the final products.

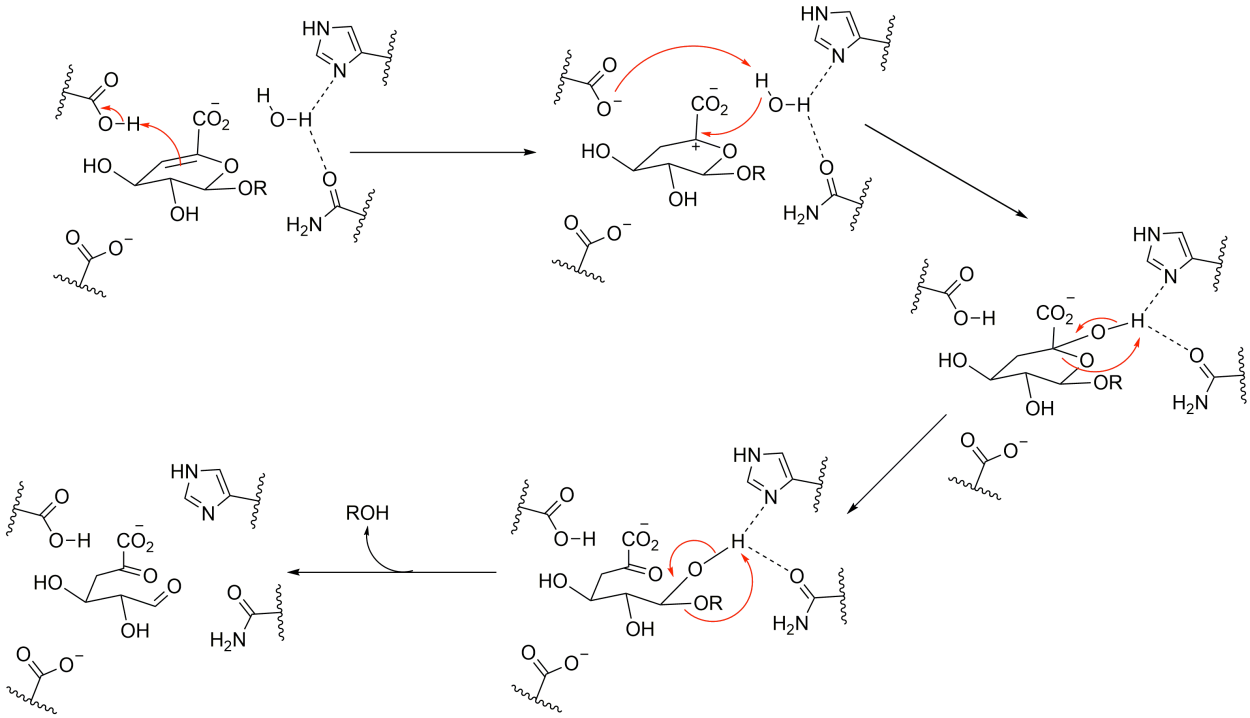


Figure 5: The postulated mechanism of UGLs from GH88 family

It should be said that this relatively simple mechanism does not explain all of the observed kinetic and crystallographic evidence and other variations on this mechanism have been proposed[7]. However, the first step of this mechanism has been established to be the hydration of the C_4 double bond[7–9], as shown in Figure 5. Moreover, observations from these studies suggest that this step is the rate-determining step of this mechanism. Interestingly, this makes UGLs promiscuous enzymes, accepting a wide range of substrates for catalysis[7,8]. This is because the anomeric bond and the aglycone do not have a substantial effect on the rate of this reaction, as they do not play a role in the rate-determining step of this mechanism. Therefore, various substrates with different aglycones are accepted for hydrolysis by these enzymes.

1.4 Thioglycosides

Thioglycosides (or S-glycosides as opposed to O-glycosides for conventional glycosides) are homologues of glycosides with one simple change: they have a sulfur atom in place of the anomeric oxygen in glycosides. This simple change however, has a profound effect on the characteristics of these molecules. One interesting example is that most glycosidases cannot hydrolyze S-glycosides[10]. Therefore S-glycosides have been used extensively as non-hydrolysable mimics of O-glycosides[12–16]. However, our present work, as well as other studies[16,17], shows that some of the conventional glycosidases are able to hydrolyze activated S-glycosides.

The full reason for why most glycosidases are not able to hydrolyze S-glycosides is unclear. Indeed, one might have thought that since anomeric C-S bonds are weaker than C-O bonds, the hydrolysis of S-glycosides should be faster than O-glycosides. However, even the rate of spontaneous hydrolysis of S-glycosides in acidic solution is usually lower compared to their O-glycoside homologues[18–20].

It has been proposed that the inability of most glycosidases to hydrolyze S-glycosides is due to the fact that with the longer bond length of C-S compared to C-O, along with the poor ability of S to form hydrogen bonds relative to O, the enzymes are not able to assist the hydrolysis of S-glycosides as much as O-glycosides[21]. Therefore, for those glycosidases that follow Koshland mechanisms, these factors may result in a change in the transition state of the rate-determining step and render enzyme catalysis ineffective for S-glycosides. Interestingly, for those enzymes that do not catalyze the hydrolysis of glycosides primarily through activation of the aglycone, thioglycosides are generally accepted as substrates for catalysis[9,22].

1.4.1 Thioglycosidases

Thioglycosidases are a class of carbohydrate active enzymes that specialize in hydrolyzing S-glycosides. Despite the fact that thioglycosidase activity has been reported for various natural organisms, such as fungi[23], bacteria[24] and mammalian cells[25], there is only one characterized family of naturally occurring thioglycosidases[26]. Therefore, it is anticipated that there may be many more families of thioglycosidases that are yet to be discovered.

The only characterized family of thioglycosidases, called myrosinases, occurs primarily in the *Brassica* genus of plants. These enzymes are responsible for hydrolyzing glucosinolates (Figure 6). Glucosinolates are a large family of thioglucosides whose hydrolysis is believed to be part of the defense system of the plants[27]. In addition to plants, myrosinases have also been identified from insects[28–32] and recently from bacteria[33]. While the enzymes from the insect sources are similar to the plant myrosinases[31], the only example of bacterial myrosinase is believed to be from a different family[33].

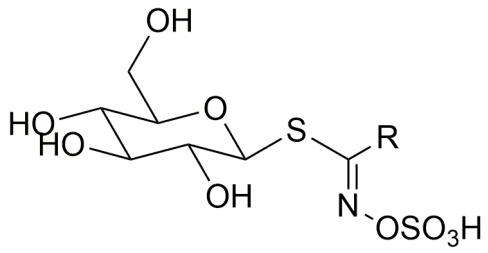


Figure 6: General structure of glucosinolates

1.5 Functional metagenomics

Functional metagenomics is a shortcut to answer questions such as 'what kind of proteins do the organisms that live in a given ecosystem produce?'. Answering this question through traditional microbiology requires culturing the living organisms of a natural environment in laboratory. Besides the fact that this would be time and resource costly, it is estimated that the vast majority of microorganisms in nature cannot be cultivated in laboratories through standard procedures[34].

Instead, in functional metagenomics, all of the DNA from an environmental sample is extracted, regardless of what organism each specific DNA is coming from. These genes then are transferred randomly into expression vectors for heterologous expression in host cells such as *E. coli*. In the resulting metagenomic libraries, the same proteins that were produced by the native organisms in natural environment, can be reproduced in the host cells in laboratories and be analyzed or screened[35] (Figure 7).

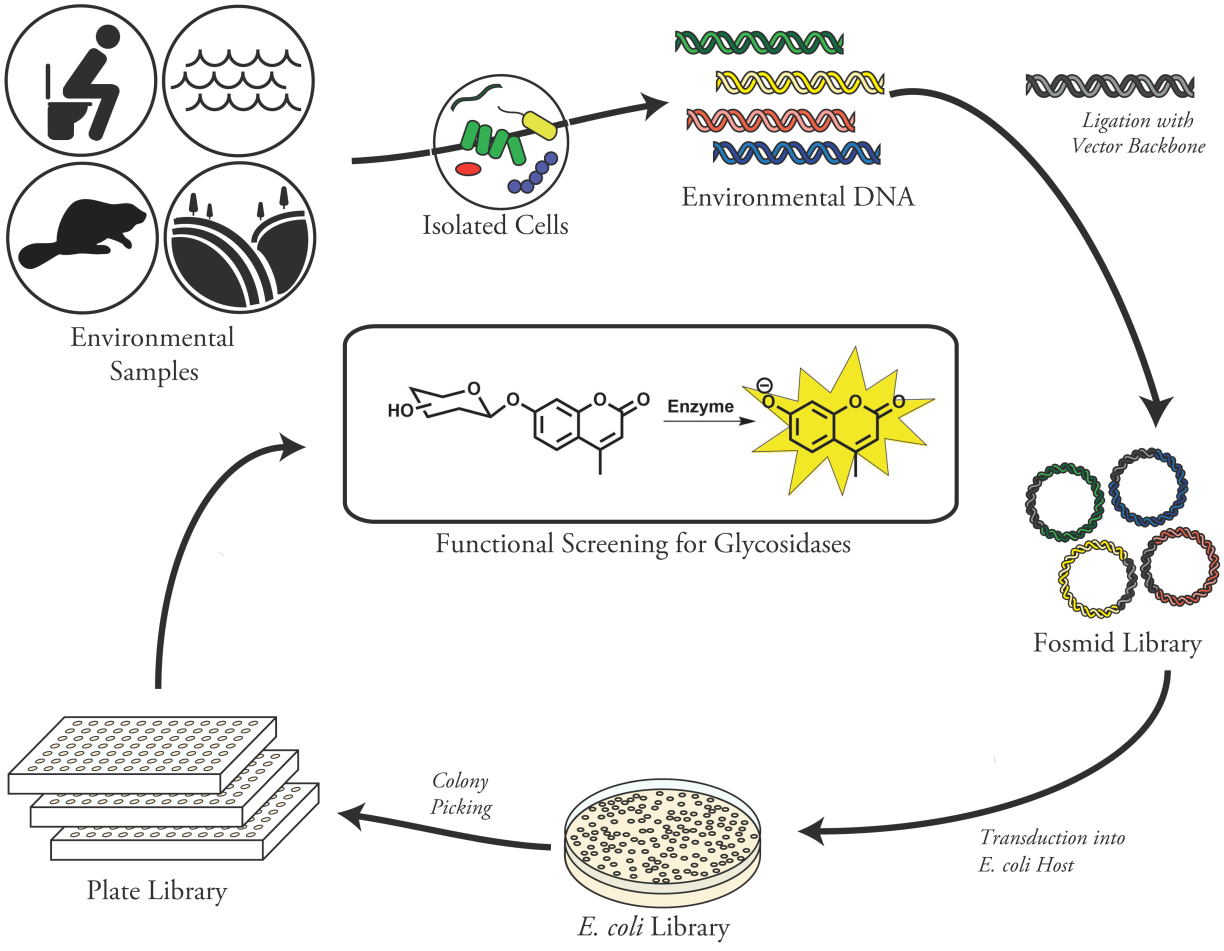


Figure 7: A simple workflow of functional metagenomic. Picture adapted from [36].

Numerous novel enzymes have been discovered from metagenomic libraries from various interesting sources. A few examples are DNA polymerases from a hot thermal pool[37], proteases from soil samples from Gobi and Death valley deserts[38] and from goat skin surface[39], amidases from soil[40] and cellulases from gut microflora of black spider flies[41].

1.5.1 Considerations on the typical procedure for screening metagenomic libraries for glycosidases

As stated before, metagenomic genes are typically expressed in host bacterial cell lines such as *E. coli*. These *E. coli* cells are typically stored in multiple 384-well plates, each well carrying cells with different genes that will result in expression of different proteins. The screening process typically involves replicating these plates and growing the bacteria overnight, followed by lysing the cells and screening the plates of cell lysates for the desired protein.

If the metagenomic library is searched for enzymes, the typical procedure for screening involves adding the appropriate reagents to each well, incubating for several hours and then searching to find out if the desired products are formed. In the case of glycosidases, the products of the enzyme catalyzed reaction are the free sugar and aglycone. For their high sensitivity and their excellent compatibility with high throughput screening, fluorogenic aglycones are most commonly used in these screenings[42].

When screening for a glycoside hydrolase, a conventional substrate will have a structure similar to the molecule in Figure 8. The sugar unit and the glycosidic linkage in this substrate will depend on the activity that the metagenomic library is screened for.

One of the most commonly used classes of fluorogenic aglycones in these substrates is 4-methylumbelliferone (MU) and its derivatives. 4-Methylumbelliferone (or 7-hydroxycoumarin), is the aglycone of the substrate shown in Figure 8. MU and its derivatives are highly fluorescent when cleaved from the sugar and in their anionic form, with limits of detections in the low μM region. In addition, they are generally stable in standard assay conditions, therefore the fluorescence signal will not be significantly lost even during the typically long assay times[42]. Tuning the aglycone to optimize the screening procedure is also easy for MU, because of its derivatizable backbone structure.

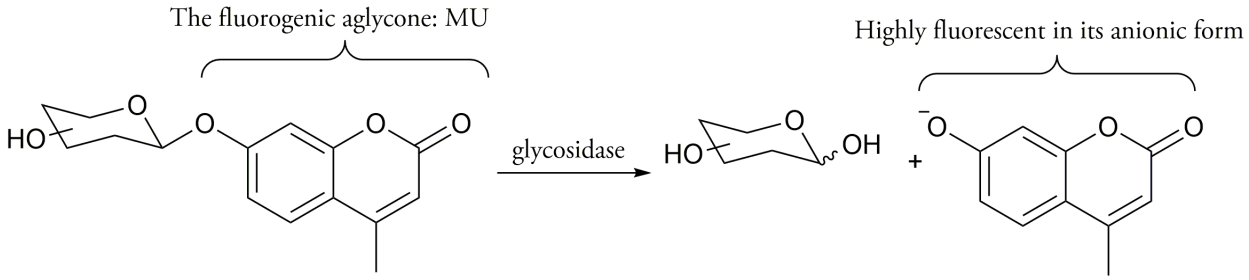


Figure 8: A general example of a conventional substrate for screening metagenomic libraries for glycoside hydrolases

Since it is the deprotonated form of the molecule that is highly fluorescent, the pKa of MU affects the sensitivity of the substrate dramatically. For example, if the screening assay is performed at pH 6, then an MU derivative with a pKa of 5 can be up to 10000 times more sensitive than a MU derivative with a pKa of 9 (assuming that the quantum yield of the deprotonated MU for both derivatives is the same).

Also, the overall stability of the glycosidic bond between the MU and the sugar is crucial. A less stable glycosidic bond will likely translate to a more sensitive substrate. Since the metagenomic enzymes are often present in very low concentrations in screening conditions, more sensitive substrates are often needed for successful screens. On the other hand, the glycosidic bond should be sufficiently stable so that the amount of background spontaneous hydrolysis does not cause problems in identifying the hits. Since the stability of the glycosidic bond is related to the leaving group ability of the aglycone and thus its pKa, the pKa of aglycone plays a crucial role here too. A lower pKa of the aglycone means a better leaving group and hence more sensitivity. Therefore, the pKa of the aglycone is generally desired to be as low as possible for good sensitivity but also high enough for sufficient stability of the glycosidic bond with negligible background hydrolysis.

One example of a MU derivative with the right pKa for screening metagenomic libraries is 6-chloro-4-methylumbelliferone (ClMU) with a pKa of 5.9. It has been shown that substrates with ClMU as their aglycones are very useful for screening metagenomic libraries for glycosidase activity[42].

1.6 Goals of the thesis

Two of the classes of enzymes that are described in the sections before, namely thioglycosidases and UGLs, are of particular interest when it comes to screening metagenomic libraries. This is because these enzymes are relatively understudied and there is a good probability that entirely new families of enzymes with these activities can be found in metagenomic libraries. As for UGLs, the two known families of enzymes with this activity, GH88 and GH105, are also relatively small and less studied. Therefore, finding more members of these families with different sequences will give us a better understanding of their characteristics and diversity. However, screening metagenomic libraries for thioglycosidases and UGLs with conventional methods faces some problems.

Since conventional MU-glycoside screening substrates are not selectively hydrolyzed by thioglycosidases, use of these substrates to screen for thioglycosidases will yield many false positives from O-glycosidases. Likewise, screening metagenomic libraries for UGLs with conventional substrates faces a similar problem. As will be shown in section 2.1 of this thesis, MU-glycoside substrates of UGLs can be hydrolyzed by β-glucuronidases (EC 3.2.1.31) too. Thus, screening for UGLs with conventional substrates will also give rise to many false positives. However, the problem for UGLs is more grave than that of thioglycosidases. As discussed in section 1.5, metagenomic libraries are usually expressed in *E. coli*. Almost all of the typical *E. coli* strains have a prominent β-glucuronidase activity. Therefore, unless a library is built exclusively for screening for UGLs and expressed in glucuronidase negative strains of *E. coli*, screening with conventional substrates will also face a serious background activity from the host cells.

The main goal of this thesis is to provide novel screening substrates to solve these problems. Focusing on the case of unsaturated glucuronidases, the first goal of this thesis was to explore the

specificity of these enzymes with conventional MU-glycoside and similar thioglycoside substrates. This requires synthesis of these compounds and comparison of the kinetics of their hydrolysis with model proteins from the two class of enzymes, UGLs and β-glucuronidases.

The second goal of this thesis is to design novel selective substrates for screening metagenomic libraries for UGLs based on the results of the previous part. Because of the problems in using a thiol-based probe, this would require use of thiol-based self-immolative linkers that will release fluorophores upon hydrolysis by enzymes. After designing and synthesizing these substrates, the practicality of screening with these substrates was to be investigated by screening a relatively small metagenomic library.

The long-term goal of this thesis is to provide a general substrate structure for screening metagenomic libraries for other classes of enzymes with the same problem. Based on the results of this study, similar substrates can be synthesized and used for screening metagenomic libraries for thioglycosidase activity, and also other possible families of enzymes for which screening metagenomic libraries faces similar problems.

Chapter 2: Results and Discussion

2.1 Testing conventional screening substrates for UGLs

We first set out to synthesize conventional MU-glycoside substrates for UGLs to determine if it is possible to carry out a selective screening with these substrates. Since it was anticipated[43] that β-glucuronidases might also hydrolyze UGL substrates, the analogous substrates for β-glucuronidases were also synthesized. These substrates are shown in Figure 9. Two different aglycones were used, MU and ClMU. Each one is attached to either glucuronic acid (GA) or unsaturated glucuronic acid (ΔGA) with an equatorial linkage.

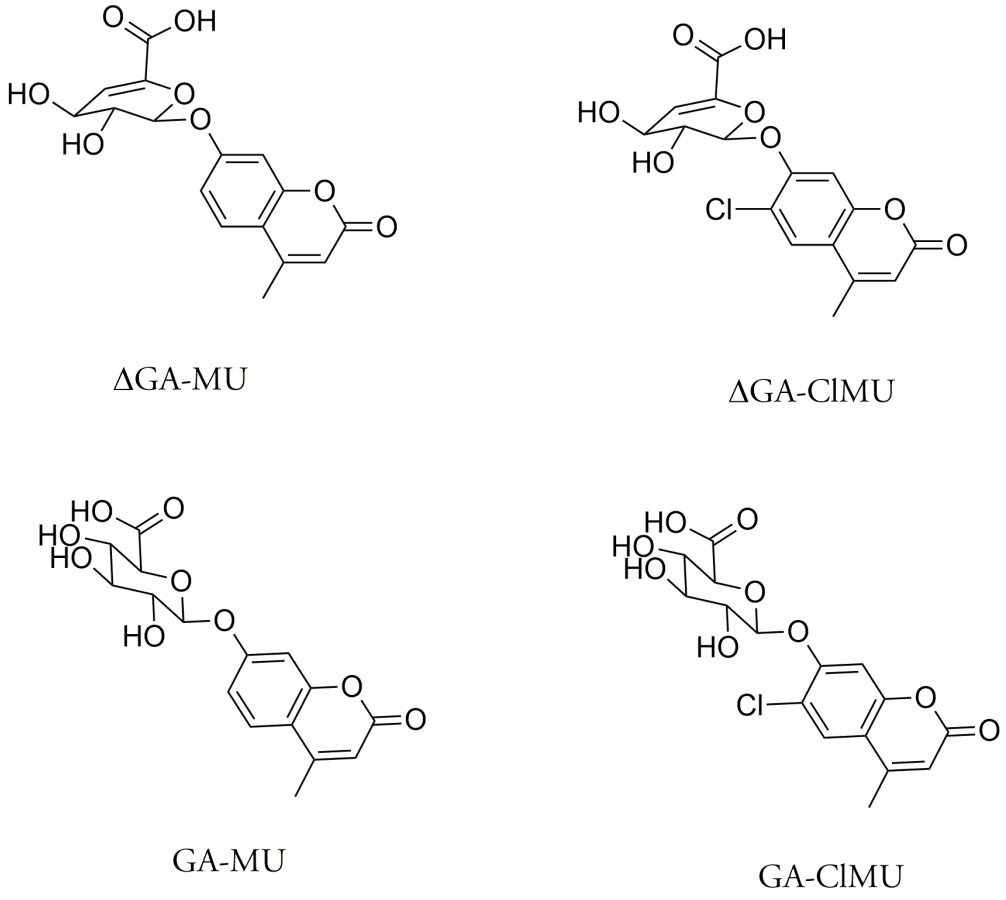
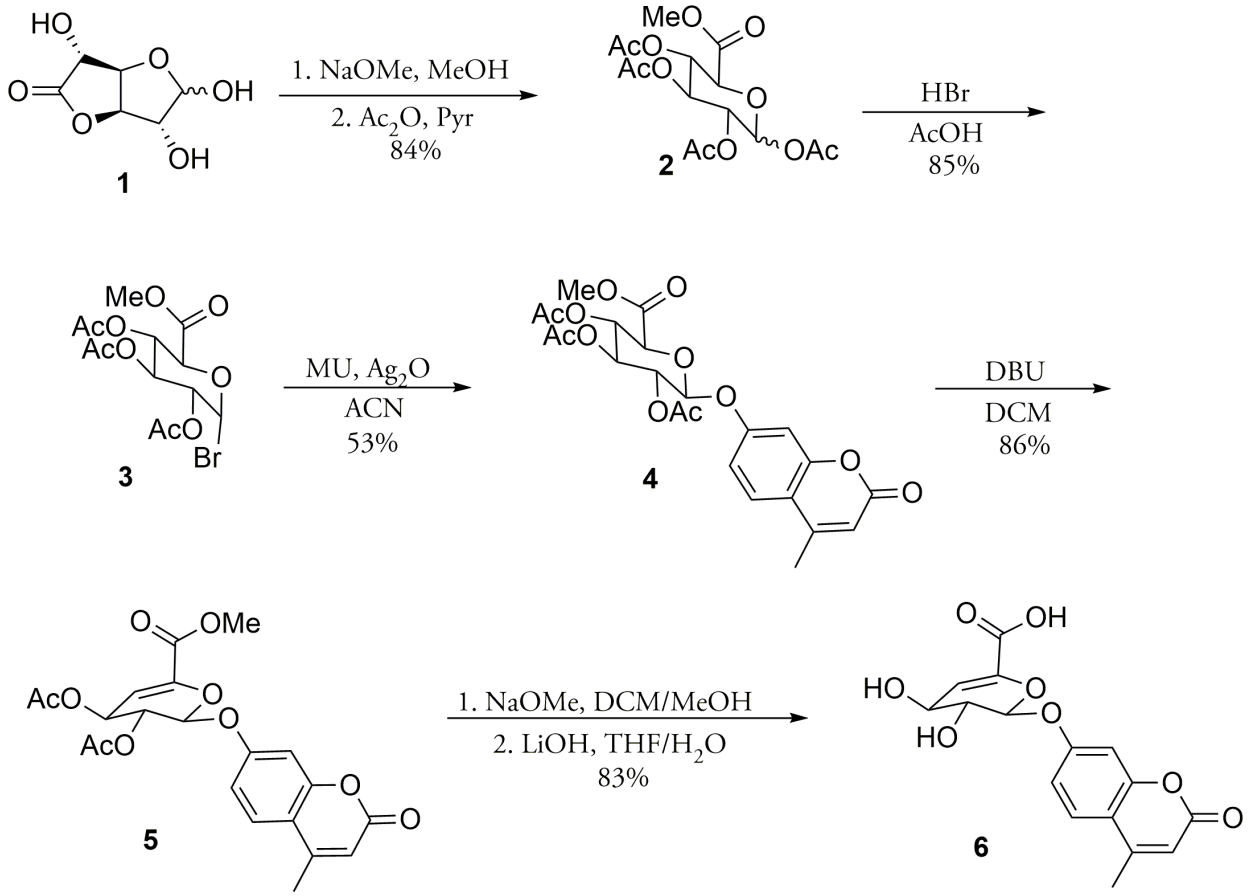


Figure 9: Substrates for kinetic analysis of β-glucuronidases and UGLs

The synthetic route for one of these substrates is shown in Scheme 1. The synthesis follows a path that has been described previously[9]. Glucuronic acid-γ-lactone (**1**) is first treated with sodium methoxide in methanol and then with acetic anhydride in pyridine to give per-*O*-acetylated methyl glucuronate (**2**). Treatment of **2** with hydrobromic acid in acetic acid then yields protected glucuronyl bromide (**3**). Koenigs-Knorr glycosylation is next employed to couple MU with the glucuronyl bromide. The next step is elimination of the C_4 acetate using DBU to form the C_4-C_5 double bond, followed by deprotection in basic methanol to yield the substrate ΔGA-MU.



Scheme 1: Synthetic scheme for ΔGA-MU

The enzymatic hydrolysis of these substrates was investigated with model UGLs and β-glucuronidases. Unsaturated glucuronidase from *Clostridium perfringens* (UGC) was used as an exemplary UGL. Two

enzymes were chosen to represent β-glucuronidases: I) β-glucuronidase from *E. coli* (BGE) and II) β-glucuronidase from bovine liver (BGB).

The Michaelis-Menten kinetic parameters for these four substrates and three enzymes were derived using the conditions reported in section 4.2.2 and are listed in Table 1.

Entry	Enzyme	Substrate	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($s^{-1}\mu M^{-1}$)
1	UGC	ΔGA - ClMU	**160 ± 21**	**10.8 ± 0.4**	**$(6.7 \pm 1.2) \times 10^{-2}$**
2		ΔGA - MU	**280 ± 25**	**8.2 ± 0.3**	**$(3.0 \pm 0.4) \times 10^{-2}$**
3	BGE	GA - ClMU	**63 ± 4**	**12.8 ± 0.2**	**$(2.0 \pm 0.2) \times 10^{-1}$**
4		GA - MU	**140 ± 12**	**13.6 ± 0.4**	**$(9.7 \pm 1.1) \times 10^{-2}$**
5		ΔGA - ClMU	**6.7 ± 0.7**	**0.048 ± 0.001**	**$(7.2 \pm 0.9) \times 10^{-3}$**
6		ΔGA - MU	**750 ± 130**	**0.037 ± 0.002**	**$(4.9 \pm 1.1) \times 10^{-5}$**
7	BGB	GA - ClMU	**109 ± 9**	**1.36 ± 0.03**	**$(1.2 \pm 0.1) \times 10^{-2}$**
8		GA - MU	**280 ± 24**	**1.39 ± 0.05**	**$(5.0 \pm 0.6) \times 10^{-3}$**
9		ΔGA - ClMU	**1260 ± 123**	**0.0146 ± 0.0008**	**$(1.2 \pm 0.2) \times 10^{-5}$**
10		ΔGA - MU	**- [a]**	**0.0015 [b]**	**- [a]**

Table 1: Kinetic parameters for conventional screening substrates at 37° C

Assay buffers: phosphate buffer pH = 6.6 for UGC and BGE and acetate buffer pH = 5 for BGB.

a) Data can't be determined under Michaelis-Menten conditions

b) Estimation of k_{cat} based on two data point in the plateau region of the Michaelis-Menten curve

Comparing the k_{cat}/K_m values, the first thing to note in the data is that β-glucuronidases are capable of catalyzing the hydrolysis of unsaturated-glucuronic acid substrates. However, as expected, this activity is 2 to 3 orders of magnitude lower than their native activity.

Also of note is that for each of the enzymes, the hydrolysis of their native ClMU-glycoside substrates is approximately 2 times faster than MU-glycoside substrates. As discussed in chapter 1.5.1,

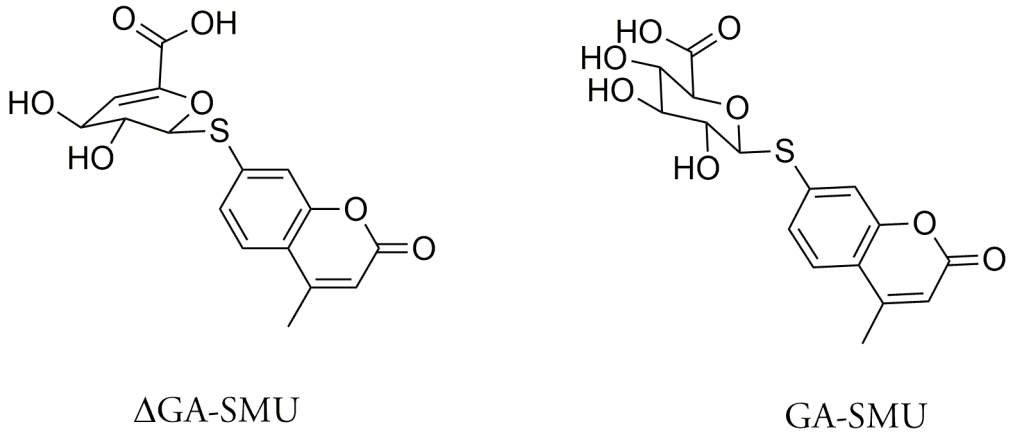
this is most probably because of the lower pKa of ClMU (pKa = 5.9[42]) compared to MU (pKa = 7.8[44]) and thus its better leaving group ability. In the context of screening, this means that there will be a higher chance for ΔGA-ClMU to give rise to false positive hits when compared to ΔGA-MU. On the other hand, this also means that ΔGA-ClMU will be much more sensitive compared to ΔGA-MU and hence more likely to find the enzymes even when they are expressed in low concentrations.

In summary, the kinetic data suggests that using a conventional substrate for screening metagenomic libraries for UGLs will be problematic. First of all, β-glucuronidases from the metagenomic genes can appear as positive hits. Secondly and more importantly, since in this study we aim to screen existing metagenomic libraries that have not been expressed in a β-glucuronidase negative strain of *E. coli*, BGE will be always present in the screening conditions and there will always be a significant background activity, making it hard to distinguish the real hits from background hydrolysis. If a metagenomic library is made exclusively for screening for UGLs the background hydrolysis can be avoided by using a β-glucuronidase negative strain. However, the β-glucuronidases from metagenomic genes will still appear as positive hits. Accordingly, more selective substrates are needed in order to carry out these screens.

2.2 Thioglycosides as substrates for screening

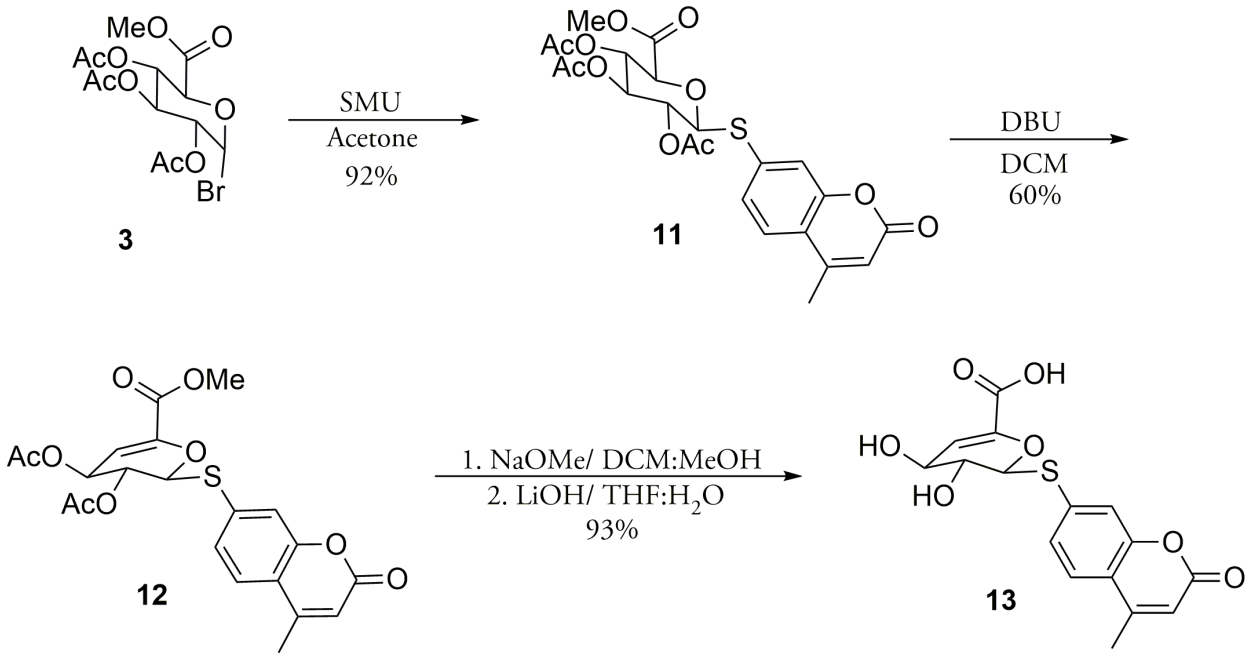
Most of the glycosidases that follow a conventional Koshland mechanism cannot catalyze the hydrolysis of thioglycosides. However, for those glycosidases that follow a different mechanism, this is not always the case. If the anomeric bond does not play a crucial role in the rate-determining step of an enzyme's mechanism, S-glycosides can potentially be hydrolyzed by the enzyme, since substitution of the anomeric oxygen with sulfur should not have any major effect on the rate of the reaction. This assumption is true unless replacing the anomeric O with S changes the rate-determining step.

Since GH2 enzymes follow a retaining Koshland mechanism, it is expected that they should not be able to effectively hydrolyze S-glycosides. However, since UGLs from GH88 family follow a different mechanism, this is not the case for them[9]. Therefore, we decided to synthesize thio-linked substrates and test to see if these substrates would be selective enough to be useful in screening. For a direct comparison to the previously synthesized substrates, 4-methyl-7-thioumbelliferone (SMU) was chosen as the aglycone. The structures of the synthesized substrates are shown in Figure 10 and the synthetic route for ΔGA-SMU is shown in Scheme 2.



ΔGA-SMU GA-SMU

Figure 10: The structures of GA-SMU and ΔGA-SMU



Scheme 2: Synthetic scheme for ΔGA-SMU

This synthesis is similar to the previous synthetic scheme with the only exception being that the glycosylation step is much easier here and there is no need for silver(I) oxide activation of the glucuronyl bromide **3**. This is because thiophenols are better nucleophiles compared to phenols. However, it should be mentioned that thiophenols are prone to oxidation and formation of dimers, so this glycosylation reaction must proceed in the strict absence of oxygen.

After synthesis and characterization of these substrates, the kinetic parameters of these two substrates with the model enzymes were determined. A summary of these new kinetic parameters along with the previous data is shown in Table 2.

Entry	Enzyme	Substrate	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($s^{-1}\mu M^{-1}$)
1	UGC	ΔGA - ClMU	160 ± 21	10.8 ± 0.4	$(6.7 \pm 1.2) \times 10^{-2}$
2		ΔGA - MU	280 ± 25	8.2 ± 0.3	$(3.0 \pm 0.4) \times 10^{-2}$
3		ΔGA - SMU	1390 ± 181	7.9 ± 0.4	$(5.7 \pm 1.0) \times 10^{-3}$
4	BGE	GA - ClMU	63 ± 4	12.8 ± 0.2	$(2.0 \pm 0.2) \times 10^{-1}$
5		GA - MU	140 ± 12	13.6 ± 0.4	$(9.7 \pm 1.1) \times 10^{-2}$
6		GA - SMU	580 ± 90	0.0055 ± 0.0004	$(9.5 \pm 2.1) \times 10^{-6}$
7		ΔGA - ClMU	6.7 ± 0.7	0.048 ± 0.001	$(7.2 \pm 0.9) \times 10^{-3}$
8		ΔGA - MU	750 ± 130	0.037 ± 0.002	$(4.9 \pm 1.1) \times 10^{-5}$
9		ΔGA - SMU	- [a]	0.00049 [b]	- [a]
10	BGB	GA - ClMU	109 ± 9	1.36 ± 0.03	$(1.2 \pm 0.1) \times 10^{-2}$
11		GA - MU	280 ± 24	1.39 ± 0.05	$(5.0 \pm 0.6) \times 10^{-3}$
12		GA - SMU	2500 ± 380	0.060 ± 0.006	$(2.4 \pm 0.6) \times 10^{-5}$
13		ΔGA - ClMU	1260 ± 123	0.0146 ± 0.0008	$(1.2 \pm 0.2) \times 10^{-5}$
14		ΔGA - MU	- [a]	0.0015 [b]	- [a]
15		ΔGA - SMU	- [a]	0.00015 [b]	- [a]

Table 2: Kinetic data for all of the substrates at 37° C

Assay buffer: phosphate buffer pH = 6.6 for UGC and BGE and acetate buffer pH = 5 for BGB.

a) Data can't be determined under Michaelis-Menten conditions

b) Estimation of k_{cat} based on two data point in the plateau region of the Michaelis-Menten curve

As expected, for the β-glucuronidases, the rate of hydrolysis of GA-SMU is several orders of magnitude slower than GA-MU (comparing the values of k_{cat}/K_m: 4 orders of magnitude for BGE (entries 5 and 6) and 2 orders of magnitude for BGB (entries 11 and 12)). For UGC however, the overall rate of hydrolysis for ΔGA-SMU is only 5 times lower than that of ΔGA-MU (entries 2 and 3). Also the rate of hydrolysis of ΔGA-SMU with β-glucuronidases is negligible.

These results may suggest that ΔGA-SMU will be a good substrate for screening metagenomic libraries for GH88 UGL activity. However, use of ΔGA-SMU faces some serious problems. First and foremost, SMU is not fluorescent. Instead, it has a relatively large extinction coefficient at 370 nm, hinting that UV-Vis screening could be used alternatively. However, UV-Vis-based screening is generally not desirable compared to fluorescence-based screening. First and foremost, UV-Vis-based screening is much less sensitive than fluorescence-based screening. Secondly, background chromogenic activities are much more common than background fluorogenic activities. Finally, although even other derivatives of SMU that are fluorescent have been reported[45], there is yet another intrinsic problem with using thiophenols as fluorescent or UV-Vis active probes and that is that thiophenols are very prone to oxidation.

2.2.2.1 Kinetics of dimerization of SMU

In order to see if a thiophenol based probe can be used in screening metagenomic libraries, the rate of dimerization reaction of SMU (Figure 11), as a model thiophenol was inspected in aqueous phase. The decay of the UV-Vis absorbance of SMU at its maximum wavelength peak (370 nm) was followed at different starting concentrations of SMU. The resulting rates were then used to determine the observed first order rate constant of the dimerization reaction (Figure 12). Based on these results, the half-life of SMU in aqueous solution at pH = 7 is about only 2 hours at room temperature.

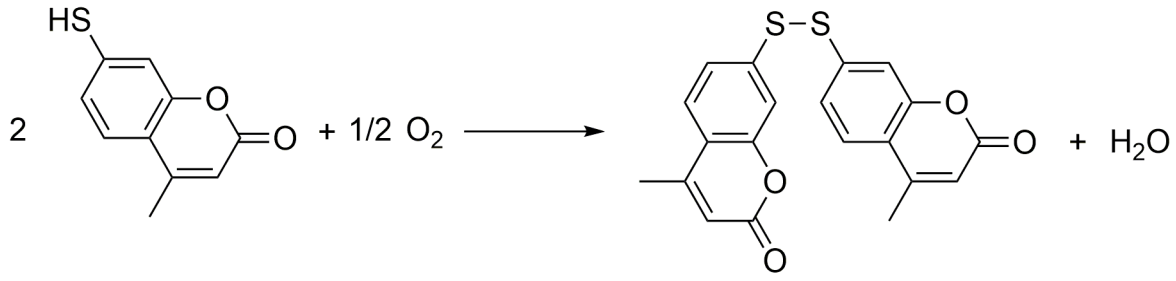


Figure 11: The dimerization reaction of SMU

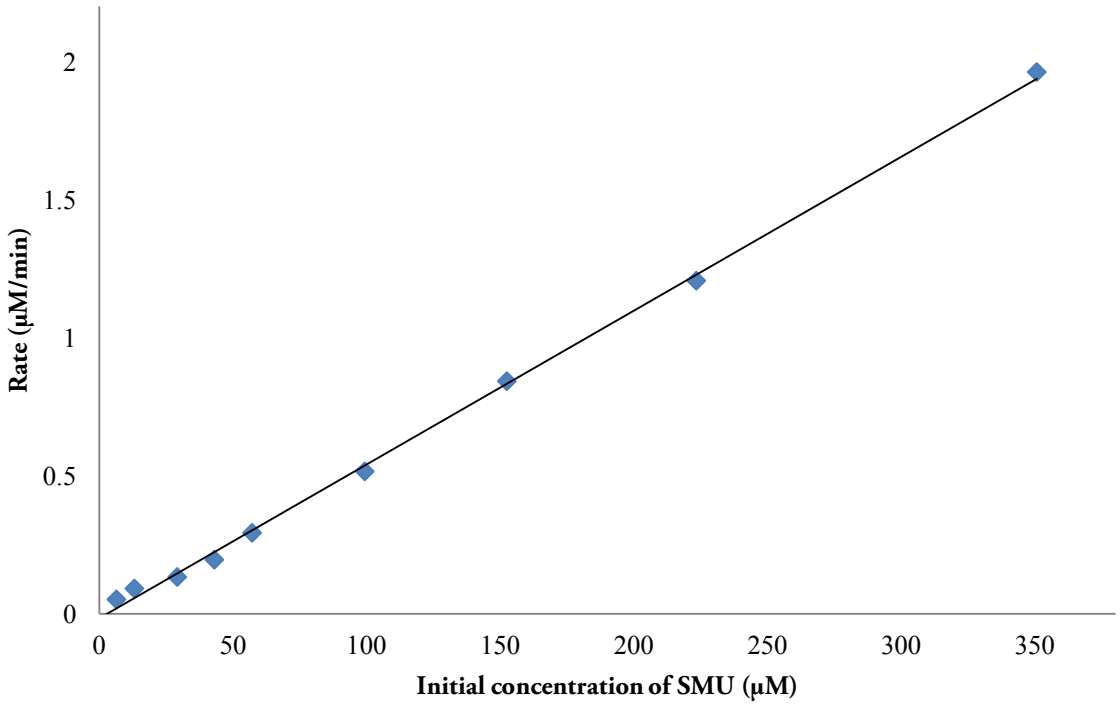


Figure 12: Observed rate of decay of SMU vs its initial concentration at pH = 7; linear regression of the data yields the equation y = 0.0056x - 0.0178 with R^2 = 0.9985

Given the fact that the standard protocol for screening metagenomic libraries usually involves incubation of the substrates with lysates for many hours, even a fluorescent derivative of SMU would not be useful for high throughput screening of these libraries, since it is not going to be stable under assay conditions.

2.3 Design and synthesis of fluorogenic substrates with self-immolative thiol linkers

When screening metagenomic libraries for UGLs, the ideal substrate will be as sensitive as ΔGA-ClMU and as selective as ΔGA-SMU. Thus, the ideal molecule will have a sulfur atom attached to the anomeric carbon and will also release a fluorescent ClMU upon hydrolysis. Such a molecule would consist of three parts: an unsaturated glucuronic acid, a thiol-based self-immolative linker that will spontaneously release the third part of the molecule, ClMU, upon cleavage of the glycosidic bond. A general structure of this molecule is shown in Figure 13.

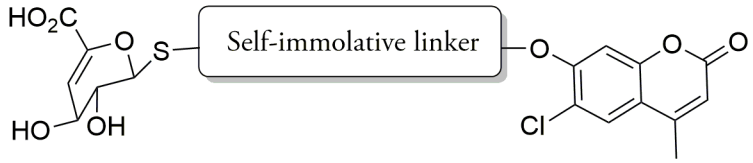


Figure 13: General structure of selective screening substrate for UGLs

The self-immolative linker of this molecule determines how effective it will be for screening purposes. First of all, the linker should be able to release the ClMU molecule quickly and quantitatively after it is cleaved from the sugar. Second, it should be small enough to be accommodated in the enzyme active site. Thirdly, it should be a good enough leaving group to be cleaved by the low concentrations of enzymes present in the metagenomic screens. Finally, it should be stable in the presence of other reagents that inevitably will be present in the screening condition (e.g. nucleophiles, reducing agents, etc. from cell lysate) so that it does not release the ClMU unless it is cleaved by the enzyme. Two substrates with two different self-immolative linkers were designed and synthesized for this purpose. The structures of these molecules are shown in Figure 14.

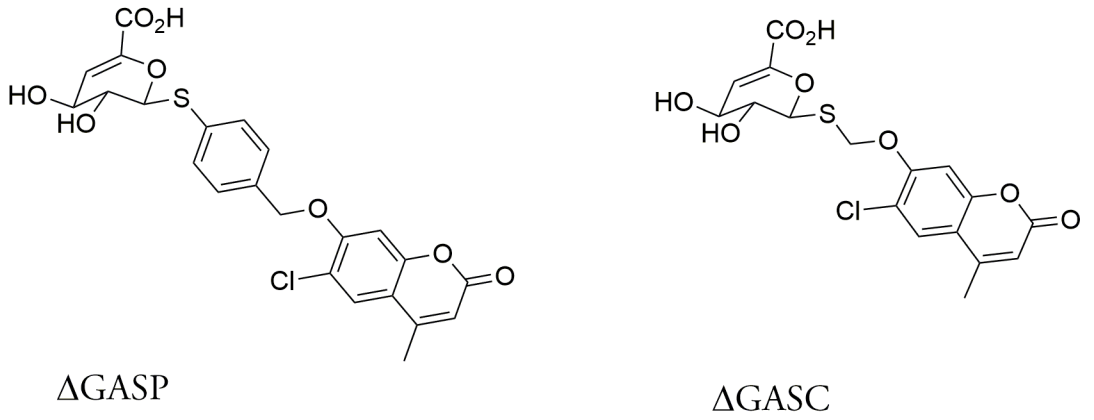


Figure 14: The structures of ΔGASP and ΔGASC

The first molecule, ΔGASP, has a thioquinone methide-generating linker. The mechanism of disintegration of ΔGASP and the structure of the generated thioquinone methide is shown in Figure 15. The generated thioquinone methide is an unstable intermediate that will react with surrounding nucleophiles or will get reduced[46].

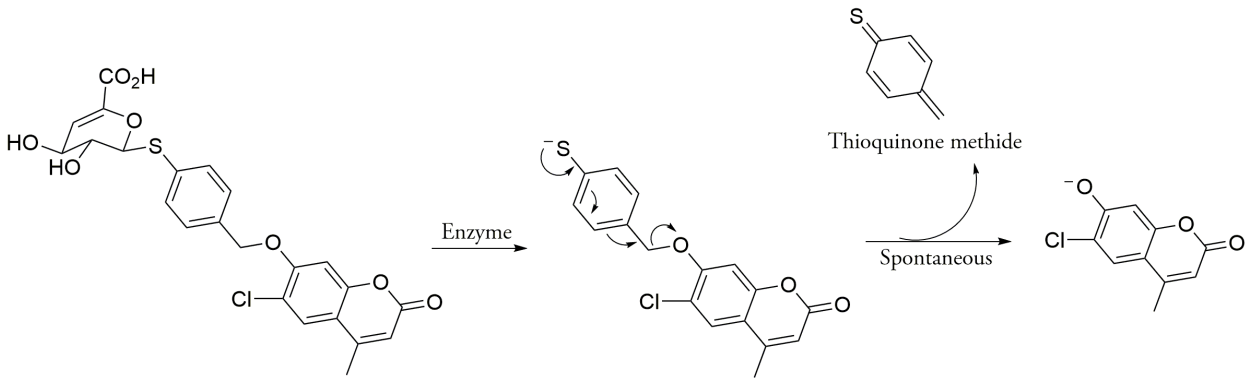


Figure 15: The mechanism of disintegration of ΔGASP

Quinone methides are well known reactive intermediates that have been used extensively as self-immolative linkers as well as in a variety of other different applications[47]. Thioquinone methides however are far less studied. Their use as intermediates in synthesis has been explored[48–51], one study reports isolation of an example of a stable thioquinone methide[52] and in other studies they are the reactive motif of mechanism-based inhibitors[53–55]. Finally, thioquinone methides as self-immolative

linkers have been reported as part of the structure of reduction-triggered prodrugs[56] and recently as the reactive linker of glutathione-sensitive sulfonate ester protecting groups[46]. The structure of ΔGASP was inspired by previous work in the Withers lab for synthesis of substrates for droplet-based fluorescence screening[57].

The linker for the second substrate is a S,O-acetal that will decompose to thioformaldehyde and release a ClMU as shown in Figure 16. Thioformaldehyde is an unstable molecule[58] that has never been isolated and is believed to spontaneously trimerize to 1,3,5-trithiane[59]. However, it is not known whether thioformaldehyde that is released in the aqueous phase will turn into 1,3,5-trithiane or not.

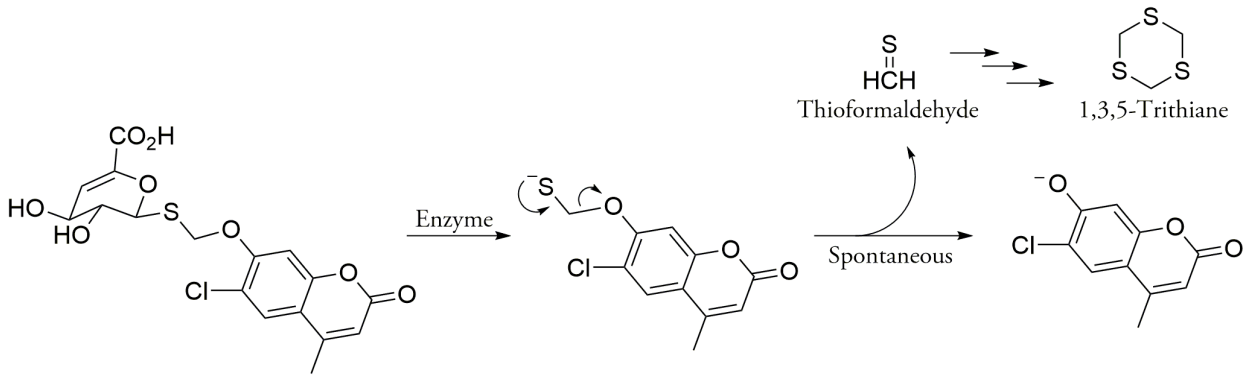


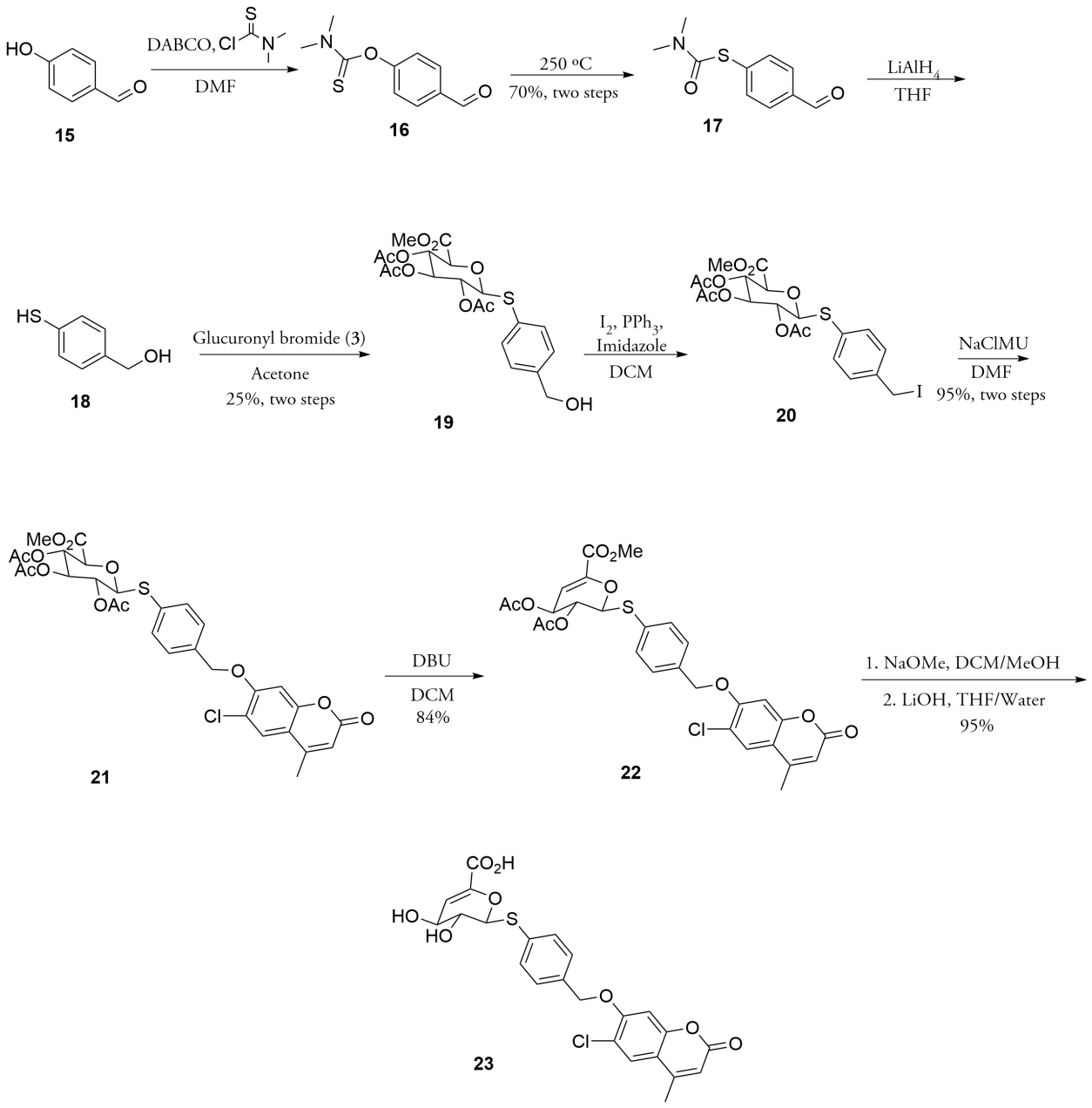
Figure 16: The mechanism of disintegration of ΔGASC

Reported uses of bis-acetal linkers in literature is nowhere near the extensive use of quinone methides. They mainly have been reported as acid labile linkers[60] for application as prodrugs[61,62] and as protein cross linkers[63,64]. Bis-thioacetal linkers have been reported as stable linkers to attach glycosides to peptides[65] and also have been isolated as part of a group of β-thioglucosides from seeds of *Afrostyrax lepidophyllus*, a genus of tree native to equatorial Africa[66]. The structure of our substrate was inspired by a bis-acetal FRET based substrate for monitoring glycosidase activity[67] and another molecule that has been tested unsuccessfully as a substrate of glycosaminidases[68]. However, these two

substrates have bis-acetal linkers and this study is the first report of a self-immolative S,O-acetal linker to the best of our knowledge.

2.3.1 Synthesis of ΔGASP

The synthetic scheme of ΔGASP is shown in Scheme 3. First 4-hydroxybenzaldehyde (**15**) is transformed to 4-thiohydroxybenzylalcohol (**18**) in three steps. This thiophenol serves as a nucleophile to displace the bromide group of glucuronyl bromide (**3**) to give compound **19**. Next, the benzyl alcohol is converted into the corresponding iodide, a good leaving group, which is subsequently displaced by ClMU in a nucleophilic substitution reaction. Finally, elimination of acetate by DBU and deprotection yields the final product.

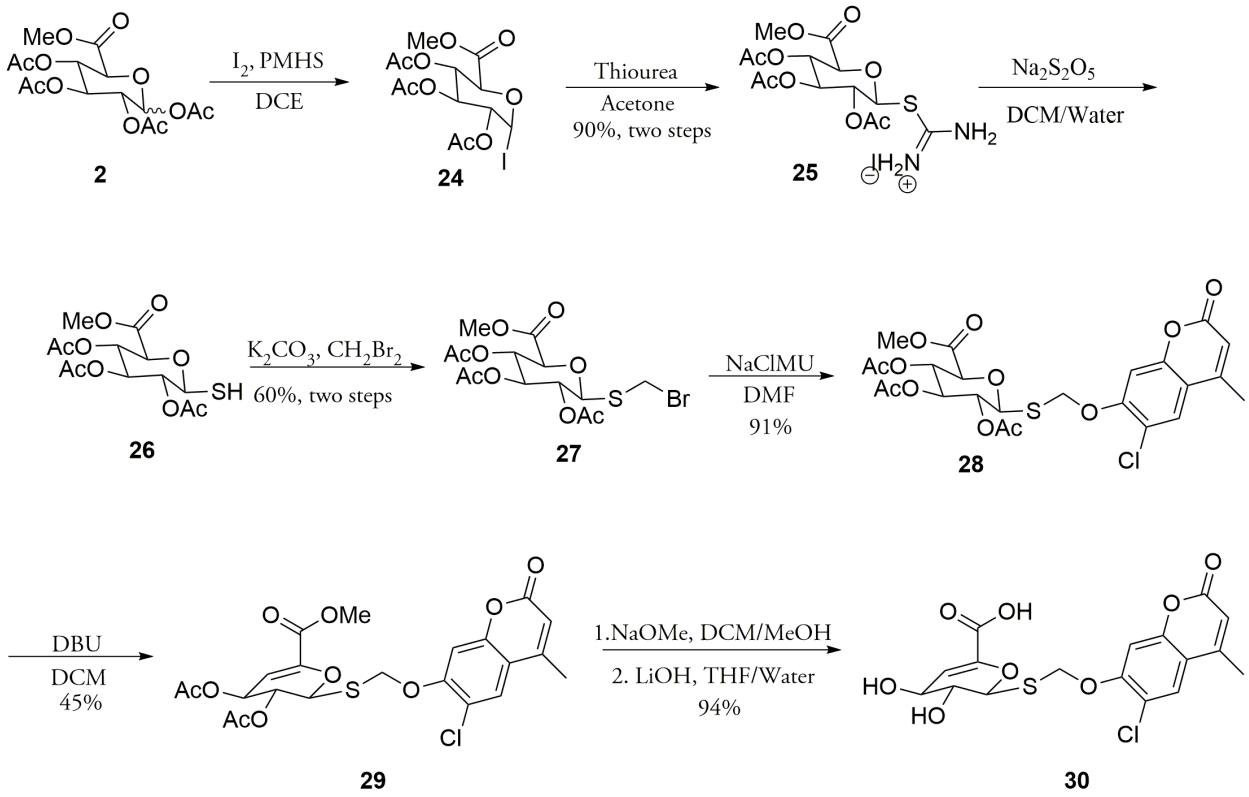


Scheme 3: Synthetic scheme for ΔGASP

2.3.2 Synthesis of ΔGASC

The synthetic scheme for ΔGASC is depicted in Scheme 4. The synthesis starts with conversion of the anomeric acetate group of **2** into the corresponding iodide. Glycosyl iodides are the more reactive analogues of commonly used glycosyl bromides. In addition to their better reactivity, their

synthesis[69] is much easier than glycosyl bromides and this may make them more favorable than glycosyl bromide for use in future synthesis. In the case of this synthesis for example, glucuronyl bromide was found to be too stable to react with thiourea. The reaction with glucuronyl iodide however, was finished within 2 hours under the same conditions. Compound **25** was then reduced to compound **26** which served as the nucleophile in the next step to attack dibromomethane to yield compound **27**. Next, the bromide was displaced by a ClMU in a substitution reaction to yield **28**. This compound was then subjected to DBU elimination of acetate and deprotection to yield the final product, ΔGASC (**30**).



Scheme 4: Synthetic scheme for ΔGASC

2.4 Testing the selective substrates

In order to imitate the standard screening procedure, usefulness of the new substrates was tested with cell lysates and not purified enzymes. *E. coli* strain BL21(DE3) carrying individually the plasmids for expressing UGC, BGE and ABG (*Agrobacterium sp*. β-glucosidase as negative control) were grown overnight in auto-induction LBE 5052 media. In addition, to test the amount of background expression of BGE, the strain that our metagenomic libraries are expressed in, *E. coli* strain EPI300, was included in this experiment. Each one of these overnight cultures was then mixed with the lysis buffer (phosphate buffer 50 mM pH = 7, containing 1% Triton X-100) containing 50 μM of one of the three substrates: ΔGA-ClMU, ΔGASC and ΔGASP. After overnight incubation of the substrates with cell lysate, the fluorescence was measured and the results are summarized in Figure 17.

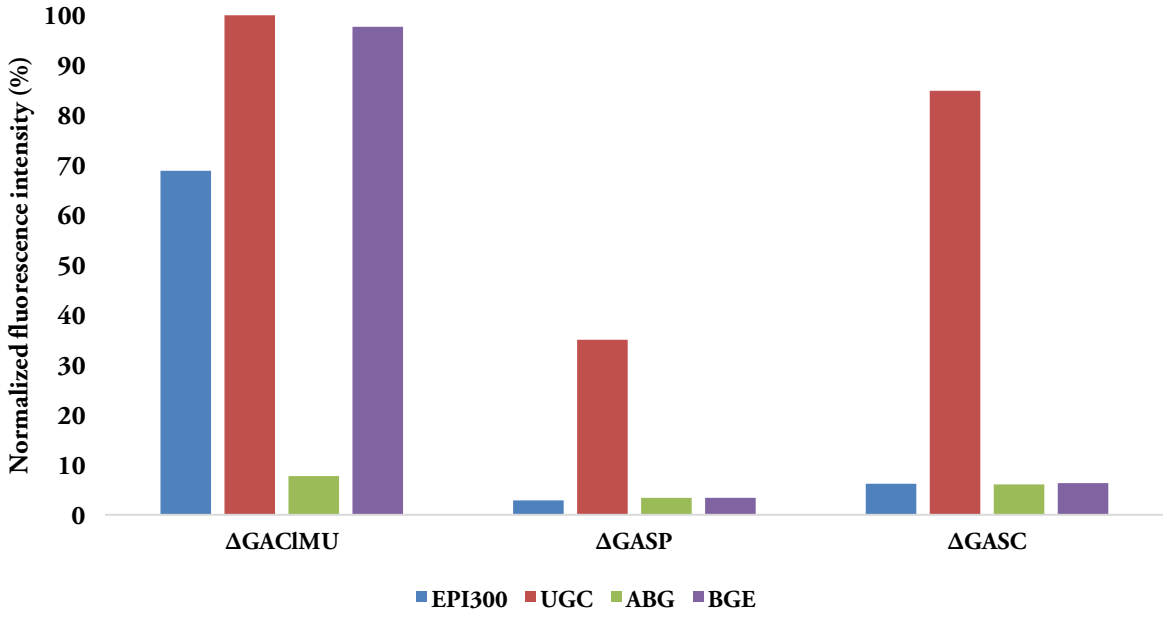


Figure 17: Fluorescent readings for the substrates after overnight incubation with cell lysates

As expected, ΔGA-ClMU does not distinguish between UGC and BGE. Moreover, the fluorescence signal for BGE is almost as high as UGC. Since the expression yields of these enzymes are different, this is most probably because the expression yield of BGE is higher compared to UGC. In addition, even the background activity of EPI300 strain shows up as a positive hit with this substrate. However, both of our selective substrates, ΔGASC and ΔGASP, only show UGC as a positive hit.

Also of note is that the resulting fluorescence signal from incubation of ΔGASP with the cell lysate is lower than that of ΔGASC. We speculate that this is because the free thiophenol after hydrolysis of ΔGASP can be oxidized to its dimer prior to forming the thioquinone methide (Figure 18). This can happen especially quickly in the screening buffer, since oxidizing agents will be present in solutions containing Triton X-100 because they are by-products of its degradation[70].

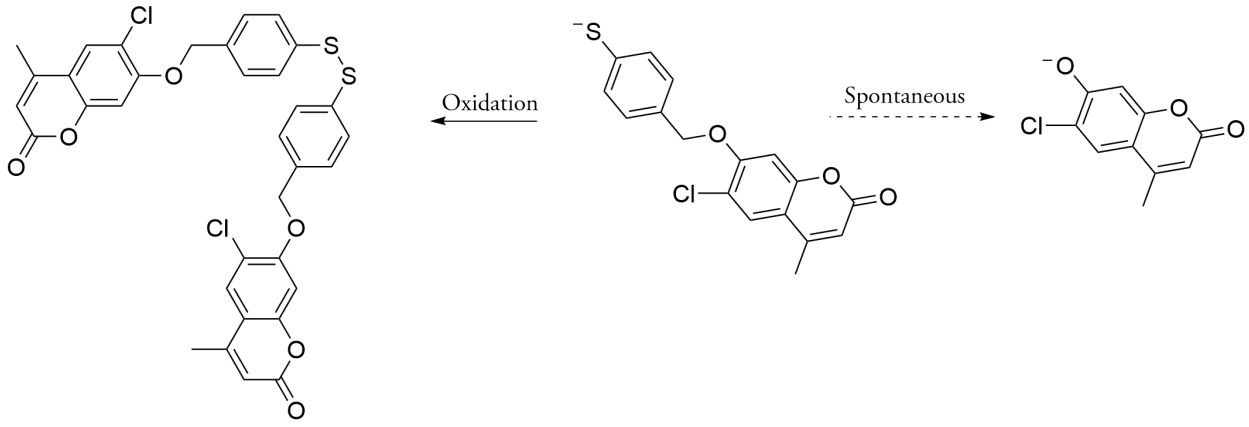


Figure 18: The speculated reason for the low fluorescence of ΔGASP

To test this assumption, one more experiment was conducted to compare the fluorescence of the substrates in the absence and presence of Triton X-100. ΔGA-ClMU, ΔGASC and ΔGASP were added (final concentrations of 50 μM) to solutions of UGC in 50 mM phosphate buffer pH = 7 in duplicates, with and without 1% Triton X-100, and incubated overnight at 37 °C. A comparison of the overnight fluorescent signals (Figure 19) first of all shows that the presence of Triton X-100 generally results in

reduction of fluorescence intensity for all substrates. More importantly, this shows that the resulting fluorescence for hydrolysis of ΔGASP is very sensitive to the presence of Triton X-100, as predicted.

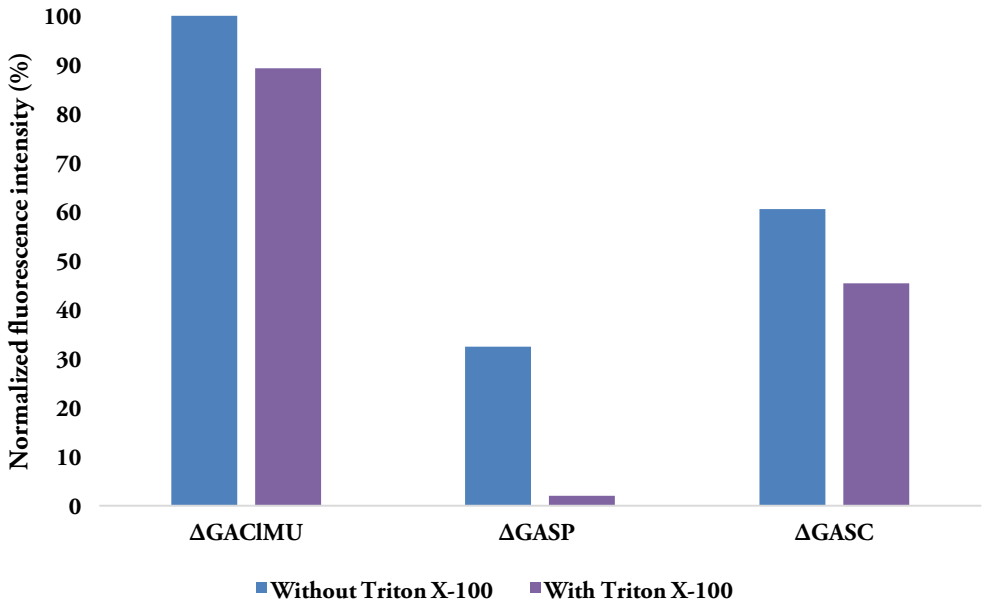


Figure 19: Comparison of fluorescence signal after overnight incubation of substrates with UGC in the absence and presence of Triton X-100

It should be noted that, although in this experiment Triton X-100 resulted in complete extinction of the fluorescence signal for hydrolysis products of ΔGASP, this is not always the case. The extent of the reduction in fluorescence signal depends on other factors, such as the concentration of enzyme and other reagents present in screening buffer (e.g. the reducing agents that are present in the cell lysate seem to alleviate this problem). In addition, instead of adding Triton X-100, lysing the cells can be achieved by other methods, such as sonication or freeze and thaw cycles. However, even in the absence of Triton X-100, the aglycone of ΔGASP is susceptible to oxidation and hence the resulting fluorescence from hydrolysis of ΔGASP has a lower signal to noise ratio, compared to ΔGASC.

2.5 Screening metagenomic libraries

Finally, a relatively small subset of two of our metagenomic libraries was screened for UGL activity. These small libraries are subsets of two metagenomic libraries that have been made previously in the Withers lab from microbial population of human and beaver feces and consist of nearly 500 *E. coli* clones each containing a fosmid of about 25 genes, thus potentially 12000 different genes.

These libraries were screened using ΔGA-ClMU, ΔGASC and ΔGASP. The resulting fluorescence readings from incubation of these substrates with the libraries were monitored after 2, 5 and 12 hours. The data for this screening is reported in terms of Z-score. The Z-score of a data point is the number of standard deviations by which the value of the data point is different from the average value of all data points. For ΔGA-ClMU, it was found that the results are the most reproducible after 5 hours of incubation, since this substrate suffers from background hydrolysis. Thus the Z-score of each well was calculated based on the fluorescence after 5 hours of incubation for ΔGA-ClMU and after 12 hours of incubation as per usual for the other two substrates.

For the human fecal library, the screening with the selective substrates (data not shown), and ΔGA-ClMU (Figure 20) does not show any hits with a Z-score of above 3. Thus, it seems that no UGL is expressed in this subset of the human fecal library.

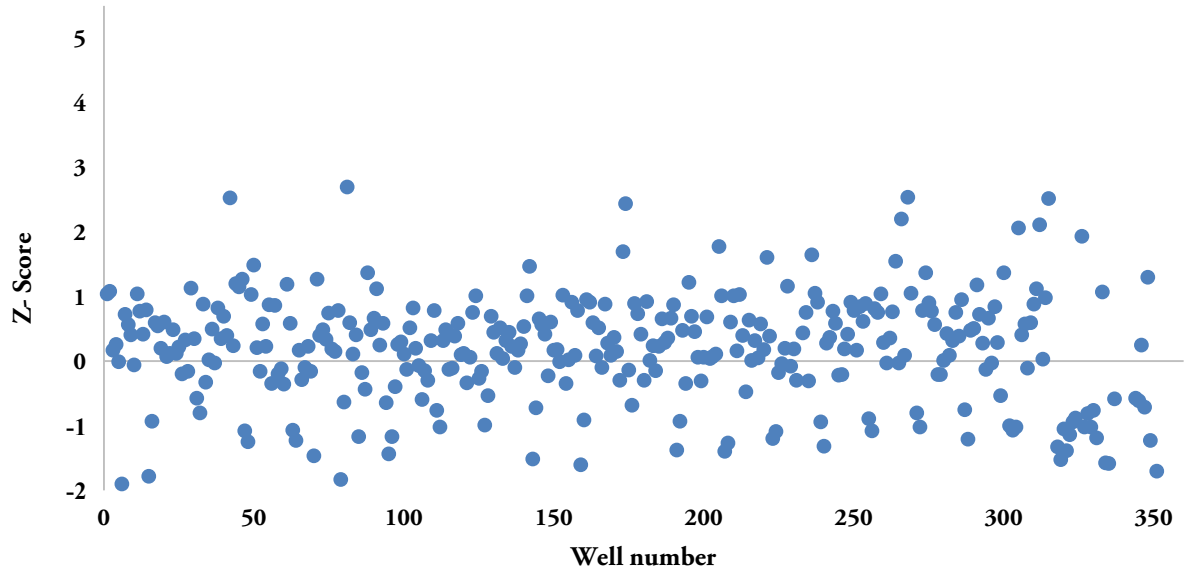


Figure 20: Z-score vs well number for screening the human fecal library with ΔGAClMU

On the other hand, for the beaver fecal library, screening with ΔGA-ClMU gives rise to two stand-out hits (wells D09, F09) (Figure 21). However, screening with ΔGASC (Figure 22) and ΔGASP (Figure 23) gives rise to only one hit (well F09). We anticipated that the one hit with our selective substrates should be a UGL, while the other hit identified from screening with ΔGA-ClMU will be a false positive.

In order to verify this, the DNA samples from the hits were prepared and sent for sequencing. The sequence of the metagenomic DNA from these hits is deposited in GenBank under the accession number SAMN03389402. Analysis of these sequences identifies the open reading frames (ORFs) shown in Figure 24. Further BLAST analysis of each of these ORFs identifies those that correspond to carbohydrate-active enzymes and the GH families to which they belong (Figure 24).

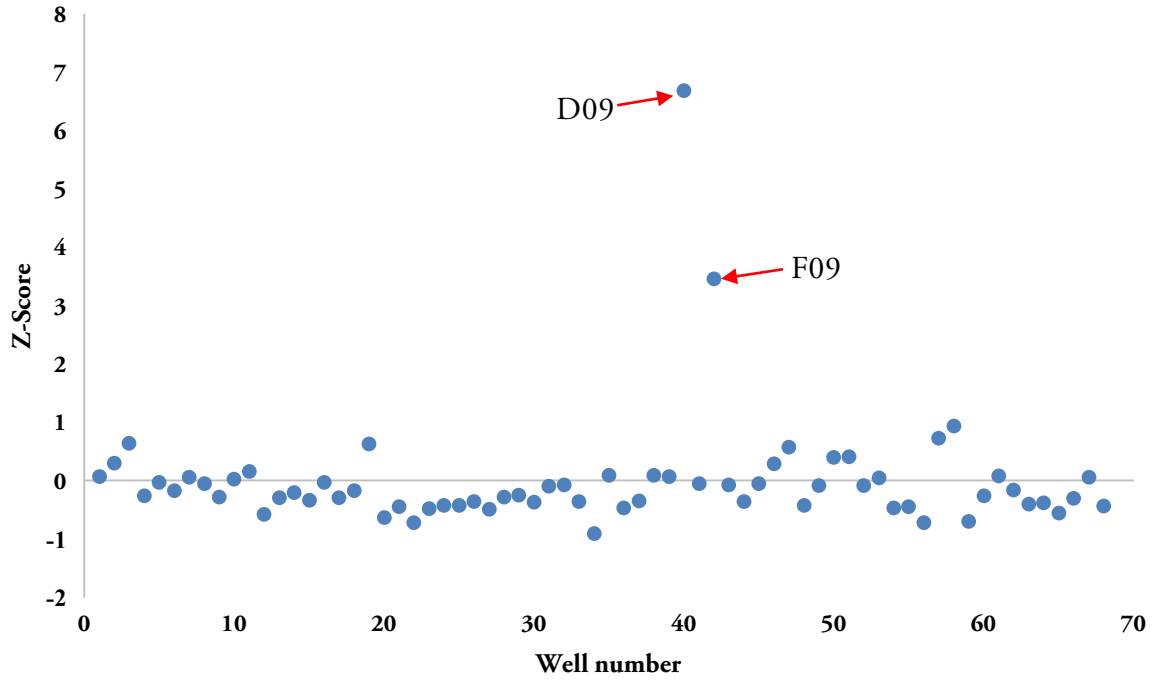


Figure 21: Z-score vs well number for screening the beaver fecal library with ΔGA-ClMU

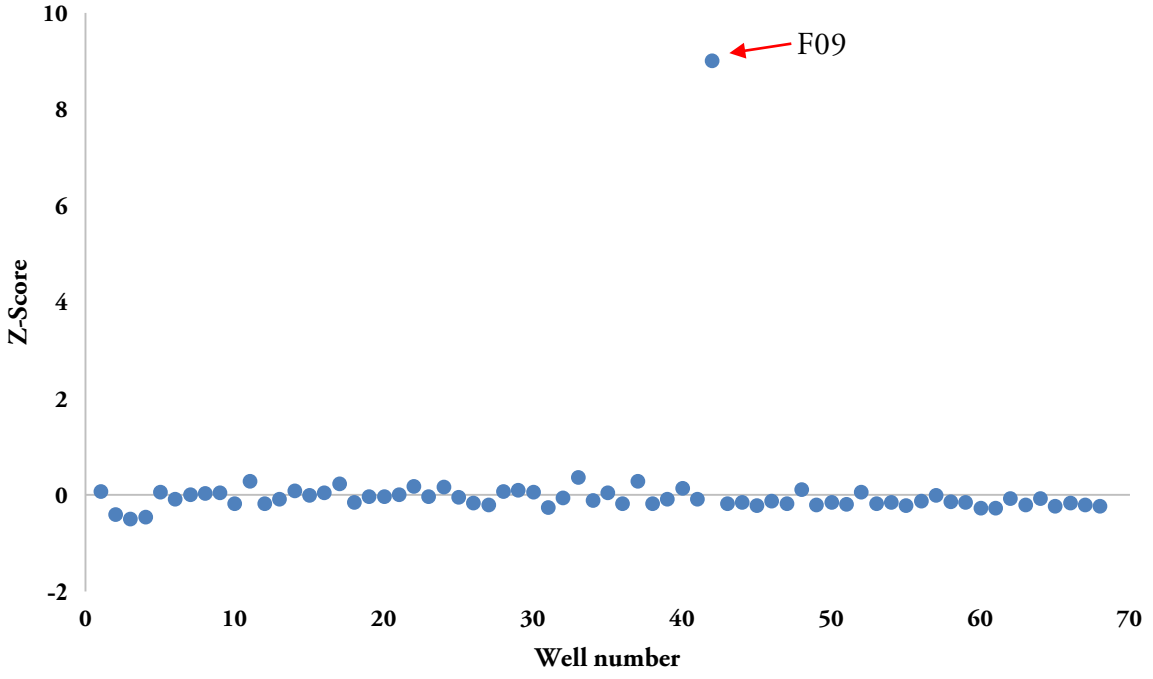


Figure 22: Z-score vs well number for screening the beaver fecal library with ΔGASC

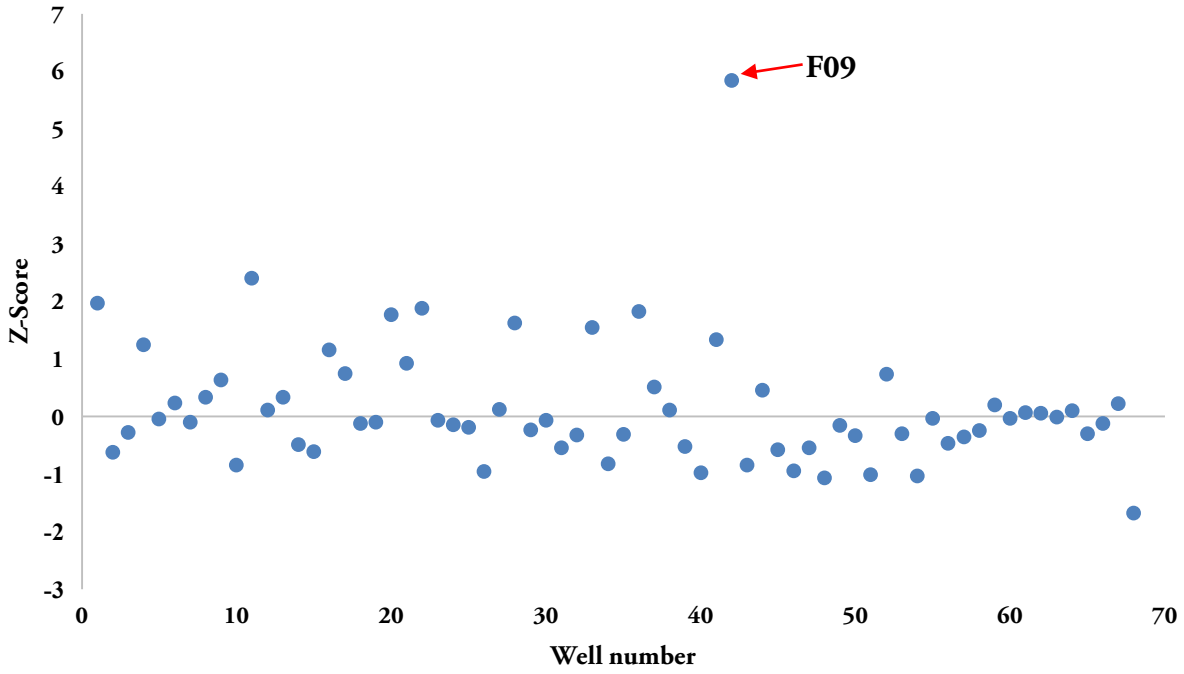


Figure 23: Z-score vs well number for screening the beaver fecal library with ΔGASP

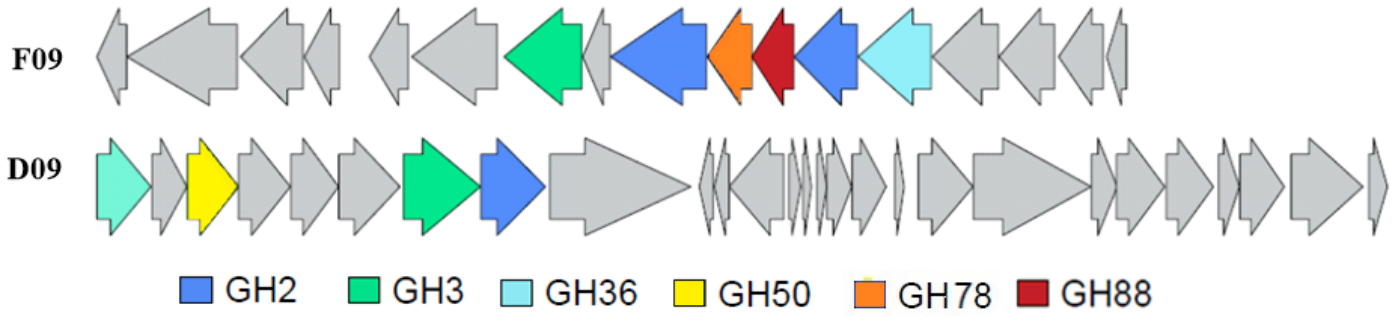


Figure 24: The genes present in the hits

Confirming the exact gene that is responsible for the observed activity will require sub-cloning of the genes present and is beyond the scope of this thesis. However, inspecting the GH genes present in these hits it can be speculated where these activities come from. The known principle activities of the GH families present in these hits are listed in Table 3.

As for the one hit that is active towards all the substrates, F09, it can be speculated that the UGL activity is from the GH88 gene present in its sequence. As for the other positive hit from screening with ΔGA-ClMU, D09, no known UGL is present in this hit. However, this hit contains a GH2 gene and the observed UGL activity can be attributed to the GH2 gene, assuming that it is a β-glucuronidase.

GH Family	Known principle activities
2	β-Galactosidase , β-Glucuronidase, β-Mannosidase
3	β-Glucosidase, β-Xylosidase
36	α-Galactosidase, α-N-Acetylgalactosaminidase
50	β-Agarase
78	α-L-Rhamnosidase
88	UGL

Table 3: The known principle activities of GH families present in the hits

Chapter 3: Conclusion

The specific aim of the work presented in this thesis, was to develop novel selective substrates for screening metagenomic libraries for UGLs. Conventional substrates for screening metagenomic libraries for UGLs can also be hydrolyzed by β-glucuronidases, giving rise to false positives hits and high background activity. Taking advantage of the differences in the mechanisms of known UGLs and β-glucuronidases, selective screening substrates were made by replacing the anomeric oxygen of the conventional substrates with sulfur.

However, since thiophenols are not stable under assay conditions, a self-immolative thiol linker that releases a fluorophore upon hydrolysis was used as part of the substrate. To this end, two substrates with different self-immolative thiol linkers were synthesized and tested with different degrees of success. One of these linkers, the S,O-acetal linker is reported for the first time in this study.

Finally, a small metagenomic library was screened with these substrates and sequences of the hits shows that these new substrates successfully identify the UGLs, while conventional substrates give rise to false hits. Larger metagenomic libraries will be screened in the future with these new substrates in a search for novel UGLs.

The broad aim of this thesis was to generate a general strategy that can be employed in making selective substrates for screening metagenomic libraries for some other glycosidase activities. An interesting example of such an activity, the screening for which faces the same problems as in screening for UGLs, is thioglycosidase activity. We believe that similar substrates can be synthesized and used for screening metagenomic libraries for thioglycosidases.

While both of the substrates described in this thesis were found to be useful in screening for known families of UGLs, it cannot be known whether the possible unknown families of enzymes can hydrolyze these substrates or not. Therefore, the limitations and the differences of these substrates should be carefully taken into consideration for future applications.

For example, the aglycone for ΔGASP is an activated thiophenol, while ΔGASC has an inactivated thiol as the aglycone. Thus, some of the enzymes that are able to hydrolyze substrates with an thioquinone methide generating linker may be not efficient in hydrolyzing substrates with an S,O-acetal linker. In order to make the S,O-acetal linker a better leaving group and achieve better sensitivity, electron-withdrawing groups such as trifluoromethyl can be added to the linker carbon to reduce the pKa of the adjacent thiol (Figure 25).

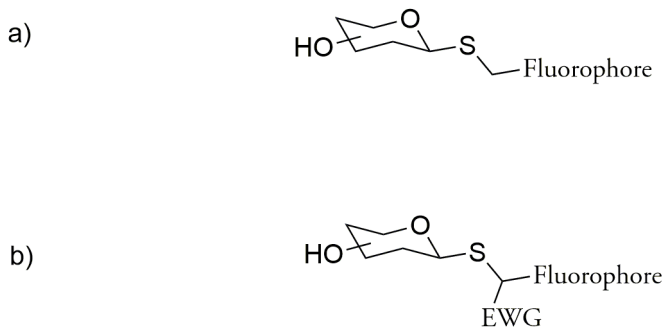


Figure 25: Summary of the discussion about S,O-acetal linker

a) The S,O-acetal generating substrate presented in this thesis

b) The possible candidate for enhancing activity of substrates with a S,O-acetal linker

On the other hand, ΔGASP is substantially larger than ΔGASC. Some of the enzymes may find it challenging to accommodate the large aglycone of a thioquinone methide linker in their active site and hence not be able to hydrolyze these substrates. However, the S,O-acetal linker of ΔGASC seems to be the smallest possible linker of its kind and should not suffer from this problem.

In addition to the problem with the large size of its aglycone that cannot be helped, ΔGASP was found to not be as sensitive as ΔGASC. This is speculated to be because the aglycone of ΔGASP is susceptible to oxidation. Thus, making the process of formation of thioquinone methides faster or making the process of dimerization of thiophenols slower could enhance the sensitivity of substrates with a thioquinone methide linker. Adding electron-donating groups on the aromatic ring makes the quinone methides more stable[71] and consequently the formation of the thioquinone methide and release of the fluorophore faster. In addition, if ortho to the mercapto group, these groups could sterically hinder formation of the dimer of the thiophenol. The number and position of these groups as well as the extent of their electron-donating ability should be determined by synthesizing these substrates and testing them (Figure 26).

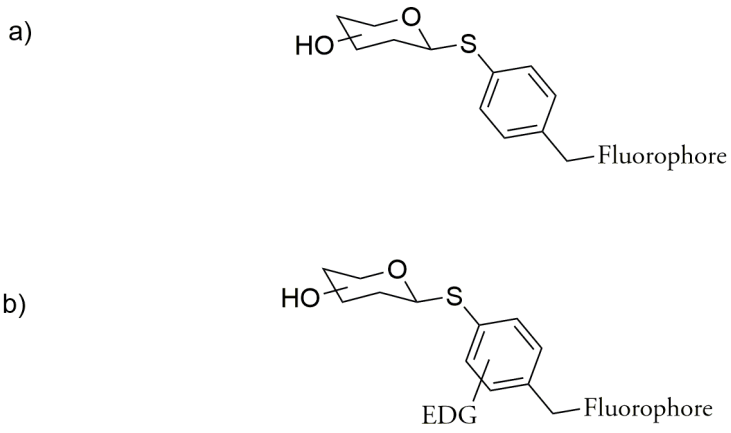


Figure 26: Summary of the discussion about thioquinone methide linker

a) The quinone methide generating substrate presented in this thesis

b) The possible candidate for enhancing activity of substrates with a thioquinone methide linker

In summary, the usefulness of a thioquinone methide-generating linker versus an S,O-acetal linker should be determined by experiment for each class of enzyme for which these thioglycosides may serve

as suitable substrates for screening purposes. In addition, the sensitivity of these substrates can be possibly enhanced by the changes in their structure suggested in this section.

Synthesis of relevant substrates for thioglycosidases and screening metagenomic libraries with the best of these substrates will be carried out in the future. Hopefully, this will lead to the discovery of unknown families of thioglycoside-cleaving enzymes with novel mechanisms.

Chapter 4: Experimental

4.1 Materials and Methods

All of the chemicals were purchased from Sigma-Aldrich or Alfa-Aesar unless otherwise noted. Solvents for anhydrous reactions were distilled under a nitrogen atmosphere; CH_2Cl_2 and acetonitrile over CaH_2 and methanol over Mg. The glassware was dried in an oven. All commercially available reagents and solvents were used without further purification. TLC was performed on silica plates 60 F_{254} aluminum sheets (Merck, Germany). TLC Spots were visualized with UV light and/or through staining with 10% ammonium molybdate in 2 M H_2SO_4. Flash column chromatography was carried out using 230-400 mesh silica gel.

^1H- and ^{13}C-NMR spectra were recorded in $CDCl_3$, $(CD_3)_2CO$, $(CD_3)_2SO$ or MeOD and are reported in δ scale in ppm and are referenced to one of the following: chloroform (δ 7.26 ppm for ^1H, δ 77.16 ppm for ^{13}C), $(CD_3)_2CO$ (δ 2.05 ppm for ^1H, δ 206.26 ppm for ^{13}C), $(CD_3)_2SO$ (δ 2.50 ppm for ^1H, δ 39.52 ppm for ^{13}C), MeOD (δ 3.31 ppm for ^1H, δ 49.00 ppm for ^{13}C). Data for the NMR spectra are reported as follows: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, br = broad signal, J = coupling constant in Hz.

Low resolution mass spectra (LRMS) were obtained using a Waters ZQ mass spectrometer equipped with ESCI ion source. High resolution mass spectra (HRMS) were recorded by the University of British Columbia mass spectrometry facility on a Waters/Micromass LCT with time of flight detection and electrospray ionization.

Enzyme kinetic measurements were performed using a Varian Cary 4000 spectrophotometer with automatic cell changer. All non-linear regressions were performed using GraFit 7.0 (Erithacus software

limited; www.erithacus.com/grafit). Kinetic measurements for the dimerization reaction of SMU and all of the screening experiments were performed using a Synergy H1 plate reader (BioTek).

4.1.1 General Procedures

4.1.1.1 General procedure for elimination using DBU

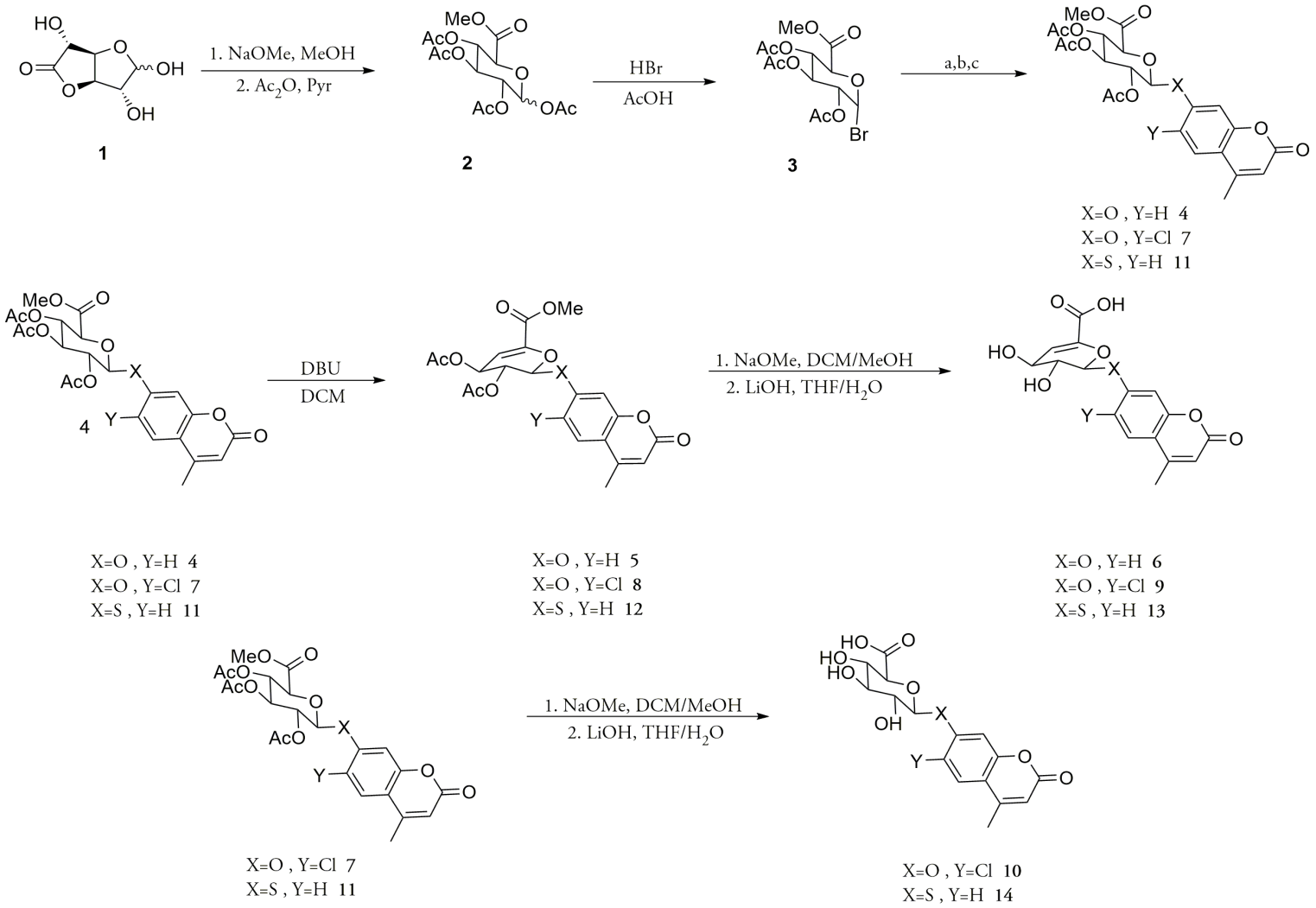
Pre-dried globally protected aryl glucuronide was dissolved in dried DCM (to give approximately 0.1 M of sugar). Molecular sieves were added and the reaction mixture was flushed with nitrogen and put under argon. DBU (1.1 eq.) was added drop-wise and the reaction mixture was stirred until TLC indicated the starting material was consumed. The reaction mixture was then loaded directly on a silica column and purified by flash column chromatography.

4.1.1.2 General procedure for Zemplén deprotection

Pre-dried globally protected aryl glucuronide (or aryl unsaturated glucuronide) was dissolved in dry 1:1 DCM/methanol mixture to approximately 0.1 M of sugar and was put under argon and cooled to 0 °C. A catalytic amount of freshly prepared sodium methoxide was then added and the reaction mixture was stirred until TLC indicated the end of the reaction.

4.1.1.3 General procedure for hydrolysis of methyl ester protecting group by aqueous lithium hydroxide

To the reaction mixture from Zemplén deprotection, after evaporation of the solvents, THF and water (3-5 mL) and 1 M LiOH (1.1 eq.) were added and the reaction mixture was stirred until TLC indicated the end of the reaction (typically 5 min). The reaction was then quenched with Amberlite IR-120(H) and filtered. The solvents were then evaporated and the product was dissolved in methanol and precipitated out of the solution by addition of diethyl ether.

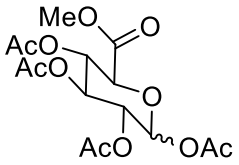


Scheme 5: Synthetic scheme for compounds 2-14

a) synthesis of **4**: MU, Ag_2O, ACN; b) synthesis of **7**: ClMU, DCM, 5% NaOH, TBAHS; c) synthesis of **11**: SMU, K_2CO_3, Acetone

4.1.2 Synthesis and characterization of compounds

Methyl 1,2,3,4-tetra-*O*-acetyl-α/β-D-glucopyranouronate (2)



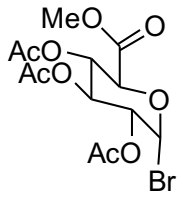
D-Glucuronic acid γ-lactone (15.2 g, 0.09 mol) was dried under vacuum and suspended in methanol (60 mL, 1.5 mol) at room temperature. Na metal (0.1 g, 0.03 mol, 0.03 eq.) was added and the reaction mixture was stirred until the reaction was complete (4 hours). The solvent was evaporated and the reaction mixture was dried under vacuum. Acetic anhydride (105 mL, 1.1 mol) and pyridine (70.5 mL, 0.87 mol) were then added and the reaction mixture was stirred at 4 °C overnight. 1 M HCl (50 mL) was then added to the reaction mixture, which was then extracted with ethyl acetate (3x). The pooled organic fractions were washed with distilled water (3x), saturated $NaHCO_3$ (3x) and brine (3x). The organic phase was then dried over Na_2SO_4, the solvent was evaporated and the product was crystallized from ethyl acetate and hexanes. The crystals were filtered, washed with hexanes and dried under vacuum. Yield: 28.44 g (0.08 mol, 84 %, The crystals are 100% β. Mother liquor was evaporated and found to be a mixture of α and β. NMR data are reported for the β anomer).

LRMS: Calcd. for $C_{15}H_{20}NaO_{11}$: 399.09 found 399.0.

^{1}H NMR (400 MHz, Acetone-d_6): δ 5.95 (d, J = 8.0 Hz, 1H, H-1), 5.45 (dd, J = 9.5 Hz, 1H, H-2), 5.16 (dd, J = 9.6 Hz, 1H, H-3), 5.08 (dd, J = 9.5, 8.0 Hz, 1H, H-4), 4.52 (d, J = 9.8 Hz, 1H, H-5), 3.70 (s, 3H, OMe), 2.08 (s, 3H, OAc), 2.00 (s, 6H, OAc), 1.97 (s, 3H, OAc).

^{13}C NMR (101 MHz, Acetone-d_6) δ 169.8, 169.6, 169.4, 169.0, 167.4, 91.6, 72.8, 71.9, 70.5, 69.6, 52.6, 20.2, 20.1, 20.0 (two peaks).

Methyl (2,3,4-tri-*O*-acetyl-α-D-glucopyranosyl bromide)uronate (3)



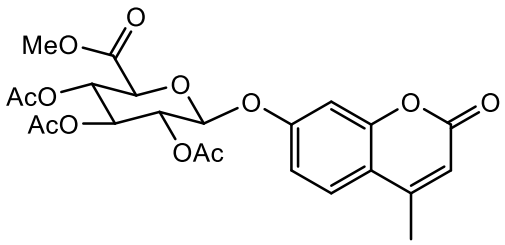
To methyl (1,2,3,4-tetra-*O*-acetyl-α/β-D-glucopyranoside) uronate (**2**) (7.0 g, 18.6 mmol) under argon and at 4 °C were added acetic acid anhydride (15 mL) and 33% HBr in AcOH (115 mL, 0.7 mol) and the mixture was stirred for 2 days. The reaction mixture was then extracted with ethyl acetate (2x), the pooled organic phases were washed with water (3x), saturated $NaHCO_3$ (3x) and brine and dried over Na_2SO_4. The solvent was evaporated and the product was crystallized from toluene/hexane, yielding yellowish crystals: 6.06 g, 0.015 mol, 85 %.

LRMS: Calcd. for $C_{13}H_{17}BrNaO_9$: 419.00 found 419.0.

^{1}H NMR (400 MHz, Acetone-d_6) δ 5.95 (d, J = 8.0 Hz, 1H, H-1), 5.45 (t, J = 9.5 Hz, 1H, H-3), 5.16 (t, J = 9.6 Hz, 1H, H-4), 5.08 (dd, J = 9.5, 8.0 Hz, 1H, H-2), 4.52 (d, J = 9.8 Hz, 1H, H-5), 3.70 (s, 3H, OMe), 2.08 (s, 3H, OAc), 2.00 (s, 3H, OAc), 1.99 (s, 3H, OAc), 1.97 (s, 3H, OAc).

^{13}C NMR (101 MHz, Acetone-d_6) δ 169.8, 169.6, 169.4, 169.0, 167.4, 91.6, 72.8, 71.9, 70.5, 69.6, 52.6, 20.2, 20.1, 20.0 (two peaks).

Methyl (4-methylumbelliferyl 2,3,4-tri-*O*-acetyl-β-D-glucopyranoside)uronate (4)



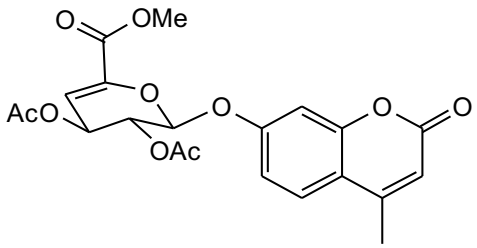
Bromide **3** (850.0 mg, 2.1 mmol), 7-hydroxy-4-methylcoumarin (519.0 mg, 2.9 mmol) and Ag_2O (926 mg, 4 mmol) were suspended in dry acetonitrile under argon and the reaction mixture stirred overnight at room temperature in the dark. The solvent was then evaporated and the resultant solid was suspended in ethyl acetate and filtered through a plug of Celite. The filtrate was then washed with $NaHCO_3$ and brine, dried over Na_2SO_4 and concentrated by evaporation of the solvent. This crude solid was purified by column chromatography (1:1 PE/EA), yielding a white powder: 522.6 mg, 0.001 mol, 53 %.

LRMS: Calcd. for $C_{23}H_{24}NaO_{12}$: 515.12 found 515.1.

^{1}H NMR (400 MHz, Acetone-d_6) δ 7.73 (d, J = 9.5 Hz, 1H, Ar), 7.06 – 7.02 (m, 2H, Ar), 6.20 (d, J = 1.3 Hz, 1H, H-3'), 5.73 (d, J = 7.8 Hz, 1H, H-1), 5.48 (t, J = 9.6 Hz, 1H), 5.28 (dd, J = 9.6, 7.8 Hz, 1H, H-2), 5.24 (t, J = 9.7 Hz, 1H), 4.71 (d, J = 9.9 Hz, 1H), 3.69 (s, 3H, OMe), 2.46 (d, J = 1.3 Hz, 3H, Me), 2.03 (s, 3H, OAc), 2.01 (s, 3H, OAc), 2.00 (s, 3H, OAc).

^{13}C NMR (101 MHz, Acetone-d_6) δ 170.4, 170.1, 167.9, 160.7, 160.3, 156.0, 153.6, 127.5, 116.5, 114.4, 113.7, 104.6, 98.7, 73.0, 72.5, 71.7, 70.2, 53.1, 20.7 (two peaks), 20.6, 18.7.

Methyl (4-methylumbelliferyl 2,3-di-*O*-acetyl-4-deoxy-α-L-threo-hex-4-enopyranosid) uronate (5)



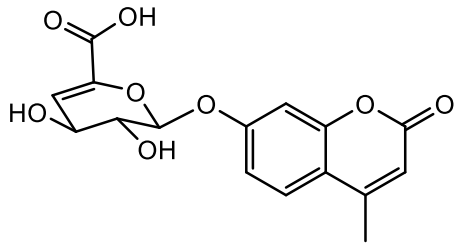
Glycoside **4** (468.2 mg, 0.95 mmol) was subjected to the DBU-mediated elimination method. The product was immediately purified by column chromatography (30:1 DCM/Acetone), yielding **5** as a colorless clear film: 355.0 mg, 0.82 mol, 86 %.

LRMS: Calcd. for $C_{21}H_{20}NaO_{10}$: 455.10 found 455.1

^{1}H NMR (400 MHz, Acetone-d_6) δ 7.72 (d, J = 9.4 Hz, 1H, Ar), 7.32 – 6.75 (m, 2H, Ar), 6.25 (dd, J = 4.1, 1.3 Hz, 1H, H-4), 6.18 (d, J = 1.3 Hz, 1H, H-3'), 6.16 (dd, J = 3.0, 0.9 Hz, 1H, H-1), 5.43 – 5.30 (m, 2H, H2&3), 3.78 (s, 3H, OMe), 2.44 (d, J = 1.3 Hz, 3H, Me), 2.13 (s, 3H, OAc), 2.13 (s, 3H, OAc), 2.09 (s, 3H, OAc).

^{13}C NMR (101 MHz, Acetone-d_6) δ 170.5, 169.9, 162.2, 160.6, 159.5, 155.8, 153.4, 143.1, 127.4, 116.4, 114.2, 113.7, 108.6, 105.0, 95.2, 68.5, 65.2, 52.9, 20.8, 20.7, 18.6.

4-Methylumbelliferyl 4-deoxy-α- L-threo-hex-4-enopyranosiduronic acid (6)



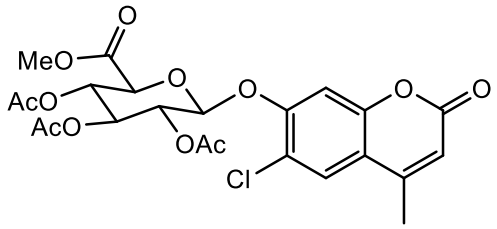
Protected glycoside **5** (335.0 mg, 0.77 mmol) was subjected to the general methods for Zemplén deprotection and hydrolysis of esters by aqueous lithium hydroxide. The product was a white powder (213.9 mg, 0.64 mmol, 83%).

HRMS: Calcd. for $C_{16}H_{15}O_8$: 335.0767; found: 335.0760

^{1}H NMR (400 MHz, Deuterium Oxide) δ 7.70 (d, *J* = 8.5 Hz, 1H, Ar), 7.19 – 7.14 (m, 2H, Ar), 6.22 (d, *J* = 1.4 Hz, 1H, H-3'), 5.99 (dd, *J* = 4.0, 0.8 Hz, 1H, H-4), 5.87 (br d, *J* = 4.9 Hz, 1H, H-1), 4.32 (br t, *J* = 4.1 Hz, 1H, H-3), 4.11 (br t, *J* = 4.3, 0.4 Hz, 1H, H-2), 2.42 (d, *J* = 1.2 Hz, 3H, Me).

^{13}C NMR (101 MHz, Deuterium Oxide) δ 168.8, 164.6, 159.1, 156.3, 153.8, 144.5, 126.7, 115.5, 114.3, 111.4, 107.5, 104.4, 97.5, 69.4, 66.0, 18.0.

Methyl (6-chloro-4-methylumbelliferyl 2,3,4-tri-*O*-acetyl-β-D-glucopyranoside) uronate (7)[42]

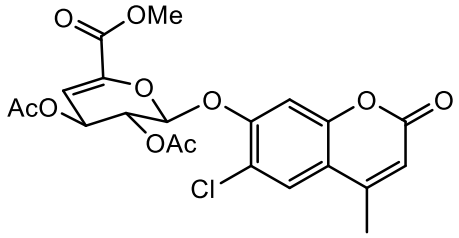


6-Chloro-4-methylumbelliferone (805 mg, 3.82 mmol), NaOH 5% (10 mL), DCM (20 mL), tetrabutylammonium hydrogen sulfate (1.43 g, 4 mmol) and **3** (1.52 g, 3.8 mmol) were mixed in a 50 mL round bottom flask and the reaction was stirred vigorously until TLC indicated the consumption of glucuronyl bromide. The reaction was then diluted with DCM. The organic phase was extracted with saturated solution of $NaHCO_3$ and brine, dried over Na_2SO_4 and purified by column chromatography (3:1 to 1:1 PE/EA). The product from the column was recrystallized from

ethyl acetate and hexanes to yield white crystals as pure product (571 mg, 1.08 mmol, 29 %). The NMR spectra was compared with the reference[42] to confirm identity of the compound.

LRMS: Calcd. for $C_{23}H_{23}ClNaO_{12}$: 549.08 found: 549.0

Methyl (6-chloro-4-methylumbelliferyl 2,3-di-*O*-acetyl-4-deoxy-α-L-threo-hex-4-enopyranoside)uronate (8)



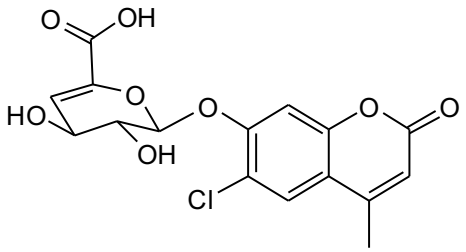
Glucuronide **7** (350.0 mg, 0.7 mmol), was subjected to the DBU-mediated elimination method. The product was immediately purified by column chromatography (30:1 DCM/Acetone), yielding **8** as a colorless clear film: 250.0 mg, 0.6 mmol, 81 %.

LRMS: Calcd. for $C_{21}H_{19}ClNaO_{10}$: 489.06 found: 489.1

^{1}H NMR (400 MHz, Acetone-d_6) δ 7.79 (s, 1H, Ar), 7.41 (s, 1H, Ar), 6.30 (dd, J = 4.6, 1.2 Hz, 1H, H-4), 6.25 (d, J = 1.4 Hz, 1H, H-3'), 6.23 (dd, J = 3.3, 0.9 Hz, 1H, H-1), 5.46 – 5.41 (m, 1H, H-2), 5.34 (ddd, J = 4.5, 2.7, 0.7 Hz, 1H, H-3), 3.79 (s, 3H, -OMe), 2.46 (d, J = 1.3 Hz, 3H, Me), 2.14 (s, 3H, Ac), 2.12 (s, 3H, Ac), 2.09 (s, 3H, Ac).

^{13}C NMR (101 MHz, Acetone-d_6) δ 170.4, 169.8, 162.2, 160.2, 154.7, 154.2, 152.8, 143.0, 127.2, 120.1, 117.0, 114.6, 108.7, 105.6, 95.6, 68.2, 64.8, 53.0, 20.9, 20.8, 18.6.

6-Chloro-4-methylumbelliferyl 4-deoxy-α-L-threo-hex-4-enopyranosiduronic acid (9)



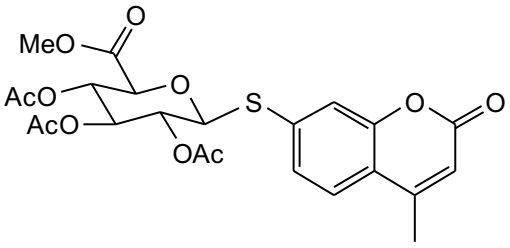
Protected glycoside **8** (125.0 mg, 0.3 mmol) was deprotected using the general method for Zemplén deprotection and the general method for saponification by aqueous lithium hydroxide. The product was obtained as a white powder (93.0 mg, 0.26 mmol, 85 %).

HRMS: Calcd. for $C_{16}H_{14}O_8Cl$: 369.0377; found: 369.0371

^{1}H NMR (400 MHz, DMSO-d_6) δ 7.85 (s, 1H, Ar), 7.43 (s, 1H, Ar), 6.32 (s, 1H, H-3'), 5.85 (d, J = 3.4 Hz, 1H, H-4), 5.82 (d, J = 6.0 Hz, 1H, H-1), 4.12 (dd, J = 5.2, 3.5 Hz, 1H, H-3), 3.76 (br t, J = 5.6 Hz, 1H, H-2), 2.40 (s, 3H, Me).

^{13}C NMR (101 MHz, DMSO) δ 163.4, 159.6, 154.3, 152.7, 152.6, 142.5, 126.1, 118.6, 115.16, 112.9, 110.9, 104.7, 98.7, 70.4, 67.0, 18.1.

Methyl (4-methylumbellifer-7-yl 1-thio -2,3,4-tri-*O*-acetyl-β-D-glucopyranoside) uronate (11)



Bromide **3** (500.0 mg, 1.26 mmol, 1.0 eq.) and K_2CO_3 (249.0 mg 1.8 mmol) were dissolved in acetone (12 mL) and flushed with nitrogen. 7-Mercapto-4-methylcoumarin (222.0 mg, 1.15 mmol)

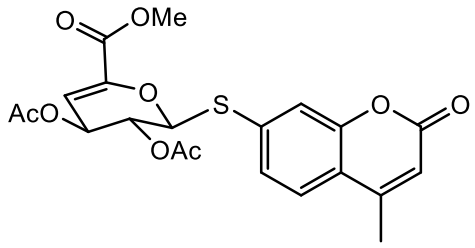
was added and the reaction mixture was stirred until the TLC indicated consumption of starting material. The product was purified by column chromatography (1:1 PE/EA), yielding **11** as a white powder: 543.0 mg, 1.1 mmol, 92 %.

LRMS: Calcd. for $C_{23}H_{24}KO_{11}S$: 547.07 found 547.1.

^{1}H NMR (400 MHz, Acetone-d_6) δ 7.73 (d, J = 8.8 Hz, 1H, Ar), 7.46 – 7.41 (m, 2H, Ar), 6.31 (d, J = 1.3 Hz, 1H, H-3'), 5.45 (t, J = 9.5 Hz, 1H, H-3), 5.44 (d, J = 10.1 Hz, 1H, H-1), 5.15 (t, J = 9.8 Hz, 1H, H-4), 5.05 (dd, J = 10.1, 9.3 Hz, 1H, H-2), 4.57 (d, J = 10.0 Hz, 1H, H-5), 3.72 (s, 3H, OMe), 2.48 (d, J = 1.3 Hz, 3H, Me), 2.05 (s, 3H, Ac), 1.99 (s, 3H, Ac), 1.96 (s, 3H, Ac).

^{13}C NMR (101 MHz, Acetone-d_6) δ 170.3, 170.0, 169.8, 167.9, 160.2, 154.6, 153.4, 138.4, 127.0, 126.5, 120.1, 118.7, 115.7, 84.9, 76.3, 73.6, 70.4, 70.2, 53.2, 20.7, 20.6 (two peaks), 18.58.

Methyl (4-methylumbellifer-7-yl 1-thio 2,3-di-*O*-acetyl-4-deoxy-α-L-threo-hex-4-enopyranoside) uronate (12)



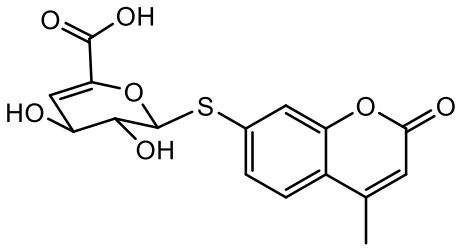
Glucuronide **11** (362.0 mg, 0.71 mmol) was subjected to the DBU-mediated elimination method. The product was immediately purified by column chromatography (20:1 DCM/Acetone), yielding **11** as a colorless film after evaporation: 191.8 mg, 0.43 mmol, 61 %.

LRMS: Calcd. for $C_{21}H_{20}NaO_9S$: 471.07 found: 471.1

^{1}H NMR (300 MHz, Acetone-d_6) δ 7.77 (d, *J* = 8.3 Hz, 1H, H-5'), 7.58 (d, *J* = 1.8 Hz, 1H, H-8'), 7.52 (dd, *J* = 8.3, 1.9 Hz, 1H, H-6'), 6.32 (d, *J* = 1.3 Hz, 1H, H-3'), 6.30 (dd, *J* = 4.7, 1.4 Hz, 1H, H-4'), 6.04 (dd, *J* = 2.7, 1.3 Hz, 1H, H-1), 5.32 (td, *J* = 2.4, 1.3 Hz, 1H, H-2), 5.22 (ddd, *J* = 4.7, 2.2, 1.3 Hz, 1H, H-3), 3.83 (s, 3H, OMe), 2.49 (d, *J* = 1.3 Hz, 3H, H-3'), 2.13 (s, 3H, OAc), 2.09 (s, 3H, OAc).

^{13}C NMR (75 MHz, Acetone-d_6) δ 170.0, 169.9, 162.6, 160.2, 154.7, 153.3, 144.1, 140.1, 126.9, 126.7, 120.2, 118.8, 115.8, 107.9, 83.6, 69.5, 64.8, 53.1, 20.8 (two peaks), 18.6.

4-Methylumbellifer-7-yl 1-thio-4-deoxy-α-L-threo-hex-4-enopyranosiduronic acid (13)



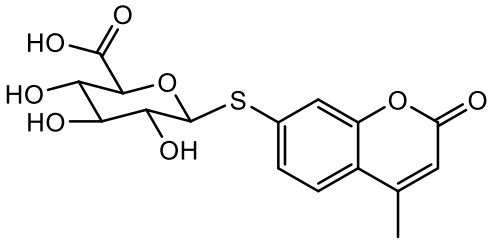
Protected glycoside **12** (167.0 mg, 0.37 mmol) was deprotected using the general methods for Zemplén deprotection and hydrolysis of methyl esters by aqueous lithium hydroxide. The product **13** was obtained as a white powder (162.0 mg, 0.44 mmol, 93%).

HRMS: Calcd. for $C_{16}H_{14}O_7SNa$: 373.0358; found: 373.0354

^{1}H NMR (400 MHz, Deuterium Oxide) δ 7.60 (d, *J* = 8.3 Hz, 1H, H-5'), 7.45 (dd, *J* = 8.1, 1.9 Hz, 1H, H-6'), 7.41 (d, *J* = 1.8 Hz, 1H, H-8'), 6.26 (d, *J* = 1.4 Hz, 1H, H-3'), 6.00 (dd, *J* = 4.3, 1.1 Hz, 1H, H-4), 5.75 (dd, *J* = 4.5, 1.1 Hz, 1H, H-1), 4.24 (ddd, *J* = 4.4, 3.5, 1.1 Hz, 1H, H-3), 4.07 (ddd, *J* = 4.6, 3.5, 1.1 Hz, 1H, H-2), 2.40 (d, *J* = 1.2 Hz, 3H, Me).

^{13}C NMR (101 MHz, Deuterium Oxide) δ 168.9, 163.9, 155.9, 152.4, 145.8, 139.6, 126.2, 125.6, 118.6, 117.2, 113.1, 106.6, 84.2, 69.7, 65.7, 17.9.

4-Methylumbellifer-7-yl 1-thio-β-D-glucopyranosiduronic acid (14)



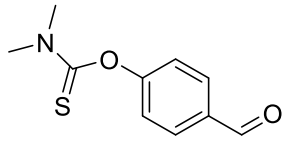
Protected glucuronide **11** (158.0 mg, 0.31 mmol) was deprotected by the general method for Zemplén deprotection and general method for hydrolysis of methyl ester by aqueous lithium hydroxide, yielding a white powder (135.5 mg, 0.37 mmol, 92 %).

HRMS: Calcd. for $C_{16}H_{17}O_8S$: 369.0644; found: 369.0641

^{1}H NMR (400 MHz, Deuterium Oxide) δ 7.67 (dd, *J* = 8.3, 1.5 Hz, 1H, Ar), 7.42 – 7.36 (m, 2H, Ar), 6.27 (br d, *J* = 1.5 Hz, 1H, H-3'), 4.78 (dd, *J* = 9.8, 1.5 Hz, 1H, H-1), 3.51 (d, *J* = 9.2 Hz, 1H, H-5), 3.37 – 3.21 (m, 2H, H3&4), 3.16 (br t, *J* = 8.7 Hz, 1H, H-2), 2.36 (s, 3H, Me).

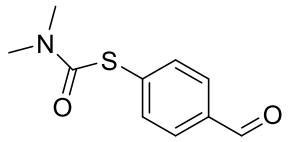
^{13}C NMR (101 MHz, Deuterium Oxide) δ 173.9, 162.0, 155.3, 153.8, 141.4, 126.8, 125.9, 118.8, 116.6, 114.3, 86.8, 79.8, 78.2, 72.8, 72.5, 19.0.

O-(4-Formylphenyl)-dimethylcarbamothioate (16)



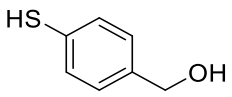
1,4-Diazabicyclo[2.2.2]octane (910 mg, 8.1 mmol, 1.5 eq) was added to 4-hydroxybenzaldehyde (660 mg, 5.4 mmol, 1 eq) in dry DMF (4 ml) under argon and the mixture was stirred at room temperature for 5 minutes, after which *N*,*N*-dimethylthiocarbamoyl chloride (1 g, 8.1 mmol, 1.5 eq) was added to the mixture in one portion and the reaction was stirred overnight. The mixture was poured on ice where a solid precipitated. The solid was collected via filtration and washed with water. The NMR spectra of the product was compared to the reference[72] and indicated pure product.

S-(4-Formylphenyl)-dimethylcarbamothioate (17)



Dried **16** was heated at 200 °C until TLC indicated the end of the reaction (4 hours). the crude product was purified by column chromatography (3:1 PE/EA) to yield 800 mg of **17** (3.8 mmol, 70% over two steps). NMR spectra was acquired and compared with the reference[72] to confirm the identity of the product.

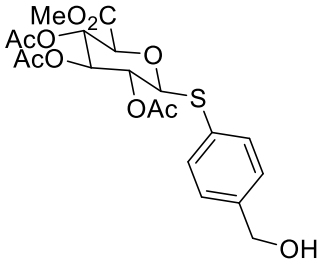
4-Mercaptobenzyl alcohol (18)



307 mg of **17** (1.47 mmol) was dissolved in dry THF. $LiAlH_4$ (115 mg, 3.03 mmol) was added to the mixture and the mixture was stirred at room temperature for 3 hours. The reaction was next refluxed for 1 hour after which the mixture was cooled down, the excess $LiAlH_4$ was quenched with ethyl acetate and then mixture was acidified with 3 M HCl and extracted with ethyl acetate three times. The pooled organic layers were dried over Na_2SO_4 and evaporated. This crude product was used without any purification in next reaction.

LRMS: Calcd. for $C_7H_7OS^-$: 139.02 found 139.1.

Methyl (4-(hydroxymethyl)phenyl 1-thio- 2,3,4-tri-*O*-acetyl-β-D-glucopyranoside)uronate (19)



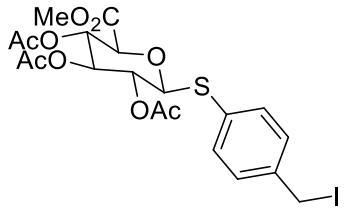
Crude **18** (180 mg) and K_2CO_3(100 mg, 0.7 mmol) was mixed with **3** (500 mg, 1.26 mmol) in acetone and the mixture was stirred overnight. The solvent was then evaporated and the mixture was suspended in ethyl acetate and extracted with water. Aqueous layer was extracted with ethyl acetate again, the pooled organic layer was dried over Na_2SO_4 and concentrated by evaporation. The crude mixture was subject to column chromatography (3:1 PE/EA) to yield the product (140 mg, 0.31 mmol, 25% with respect to **3**) along with the dimer of the thiophenol as the major by-product.

LRMS: Calcd. for $C_{20}H_{24}NaO_{10}S$: 479.10 found 479.1.

^{1}H NMR (400 MHz, Acetone-d_6) δ 7.50 (d, J = 8.3 Hz, 2H, Ar), 7.35 (d, J = 8.4 Hz, 2H, Ar), 5.38 (t, J = 9.4 Hz, 1H, H-3), 5.14 (d, J = 10.1 Hz, 1H, H-1), 5.07 (t, J = 9.8 Hz, 1H, H-4), 4.92 (dd, J = 10.1, 9.2 Hz, 1H, H-2), 4.64 (d, J = 5.6 Hz, 2H, CH_2), 4.41 (d, J = 10.0 Hz, 1H, H-5), 4.29 (t, J = 5.8 Hz, 1H, OH), 3.71 (s, 3H, OMe), 2.04 (s, 5H, OAc), 1.97 (s, 3H, OAc), 1.94 (s, 3H, OAc).

^{13}C NMR (101 MHz, Acetone-d_6) δ 170.2, 170.0, 169.7, 168.0, 144.0, 133.7, 130.6, 128.0, 86.2, 76.3, 73.8, 70.6, 70.3, 64.2, 53.0, 20.8, 20.6, 20.5.

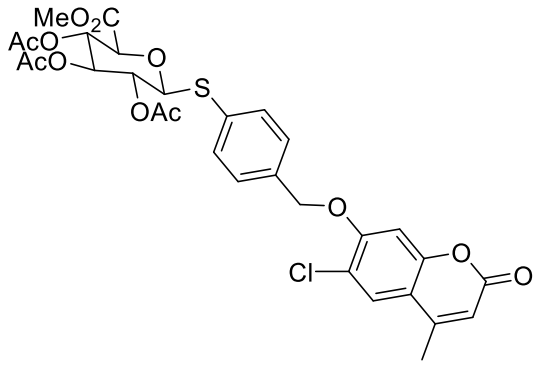
Methyl (4-(iodomethyl)phenyl 1-thio- 2,3,4-tri-*O*-acetyl-β-D-glucopyranoside)uronate (20)



Alcohol **19** (40 mg, 0.088 mmol) was dissolved in dry DCM and added to a stirred mixture of iodine (44 mg, 0.175 mmol, 2 eq), triphenylphosphine (46 mg, 0.176 mmol, 2 eq) and imidazole (30 mg, 0.44 mmol, 5 eq) in dry DCM. The solvent was evaporated when TLC indicated the end of reaction (30 min) and the crude product was subjected to a short silica column (6:1 PE/EA) to yield 40 mg of mostly pure product. This compound was not stable and was used directly in the next step.

LRMS: Calcd. for $C_{20}H_{23}INaO_9S$: 589.0 found 589.1.

Methyl (4-(6-chloro-4-methylumbelliferyl-methyl) phenyl 1-thio-2,3,4-tri-*O*-acetyl-β-D-glucopyranoside)uronate (21)



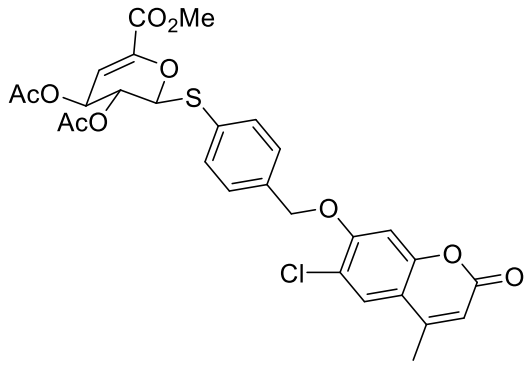
Iodide **20** (27 mg, 0.047 mmol) and NaClMU (6-chloro-4-methylcoumarin sodium salt) (11 mg, 0.047 mmol, 1.1 eq) were dissolved in dry DMF (3 mL) under Argon at room temperature and the mixture was stirred until TLC indicated the end of the reaction. The solvent was then evaporated and the residue was dissolved in ethyl acetate and extracted with a saturated aqueous solution of $NaHCO_3$. The crude was then purified by column chromatography (1:1 PE/EA) to give product **21** as a colorless film: 29 mg, 45 mmol, 95%.

LRMS: Calcd. for $C_{30}H_{29}ClNaO_{12}S$: 671.1 found 671.2.

^{1}H NMR (400 MHz, Acetone-d_6) δ 7.79 (s, 1H, Ar), 7.62 – 7.58 (m, 2H, Ar), 7.57 – 7.54 (m, 2H, Ar), 7.16 (s, 1H, Ar), 6.20 (d, J = 1.3 Hz, 1H, H-3'), 5.42 (t, J = 9.4 Hz, 1H, H-3), 5.39 (s, 2H, CH2), 5.24 (d, J = 10.1 Hz, 1H, H-1), 5.11 (t, J = 9.8 Hz, 1H, H-4), 4.98 (dd, J = 10.1, 9.3 Hz, 1H, H-2), 4.46 (d, J = 10.0 Hz, 1H, H-5), 3.71 (s, 3H, OMe), 2.46 (d, J = 1.3 Hz, 3H, Me), 2.04 (s, 3H, Ac), 1.97 (s, 3H, Ac), 1.95 (s, 3H, Ac).

^{13}C NMR (101 MHz, Acetone-d_6) δ 170.3, 170.0, 169.8, 168.0, 160.5, 157.5, 154.8, 153.1, 137.2, 133.2, 129.3, 129.2, 126.9, 119.5, 115.2, 113.7, 103.0, 86.0, 76.3, 73.7, 71.4, 70.7, 70.3, 53.1, 20.8, 20.6 (two peaks), 18.6.

Methyl (4-(6-chloro-4-methylumbelliferyl-methyl) phenyl 1-thio-2,3-di-*O*-acetyl-4-deoxy-α-L-threo-hex-4-enopyranoside)uronate (22)



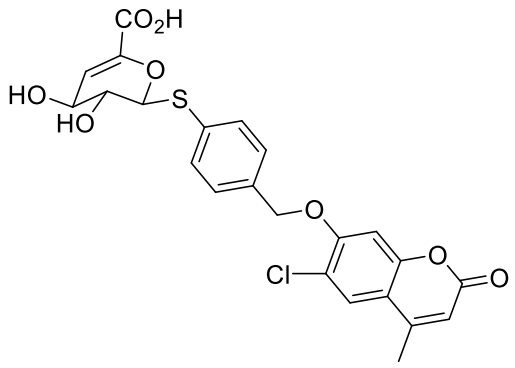
Glucuronide **21** (25 mg, 0.038 mmol) was subjected to the general method for elimination of OAc using DBU to yield 19 mg (0.032 mmol, 84%) of product after purification by flash column chromatography (50:1 DCM/Acetone).

LRMS: Calcd. for $C_{28}H_{25}ClNaO_{10}S$: 611.08 found 611.2.

^{1}H NMR (400 MHz, Acetone-d_6) δ 7.79 (s, 1H, Ar), 7.67 (d, *J* = 8.3 Hz, 2H, Ar), 7.59 (d, *J* = 8.5 Hz, 2H, Ar), 7.16 (s, 1H, Ar), 6.27 (dd, *J* = 4.7, 1.4 Hz, 1H, H-4), 6.20 (d, *J* = 1.4 Hz, 1H, H-3'), 5.82 (dd, *J* = 2.8, 1.4 Hz, 1H, H-1), 5.40 (s, 2H, CH2), 5.30 (ddd, *J* = 2.7, 2.3, 1.4 Hz, 1H, H-2), 5.19 (ddd, *J* = 4.7, 2.3, 1.4 Hz, 1H, H-3), 3.82 (s, 3H, OMe), 2.46 (d, *J* = 1.3 Hz, 3H, Me), 2.11 (s, 3H, Ac), 2.07 (s, 3H, Ac).

^{13}C NMR (101 MHz, Acetone-d_6) δ 170.1, 169.9, 162.7, 160.5, 157.4, 154.8, 153.1, 144.2, 137.3, 135.0, 133.2, 129.3, 126.9, 119.5, 115.2, 113.6, 107.7, 103.0, 84.7, 71.4, 69.6, 65.0, 53.0, 20.9, 20.8, 18.7.

4-(6-Chloro-4-methylumbelliferyl-methyl) phenyl 1-thio 4-deoxy-α-L-threo-hex-4-enopyranosiduronic acid (23)



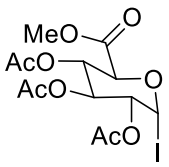
Protected glycoside **22** (19 mg, 0.032 mmol) was subjected to the general method for Zemplén deprotection (section 4.1.1.2) and deprotection of methyl ester (section 4.1.1.3). The crude product was purified by precipitation of solid via addition of diethyl ether to solution of **23** in MeOH. (15 mg, 0.031mmol, 95%).

HRMS: Calcd. for $C_{23}H_{18}O_8SCl$: 489.0411; found: 489.0403

^{1}H NMR (400 MHz, Methanol-d_4) δ 7.76 (s, 1H, Ar), 7.66 (d, J = 8.2 Hz, 2H, Ar), 7.46 (d, J = 8.3 Hz, 2H, Ar), 7.10 (s, 1H, Ar), 6.19 (d, J = 1.3 Hz, 1H, H-3'), 6.16 (dd, J = 4.1, 1.0 Hz, 1H, H-4), 5.47 (dd, J = 4.8, 1.1 Hz, 1H, H-1), 4.10 (ddd, J = 4.0, 3.9, 1.1 Hz, 1H, H-3), 3.88 (ddd, J = 4.7, 3.7, 1.0 Hz, 1H, H-2), 2.41 (d, J = 1.3 Hz, 3H, Me).

^{13}C NMR (101 MHz, Methanol-d_4) δ 165.9, 162.8, 158.1, 154.7, 154.6, 143.7, 136.8, 136.1, 133.4, 129.1, 126.9, 120.7, 115.3, 113.3, 112.0, 102.9, 88.0, 71.8 (two peaks), 67.4, 18.6.

Methyl (2,3,4-tri-*O*-acetyl-α-D-glucopyranosyl iodide)uronate (24)



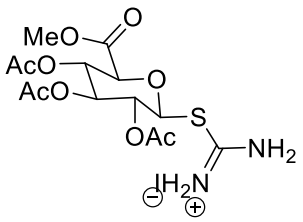
Globally protected glucuronic acid **2** (1.130 g, 3.0 mmol) and iodine (1.068g, 4.2 mmol, 1.4 eq) and poly(methylhydrosiloxane) (255 μL, 1.4 eq) were dissolved in 1,2-dichloroethane (15 mL) and refluxed for 30 minutes. The mixture was then cooled to room temperature, diluted with DCM and extracted with an aqueous solution of $NaHCO_3$ and $Na_2S_2O_3$. The aqueous portion was then extracted with DCM two times. The pooled organic solvents were then extracted with brine and dried over Na_2SO_4. The mixture was evaporated to give 1.45 g of crude product. This crude product was used in subsequent reactions without any purification. A separate sample of the crude product was purified by flash column chromatography (2:1 PE/EA) to give pure product **24**.

LRMS: Calcd. for $C_{13}H_{17}INaO_9$: 466.98 found:

^{1}H NMR (400 MHz, Acetone-d_6) δ 7.17 (d, J = 4.2 Hz, 1H, H-1), 5.48 (t, J = 9.6 Hz, 1H, H-3), 5.30 (t, J = 9.9 Hz, 1H, H-4), 4.39 (dd, J = 9.8, 4.2 Hz, 1H, H-2), 4.30 (d, J = 10.2, 1H, H-5), 3.72 (s, 3H, OMe), 2.06 (s, 3H, OAc), 2.02 (s, 3H, OAc), 2.01 (s, 3H, OAc).

^{13}C NMR (101 MHz, Acetone-d_6) δ 170.2, 170.0 (two peaks), 167.2, 76.0, 74.2, 71.8, 70.6, 68.9, 53.4, 20.8, 20.6, 20.5.

Methyl (2,3,4-tri-O-acetyl-β-D-glucopyranosyl 1-isothiouronium iodide)uronate (25)

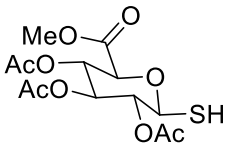


1.45 of crude **24** and thiourea (342 mg, 4.5 mmol, 1.5 eq) were dissolved in dry acetonitrile. The mixture was stirred at 60 º C for 2 hours, then cooled to room temperature. Pure product was filtered from the mixture as a white precipitate (1.40 g, 2.7 mmol, 90% over two steps).

^{1}H NMR (400 MHz, Deuterium Oxide) δ 5.59 (d, *J* = 9.9 Hz, 1H, H-1), 5.56 (t, *J* = 9.1 Hz, 1H-H-3), 5.38 (dd, *J* = 9.9, 9.0 Hz, 1H, H-2), 5.33 (t, *J* = 9.6 Hz, 1H, H-4), 4.63 (d, *J* = 9.8 Hz, 1H, H-5), 3.80 (s, 3H, OMe), 2.17 (s, 3H, OAc), 2.14 (s, 3H, OAc), 2.12 (s, 3H, OAc).

^{13}C NMR (101 MHz, Deuterium Oxide) δ 172.9, 172.8, 172.5, 168.9, 167.4, 81.2, 75.1, 72.2, 69.0, 68.9, 53.9, 20.20 (two peaks), 20.1.

Methyl (2,3,4-tri-*O*-acetyl-1-β-thio-D-glucopyranoside) uronate[73] (26)

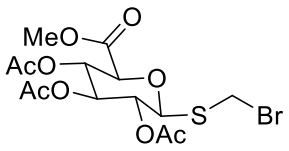


Thioglycoside **25** (175 mg, 0.34 mmol) was dissolved in DCM (4 mL). $Na_2S_2O_5$ (300 mg, 1.58 mmol) was dissolved in water and added to the solution of starting material in DCM. The biphasic mixture was refluxed at 60 ºC for one hour. The mixture was then diluted with DCM and extracted with water. The aqueous layer was back-extracted with DCM and the pooled organic solvents were

dried over Na_2SO_4 and evaporated. This crude compound was used in the next step without any purification.

LRMS: Calcd. for $C_{13}H_{18}NaO_9S$: 373.06 found 373.0.

Methyl (bromomethyl 1-thio 2,3,4-tri-*O*-acetyl-β-D-glucopyranoside)uronate (27)



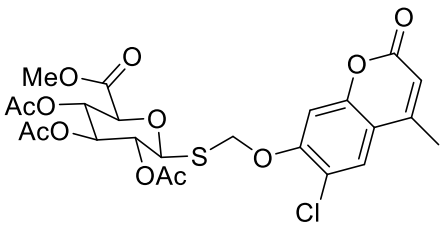
Crude **26** (120 mg) and K_2CO_3 (100 mg, 0.72 mmol) were added to dried CH_2Br_2 (14 ml, 0.2 mol) under Ar. The mixture was stirred at room temperature overnight and then filtered, evaporated and purified by flash column chromatography (2:1 PE/EA) to give 90 mg of pure product. (0.20 mmol, 60% over two steps, >90% β anomer).

LRMS: Calcd. for $C_{14}H_{19}BrNaO_9S$: 464.98 found 464.9.

^{1}H NMR (400 MHz, Acetone-d_6) δ 5.46 (t, J = 9.3 Hz, 1H, H-3), 5.16 (t, J = 9.8 Hz, 1H, H-4), 5.16 (d, J = 10.2 Hz, 1H, H-1), 5.06 (dd, J = 10.2, 9.1 Hz, 1H, H-2), 4.90 (d, J = 2.4 Hz, 2H, CH_2), 4.47 (d, J = 10.0 Hz, 1H, H-5), 3.70 (s, 3H, OMe), 2.01 (s, 3H, OAc), 1.99 (s, 3H, OAc), 1.97 (s, 3H, OAc).

^{13}C NMR (101 MHz, Acetone-d_6) δ 206.2, 170.1, 169.9, 169.8, 167.8, 83.0, 76.4, 73.3, 70.4, 70.2, 52.9, 33.7, 20.6, 20.5, 20.4.

Methyl ((6-chloro-4-methylumbelliferyl) 1-thio 2,3,4-tri-*O*-acetyl-β-D-glucopyranoside)uronate (28)



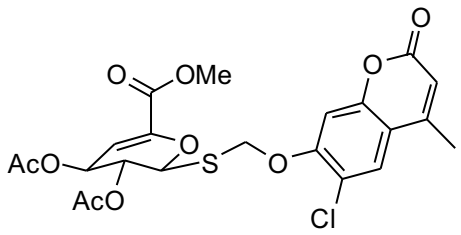
NaClMU (44 mg, 0.2 mmol, 2 eq) and **27** (72 mg, 0.16 mmol) were dissolved in dry DMF under Ar and the mixture was stirred overnight at room temperature. The solvent was then evaporated and the mixture was purified by flash column chromatography (30:1 DCM/Acetone) to give 85 mg of product (0.15 mmol, 91%, >90% β anomer).

LRMS: Calcd. for $C_{24}H_{25}ClNaO_{12}S$: 595.07 found 595.0.

^{1}H NMR (400 MHz, Acetone-d_6) δ 7.81 (s, 1H, Ar), 7.22 (s, 1H, Ar), 6.23 (d, J = 1.3 Hz, 1H, H-3'), 5.75 (d, J = 12.1 Hz, 1H, CH_2), 5.66 (d, J = 12.1 Hz, 1H, CH_2), 5.38 (t, J = 9.4 Hz, 1H, H-3), 5.22 (d, J = 10.2 Hz, 1H, H-1), 5.13 (t, J = 9.8 Hz, 1H, H-4), 5.02 (dd, J = 10.2, 9.3 Hz, 1H, H-2), 4.40 (d, J = 10.0 Hz, 1H, H-5), 3.69 (s, 3H, OMe), 2.47 (d, J = 1.3 Hz, 3H, Me), 1.97 (s, 3H, OAc), 1.95 (s, 3H, OAc), 1.86 (s, 3H, OAc).

^{13}C NMR (101 MHz, Acetone-d_6) δ 170.3, 170.0, 169.8, 167.9, 160.4, 155.6, 154.4, 153.0, 127.0 (two peaks), 120.5, 116.0, 114.1, 104.7, 83.2, 76.6, 73.6, 71.1, 70.7, 70.4, 53.0, 20.6 (two peaks), 18.6.

Methyl ((6-chloro-4-methylumbelliferyl) 1-thio-2,3-di-*O*-acetyl-4-deoxy-α-L-threo-hex-4-enopyranosid)uronate (29)



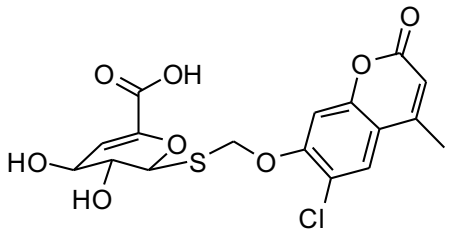
Thioglucuronide **28** (36 mg, 0.063 mmol) was subjected to the general method for elimination of OAc using DBU to yield 15 mg (0.029 mmol, 45%, >90% β anomer) of product **29** after purification by flash column chromatography (1:1 PE/EA).

LRMS: Calcd. for $C_{22}H_{21}ClNaO_{10}S$: 535.04 found 534.9.

^{1}H NMR (400 MHz, Acetone-d_6) δ 7.82 (s, 1H, Ar), 7.25 (s, 1H, Ar), 6.24 (d, J = 1.3 Hz, 1H, H-3'), 6.22 (dd, J = 4.4, 1.8 Hz, 1H, H-4), 5.94 (t, J = 1.9 Hz, 1H, H-1), 5.73 (d, J = 12.0 Hz, 1H, CH_2), 5.64 (d, J = 12.0 Hz, 1H, CH_2), 5.17-5.06 (m, J = 4.9, 1.9 Hz, 2H, H-2,3), 3.80 (s, 3H, OMe), 2.48 (d, J = 1.3 Hz, 3H, Me), 2.06 (s, 3H, OAc), 2.04 (s, 3H, OAc).

^{13}C NMR (101 MHz, Acetone-d_6) δ 170.0, 169.8, 162.6, 160.4, 155.6, 154.4, 153.0, 144.2, 127.1, 120.5, 116.1, 114.2, 107.6, 104.7, 79.4, 71.1, 69.7, 64.6, 53.0, 20.8 (two peaks), 18.7.

(6-Chloro-4-methylumbelliferyl) 1-thio-4-deoxy-α-L-threo-hex-4-enopyranosiduronic acid (30)



Protected thioglycoside **29** (15mg, 0.029 mmol) was subjected to the general method for Zemplén deprotection (section 4.1.1.2) and deprotection of methyl esters(section 4.1.1.3) to yield **30** as a white precipitate (11.3 mg, 0.027 mmol, 94 %, >90 % β anomer).

HRMS: Calcd. for $C_{17}H_{14}O_8SCl$: 413.0098; found: 413.0103.

^{1}H NMR (400 MHz, Deuterium Oxide) δ 7.81 (s, 1H, Ar), 7.19 (s, 1H, Ar), 6.33 (br d, *J* = 1.4 Hz, 1H, H-3'), 6.30 (dd, *J* = 4.5, 1.2 Hz, 1H, H-4), 5.80 (d, *J* = 3.9 Hz, 1H, H-1), 5.74 (d, *J* = 12.1 Hz, 1H, CH_2), 5.56 (d, *J* = 12.0 Hz, 1H, CH_2), 4.40 – 4.11 (m, 1H, H-3), 4.15 – 3.85 (m, *J* = 4.6, 2.1 Hz, 1H, H-2), 2.48 – 2.42 (br d, 3H, Me).

^{13}C NMR (101 MHz, Deuterium Oxide) δ 165.8, 163.6, 155.8, 155.7, 153.6, 142.6, 127.5, 121.2, 116.0, 113.5, 112.6, 104.5, 82.6, 71.7, 71.0, 65.6, 19.4.

4.2 Biochemistry

4.2.1 Expression and purification of UGC

Previously reported procedure for expression and purification of UGC[9] was followed with minor changes. Briefly, the *E. Coli* cells strain BL21(DE3) carrying the plasmid pET28a::UGL were grown overnight in LB media (5 ml LB, 50 μg/mL kanamycin). The overnight culture was sub-cultured (500 μL into 500 mL) into TYP media with 50 μg/mL kanamycin and shaken at 37° C for 5 hours (200 rpm) at which point the culture was induced with IPTG (final concentration 0.1 mM) and shaken at 37 °C overnight (200 rpm). Cells were pelleted by centrifugation at 4 °C for 30 min (5000 rpm) then homogenized three times by EmusliFlex-C3 homogenizer (Avestin Inc., Ontario) in 10 mL of lysis buffer (20 mM Tris.HCl pH 8, 20 mM imidazole, 25 mM NaCl, 1mM DTT, benzonase 0.5 μL, one tablet of EDTA-free protease inhibitor cocktail (Roche, Switzerland)). The extract was then clarified by centrifugation at 15000 rpm for 30 min. The supernatant was then loaded on a 1 mL HistrapFF column, washed with 20 mL of buffer A (20 mM Tris.HCl pH 8, 20 mM imidazole, 25 mM NaCl, 1 mM DTT), 15 mL 2.5% buffer B (20 mM Tris.HCl pH 8, 400 mM imidazole 25 mM NaCl, 1 mM DTT) in buffer A, 15 mL 40% buffer B in buffer A and 5 mL buffer B. The fractions were analyzed by SDS-PAGE. The fractions with the pure protein were pooled, concentrated and exchanged to the storage buffer (20 mM Tris pH 8, 1mM DTT) using a 30 kDa cut-off centrifugal filter and then stored at 4 °C. The final concentration of the protein was determined based on the UV-vis absorbance at 280 nm to be 190 μM. (ε_{280} of 106230 1/M.cm calculated based on the sequence (http://web.expasy.org/protparam/)).

4.2.2 Michaelis-Menten Kinetics

The kinetic parameters of hydrolysis of GA-MU with BGE was determined by fluorescence stopped assay. For the rest of substrates, the kinetic parameters were determined by measuring the increase in the UV-vis absorbance of the hydrolyzed aglycones at 37 °C. For UGC and BGE the assay buffer was 50 mM sodium phosphate buffer pH 6.6 with 1 mg/mL BSA. For BGB the assay buffer was 50 mM sodium acetate buffer pH 5.0 with 1 mg/mL BSA. The latter pH was chosen in spite of the fact that the optimum pH for BGB is 4.5[74]. The reason is that at pH 4.5 the extinction coefficient for the aglycones is so low that monitoring the rates of reactions will be troublesome. These extinction coefficients are larger at pH 5 and also the activity of BGB at pH 5 is still nearly 90% of the optimal activity.

The substrates that are used in these experiments are shown in Figure 27. For ΔGASMU and GASMU, the assay buffer also contained 10 mM TCEP to prevent dimerization of SMU.

The extinction coefficients for the aglycones were calculated by reading the absorbance of known concentrations of the free aglycones in the assay buffers. The change in absorbance was monitored by a Cary 4000 UV-Vis Spectrometer, Agilent Technologies at 370 nm for MU and SMU and 368 nm for ClMU.

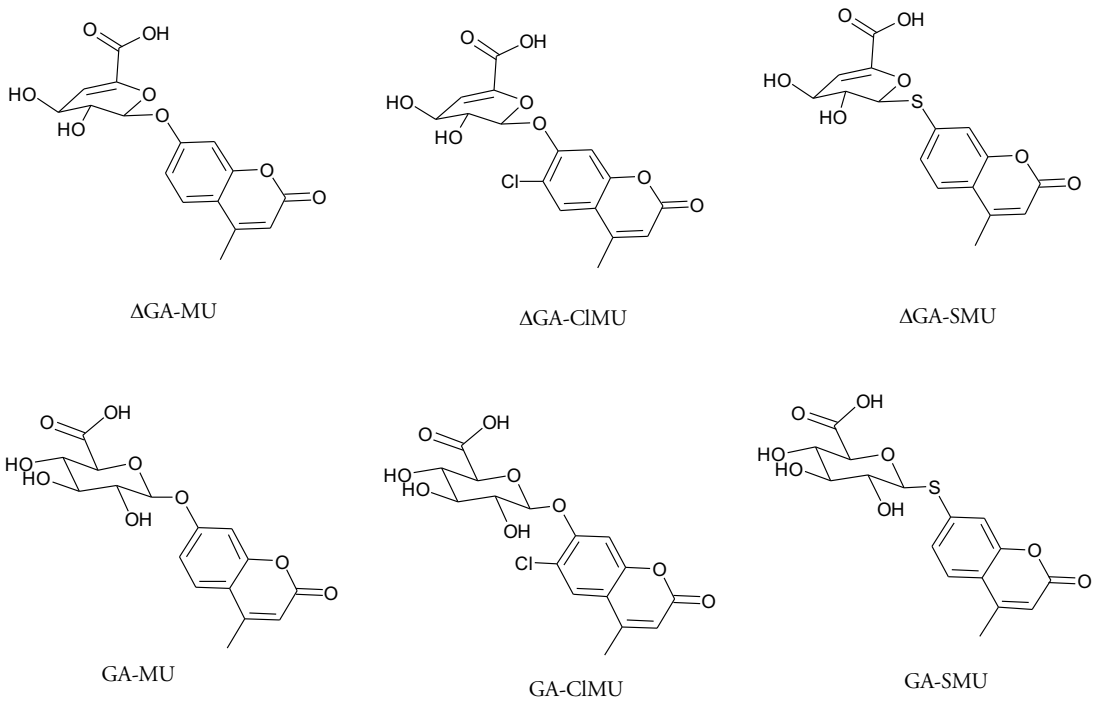


Figure 27: The structure of the substrate used in kinetic analysis

UGC was expressed and purified as described in section 4.2.1. BGE was purchased from Sigma Aldrich (G7396) and dissolved in the storage buffer (20 mM HEPES, pH 7.4, 100 mM NaCl, and 0.5 mM TCEP). BGB was purchased from Sigma Aldrich (G0251) and dissolved in 0.2% (w/v) NaCl according to the manufacturer's protocol. The final concentration of UGC and BGB was determined based on absorbance at 280 nm and the concentration of BGE was calculated using the manufacturer's specification of the protein content in the solid enzyme powder.

The initial rate of hydrolysis was monitored for different concentrations of each substrate ranging from $K_M/7$ to $K_M \times 7$. However, due to the poor solubility of some substrates in higher concentrations and also to satisfy the Michaelis-Menten conditions ([S] >> [E]) in lower concentrations this was not always possible. The concentration of the enzyme was kept constant for each substrate.

The obtained data were plotted according to the Michaelis-Menten equation (E.1) with GraFit 7.0 and the plots are represented in the attachment.

$$V_0 = (k_{cat}[E]_0[S])/ (K_M + [S]) \qquad E.1$$

Since the activity of BGE towards ΔGA-MU is too low be measurable through standard UV-Vis method, the kinetic parameters of this substrate-enzyme pair was determined by a fluorescence stopped assay. Aliquots of the hydrolysis reaction of ΔGA-MU with BGB was transferred to a glycine buffer (100 mM, pH = 10) at different time points and the resulting fluorescence signals were measured immediately with a Synergy H1 plate reader (BioTek). The initial rate of the hydrolysis reaction was measured through this method using different starting concentrations of ΔGA-MU and the Michaelis-Menten curve was constructed as described before.

For the rest of the substrates for which the enzyme activity was too low to be measured appropriately at low substrates concentrations and for which fluorescence stopped assay was not possible, the values of V_{max} was estimated by performing two to three assays with the maximum possible concentrations of substrate and with the maximum possible concentration of enzyme. The rates were measured as described above, using Cary 4000 UV-Vis Spectrometer. Since the observed rates for these substrates concentrations were similar, the measured conditions were assumed to correlate to the plateau region of the Michaelis-Menten curve and k_{cat} was calculated based on the assumption that the average of the rates is the V_{max}.

4.2.3 Kinetics of dimerization of SMU

SMU from a freshly made stock solution in DMSO with a known concentration was added to phosphate buffer 100 mM pH = 7 at different concentrations in a 96-well plate. The absorbance of these solutions was monitored using a Synergy H1 plate reader (BioTek) in 25 °C at 370 nm. The rate of the decrease in the absorbance was calculated for each starting concentration based on the

extinction coefficient of SMU at 370 nm. These rates were then plotted versus the initial concentration according to the first order rate equation (E.2) to give the observed first order rate.

$$\text{Rate} = -\frac{\Delta[SMU]}{\Delta t} = k_{obs}[SMU]_{tot.} \qquad \text{E.2}$$

The resulting graph is shown in Figure 12. It should be said that since the rate of this reaction is highly dependent on the oxygen content of the buffer, the assay was done under the same conditions and with the same buffer that is used in screening procedure.

4.2.4 Screening the metagenomic library

The screening procedure is based on that reported[36] with minor changes.

First the needed number of plates were filled with (50 µL/well) LB media that contains 100 µg/mL arabinose and 12.5 µg/mL chloramphenicol. These plates were next inoculated with bacteria from the metagenomic libraries described in section 2.5. The plates were then incubated overnight at 37°C. Next, 50 µL of screening buffer (100 mM phosphate buffer pH = 7 and 2% Triton X-100 containing 100 µM of one of the substrates ΔGAClMU, ΔGASP, ΔGASC) was added to each well and the plates were sealed and again incubated at 37 °C in a closed container. The fluorescence of each well was measured after 2 and 5 hours and overnight incubation using a Synergy H1 plate reader (BioTek) with an excitation wavelength of 365 nm and an emission wavelength of 450 nm. The Z-score was then calculated for each well on a per-plate basis based on E.3:

$$\text{Z-score} = \frac{\text{Fluorescence} - \text{Average of fluorescence for all of the data}}{\text{Standard deviation of all the data}} \qquad \text{E.3}$$

It was noted that the results at 5 hours are significantly different from overnight results for the screens with ΔGA-ClMU. When this experiment was done in triplicate for one of the plates, results at

5 hours were reproducible, while overnight results were not. Therefore, the fluorescence values after 5 hours of incubation were used for calculating the Z-score and selecting the hits.

This difference is most probably a result of the background hydrolysis caused by BGE. BGE is present in all of the wells from the host *E. Coli* cells, although apparently in low concentrations. After 5 hours, the level of background hydrolysis is still low enough for reliable readings to be made. However, after overnight incubation the amount of background hydrolysis will be significant and hence obscuring any activity that is due to other enzymes.

The DNA for the hits with highest Z scores were prepared and sent for end-sequencing and compared with the full sequences of the hits to confirm the identity of the hits. The found hits had been sequenced previously, since they were also found in separate screens due to the fact that they contain multiple GH genes.

References

1. Chothia C, Lesk AM. The relation between the divergence of sequence and structure in proteins. *EMBO J*. 1986;5(4):823-826. doi:060 fehlt.
2. Henrissat B, Vegetales M, Grenoble F. A classification of glycosyl hydrolases based sequence similarities amino acid. *Biochem J*. 1991;280((Pt 2)):309-316. doi:10.1007/s007920050009.
3. Lairson LL, Withers SG. Mechanistic analogies amongst carbohydrate modifying enzymes. *Chem Commun (Camb)*. 2004:2243-2248. doi:10.1039/b406490a.
4. Koshland DE. Stereochemistry and The Mechanism of Enzymatic Reactions. *Biol Rev*. 1953;28(February):416-436. doi:10.1111/j.1469-185X.1953.tb01386.x.
5. Wallace BD, Wang H, Lane KT, et al. Alleviating Cancer Drug Toxicity by Inhibiting a Bacterial Enzyme. *Science (80-)*. 2010;330(6005):831-835. doi:10.1126/science.1191175.
6. Ernst S, Langer R, Cooney CL, Sasisekharan R. Enzymatic Degradation of GlycosaminogIycans. *Crit Rev Biochem Mol Biol*. 1995;30(5):387-444. doi:10.3109/10409239509083490.
7. Jongkees SAK, Yoo H, Withers SG. Mechanistic investigations of unsaturated glucuronyl hydrolase from clostridium perfringens. *J Biol Chem*. 2014;289(16):11385-11395. doi:10.1074/jbc.M113.545293.
8. Jongkees SAK, Yoo H, Withers SG. Mechanistic insights from substrate preference in unsaturated glucuronyl hydrolase. *ChemBioChem*. 2014;15(1):124-134. doi:10.1002/cbic.201300547.
9. Jongkees SAK, Withers SG. Glycoside Cleavage by a New Mechanism in Unsaturated

Glucuronyl Hydrolases. *J Am Chem Soc*. 2011;133(48):19334-19337. doi:10.1021/ja209067v.

10. Driguez H. Thiooligosaccharides as tools for structural biology. *Chembiochem*. 2001;2(5):311-318. doi:10.1002/1439-7633(20010504)2:5<311::AID-CBIC311>3.0.CO;2-L.

11. De Leon CA, Levine PM, Craven TW, Pratt MR. The Sulfur-Linked Analogue of O-GlcNAc (S-GlcNAc) Is an Enzymatically Stable and Reasonable Structural Surrogate for O-GlcNAc at the Peptide and Protein Levels. *Biochemistry*. 2017;56(27):3507-3517. doi:10.1021/acs.biochem.7b00268.

12. Richard JP, Toteva MM, Crugeiras J. Structure-reactivity relationships and intrinsic reaction barriers for nucleophile additions to a quinone methide: A strongly resonance-stabilized carbocation. *J Am Chem Soc*. 2000;122(8):1664-1674. doi:10.1021/ja9937526.

13. Driguez H. Thiooligosaccharides in glycobiology. *Top Curr Chem*. 1997;187:85-116. doi:10.1007/BFb0119254.

14. Czjzek M, Cicek M, Zamoni V, et al. Crystal structure of a monocotyledon (maize ZMGlu1) β-glucosidase and a model of its complex with p-nitrophenyl β-d-thioglucoside. *Biochem J*. 2001;354(1):37-46. doi:10.1042/bj3540037.

15. Parsiegla G, Reverbel-Leroy C, Tardif C, Belaich JP, Driguez H, Haser R. Crystal structures of the cellulase Cel48F in complex with inhibitors and substrates give insights into its processive action. *Biochemistry*. 2000;39(37):11238-11246. doi:10.1021/bi001139p.

16. Shen H, Byers LD. Thioglycoside hydrolysis catalyzed by Beta-glucosidase. *Biochem Biophys Res Commun*. 2007;362(3):717-720. doi:10.1016/j.bbrc.2007.08.043.

17. Narine AA, Watson JN, Bennet AJ. Mechanistic requirements for the efficient enzyme-catalyzed hydrolysis of thiosialosides. *Biochemistry*. 2006;45(30):9319-9326. doi:10.1021/bi0607507.

18. Capon B. Mechanism in carbohydrate chemistry. *Chem Rev*. 1969;69(4):407-498. doi:10.1021/cr60260a001.

19. Capon B, Overend WG. Overend: 989. Reactions at Position 1. 1962;(5138):5138-5141.

20. De KK, Timell TE. The acid hydrolysis of glycosides. *Carbohydr Res*. 1967;4(2):177-181. doi:10.1016/S0008-6215(00)82576-8.

21. Macauley MS, Stubbs KA, Vocadlo DJ. O-GlcNAcase catalyzes cleavage of thioglycosides without general acid catalysis. *J Am Chem Soc*. 2005;127(49):17202-17203. doi:10.1021/ja0567687.

22. Yip VLY, Withers SG. Family 4 glycosidases carry out efficient hydrolysis of thioglycosides by an α,β-elimination mechanism. *Angew Chemie - Int Ed*. 2006;45(37):6179-6182. doi:10.1002/anie.200601421.

23. Reese ET, Clapp RC, Mandels M. A thioglucosidase in fungi. *Arch Biochem Biophys*. 1958;75(1):228-242. doi:10.1016/0003-9861(58)90412-0.

24. Meulenbeld GH, Hartmans S. Thioglucosidase activity from Sphingobacterium sp. strain OTG1. *Appl Microbiol Biotechnol*. 2001;56(5-6):700-706. doi:10.1007/s002530100726.

25. Goodman I, Fouts JR, Bresnick E, Menegas R, Hitchings GH. A Mammalian Thioglycosidase. *Science (80-)*. 1959;130(3373):450-451. doi:10.1126/science.130.3373.450.

26. Burmeister WP, Cottaz S, Rollin P, Vasella A, Henrissat B. High Resolution X-ray

Crystallography Shows That Ascorbate Is a Cofactor for Myrosinase and Substitutes for the Function of the Catalytic Base. *J Biol Chem*. 2000;275(50):39385-39393. doi:10.1074/jbc.M006796200.

27. Botti MG, Taylor MG, Botting NP. Studies on the mechanism of myrosinase: Investigation of the effect of glycosyl acceptors on enzyme activity. *J Biol Chem*. 1995;270(35):20530-20535. doi:10.1074/jbc.270.35.20530.

28. Jones AME, Winge P, Bones AM, Cole R, Rossiter JT. Characterization and evolution of a myrosinase from the cabbage aphid Brevicoryne brassicae. *Insect Biochem Mol Biol*. 2002;32(3):275-284. doi:10.1016/S0965-1748(01)00088-1.

29. Jones AME, Bridges M, Bones AM, Cole R, Rossiter JT. Purification and characterisation of a non-plant myrosinase from the cabbage aphid Brevicoryne brassicae (L.). *Insect Biochem Mol Biol*. 2001;31(1):1-5. doi:10.1016/S0965-1748(00)00157-0.

30. Pontoppidan B, Ekbom B, Eriksson S, Meijer J. Purification and characterization of myrosinase from the cabbage aphid (Brevicoryne brassicae), a brassica herbivore. *Eur J Biochem*. 2001;268(4):1041-1048. doi:10.1046/j.1432-1327.2001.01971.x.

31. Husebye H, Arzt S, Burmeister WP, et al. Crystal structure at 1.1 Å resolution of an insect myrosinase from Brevicoryne brassicae shows its close relationship to β-glucosidases. *Insect Biochem Mol Biol*. 2005;35(12):1311-1320. doi:10.1016/j.ibmb.2005.07.004.

32. Francis F, Lognay G, Wathelet JP, Haubruge E. Characterisation of aphid myrosinase and degradation studies of glucosinolates. *Arch Insect Biochem Physiol*. 2002;50(4):173-182. doi:10.1002/arch.10042.

33. Albaser A, Kazana E, Bennett MH, et al. Discovery of a Bacterial Glycoside Hydrolase Family

3 (GH3) β-Glucosidase with Myrosinase Activity from a Citrobacter Strain Isolated from Soil. *J Agric Food Chem*. 2016;64(7):1520-1527. doi:10.1021/acs.jafc.5b05381.

34. Hugenholtz P, Goebel BM, Pace NR. Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. *J Bacteriol*. 1998;180(24):6793. doi:0021-9193/98/$04.00+0.

35. Berini F, Casciello C, Marcone GL, Marinelli F. Metagenomics: Novel enzymes from non-culturable microbes. *FEMS Microbiol Lett*. 2017;364(21):1-19. doi:10.1093/femsle/fnx211.

36. Armstrong Z, Rahfeld P, Withers SG. *Discovery of New Glycosidases From Metagenomic Libraries*. Vol 597. 1st ed. Elsevier Inc.; 2017. doi:10.1016/bs.mie.2017.06.001.

37. Moser MJ, DiFrancesco RA, Gowda K, et al. Thermostable DNA polymerase from a viral metagenome is a potent RT-PCR enzyme. *PLoS One*. 2012;7(6). doi:10.1371/journal.pone.0038371.

38. Neveu J, Regeard C, Dubow MS. Isolation and characterization of two serine proteases from metagenomic libraries of the Gobi and Death Valley deserts. *Appl Microbiol Biotechnol*. 2011;91(3):635-644. doi:10.1007/s00253-011-3256-9.

39. Pushpam PL, Rajesh T, Gunasekaran P. Identification and characterization of alkaline serine protease from goat skin surface metagenome. *AMB Express*. 2011;1(1):1-10. doi:10.1186/2191-0855-1-3.

40. Gabor EM, De Vries EJ, Janssen DB. Construction, characterization, and use of small-insert gene banks of DNA isolated from soil and enrichment cultures for the recovery of novel amidases. *Environ Microbiol*. 2004;6(9):948-958. doi:10.1111/j.1462-2920.2004.00643.x.

41. Lee CM, Lee YS, Seo SH, et al. Screening and characterization of a novel cellulase gene from

the gut microflora of Hermetia Illucens using metagenomic library. *J Microbiol Biotechnol*. 2014;24(9):1196-1206. doi:10.4014/jmb.1405.05001.

42. Chen H-M, Armstrong Z, Hallam SJ, Withers SG. Synthesis and evaluation of a series of 6-chloro-4-methylumbelliferyl glycosides as fluorogenic reagents for screening metagenomic libraries for glycosidase activity. *Carbohydr Res*. 2016;421:33-39. doi:10.1016/j.carres.2015.12.010.

43. Azoulay M, Ivastchenko N, Florent J-C, Claude Monneret. p-Nitrophenyl beta-D-glucopyranosiduronic analogs as potential substrates for beta-glucuronidase. *Bioorganic Med Chem Lett*. 1997;7(10):1307-1310.

44. Adamczyk M, Cornwell M, Huff J, Rege S, Rao TVS. Novel 7-hydroxycoumarin based fluorescent labels. *Bioorganic Med Chem Lett*. 1997;7(15):1985-1988. doi:10.1016/S0960-894X(97)00345-4.

45. Reddie KG, Humphries WH, Bain CP, Payne CK, Kemp ML, Murthy N. Fluorescent Coumarin Thiols Measure Biological Redox Couples. *Org Lett*. 2012;14(3):680-683. doi:10.1021/ol203105c.

46. Choi A, Miller SC. Reductively-labile sulfonate ester protecting groups that are rapidly cleaved by physiological glutathione. *Org Biomol Chem*. 2017;15(6):1346-1349. doi:10.1039/C7OB00063D.

47. Rokita S. Frontmatter. In: *Quinone Methides*. Hoboken, NJ, USA: John Wiley & Sons, Inc.; 2009:i-xviii. doi:10.1002/9780470452882.fmatter.

48. Hortmann AG, Aron AJ, Bhattacharya AK. 3H-1,2-Benzodithiole Oxides: Studies Directed toward the Generation of o-Thiobenzoquinone Methide and Benzo[b]thiete. *J Org Chem*.

1978;43(17):3374-3378. doi:10.1021/jo00411a024.

49. Okazaki R, Ishii F, Inamoto N. Reactions of Thioacylketene Thioacetals and o -Thioquinone Methides with Enamines. *Bull Chem Soc Jpn*. 1978;51(1):309-312. doi:10.1246/bcsj.51.309.

50. Kang K-T, Okazaki R, Inamoto N. Synthesis and Properties of o -Thioquinone Methides Having Ketene Aminal, Ketene Acetal, Ketene Monothioacetal, or Ketene Dithioacetal Group. *Bull Chem Soc Jpn*. 1979;52(12):3640-3646. doi:10.1246/bcsj.52.3640.

51. Okazaki R, Ishii F, Sunagawa K, Inamoto N. Unusual Mode of Dimerization of o -Thioquinone Methide. Isolation of Head-to-Head and Head-to-Tail Dimers. *Chem Lett*. 1978;7(1):51-54. doi:10.1246/cl.1978.51.

52. Itoh T, Fujikawa K, Kubo M. *p* -Thioquinone Methides: Synthesis and Reaction. *J Org Chem*. 1996;61(23):8329-8331. doi:10.1021/jo9609541.

53. Yaouancq L, Anissimova M, Badet-Denisot MA, Badet B. Design and evaluation of mechanism-based inhibitors of D-alanyl-D-alanine dipeptidase VanX. *European J Org Chem*. 2002;(21):3573-3579.

54. Massiere F, Badet-Denisot MA, Rene L, Badet B. Design, synthesis, and evaluation of the first mechanism-based inhibitor of glucosamine 6-phosphate synthase. *J Am Chem Soc*. 1997;119(24):5748-5749. doi:10.1021/ja970254t.

55. Aráoz R, Anhalt E, René L, Badet-Denisot MA, Courvalin P, Badet B. Mechanism-based inactivation of VanX, a D-alanyl-D-alanine dipeptidase necessary for vancomycin resistance. *Biochemistry*. 2000;39(51):15971-15979. doi:10.1021/bi001408b.

56. Senter PD, Pearce WE, Greenfield RS. Development of a Drug-Release Strategy Based on the Reductive Fragmentation of Benzyl Carbamate Disulfides. *J Org Chem*. 1990;55(9):2975-

2978. doi:10.1021/jo00296a082.

57. Ma F, Fischer M, Han Y, Withers SG, Feng Y, Yang GY. Substrate Engineering Enabling Fluorescence Droplet Entrapment for IVC-FACS-Based Ultrahigh-Throughput Screening. *Anal Chem*. 2016;88(17):8587-8595. doi:10.1021/acs.analchem.6b01712.

58. Clouthier DJ, Ramsay DA. The Spectroscopy of Formaldehyde and Thioformaldehyde. *Annu Rev Phys Chem*. 1983;34(1):31-58. doi:10.1146/annurev.pc.34.100183.000335.

59. DuBois MR. Homogeneous Hydrogenolysis of Carbon Disulfide Catalyzed by a Molybdenum Dimer with Sulfido Ligands. *J Am Chem Soc*. 1983;105(11):3710-3711. doi:10.1021/ja00349a060.

60. Liu B, Thayumanavan S. Substituent Effects on the pH Sensitivity of Acetals and Ketals and Their Correlation with Encapsulation Stability in Polymeric Nanogels. *J Am Chem Soc*. 2017;139(6):2306-2317. doi:10.1021/jacs.6b11181.

61. Gillies ER, Goodwin AP, Fréchet JMJ. Acetals as pH-sensitive linkages for drug delivery. *Bioconjug Chem*. 2004;15(6):1254-1263. doi:10.1021/bc049853x.

62. Tietze LF, Neumann M, Möllers T, et al. Proton-mediated Liberation of Aldophosphamide from a Nontoxic Prodrug: A Strategy for Tumor-selective Activation of Cytocidal Drugs. *Cancer Res*. 1989;49(15):4179-4184.

63. Srinivasachar K, Neville DM. New Protein Cross-Linking Reagents That Are Cleaved by Mild Acid. *Biochemistry*. 1989;28(6):2501-2509. doi:10.1021/bi00432a023.

64. Neville DM, Srinivasachar K, Stone R, Scharff J. Enhancement of immunotoxin efficacy by acid-cleavable cross-linking agents utilizing diphtheria toxin and toxin mutants. *J Biol Chem*. 1989;264(25):14653-14661.

65. Zhu X, Pachamuthu K, Schmidt RR. Synthesis of novel S-neoglycopeptides from glycosylthiomethyl derivatives. *Org Lett*. 2004;6(7):1083-1085. doi:10.1021/ol036186z.

66. Ngane AN, Lavault M, Séraphin D, Landreau A, Richomme P. Three 1-thio-β-D-glucopyranosides from the seeds of Afrostyrax lepidophyllus Mildbr. *Carbohydr Res*. 2006;341(17):2799-2802. doi:10.1016/j.carres.2006.09.021.

67. Cecioni S, Vocadlo DJ. Carbohydrate Bis-acetal-Based Substrates as Tunable Fluorescence-Quenched Probes for Monitoring exo-Glycosidase Activity. *J Am Chem Soc*. 2017;139(25):8392-8395. doi:10.1021/jacs.7b01948.

68. Chibba A, Dasgupta S, Yakandawala N, Madhyastha S, Nitz M. Chromogenic carbamate and acetal substrates for glycosaminidases. *J Carbohydr Chem*. 2011;30(7-9):549-558. doi:10.1080/07328303.2011.610543.

69. Giordano M, Iadonisi A. Polymethylhydrosiloxane (PMHS): A convenient option for synthetic applications of the iodine/silane combined reagent - Straightforward entries to 2-hydroxyglycals and useful building-blocks of glucuronic acid and glucosamine. *European J Org Chem*. 2013;(1):125-131. doi:10.1002/ejoc.201201084.

70. Ashani Y, Catravas GN. Highly reactive impurities in Triton X-100 and Brij 35: Partial characterization and removal. *Anal Biochem*. 1980;109(1):55-62. doi:10.1016/0003-2697(80)90009-3.

71. Weinert EE, Dondi R, Colloredo-Melz S, et al. Substituents on quinone methides strongly modulate formation and stability of their nucleophilic adducts. *J Am Chem Soc*. 2006;128(36):11940-11947. doi:10.1021/ja062948k.

72. Ge C, Wang H, Zhang B, et al. A thiol–thiosulfonate reaction providing a novel strategy for

turn-on thiol sensing. *Chem Commun*. 2015;51(80):14913-14916. doi:10.1039/C5CC05390K.

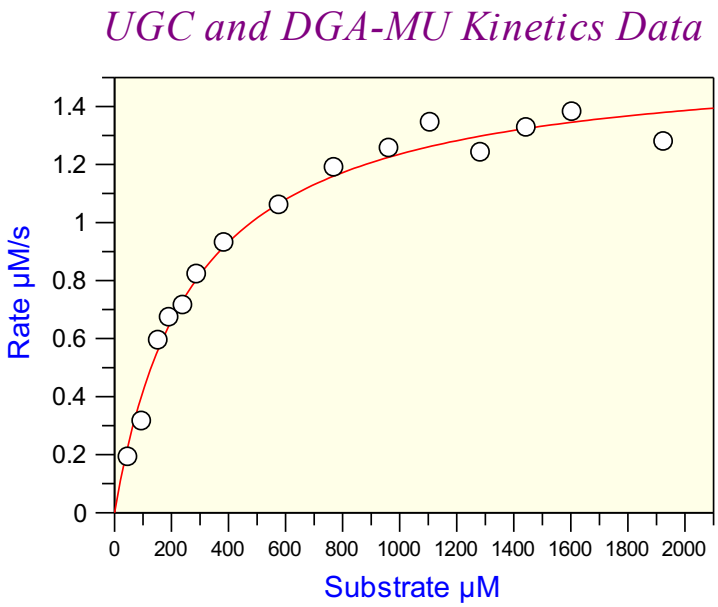
73. Pilgrim W, Murphy P V. SnCl 4 - and TiCl 4 -Catalyzed Anomerization of Acylated O - and S -Glycosides: Analysis of Factors That Lead to Higher α:β Anomer Ratios and Reaction Rates. *J Org Chem*. 2010;75(20):6747-6755. doi:10.1021/jo101090f.

74. Himeno M, Hashiguchi Y, Kato K. β-Glucuronidase of Bovine Liver. *J Biochem*. 1974;76(6):1243-1252. doi:10.1093/oxfordjournals.jbchem.a130677.

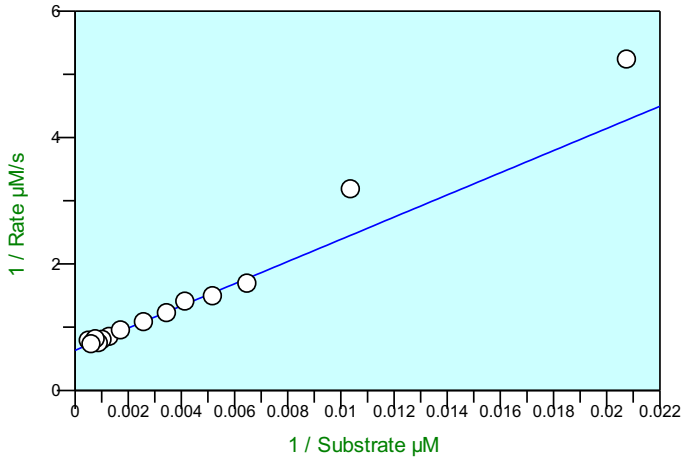
Appendices

Appendix A Plots for Michaelis-Menten kinetics

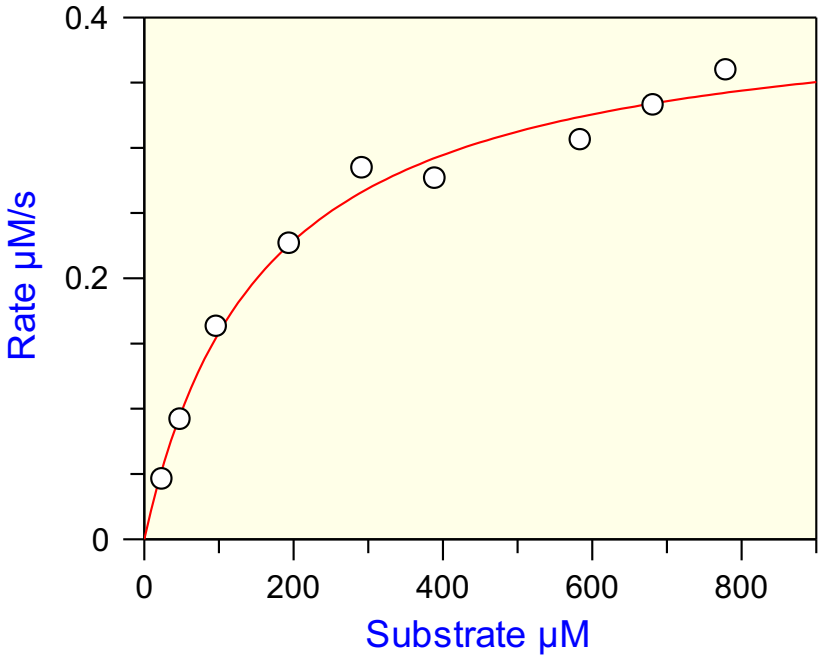
A.1 Plots for UGC



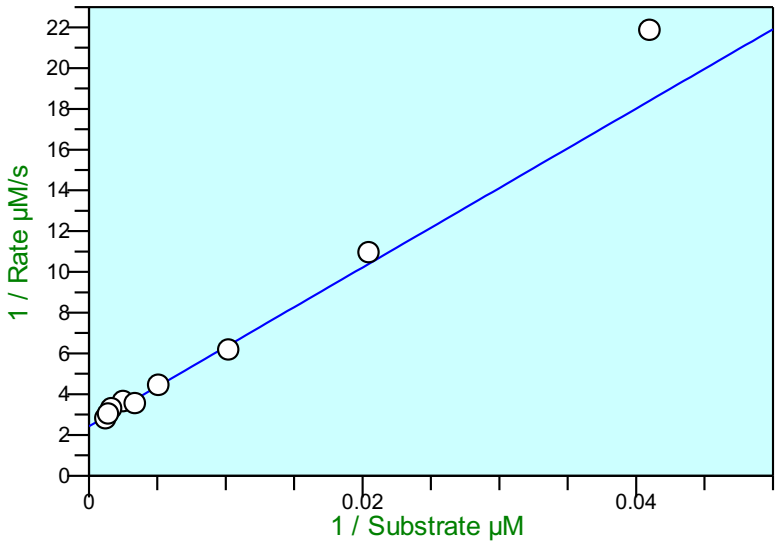
Parameter	Value	Std. Error
Vmax	1.5779	0.0420
Km	276.9117	24.8375



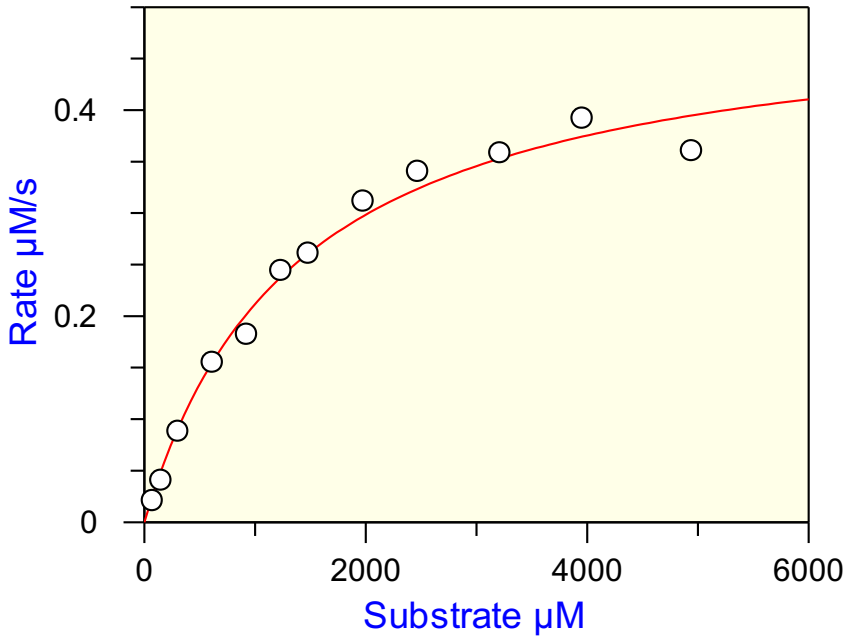
UGC and DGA-ClMU Kinetics Data



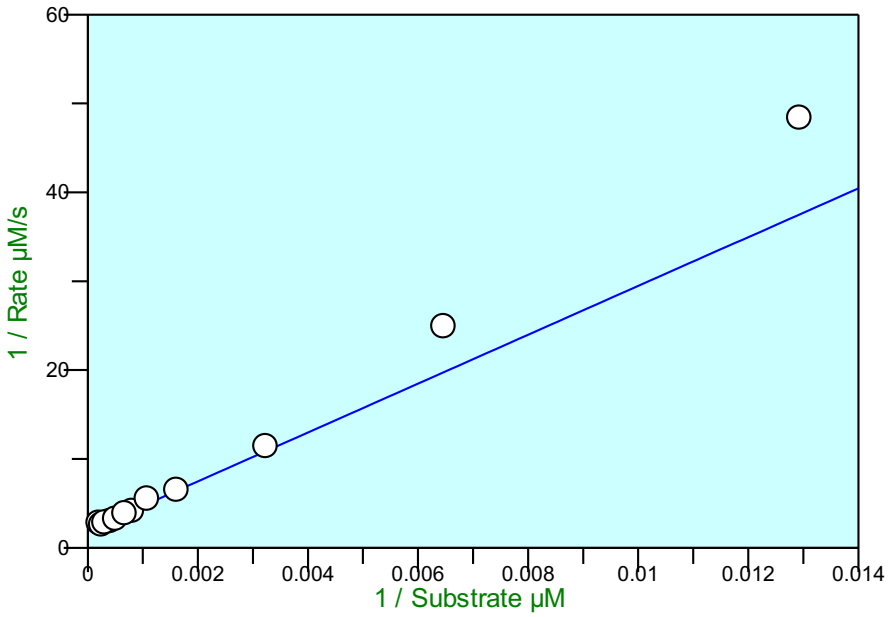
Parameter	Value	Std. Error
Vmax	0.4130	0.0170
Km	160.9836	21.2062



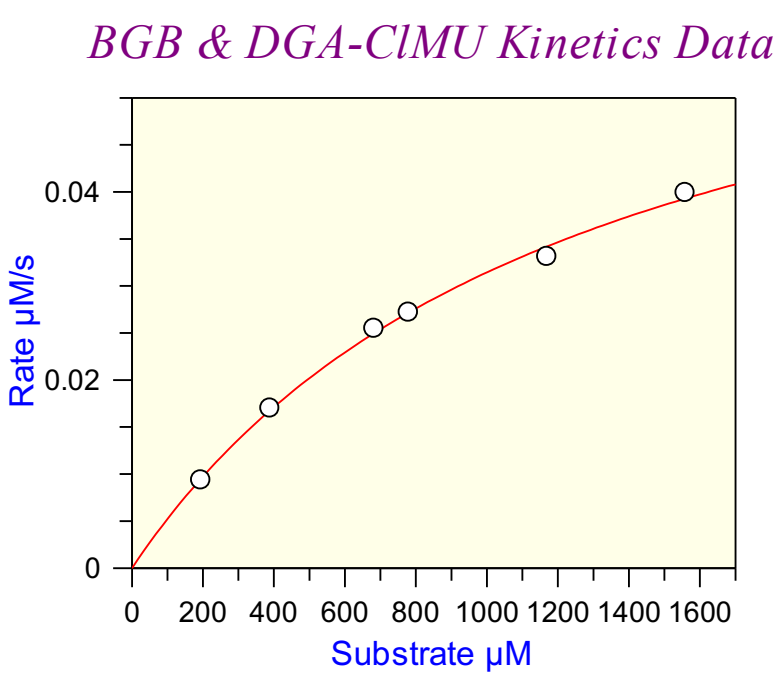
UGC & DGA-SMU Kinetics Data



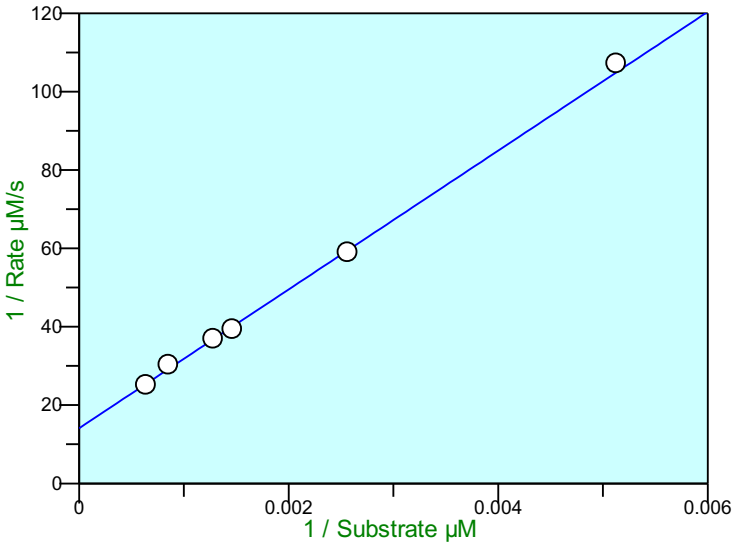
Parameter	Value	Std. Error
Vmax	0.5057	0.0256
Km	1389.5530	180.7314



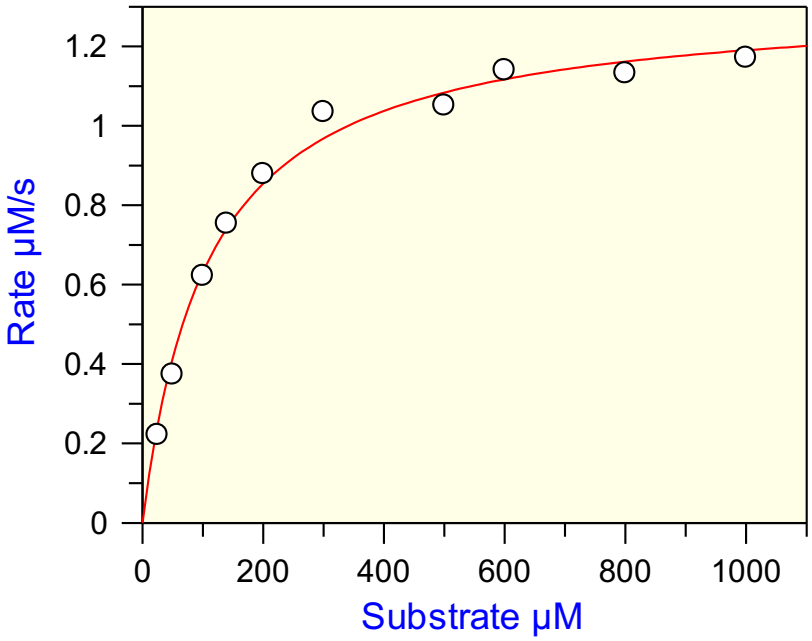
A.2 Plots for BGB



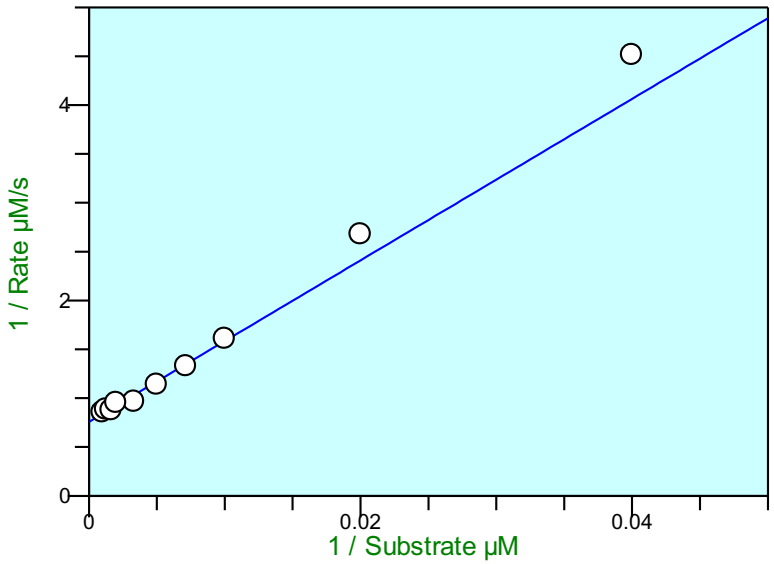
Parameter	Value	Std. Error
Vmax	0.0710	0.0039
Km	1256.3924	123.4729



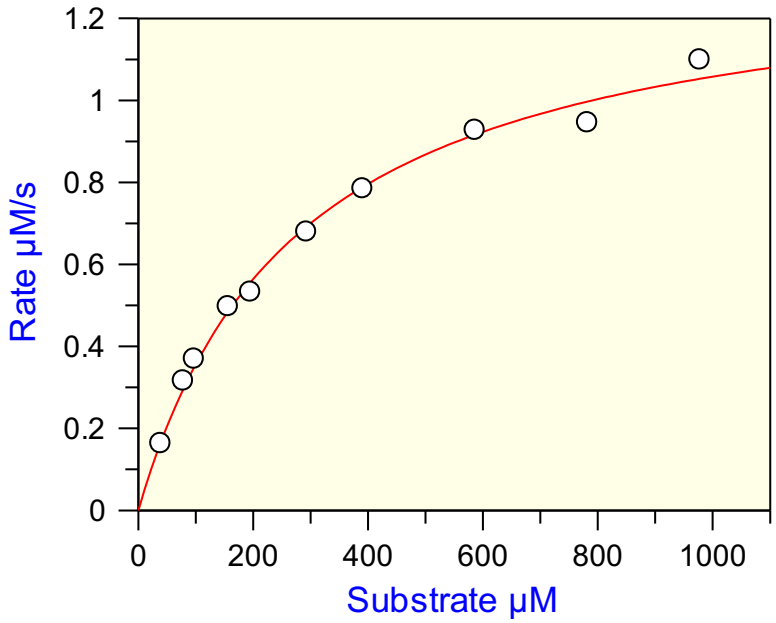
BGB & GA-ClMU Kinetics Data



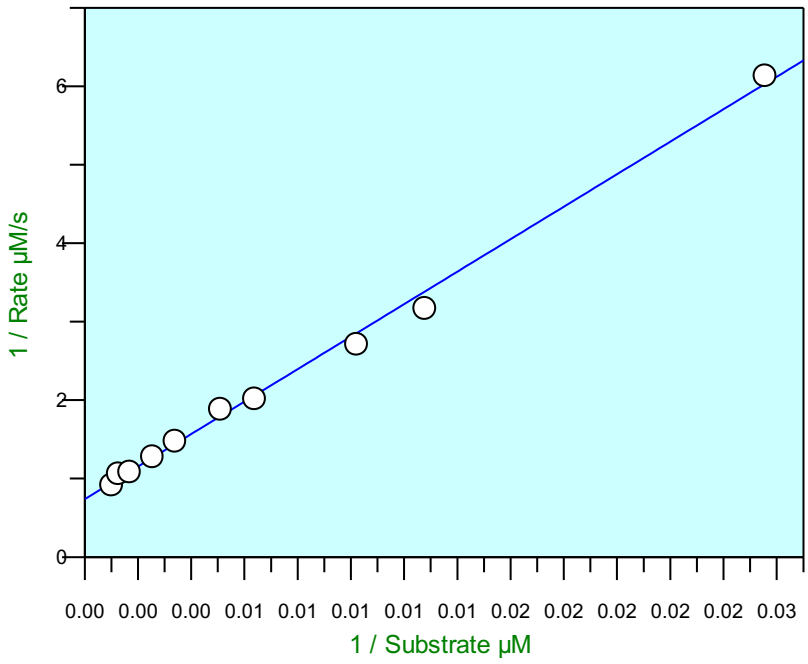
Parameter	Value	Std. Error
Vmax	1.3201	0.0303
Km	109.0818	9.1309



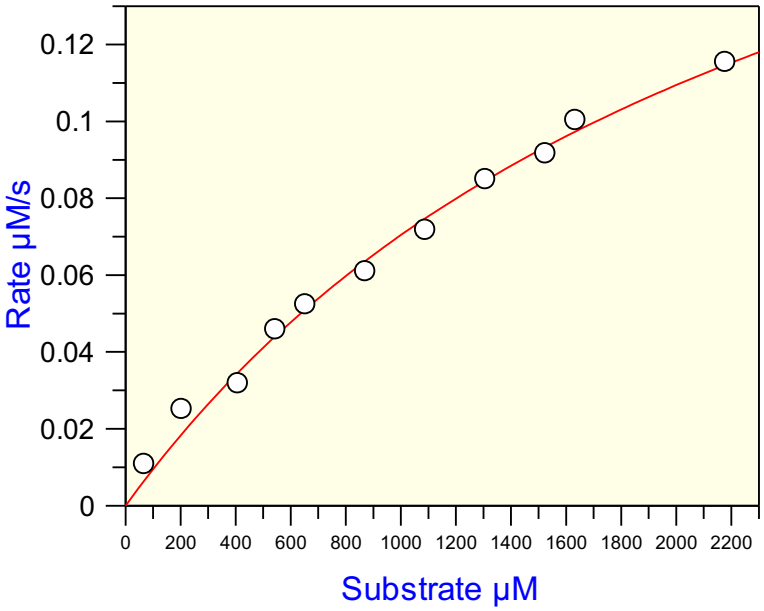
BGB & GA-MU Kinetics Data



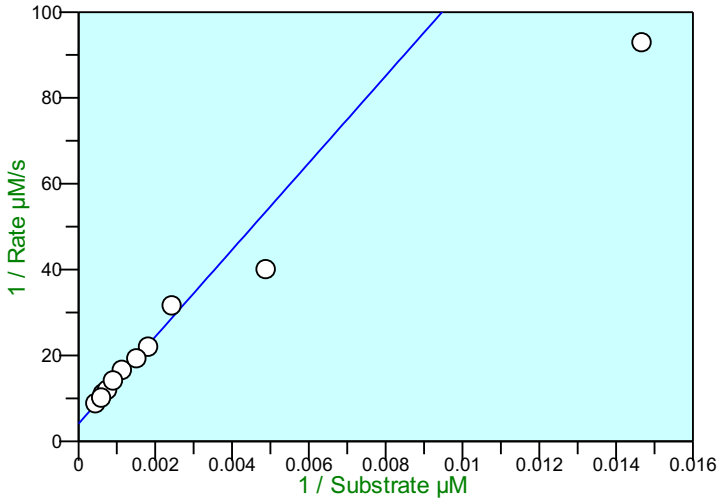
Parameter	Value	Std. Error
Vmax	1.3547	0.0459
Km	280.4581	23.5972



BGB & GA-SMU Kinetics Data

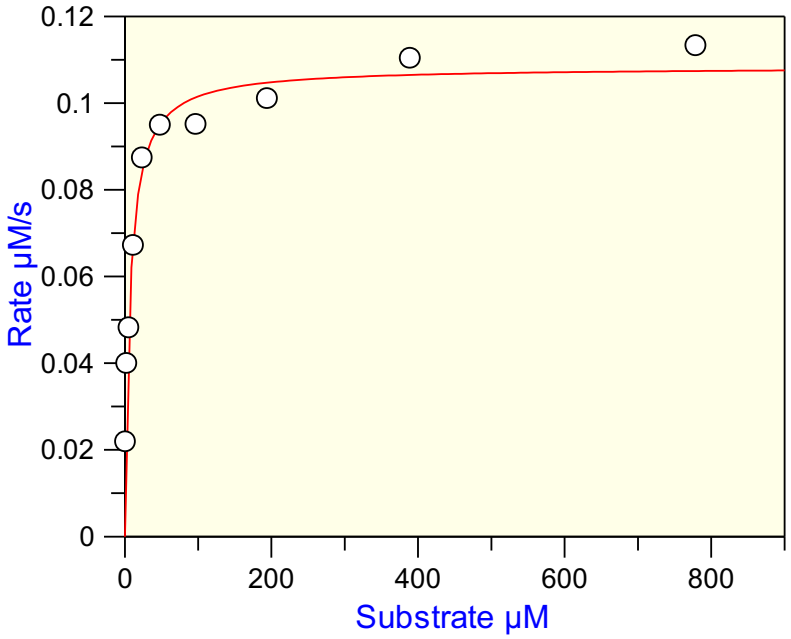


Parameter	Value	Std. Error
Vmax	0.2456	0.0240
Km	2488.4068	375.6549

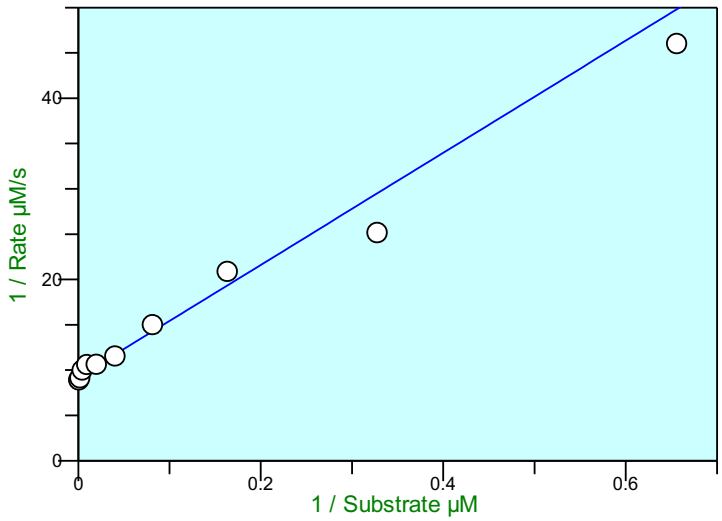


A.3 Plots for BGE

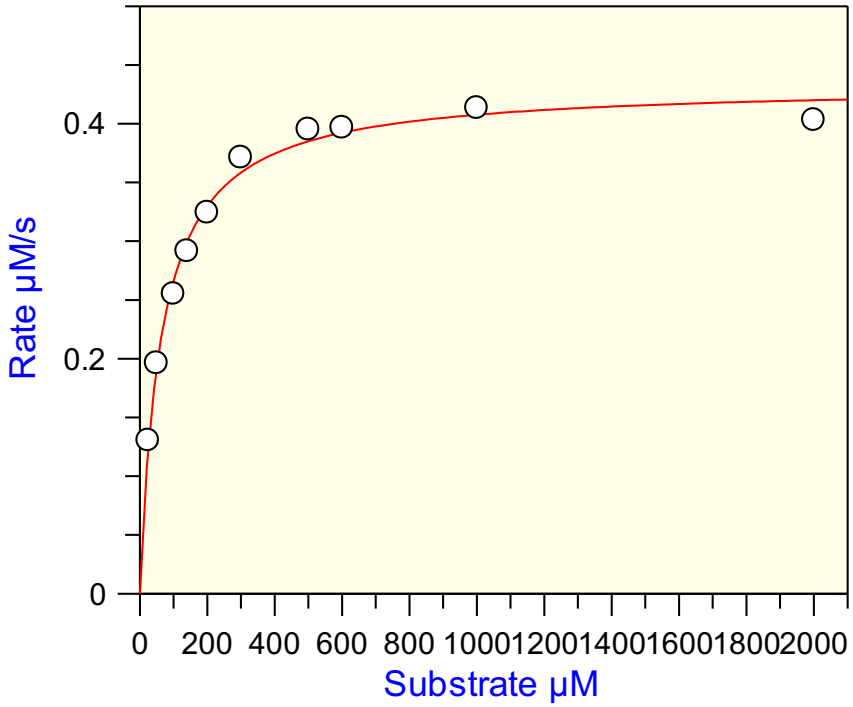
BGE & DGA-ClMU Kinetics Data



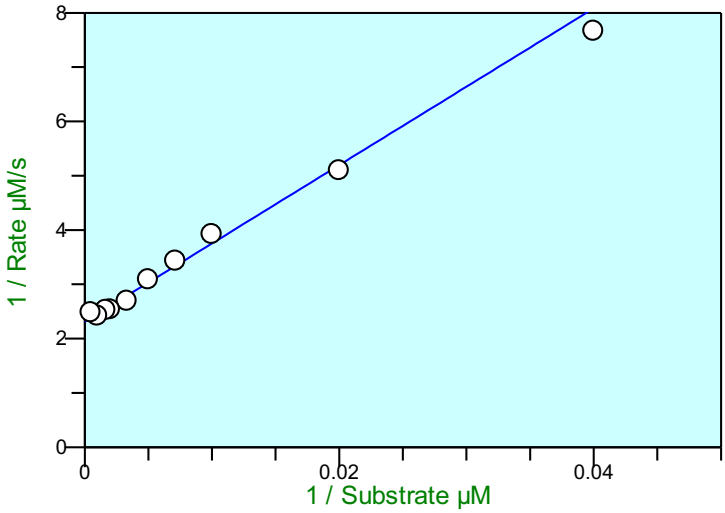
Parameter	Value	Std. Error
Vmax	0.1083	0.0023
Km	6.6999	0.7334



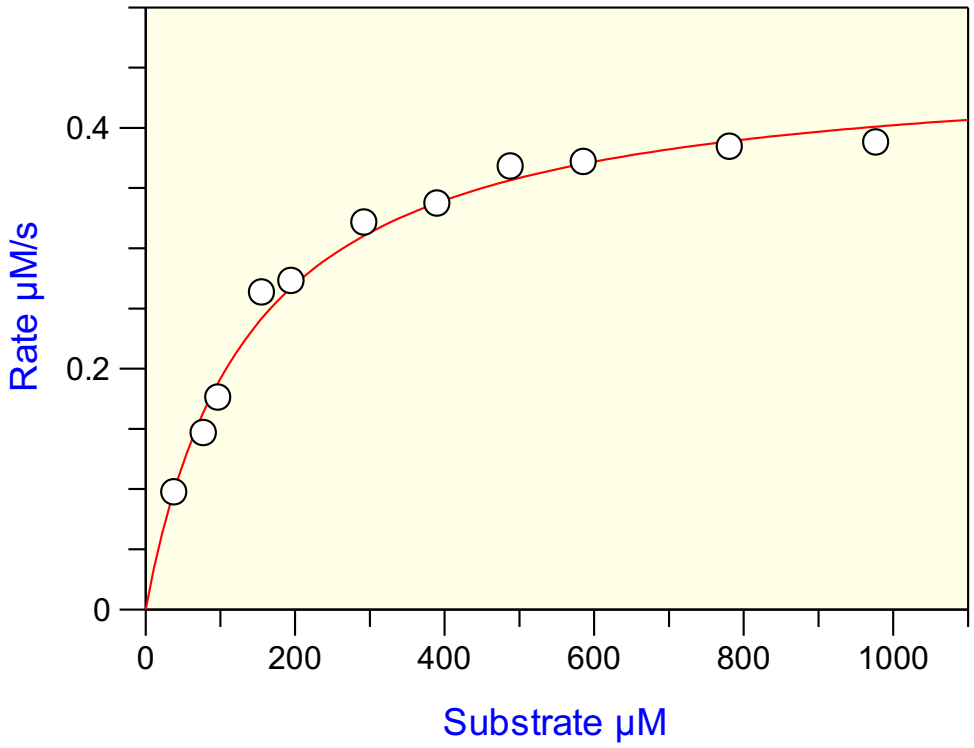
BGE & GA-ClMU Kinetics Data



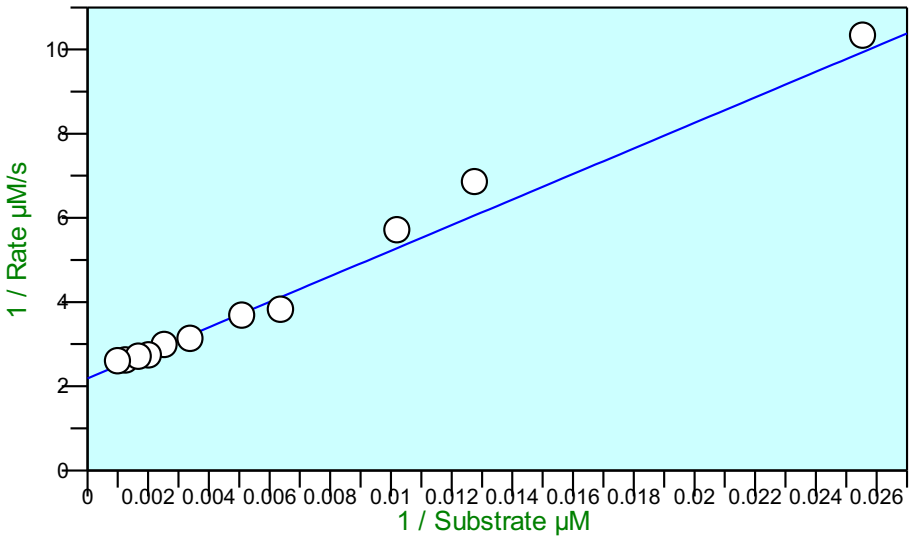
Parameter	Value	Std. Error
Vmax	0.4331	0.0067
Km	62.5083	4.4109



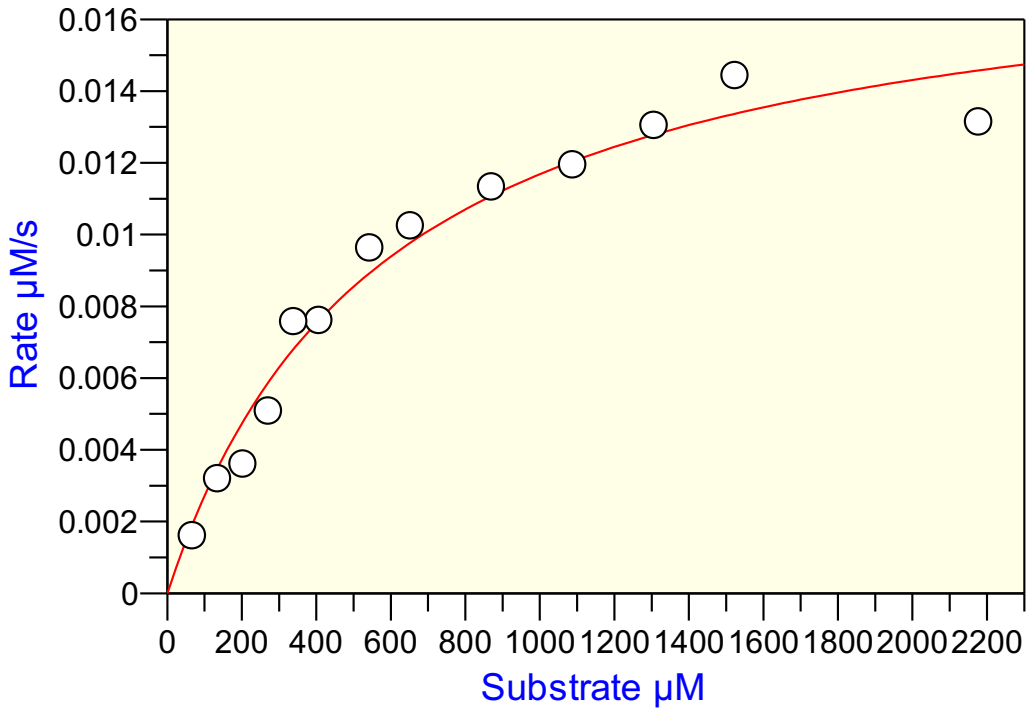
BGE & GA-MU Kinetics Data



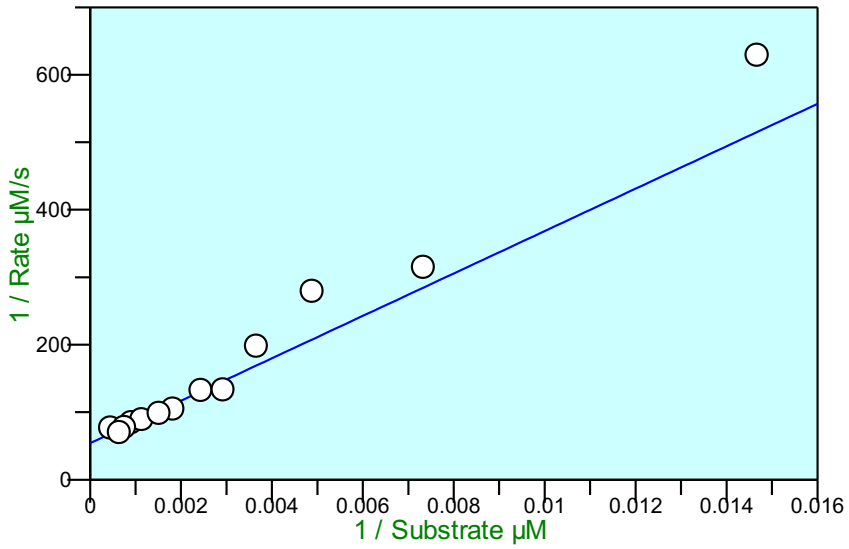
Parameter	Value	Std. Error
Vmax	0.4581	0.0119
Km	139.0909	11.9405



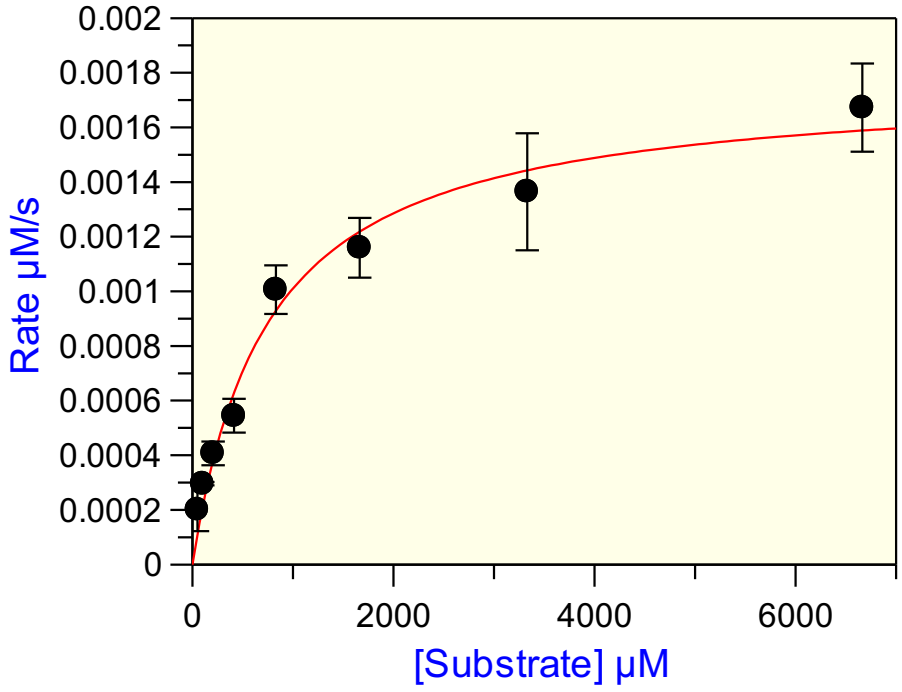
BGE & GA-SMU Kinetics Data



Parameter	Value	Std. Error
Vmax	0.0185	0.0012
Km	579.8402	90.2192



BGE & DGA-MU Kinetics Data



Parameter	Value	Std. Error
Vmax	0.0018	9.57933e-005
Km	749.7612	129.4362