Mechanistic Investigation of Families of Retaining Glycosyl Hydrolases

by

Saswati Chakladar

B.Sc. (Hons., Chemistry), University of Delhi, India, 2001
M.Sc., University of Delhi, India, 2003

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in the

Department of Chemistry

Faculty of Science

SIMON FRASER UNIVERSITY

Spring 2012

Approval

Name: **Saswati Chakladar**

Degree: **Doctor of Philosophy (Chemistry)**

Title of Thesis: ***Mechanistic Investigation of Families of Retaining Glycosyl Hydrolases***

Examining Committee:

Chair: **Dr. David Vocadlo**
Professor

Dr. Andrew J. Bennet
Professor, Department of Chemistry
Senior Supervisor

Dr. Robert A. Britton
Associate Professor, Department of Chemistry
Supervisor

Dr. George R. Agnes
Professor, Department of Chemistry
Supervisor

Dr. B. Mario Pinto
Professor, Department of Chemistry
Internal Examiner

Dr. David Palmer
Associate Professor, Department of Chemistry
University of Saskatchewan
External Examiner

Date Defended/Approved: April 24, 2012

Partial Copyright Licence

Abstract

Enzymes that catalyze the removal of carbohydrate linkages from biological molecules are called glycosyl hydrolases (GH). These enzymes have been catagorized into more than 130 different families. Family-4 glycosyl hydrolases catalyze hydrolysis *via* an unusual oxidation-reduction mechanism that requires NAD^+, Mn^{2+} and reducing conditions for catalytic activity, a situation that contrasts the classical nucleophilic substitution mechanism of most glycosyl hydrolases. Most of this thesis is directed towards detailed mechanistic evaluations of the GH4 α-galactosidase from *Citrobacter freundii* and the GH109 α-*N*-acetylgalactosaminidase, which has a bound active site NAD^+, from *Elizabethkingia meningosepticum*. Experimentally, this involved cloning the gene, expressing and purifying the recombinant protein, synthesizing substrates, measuring pH-rate profiles, Brønsted parameters and deuterium kinetic isotope effects (KIEs).

The MelA α-galactosidase (GH4, *C. freundii*) possesses a substrate specificity that is limited to non-phosphorylated α-D-galactosides. The measured primary deuterium KIEs show that the first irreversible step involves a concerted oxidation of the C-3 hydroxyl group and proton abstraction at C-2. Subsequent elimination of the leaving group is kinetically silent.

The GH109 α-*N*-acetylgalactosaminidase (*E. meningosepticum*) displayed a significant degree of substrate promiscuity as it catalyzed the hydrolysis of aryl 2-acetamido-2-deoxy-α-D-galactosides and aryl α-D-galactosides efficiently. Notably, Brønsted analysis on both aryl 2-acetamido-2-deoxy-D-galactosides and aryl α-D-galactosides suggests that leaving group departure is not rate-limiting. The measured deuterium KIEs kinetic isotope effects using the "non-natural" phenyl α-D-galactoside substrate revealed that both C-3 oxidation and C-2 proton abstraction are kinetically significant.

A bicyclo[4.1.0]heptyl mimic of an aryl α-D-galactopyranoside, (1*R*,2*S*,3*S*,4*R*,5*S*,6*S*)-5-(3,5-difluorophenoxy)-1-(hydroxymethyl)bicyclo[4.1.0]heptan-

2,3,4-triol, was made and shown to be an irreversible inhibitor of the GH36 α-galactosidase from *Thermotoga maritima*. In contrast, the diastereomeric L-*altro* isomer failed to inactivate the enzyme, while the analogous cyclohexene was shown to be a tight-binding competitive inhibitor. The measured pH-rate profiles for inhibition and reactivation as well as the corresponding catalytic and inhibitory proficiencies suggested that inhibition results from the formation of an oxacarbenium ion in the enzymatic site that is trapped rapidly by an active site enzymatic residue.

Keywords: Alpha galactosidase; Alpha galactosaminidase; Kinetic Isotope Effects; Irreversible inhibitor; Cloning; Expression; Kinetics

Dedication

To my family especially my husband Krishnendu

Acknowledgements

I would like to express my sincerest gratitude to my senior supervisor, Prof. Andrew Bennet for the opportunity to work in his laboratory. The supervision and support that he gave me truly helped the progression and smoothness of the doctoral program. Prof. Bennet is an excellent teacher with vast and diverse knowledge in the field of chemistry and biochemistry. His humourous nature and friendly behaviour throughout the graduate studies made the journey of five years, an enjoyable learning period. One simply could not wish for a friendlier supervisor. I would also like to thank him for the training and support he provided in order to excel in scientific writing and giving me the opportunity to participate in several national and international conferences.

I would like to thank Prof. Rob Britton for serving on my supervisory committee. I whole-heartedly appreciate his insightful comments on my research projects during my committee meetings which helped me in improving on my in-depth understanding of the mechanistic aspects of organic chemistry. Prof. George Agnes has offered much advice and insight throughout my doctoral studies, in both academic and personal development. I truly thank him for giving me the opportunity to serve on the graduate student panel offered by Dean of Graduate Studies, facilitating undergraduates in persuing graduate studies.

I would like to extend my gratitude to Prof. B. Mario Pinto for agreeing to be my internal examiner amidst his busy schedule and Prof. David Vocadlo for chairing my thesis procedure. In particular, I would like to thanks Prof. Vocadlo for his valuable suggestions and collaborations. I am thankful to Prof. David Palmer for participating as the external examiner of this thesis.

In my daily work I have been blessed with a friendly and cheerful group of labmates. I would like to begin by thanking my past labmates: Dr. Ivan Hemeon, Dr. Deepani Indurugalla, Yi Wang and Zoran Cikojevic for their valuable advice and helping me get started in Bennet lab. Thanks to Dr. Jefferson Chan, who recently graduated from our lab. I am extremely indebted to my present labmates, especially Lydia Cheng for being a great friend and also for teaching me molecular biology and biochemistry from

scratch and, Fahimeh Shidmoossavee for her continued support throughout my thesis submission process and for being such a wonderful friend. All the very best to both of you for your thesis! My sincere wishes to other lab members for their scientific career: Kobi Khazaei, Viviana Cerda and Saeideh Shamsi. I had the pleasure to work with several undergraduate students and I would like to wish them a bright future: Mary Choi, Timothy Teoh, Ariel Tang, Gurtej Sandhu and April Lu.

I am grateful to Dr. Tracey Gloster, for her valuable advice for my research work in molecular biology and scientific discussions. I also appreciate the help of Thomas Clark, Dr. Anuj Yadav and David Shen in scientific consultation.

The skilled technical assistance of Colin Zhang (NMR technician), Hongwen Chen (Mass Spectrometer Technician), Frank Haftbaradaran (Elemental Analysis), Lynn Wood (Graduate Secretary) and Fred Chin (Systems Consultant) is highly appreciated.

I would like to acknowledge the help and support I have got from my friends, especially Swati Vartak, Brinda Prasad and Sudeshna Ghosh.

Last but not the least, my special thanks to my family without whom I could not have come so far in my educational career. I consider myself fortunate to have such a loving and supportive family-my parents, Mr. Subimal and Mrs. Jayati Chakladar, my sisters-Bhaswati and Arpita, my brother–in-law Debasish and my little nephew Rupanjan. I would also like to thank my parent-in-law, Mr. and Mrs Mukherjee and brother-in-law Ramendu for their continued encouragement. My heartfelt thanks to my husband Krishnendu, without whose love and support I could not have accomplished what I have today. His patience and inspiration, particularly in the last few months of thesis writing and submission is greatly appreciated.

Table of Contents

List of Figures

List of Schemes

List of Tables

Glossary

Tm Gal	*Thermotoga maritima* α-galactosidase
°C	Degree Celcius
α-GAL	α-galactosidase
α-NAGAL	α-*N*-acetylgalactosaminidase
Å	Angstrom
BSA	Bovine Serum Albumin
CAZy	Carbohydrate-Active enzymes database
Da	Dalton
DCM	Dichloromethane
DMF	*N, N*-Dimethylformamide
DTT	Dithiothreitol
E.coli	*Escherichia coli*
E.Meningosepticum	*Elizabethkingia Meningosepticum*
$E1_{CB}$	Elimination unimolecular *via* conjugate base
EC	Enzyme Classification number
ES-MS	Electro-spray mass spectrometry
ERT	Enzyme Replacement Therapy
GAC	General Acid Catalysis
Gal	α-D-Galactopyranosides
GalNAc	2-Acetamido-2-deoxy-D-galactopyranose
GBC	General Base Catalysis
GH	Glycosyl Hydrolase
h	Hour
HEPES	4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid
His	Histidine
HRMS	High Resolution Mass Spectrometry
IC_{50}	The half of maximal inhibitory concentration
J	Spin-spin coupling constant
Kb	Kilobase
k_{cat}	Turnover number

k_{cat}/K_m	Catalytic Efficiency
$k_{cat}/K_m \times 1/k_{uncat}$	Catalytic Proficiency
K_i	Inhibition constant
KIE	Kinetic Isotope Effect
K_m	Michaelis Constant
LSD	Lysosomal Storage Disorder
m	Multiplet
Mel	Melibiose
MES	2-(*N*-morpholino)ethanesulfonic acid
MHz	MegaHertz
min	Minute
mL	Millilitre
mM	Millimolar
Mn	Manganese
MS	Mass Spectrometry
NAD	Nicotinamide Adenine Dinucleotide
NADH	Nicotinamide Adenine Dinucleotide (reduced form)
NADP	Nicotinamide Adenine Dinucleotide Phosphate
NMR	Nuclear Magnetic Resonance
nOe	Nuclear Overhauser Effect
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
Pyr	Pyridine
R*f*	Retention Factor
rpm	Revolutions per minute
s	Singlet
t	Triplet
TBAF	tera-n-butylammonium fluoride
TCEP	*tris*(2-carboxyethyl)phosphine
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
TLC	Thin Layer Chromatography

TS	Transition State
U	Unit of enzyme activity
V_{max}	Maximum reaction rate of an enzyme
ZPE	Zero-point energy
IPTG	2-propyl 1-thio-β-D-galactoside

1. General Introduction

1.1. Carbohydrates

Carbohydrates are the most abundant biomolecules produced by living organisms, in which they can be fuels, metabolic intermediates and energy stores. Sugar and starch are dietary staples in many parts of the world and the oxidation of these carbohydrates is the central energy-yielding pathway for non-photosynthetic cells. Polysaccharides such as cellulose and chitin comprise the primary structural component for plants and insects. Historically, carbohydrates were defined as polyhydroxy aldehydes (aldoses) or ketones (ketoses), or substances that yield those compounds upon hydrolysis. The most commonly accepted empirical formula of carbohydrates is $C_n(H_2O)_m$. The term "monosaccharides" refers to carbohydrates containing a single sugar unit whereas "polysaccharides" refers to multiple monosaccharide units joined together by acetal or ketal linkages. Common monosaccharides like D-glucose (n = m = 6) **1-1a** and D-galactose (n = m = 6) **1-1b** or the disaccharide sucrose (n = 12, m = 11) are examples of $C_n(H_2O)_m$ sugars. Four of the six carbon atoms in these hexoses ($C_n(H_2O)_m$ where n = 6) are stereogenic centres, based on which there are 16 possible hexoses. A series of eight common names with the prefixes D (dextrorotatory) or L (levorotatory) are used to denote these 16 hexoses. A change in stereochemical configuration of a single carbon atom in sugar is referred to as epimerization, for example D-glucose and D-galactose are C-4 epimers (**Figure 1-1**). In the open chain form, the carbonyl group of the aldehyde or ketone in aqueous solution is in equilibrium with a hydrated form (**Figure 1-2**). Intramolecular cyclization due to the attack of a nucleophilic hydroxyl (OH) group onto the carbonyl centre, primarily results in two types of cyclic hemiacetals, pyranose (6-membered ring) and furanose (5-membered ring) (**Figure 1-2**). The newly formed stereocentre resulting from hemiacetal formation, exist as an equilibrium mixture of two diastereomers, differing only in their configuration at the masked carbonyl

group of a sugar and these are termed as anomers. Anomers are categorized by their designators α and β. (**Figure 1-2**):

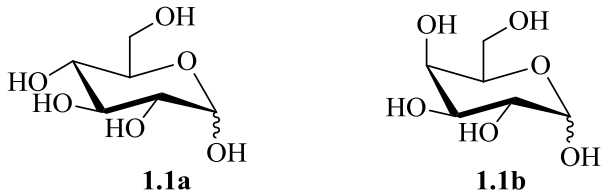


Figure 1-1. Commonly known monosaccharides D-glucose (**1.1a**) and D-galactose (**1.1b**).

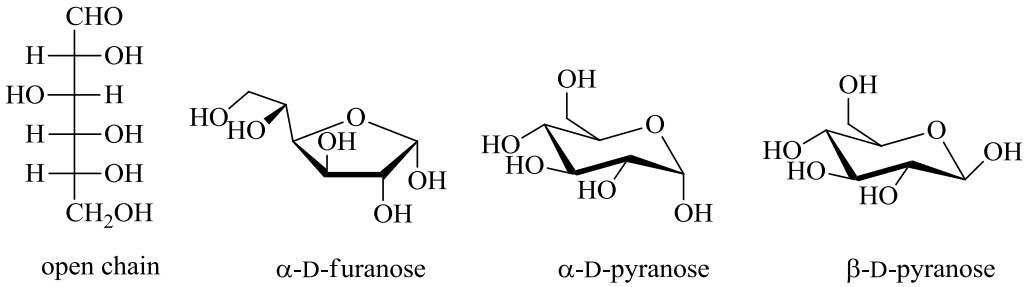


Figure 1-2. The different forms of D-glucose which exists in the aqueous solution, open chain form, 5-membered ring (furanose form) and both the anomers for pyranose form (α and β).

In addition to the classically defined monosaccharides, naturally occurring aminosugars are known, for example D-glucosamine, D-galactosamine and D-mannosamine, where the 2-OH group has been replaced with a -NH_2 functionality (**Figure 1-3a**). In nature, these sugars are commonly found as the acetamido derivative (*N*-acetyl), for example 2-acetamido-2-deoxy-D-galactose (**1.3b**). Other modifications to the hydroxyl groups of carbohydrate that are often encountered include inorganic esters such as nitrate, phosphate or sulphate.

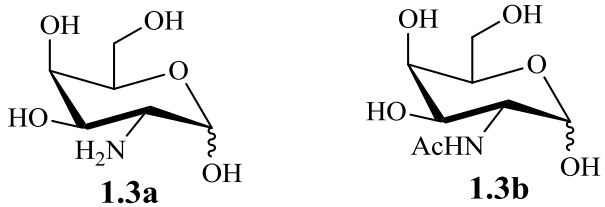
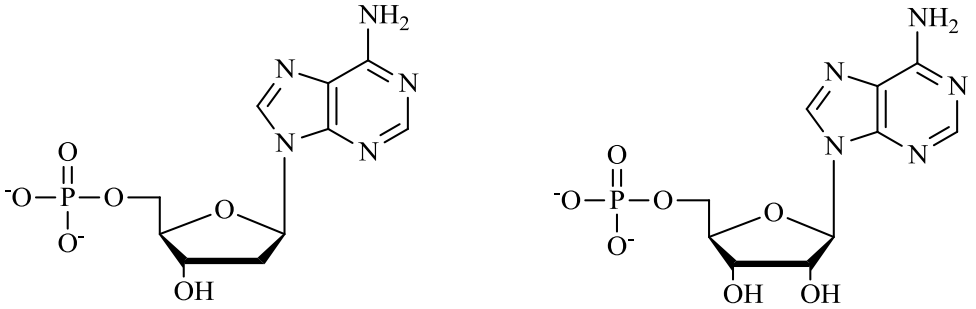


Figure 1-3. Structures of naturally occurring D-galactosamine **1.3a** and its *N*-acetylated amide form 2-acetamido-2-deoxy-D-galactose (GalNAc) **1.3b**.

All sugars that contain an unbranched carbon chain are commonly referred to as either one of these; trioses (3 carbon unit), tetroses (4 carbons), pentoses (5 carbons) or hexoses (6 carbons). In living organisms, the pentoses – 2-deoxy-D-ribose and D-ribose are part of the key structural component of the nucleic acids DNA (**1.4a**) and RNA (**1.4b**), respectively.



1.4a **1.4b**

Figure 1-4. Structurally important pentoses as key components in deoxyadenosine 5'-monophosphate **1.4a** and adenosine 5'-monophosphate **1.4b** that are constituents of the biological polymers DNA and RNA, respectively.

Composite structures that contain carbohydrates covalently linked to either proteins or lipids are commonly called glycoconjugates. These materials are involved in a multitude of cellular communication and signalling events.(*1*) The majority of glycoconjugates can be sub-classified into those attached to lipids (glycolipids) and those attached to proteins (glycoproteins). Glycoproteins are further categorized into two subtypes (*N*-linked or *O*-linked), a classification system that is based on the linkage between the carbohydrate and the peptide backbone. Specifically, the linkage is either to; a) the side chain nitrogen atom of an asparagine (Asn) (*N*-linked glycoproteins) **1.5a**; or to b) the oxygen atom of hydroxyl side chain from either a serine (Ser) or a threonine (Thr) residue (*O*-linked glycoproteins) **1.5b**.

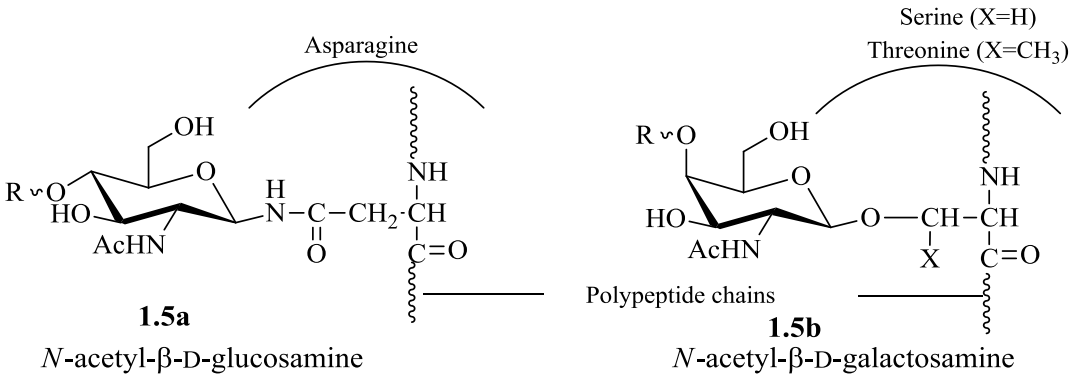


Figure 1-5. General structures of the two types of glycoproteins, linked through the amide side chain of an asparagine (*N*-linked glycoproteins) **1.5a** or to the oxygen atom of a serine or threonine residue (*O*-linked glycoprotein) **1.5b**, where R= oligosaccharide chain.

The sugar components of glycoconjugates - called glycans are often comprised of the hexoses in their pyranose form, commonly in their D-configuration. The glycosidic linkages between two monosaccharides can possess either: a) α-configuration or; b) β-configuration. Hexoses which are more frequently found in glycans comprise of D-glucose and its conventional structural modifications, either through single epimerization (D-mannose and D-galactose) or by oxidation of the C-6 CH_2OH (D-glucuronic acid) (**Figure 1-6**).

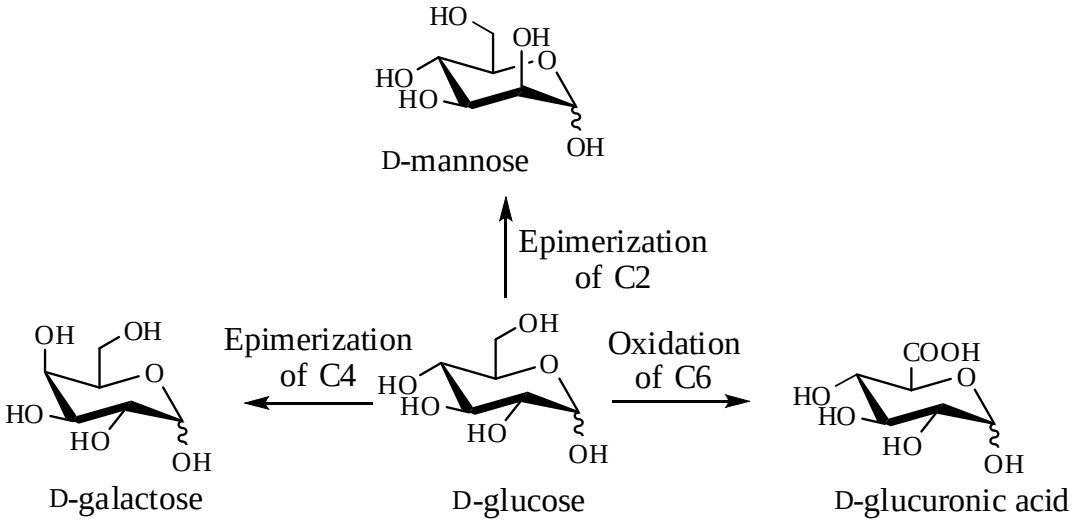


Figure 1-6. Structural representation of some common structural modifications of hexoses (here shown for D-glucose).

The challenging field of glycobiology encompasses the understanding of the various relationships between the biologically relevant carbohydrates and their interactions with proteins and other biomolecules. One of the most widely studied interactions being between that of a carbohydrate and the corresponding enzyme that processes it. The biological functions that primarily involve carbohydrate processing enzymes are: i) catabolism leading to production of energy; ii) linking carbohydrates to and hydrolysing them from oligosaccharides; iii) modification of the various functional groups attached to carbohydrates; and iv) epimerization of various stereogenic centres within the carbohydrate. There are more than 250 different enzymes involved in the biosynthesis of the elaborate oligosaccharide structures found on glycoconjugates, and more than 30 different enzymes participating directly in the synthesis of a single complex *N*-linked oligosaccharide. These enzymes include: i) glycosyltransferases, which are responsible for transferring specific sugar units; and ii) glycosidases or glycosyl hydrolases, which are responsible for trimming specific monosaccharides from glycoconjugates or oligosaccharides (**Figure 1-7**). The two major classes of glycosidases are i) *exo*-glycosidases, which catalyze the step-wise hydrolysis of sugars from the non-reducing termini of carbohydrate polymers; and ii) *endo*-glycosidases, which catalyze the cleavage of internal glycosidic bonds. Understanding the mechanism of action of these glycosyltransferases and glycosyl hydrolases, which are implicated in numerous human ailments such as viral and bacterial infections, has been an on-going area of research in glycobiology.

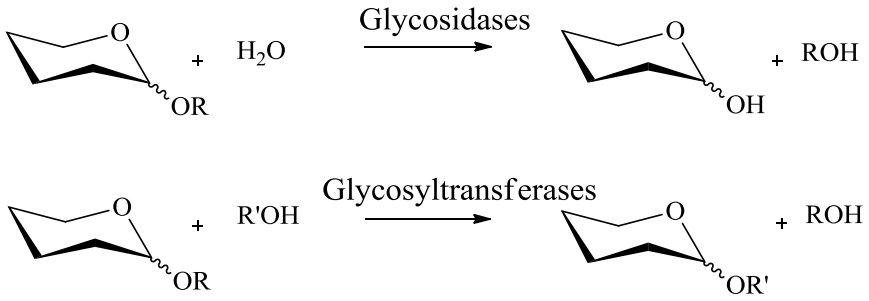


Figure 1-7. General reaction schemes for glycosidases and glycosyltransferases.

A more detailed description of the plethora of biological functions that are mediated by these glycoconjugates is beyond the scope of this thesis, but the primary focus of this

work will be to discuss detailed mechanistic studies on a few select biologically relevant retaining glycosidases.

1.2. D-Galactose

A biologically important hexose that occurrs in mammals is the C-4 epimer of D-glucose - D-galactose. The biological interconversion of galactose and glucose takes place only by Leloir pathway, which requires the three enzymes- galactokinase, galactose-1-P uridylyltransferase, and UDP-galactose 4-epimerase.(*2*) The Leloir pathway illustrating the metabolic pathway for the catabolism of D-galactose is shown in **Figure 1-8**.(*2*)

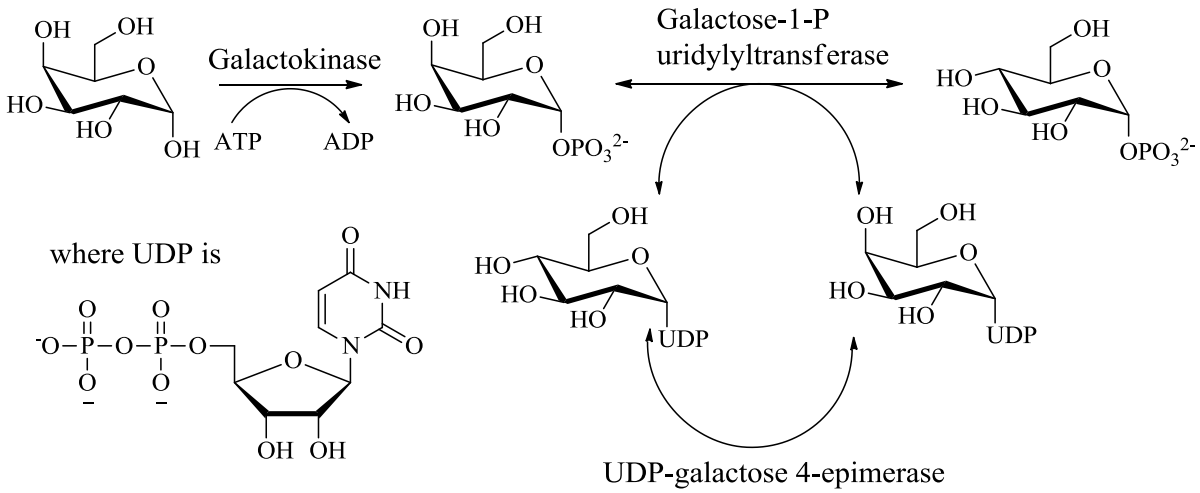
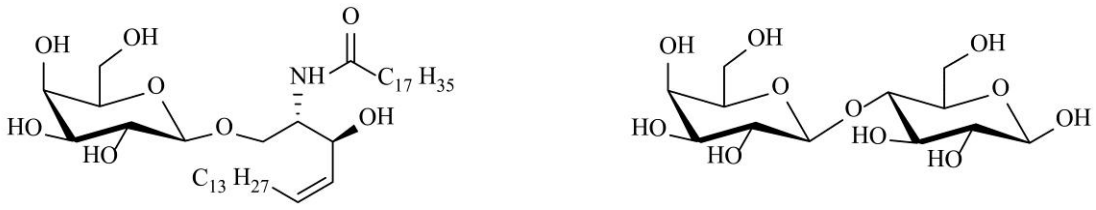


Figure 1-8. The Leloir Pathway for catabolism of D-galactose.(*2*)

Popularly known as the brain sugar,(*3*) D-galactose is the key structural component of cerebrosides (called galactocerebrosides) (**1.9a).** Galactosyl ceramides and their metabolites have been shown to possess important functions in promoting the regulation of nerve cells,(*3, 4*) regulating protein kinase C activities,(*5*) and modulating the function of hormone receptors.(*6*) The disaccharide lactose ((β-D-galactopyranosyl-(1→4)-D-glucose) (**1.9b**), which is a key carbohydrate component in milk products, is another well-known derivative of D-galactose that has a D-glucose residue at the reducing terminus. A genetic metabolic disorder called galactosemia caused by low galactose 1-phosphate uridylyltransferase activity

and this results in inability to break down the galactose formed from the hydrolysis of lactose, a situation that leads to an over-accumulation of galactose in the body.(*7, 8*) Also, blood group antigens demonstrate the biological importance of D-galactose. Blood group B antigen differs from the blood group O antigen by a terminal α-1,3 linked D-galactose moiety (**Figure 1-10**).(*9*)



1.9a β-D-galactosyl ceramide **1.9b** lactose (β-D-galactopyranosyl-(1→4)-D-glucopyranose

Figure 1-9. Importance of D-galactose in biological systems; a) galactocerebrosides **1.9a**, and b) milk sugar-lactose **1.9b**.

One of the most important derivatives of D-galactose is 2-amino-2-deoxy-D-galactose **1.3a.** Structurally, this amino sugar is a constituent of hormones such as follicle stimulating hormone (FSH) and luteinizing hormone (LH). However, the predominant form that occurs in nature is the *N*-acetylated derivative, which is abbreviated as GalNAc **1.3b**. The terminal sugar component on the blood types A antigen is a 1,3-linked 2-acetamido-2-deoxy-α-D-galactose residue, the immunodeterminant for blood group A specificity (**Figure 1-10**).(*10, 11*)

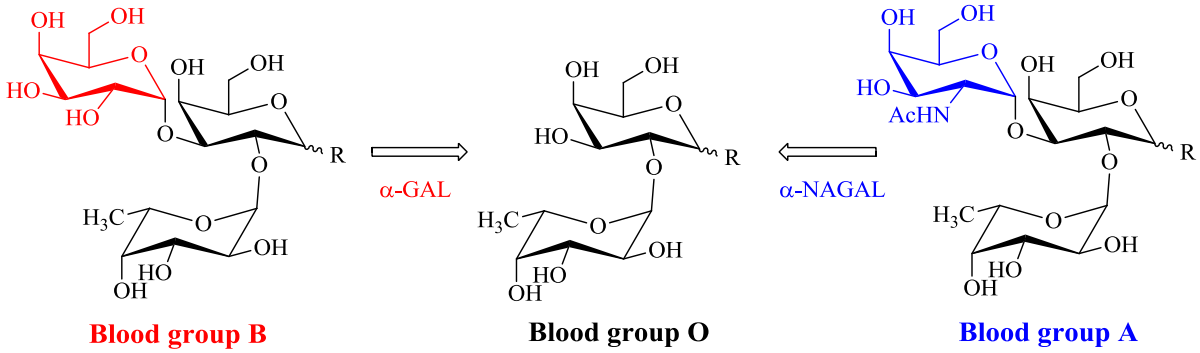


Figure 1-10. (ABO) blood group specificity determined by its terminal sugar moiety and the corresponding *exo*-glycosidases which can potentially effect the conversion to "universal donor" blood group O, α-galactosidase (α-GAL; Blood group B to O conversion) and α-*N*-acetylgalactosaminidase (α-NAGAL; blood group A to O conversion).(*11, 12*)

Importantly, in one form of glycoproteins called mucins, *O*-glycans (**Figure 1-5b)** are covalently *α*-linked via an *N*-acetylgalactosamine (GalNAc) moiety to the -OH of a serine or a threonine residue by an *O*-glycosidic bond to give the core structure that is commonly referred to as Tn antigen (**Figure 1-11**).(*13*) Mucin glycoproteins are ubiquitous in mucous secretions that can be found on cell surfaces and in body fluids. Compared to *O*-glycans from normal mucins, cancer associated *O*-glycans are often truncated and commonly contain the Tn and T (Core 1, Gal β-(1→3)-GalNAcα1-Ser/Thr) antigens and their sialylated version.(*13, 14*) More specifically, these mucin *O*-linked antigens like, T, Tn and sialosyl-Tn antigens have been shown to be the onco-developmental cancer-associated antigens in the colon.(*14, 15*)

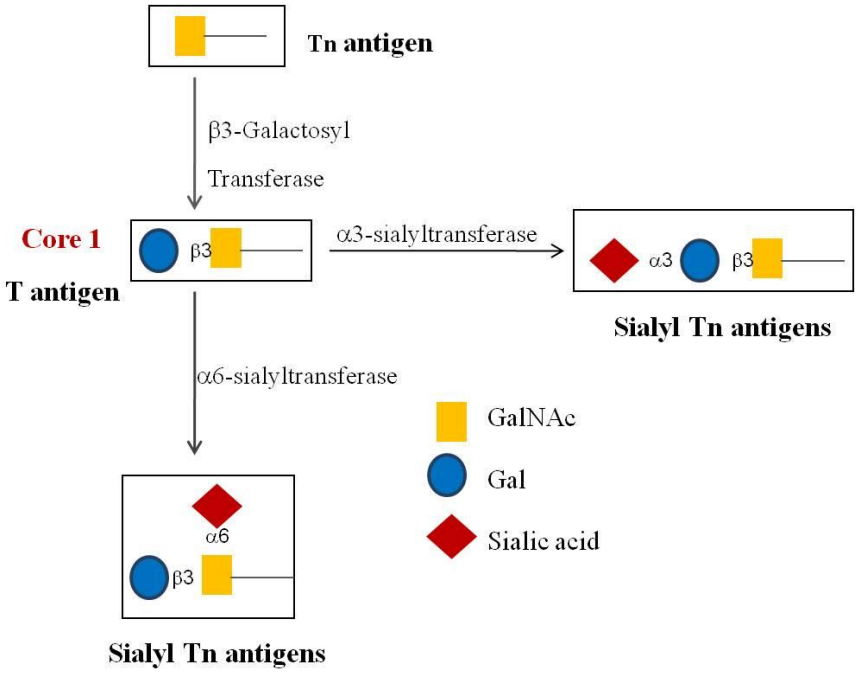


Figure 1-11. Mucin *O*-linked glycoproteins; Biosynthesis of T antigens and the role of GalNAc in the core building block of *O*-glycans.(*15*)

1.3. Galactosidases

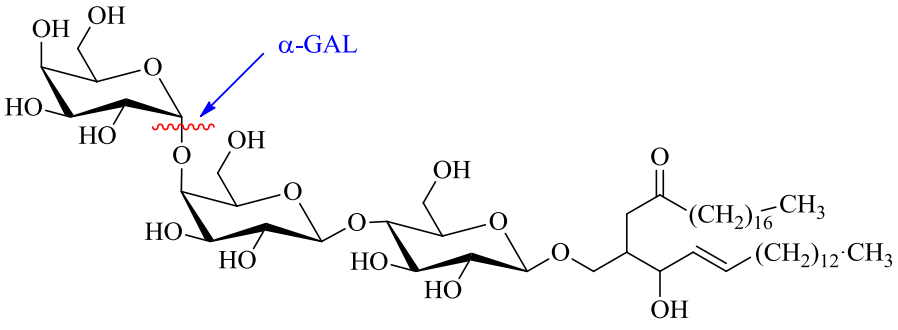
The class of enzymes that catalyze the cleavage of terminal galactopyranosyl residues from various glycoconjugates are called *exo*-galactosidases. These enzymes can be divided in two subgroups based on their specificity towards the two anomeric forms of galactosides, that is: α-galactosidase (EC 3.2.1.22) and β-galactosidase (EC 3.2.1.23). Both, α- and β-

galactosidases are widespread in microorganisms, animals and plants and they often exhibit pharmacologically significant roles. In this thesis, we will be focussing primarily on the mechanistic evaluations of the two pharmaceutically important α-processing enzyme families: α-galactosidase and α-*N*-acetylgalactosaminidase. Interestingly, in the human genome, the *NAGA* gene (encoding α-*N*-acetylgalactosaminidase, EC 3.2.1.49) is closely associated with the *GLA* gene (encoding α-galactosidase A, EC 3.2.1.22).(*16*) These two genes are believed to have evolved from the same ancestral precursor.(*17*) The corresponding proteins, α-*N*-acetylgalactosaminidase (α-NAGAL) and α-galactosidase (α-GAL) share 46% amino acid sequence identity.(*16*)

1.4. α-Galactosidase (EC 3.2.1.22)

Glycosyl hydrolases (GH) are currently classified into approximately 130 different families based on protein sequence and structure.(*18*) The evolutionarily, structurally, and mechanistically related families are grouped together in a higher hierarchical level termed "clans".(*19*) As detailed in the CAZy (**C**arbohydrate **A**ctive en**Zy**mes) database, EC 3.2.1.22 encompasses α-galactosidases from GH families-4, 27, 36, 57, 97 and 110. Human α-galactosidase A (α-GAL), which belongs to GH27, is a lysosomal enzyme that catalyzes the hydrolysis of terminal α-D-galactopyranosyl residues (**1.12**) from various glycoconjugates in the lysosome, including glycoproteins, glycolipids and polysaccharides.(*20*) The 1.45 Kb α-galactosidase A gene (*GLA A*) is localized in the X-chromosome region and encodes a glycoprotein of 429 amino acids.(*21*) In humans, inherited mutations in this *GLA A* gene elicits a lysosomal storage disorder (LSD), called Fabry's disease.(*22, 23*) The molecular basis for two types of Fabry's disease has been identified and these involve: a) mutations in the active site of the protein; or b) mutations that reduce the kinetic stability of the folded protein, thereby disrupting the trafficking to the lysosome.(*20*) Furthermore, a deficiency of active catabolic enzymes (α-galactosidase in this case) in the lysosome could be an outcome of enzyme instability either in the acidic conditions of lysosome (~pH 5) or in the more neutral cellular conditions before it is trafficked to the lysosome. As a consequence of such an insufficiency of active enzyme, the over-accumulation of unprocessed substrates

(primarily globotriaosylceramide, GL3, **Figure 1-12**) in the cells lining blood vessels and in the kidneys, heart, and nervous system leads to an impairment of their proper functioning.(*22, 24*) Fabry's and Gaucher's (a deficiency of β-glucosidase) diseases are amongst the most well-known lysosomal storage disorders.(*23, 25*) Of note, deficiency of lysosomal α-NAGAL (EC 3.2.1.49) in humans, which is responsible for the degradation of glycolipids, glycoproteins and oligosaccharides, also leads to the lysosomal storage disorders known as Schindler and Kanzaki diseases.(*26, 27*) In Schindler disease, over accumulation of glycopeptides/glycoproteins leads to neurological, skin and cardiac disorders.(*28, 29*) While there is no currently available treatment for Schindler disease, the only FDA approved treatment for Fabry patients is enzyme replacement therapy (ERT) in which recombinant human α-GAL, produced in engineered human cell lines, is intravenously administered as a means to augment the abnormally low α-GAL activity.(*30*) However, this treatment option is not ideal due to the following reasons: a) the large quantity of recombinant enzyme needed and the associated high production costs; and b) approximately 88% of the patients develop an immune response to the administered recombinant enzyme.(*31, 32*) More recently, an alternative approach to cope with the immune response in ERT has been proposed by Tomasic *et al.*, based on protein engineering.(*33*) Enzymatic specificity of the two human lysosomal enzymes (α-GAL and α-NAGAL, both from GH27) has been altered by replacement of active site residues. The newly designed enzymes have been proposed as non-immunogenic alternatives in ERT for Fabry and Schindler diseases.(*33*)

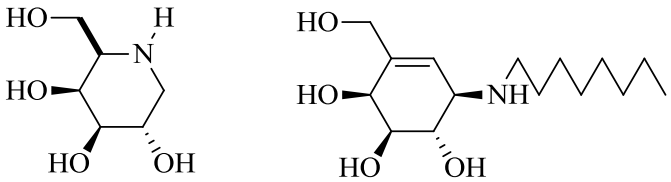


Globotriaosylceramide (GL3)

(The shingolipid which primarily accumulates in a Fabry patient)

Figure 1-12. Role of α-GAL in Fabry disease, in which a X-chromosome linked genetic defect causes the loss of functional α-GAL resulting in over-accumulation of neutral substrates in lysosomes.(*20*)

A new approach to the design of therapeutics for Fabry disease uses 'chemical chaperones', a stratagem that was conceptualized by Sato. S. *et al.*(*34*) and later on successfully implemented by Fan *et al.* in 1999.(*35*) Chemical chaperone therapy has two major advantages over enzyme replacement therapy, and they are: a) oral administration; and b) accessibility to the brain. Because these chemical chaperones are low molecular weight molecules, it is anticipated that they may have better oral availability with broad bio-distribution, including the central nervous system (CNS).(*34, 35*) The use of potent competitive inhibitor of α-GAL A, 1-deoxy-galactonojirimycin **1.13a** (IC_{50} of 4.7 nm with 4-nitrophenyl α-D-galactoside) as a 'chemical chaperone' has shown to elevate the α-GAL A activity in Fabry lymphoblasts.(*35*) Such pharmacological chaperones selectively bind and stabilize mutant forms of α-Gal A in the endoplasmic reticulum, facilitating proper protein folding and trafficking through the Golgi apparatus, resulting in increasing lysosomal enzyme activity.(*35*) Four years later in 2003, motivated by the same concept, Matsuda *et al.* designed and tested a galactose derivative, *N*-octyl-4-epi-β-valienamine **1.13b**, for a chemical chaperone therapy of a human neurogenetic disease, β-galactosidosis (GM1-gangliosidosis and Morquio B disease).(*36*) **1.13b** has been shown to be a potent inhibitor of lysosomal β-galactosidase *in vitro*.(*36*)



1.13a **1.13b**

Figure 1-13. Structures of selected inhibitors, where 1-deoxy-galactonojirimycin **1.13a**(*35*) was the first inhibitor tested as a chemical chaperone therapy to treat α-galactosidase A deficiency and *N*-octyl-4-epi-β-valienamine **1.13b** was shown to be a potent inhibitor of lysosomal β-galactosidase.(*36*)

α-GAL is also one of the most studied enzymes in blood group conversions.(*37*) That is, removal of the terminal D-galactose or *N*-acetyl D-galactosamine would potentially eliminate the complications and risk associated with ABO-incompatible blood transfusions by creating a universal blood group. ABO blood group specificity is determined by the terminal carbohydrate moiety linked to the oligosaccharides **Figure 1-10**. The continued

search for efficient *exo*-glycosidases which possess: a) high substrate specificity; b) neutral pH optimum; and c) high efficiency, led to the identification of GH109 and GH110 families of hydrolases, which reinvigorated this field of research.(*11, 38*) In 1982, Goldstein and co-workers pioneered the concept of blood group interconversions by using *coffee bean* α-galactosidase to covert B erythrocytes to blood group O cells, without compromising on their ability to survive post- transfusion.(*37*) However, because this enzyme has poor kinetic properties and low pH optimum, the process was not economically viable.(*37*) Twenty five years later, in order to improve upon Goldstein's research, Liu *et al.* initiated a search for α-galactosidase possessing improved efficiency with the result that an enzyme in a new GH family (GH110), was discovered in 2007.(*11, 38*) Moreover, these enzymes could be expressed in *E. coli* in high yield (1 mg/L).(*11*) The efficiency of GH110 α-GAL, in converting one unit of RBC (blood group B to O), has been demonstrated to be ~1000 fold(*11*) (2 mg recombinant enzyme required for 1 unit of blood) higher than *coffee bean* α-galactosidase.(*37*) This conversion was achieved after 60 min incubation in a buffer system containing 200 mM glycine (pH 6.8) and 3 mM NaCl at 26 °C (**Figure 1-10**).(*11*)

1.4.1. Active site architecture of α-D-galactosidases

According to the glycosyl hydrolase enzyme classification, which is based on amino acid sequences (CAZy database, http://www.cazy.org), α-D-galactosidases (α-GAL) belong to hydrolase families 4,(*39*) 27,(*40*) 36,(*41*) 57,(*42*) (retaining glycosidases) and 110 (inverting glycosidases),(*38*) with the exception of GH97 which contains both retaining and inverting glycosidases.(*43*) α-D-Galactosidases from eukaryotic organisms are grouped predominantly into family GH27, whereas those from microbial sources are grouped into families GH4, GH36, GH57 and GH97. Both mechanistic and structural evaluations of α-GAL enzymes from families GH27(*44*) and GH36(*41*) have been performed. Prior to this thesis, no reports have been submitted for mechanistic elucidation of α-GAL from GH4 hydrolases. Herein, we report the detailed characterization of GH4 α-GAL from *Citrobacter freundii* (Chapter 2), focussing on kinetic isotope effects obtained using various substrate isotopologues.(*45*) The X-ray single crystal diffraction structures for the α-GAL from *T. maritima* (GH36) [PDB 1ZY9](*41*) and human (GH27) [PDB 3H53](*21*) indicate that these

enzymes function by use of a classical retaining catalytic mechanism that was originally proposed by Koshland.(*46*) All of these crystal structures contain an N-terminal $(\beta/\alpha)_8$ barrel and a C-terminal domain with antiparallel β strands. In each structure, the catalytic domain is at the C terminal ends of the β strands in the $(\beta/\alpha)_8$ barrel.(*21, 41*) Two aspartate residues, D170 (the catalytic nucleophile) and D231 (the general acid/base residue) have been identified as the important catalytic residues for GH27 human α-GAL.(*21*) Similarly, D327 and D387 have been designated as the catalytic nucleophile and general acid/base residue, respectively, for *T. maritima* GH36 α-GAL.(*41*) The overall hydrolysis reaction proceeds in two sequential steps, each involving an inversion of stereochemistry at the anomeric centre such that the reaction proceeds with net retention of stereochemistry (**Figure 1-15**). In retaining glycosidases, the distance between the two catalytic carboxylate residues is typically between 5-6 Å, while for inverting glycosidases the same distance increases to 9–11 Å.(*47*) In *T. maritima* GH36 α-GAL, the X-ray crystal structures have revealed that the catalytic residues (D327 and D387) are separated by approximately 6.8 ± 0.5 Å, which is consistent with an average separation for GH27 human α-GAL catalytic residues of 6.5 Å.(*21, 41*) Of particular interest is the α-GAL from GH4 family of hydrolases, which operates via an oxidation-reduction mechanism that involves two cofactors, NAD^+ and Mn^{2+}, which are intimately involved in the mechanism.(*45*) A detailed analysis of GH4 α-GAL from *Citrobacter freundii* will be discussed in detail in Chapter 2.

1.4.2. α-*N*-Acetylgalactosaminidase (EC 3.2.1.49)

The enzyme α-*N*-acetylgalactosaminidase (α-NAGAL) removes terminal α-D-GalNAc (**1.14b**) monosaccharide residues from glycoconjugates, which are primarily *O*-linked sugars that are attached to serine and threonine residues. According to the CAZy database, these enzymes are found in three families- GH27, 36 and 109. Among these, GH27 and 36 are further grouped into Clan D, on the basis of their protein folding $(\beta/\alpha)_8$) and mechanism of action.(*48*) The two most important and widely studied roles of α-NAGAL are the following: a) in lysosomal storage disorders (GH27, human α-NAGAL);(*49*) and in b) hydrolysis of blood group A immunodeterminants from erythrocytes **Figure 1-10** (both GH27 and GH109 (bacterial α-NAGAL)).(*10, 11*) The immunodominant monosaccharide

that determines blood group A specificity is a terminal α-1,3-linked *N*-acetyl-D-galactosamine (GalNAc, **1.14b**). Hence, the *exo*-glycosidase that can catalyze the substrate specific cleavage of the A antigen to convert it to the H (O) antigen is α-NAGAL.(*10*) To date, α-NAGAL has been isolated from eukaryotic sources like *homo sapiens* (GH27),(*49*) bovine,(*50*) porcine liver,(*51*) snail(*52*) and starfish(*53*) to name a few and also from prokaryotic sources, such as *C. perfringens*(*54*) and *E. meningosepticum* GH109.(*11*) Nevertheless, most of the α-NAGAL enzymes isolated from these eukaryotic sources possess residual α-GAL activity. This observation unveils the only difference in the active site of α-GAL and α-NAGAL. In the crystal structure of GH27 human α-GAL,(*21*) the residues near the 2-OH on the substrate include larger sized glutamate and leucine, whereas in crystal structure of GH27 human α-NAGAL,(*49*) the larger substitution 2-*N*-acetyl group on the substrate interacts with two smaller residues, serine and alanine in this case.(*49*) Interestingly, the α-NAGAL isolated from bacteria *C. perfringens*, not only exhibited strict substrate specificity (α-GalNAc) but also required reducing conditions (DTT) for its catalytic activity.(*54*) GH109 is the only other bacterial family reported to date, which encodes for α-NAGAL enzymes.(*11*) Of note, these enzymes belonging to GH109 have been shown to modify efficiently the blood group A antigen under optimum physiological conditions.(*11*) The crucial and thought provoking observation from the X-ray crystal structure of GH109 α-NAGAL from *E. meningosepticum* [PDB 2IXA], was the detection of NAD^+ in the enzyme active site.(*11*) Another known family of hydrolases that requires NAD^+ for catalytic activity are the GH4 glycosidases.(*39, 55*) The sequence-based relationship with oxido-reductases,(*11, 56*) and the proven dependence on three important cofactors for catalysis - NAD^+, Mn^{2+} and reducing conditions (DTT/ TCEP), together made the GH4 glycosidases mechanistically conspicuous.(*45, 57, 58*) On the other hand, deficiency in active GH27 human lysosomal hydrolase, α-NAGAL, leads to lysosomal storage disorders that are called Schindler and Kanzaki diseases.(*26, 28, 59*) Due to the dearth of active α-NAGAL in lysosome, over accumulation of glycolipids and glycopeptides results in wide range of symptoms, including neurological, skin and cardiac abnormalities.(*26, 30, 60*)

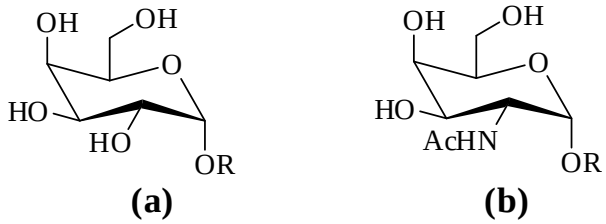


Figure 1-14. a) α-D-Galactopyranoside residue and b) 2-acetamido-2-deoxy-α-D-galactopyranoside residue.

1.4.2.1. α-NAGAL and lysosomal storage disorder

Human lysosomal α-NAGAL from GH27 is a retaining enzyme that operates by employing a standard double-displacement mechanism that has attracted considerable interest given that it is a lysosomal hydrolase. Importantly, deficiency of lysosomal α-NAGAL (EC 3.2.1.49) in humans, which is responsible for the degradation of glycolipids, glycoproteins and oligosaccharides, leads to the lysosomal storage disorders known as Schindler and Kanzaki diseases.(*26, 27*) In Schindler disease, over accumulation of glycopeptides/glycoproteins leads to neurological, skin and cardiac disorders.(*28, 29*) Kanzaki disease, often referred to as type II Schindler disease, displays mild cognitive impairments and a characteristic skin lesion, angiokeratoma.(*27, 60*) Based on the recent characterization of substrate accumulation that results from α-NAGAL deficiency in the lysosomes, instead of the expected α-D-GalNAc containing saccharides, the isolated substrates have been shown to contain sialic-acid- and galactose-terminal saccharides) (**Figure1-11**).(*29*) The deficiency of active α-GAL, which originates from either: a) mutations to the enzymatic active site; or b) protein misfolding, leads to another lysosomal storage disorder, named Fabry's disease.(*22*) It is an X-linked chromosomal disorder affecting one in 40,000 males with premature myocardial infarction, strokes and renal failure.(*22, 24*) Schindler disease is a relatively rare disorder because mutations in both copies of the autosomal *NAGA* gene are required before symptoms develop (recessive) as opposed to Fabry's disease where only a single mutation is needed in the X-chromosomal *GLA* gene.(*28, 59*) The only FDA approved therapy to treat Fabry's disease is enzyme replacement therapy, while there is no currently available treatment for Schindler disease.(*30*) A major problem with enzyme replacement therapy is that an immune response is developed in 88% patients against the injected glycoprotein.(*28*) Notably, Tomasic *et al.*

successfully employed a protein engineering approach in order to convert the enzymatic specificity of α-NAGAL to α-GAL, where the mutant protein possess a similar structure to those produced in the patient for hydrolysis of α-D-GalNAc residues, an approach that is expected to lead to non-immunogenic alternatives for enzyme replacement therapy in Fabry's patients.(*33*)

1.4.2.2. Active site architecture of α-*N*-acetylgalactosaminidase

X-ray crystal structures have been reported for α-NAGAL from: a) chicken (GH27) [PDB IKTB];(*16*) b) human (GH27) [PDB 3H53];(*49*) and c) bacteria (GH109) [PDB 2IXA].(*11*) The active sites have been shown to be remarkably conserved in both the α-NAGAL enzymes from GH27. That is, in the active site of chicken and human α-NAGAL (GH27), two catalytically important aspartate residues have been located, and these are the catalytic nucleophile and the general acid/base. In chicken α-NAGAL, D140 (catalytic nucleophile) and D201 (general acid/base) are poised on the opposite sides of the glycosidic linkage at a distance of 6.5 Å,(*16*) a distance that is slightly larger than in most retaining glycosidases (~5.5 Å). These enzymes operate by a standard double-displacement mechanism **Figure 1-15**, where two successive nucleophilic attacks occur at the α-anomeric centre to result in an overall retention of anomeric configuration.(*61*) Notably, the crystal structure of human α-NAGAL (GH27), when soaked with the substrate or the product (α-GalNAc), showed that a binding conformational transition had occurred from low energy chair conformation to a higher energy $^{1}S_{3}$ skew boat.(*49*) The corresponding ligand binding studies with α-D-GalNAc showed the characteristic binding residues to be- S188, A191, and R213, with the *N*-acetyl group at C-2 **1.14b.**(*16*) Interestingly, although the crystal structure of GH109 α-NAGAL from *E. meningosepticum*(*11*) exhibited a bound NAD^{+} in the active site (like GH4 hydrolases), it did not show a metal binding site. Furthermore, the experimental conditions in which the enzyme was active indicated that reducing conditions were not necessary for maintaining the catalytic activity (unlike GH4 glycosidases).(*11, 39*) Chapter 4 details a mechanistic evaluations of the GH109 α-NAGAL enzyme from *E meningosepticum*.

1.5. Catalytic mechanism of classical glycosyl hydrolases

Based on the anomeric stereochemistry of the substrate and product, glycosidases can be categorized into retaining (**Figure 1-15**) or inverting (**Figure 1-16**) enzymes. Classically, the glycosidase function by way of either a double-displacement or a single displacement mechanism, originally outlined by Koshland in 1953,(*46*) a proposal that is now universally accepted.(*18, 61*) With no known exception to date, the molecular mechanism appears conserved among the members of a given sequence-based family.(*19, 48*)

1.5.1. Retaining glycosidases

For the retaining glycosidases, three different mechanisms of catalysis have been elucidated, and these are: a) nucleophilic catalysis using an enzyme active site residue;(*42, 62*) b) NAD^+ dependent oxidation/reduction catalysis;(*39, 63*) and c) intramolecular neighbouring group assisted catalysis using an acetamido group.(*64*) For the first mechanism, catalysis is typically provided by two active site carboxylate groups- aspartate and glutamate.(*61*) The main exception is for *exo*-sialidases(*65, 66*) and *trans*-sialidases(*67, 68*) where an active site tyrosine is the catalytic nucleophile. This double-displacement or ping-pong mechanism entails two steps.(*46*) In the first step, one of the two catalytic carboxylic acids acts as a general acid catalyst (GAC) to assist departure of the aglycone while nucleophilic attack by the other carboxylate occurs at the anomeric centre leading to the formation of a covalently bound glycosyl-enzyme intermediate. Next, the acid/base carboxylate residue now acts as a general base catalyst (GBC) to deprotonate a nucleophilic water molecule as it attacks the glycosyl-enzyme intermediate to give the final hydrolyzed product(*46*) (as shown in **Figure 1-15**). The transition state in both steps involves considerable positive charge development on the carbohydrate residue and this is often referred to as oxacarbenium ion-like character. The distance between the general acid/base residue and the catalytic nucleophilic residue in retaining glycosidases have been found to be in the range of 5.5-6.5 Å, fairly consistently.(*69*)

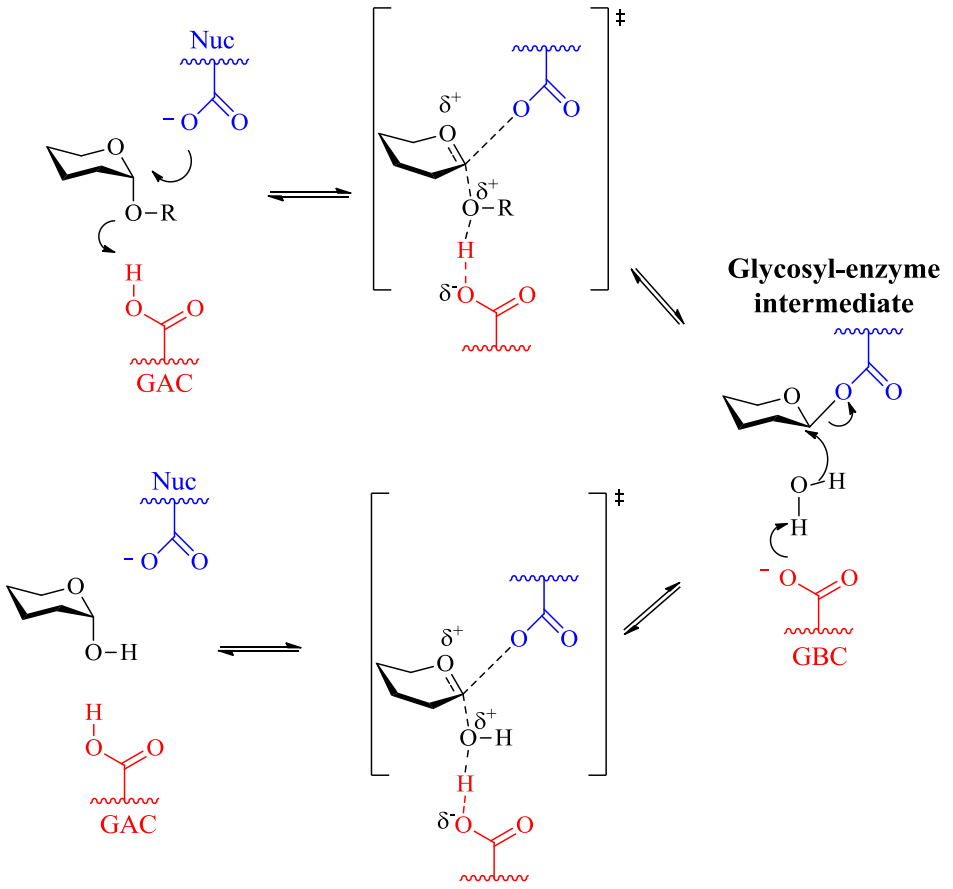


Figure 1-15. Double-displacement mechanism for retaining glycosidases involving two active site carboxylate (aspartate or glutamate) residues.(*69*)

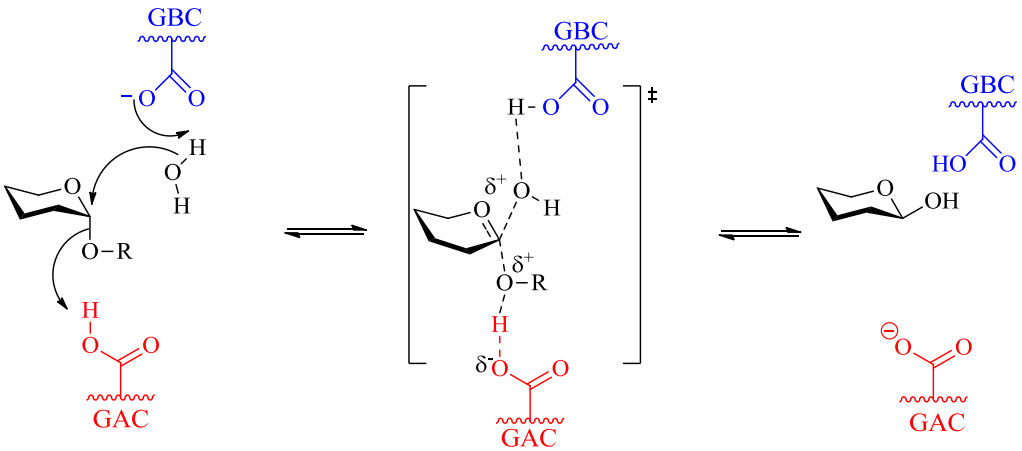


Figure 1-16. Catalytic mechanism for inverting glycosyl hydrolases showing two carboxylate residues in the enzyme active site acting as GAC and GBC.(*61*)

In 1988, Withers *et al.* reported the first trapping of the glycosyl-enzyme intermediate using a fluorine substituted inhibitor, 2-deoxy-2-fluoro β-D-glucopyranosyl fluoride, to inactivate β-glucosidase from *Alcaligenes faecalis*.(*70*) Substitution of the hydroxyl (OH) group by a fluorine atom at C-2 leads to two key alterations, and they are: a) destabilization of the partial positively charged oxacarbenium ion-like transition state by the electron withdrawing fluorine atom; and b) removal of a key hydrogen bonding interaction existing between the 2-OH in the enzymatic transition state.(*70*) These two factors result in the accumulation of the glycosyl-enzyme intermediate, which was further affirmed by the following experimental techniques: i) electrospray ionization mass spectroscopy; and ii) ^{19}F nuclear magnetic resonance spectroscopy (^{19}F NMR).

A fundamentally different glycosidase mechanism has been unveiled through structure determination and kinetic analyses of the NAD^+ and divalent metal ion dependent glycosidases. The mechanism of these Family 4 hydrolases involves oxidation of 3-OH group, followed by an $E1_{CB}$ elimination reaction.(*45, 56, 63*) Hence, compounds that are designed to mimic the glycopyranosylium-ion like transition state are not expected to inhibit these enzymes. This family of hydrolases include both α- and β-glycosidases.(*56, 57*) The mechanistic scheme is shown in **Figure 1-17.**

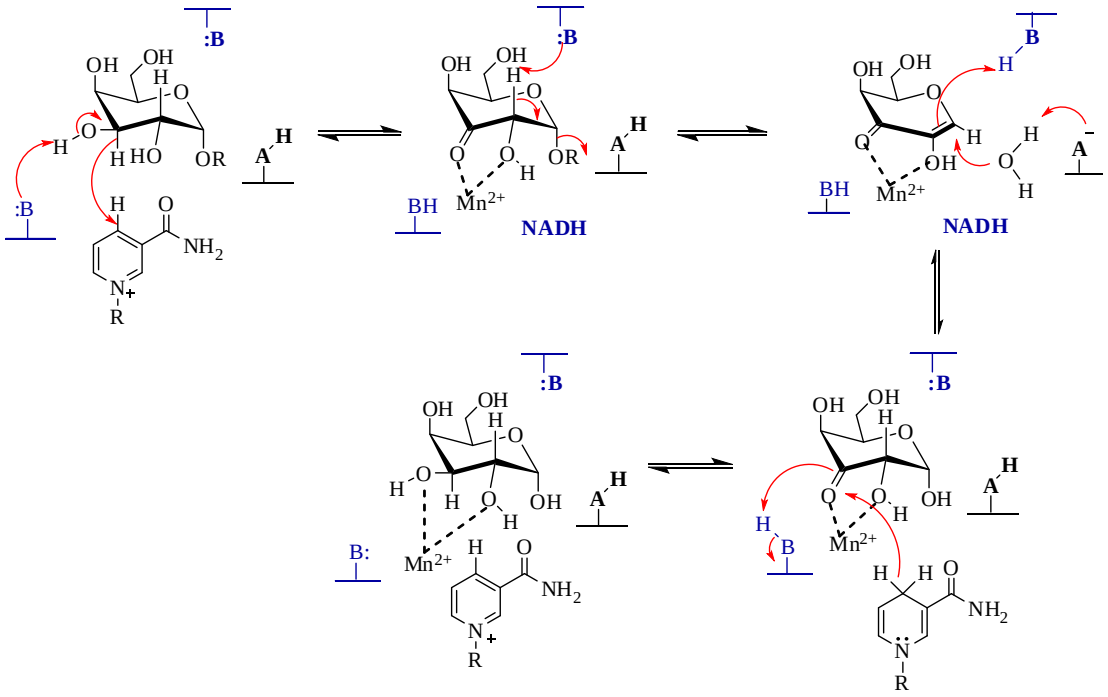


Figure 1-17. General scheme for the catalytic mechanism of GH4 family of hydrolases, involving NAD^+ and a divalent metal ion, Mn^{2+}.(*45*)

Substrate assisted catalysis, the catalytic mechanism for glycosyl hydrolases from GH 18,(*71*) 20,(*72*) 56(*73*) and 84(*74*) involves anchimeric assistance from the 2-acetamido group of the substrate. The first step of the reaction involves a cyclization that occurs *via* attack of the 2-acetamido carbonyl oxygen on the anomeric centre to form a covalently linked bicyclic oxazoline intermediate. Two correctly positioned carboxylate groups act as: a) general base catalyst –assisting in polarization of the 2-acetamido group; and b) general acid catalyst-assisting in departure of the aglycone. In the next step, ring opening of the oxazoline intermediate proceeds by a general base catalyzed attack of a water molecule on the anomeric centre with the general acid catalyzed ring-opening to regenerate the amide group (**Figure 1-18**).

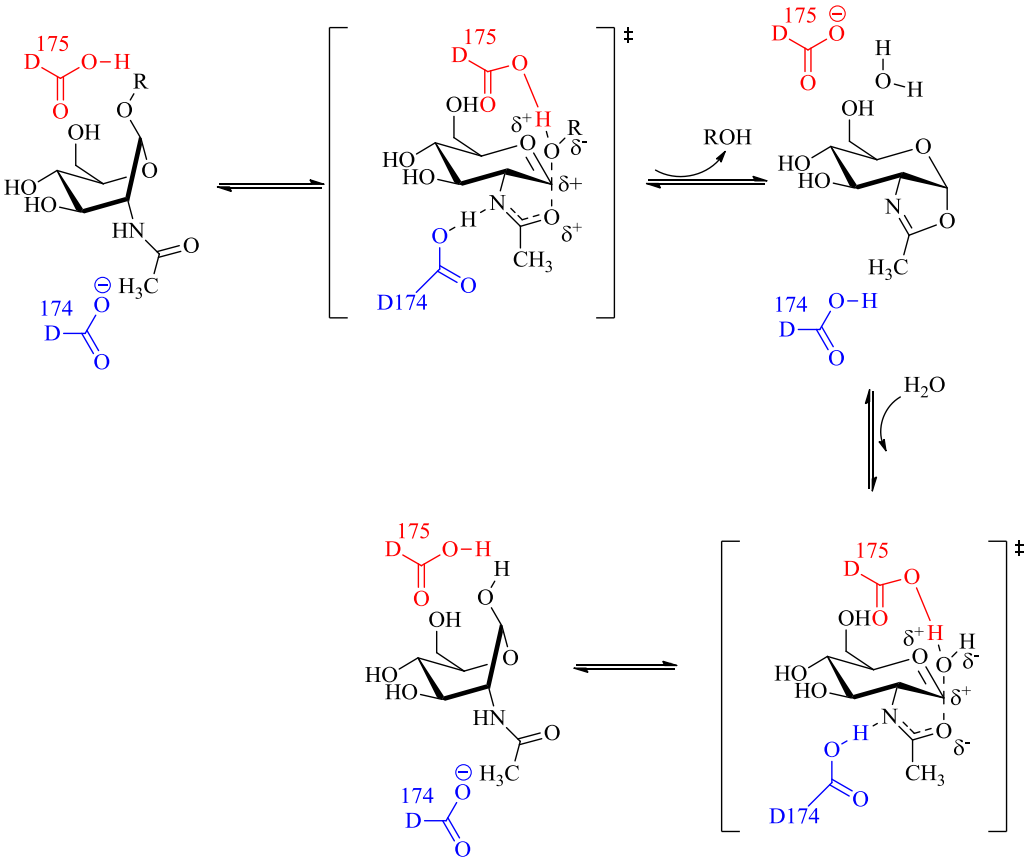


Figure 1-18. The catalytic mechanism of GHs using substrate assisted catalysis, clearly showing the transient oxazoline ring formation. The two strategically positioned carboxyl residues, on opposite sides of the glycosidic linkage, play the key catalytic roles, firstly, by polarising the 2-acetamido group on the substrate to facilitate the nucleophilic attack by the carbonyl oxygen at the anomeric centre and secondly, by sequentially acting as GAC/GBC.(*64*, *74*)

1.5.2. Inverting glycosidases

In inverting glycosidases, the two carboxylate residues (acting as general acid/base catalyst) are poised at a distance of 9-10 Å, on the opposite sides of the glycosidic linkage, to accommodate both a water molecule and the substrate into the active site for the catalytic mechanism (**Figure 1-16**).(*69*) Unlike retaining enzymes, which operates *via* a double-displacement mechanism, inverting glycosidases operate *via* a single-displacement mechanism, through an oxacarbenium ion-like transition state (**Figure 1-16**).(*61*, *69*) One of the carboxylate residue acts as a GAC to protonate the aglycone moiety. Concomitantly, the nucleophilic water molecule attacks the anomeric centre, from the opposite face to that occupied by the aglycone, with assistance from the carboxylate residue that acts as a GBC.

As a result, the hydrolyzed product has an inverted stereochemistry at the anomeric centre (**Figure 1-16**).

An impressive piece of evidence for the inverting mechanism was procured from the work of Hehre using the "wrong" glycosyl fluoride as substrates.(*75*) In 1979, Hehre *et al.* reported that β-amylase hydrolyzed β-D-maltosyl fluoride, the opposite anomer of the normal glycosidic linkage hydrolyzed, into maltose and fluoride ions.(*76*) **Figure 1-19** illustrates an example showing the hydrolysis of α-D-cellobiosyl fluoride by inverting Cel6 enzymes following Hehre mechanism.(*77*) Later on, various inverting enzymes were shown to hydrolyze the "wrong" glycosyl fluorides(*78-80*) suggesting that the reaction is common among inverting GHs (**Figure 1-19**). Instead of simple hydrolysis, this reaction mechanism consists of two steps. In the first step, a new glycoside of the correct anomer is formed from the wrong glycosyl fluoride and an acceptor by Walden inversion. The new glycoside is then hydrolyzed with anomeric inversion at the same site on the enzyme before it is released from the active centre, which is the normal reaction of an inverting GH.(*76, 81, 82*)

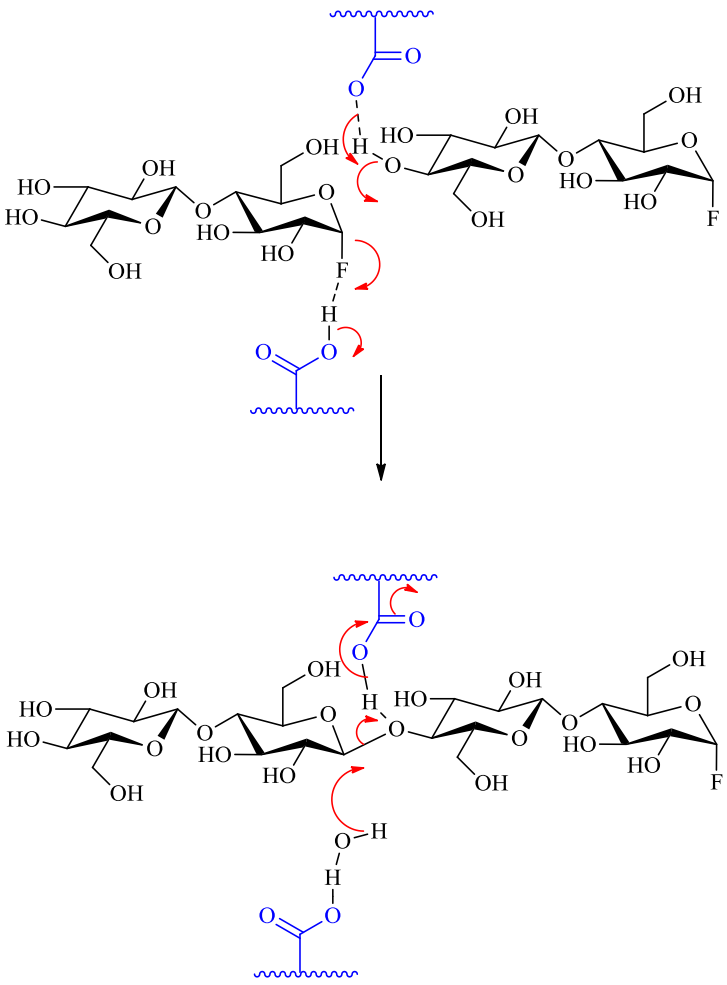
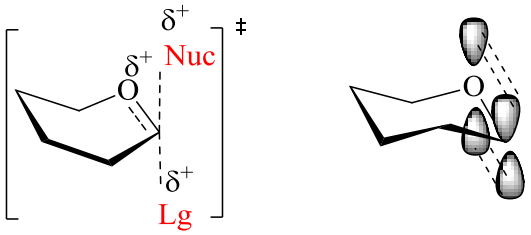


Figure 1-19. Hehre mechanism for the transformation of α-D-cellobiosyl fluoride by Cel6 enzymes (cellobiohydrolase II enzymes) from *Humicola insolens* and *Trichoderma reesei*.(*77*) This shows that inverting glycosidases could hydrolyse the glycosyl fluorides of opposite configuration to the substrate by a re-synthesis-hydrolysis mechanism involving two molecules of the `wrong' fluoride.

1.5.3. Oxacarbenium ion character at enzymatic transition state

Enzymatic mechanisms reflect an interplay between intrinsic reactivity of the substrate and the catalytic strategies that result in a lowering of the activation energy. Regardless of whether the enzyme catalyzed glycosylation and deglycosylation are dissociative (D_N*A_N)(*83, 84*) or "exploded" associative (A_ND_N),(*84, 85*) there is a substantial cationic character on the sugar component in the enzymatic transition state,(*86*) which is commonly referred to as the oxacarbenium ion-like transition state. In such transition states, a lone pair of electrons from the endocylic oxygen atom assists in stabilizing positive charge development at the anomeric centre and transiently enforces a planar geometry for the C-5,

O-5, C-1, and C-2 atoms (**Figures 1-20a**, and the corresponding orbital overlap is shown in **1-20b**). Other structural characteristics of such transition states include; a largely broken glycosidic bond, partial bond development between the enzymatic nucleophile and the anomeric carbon atom and lastly, a distorted pyranose ring conformation that enables an improved delocalization of the positive charge on the carbon atom with that of the co-planar endocyclic oxygen atom.(*69*) To summarize, the anomeric centre in the transition state(s) for enzymatic glycoside hydrolyses are sp^2 like with a partial positive charge that likely is predominantly located across the C1-O5 bond.



1-20a **1-20b**

Figure 1-20. Stabilization of the developing positive charge at the anomeric centre developed at the transition state by the lone pair π-electrons from the endocyclic oxygen atom favoured by the co planarity of the C1 and O1 atoms.

Sinnott was perhaps the first to emphasize the conformational implications, based on series of secondary α-deuterium and β-deuterium KIE experiments.(*87*) That is, distortion of the glycoside to a half-chair ($^4H_3$ and $^3H_4$ or their equivalent ^4E and ^3E envelope forms) or boat ($^{2,5}B$ or $B_{2,5}$) conformations is necessary for the transition state to have a delocalized positive charge **Figure 1-21**. Recent reports have indicated that different transition state conformations are adopted by different enzymes as shown in **Figure 1-22**.

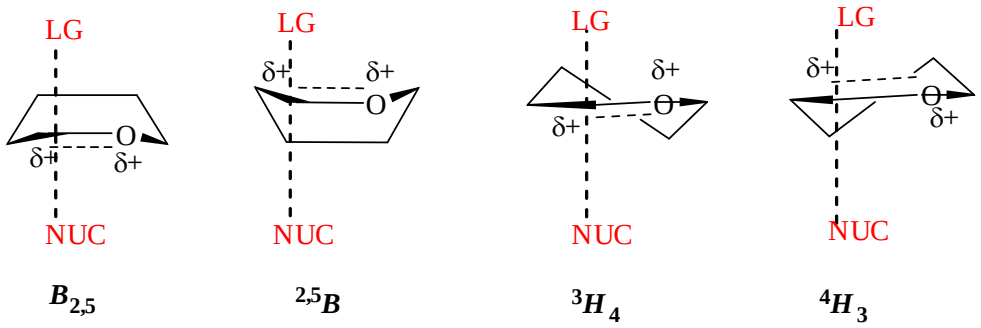


Figure 1-21. Possible conformations adopted by the pyranose ring for the transition states that resemble an oxacarbenium ion. Also shown are the relative positions of the leaving group (LG) and the nucleophile (NUC).(*61*)

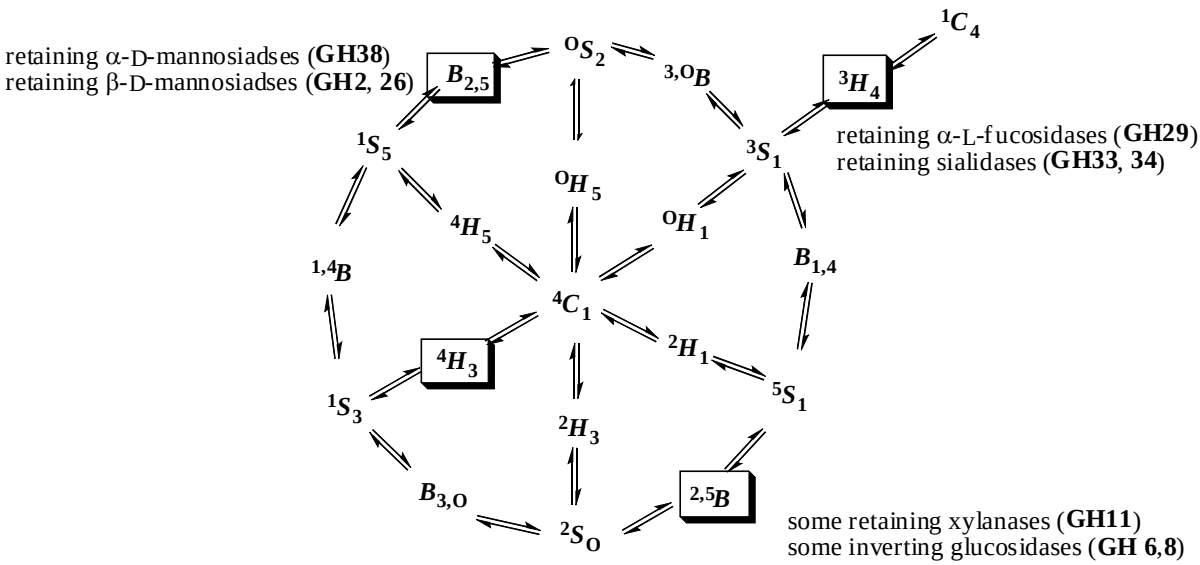


Figure 1-22. Stoddard's pseudorotational itinerary showing the conformational itineraries followed by various GH families. Stereochemical considerations demand distortion to $^{4}H_{3}$ or $^{3}H_{4}$ half-chair (or their equivalent ^{4}E and ^{3}E envelope forms) or, $^{2,5}B$ or $B^{2,5}$ boat conformations (shown boxed) at the transition state.(*61*)

The importance of understanding conformational changes that occur on going from bound substrate to product cannot be overstated if one wishes to design inhibitors that mimic transition state properties, like that of charge, geometry and conformation. As the substrate progresses from the Michaelis complex to products, chemical changes occur and these are caused by the enzyme *via* events such as protonation, proton abstraction, electron transfer and geometric perturbation. These possible changes must be taken into consideration while

performing mechanistic studies. Of note, *ab initio* and molecular modelling studies of the enzyme-substrate complex can also provide useful tips and mechanistic insights into the origins of catalysis by glycosyl hydrolases.

1.6. Catalytic mechanism of GH4 α-galactosidase and GH109 α-galactosaminidase

In contrast to the normal mechanisms of action (**Figure 1-15** and **Figure 1-16**), the activity of the GH4 family of retaining hydrolases is dependent on NAD^+. In 2004, Yip *et al.* experimentally determined the redox mechanism utilized by GH4 hydrolases and concluded that this family of glycosidases are dependent on three important cofactors - NAD^+, a divalent metal ion (Mn^{2+}) and reducing conditions (β-mercaptoethanol), for their catalytic activity (**Figure 1-17**).(*39*) Notably, GH4 family members possess a remarkable degree of substrate promiscuity, that is, this family includes both α- and β-glycosidases, where the substrates could be either phosphorylated or non-phosphorylated.(*45, 56, 57*) The structure of GlvA α-glucosidase showed that a bound NAD^+ moiety was positioned just below the C-3 atom of the sugar, an arrangement that facilitates hydride abstraction at C-3 (**Figure 1-23**).(*55*) Furthermore, kinetic isotope effect measurements demonstrated that abstraction of a hydride from C-3 and a proton from C-2, were both partially rate limiting.(*39, 56, 57*) Experimentally, GH4 6-phospho-β-glucosidase from *Thermotoga maritima* upon reacting with CD_3OD in D_2O yielded trideuteriomethyl 6-phospho-2-(^{2}H)-β-D-glucoside, the NMR spectral assignment of which, established two key inferences: a) the proton on C-2 is rendered acidic during the enzymatic process; and b) the catalytic mechanism involves retention of anomeric configuration.(*39, 56, 57*) In a nutshell, the proposed mechanism by Yip *et al.* for GH4 6-phospho β-glucosidase from *BglT*, involves hydride abstraction at C-3 in the first step, that results in oxidation of 3-OH to a ketone.(*39*) Such an oxidation, in turn acidifies the C-2 proton, allowing deprotonation by an enzymatic base, with a concomitant acid-catalyzed elimination of the glycosidic oxygen, forming a 1,2-unsaturated intermediate. The resultant α,β-unsaturated Michael acceptor then undergoes base-catalyzed attack of a water molecule to generate a 3-keto glucose derivative. The final step is the reduction of the C-3 keto group by the bound NADH to complete the catalytic cycle (**Figure 1-17**).

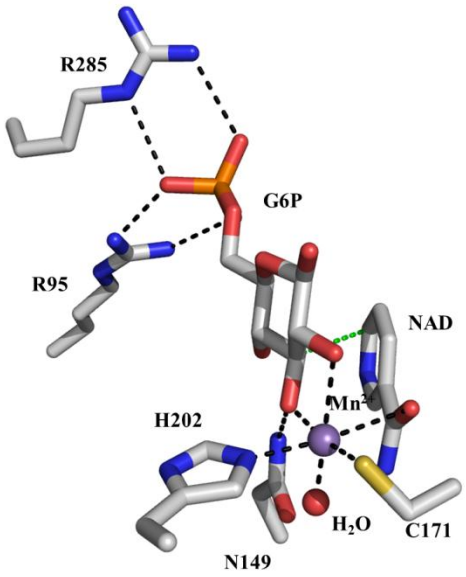


Figure 1-23. Active site structure of GH4 GlvA α-glucosidase revealing the coordination to the two most important cofactors for GH4 glycosyl hydrolases, nucleotide NAD^+ and divalent metal ion Mn^{2+}.(*55*) (Thanks to Dr. Tracey Gloster for providing the representation of the active site)

In 2007, Liu *et al.* discovered a new family of hydrolases from prokaryotes, GH109 α-*N*-acetylgalactosaminidases (α-NAGAL, EC 3.2.1.49), and various members from this family have proven efficient in removal of blood group A immunodeterminants from erythrocytes.(*11*) Of note, this GH family shares a close relationship with that of GH4 glycosyl hydrolases.(*11, 57*) The 2.3 Å resolution X-ray crystal structure of the *E. meningosepticum* α-NAGAL (GH109), revealed a bound NAD^+ in the active site (**Figure 1-24**). This observation led to the assumption of a mechanistic similarity between GH109 and that of GH4 hydrolases.(*11, 55*) The α-NAGAL studied from GH109 differs from the enzymes in GH4 in two crucial aspects, that is for activity: a) the GH109 enzyme does not depend on a bound divalent metal ion; and 2) it does not require reducing conditions (DTT, TCEP or β-mercaptoethanol).(*11*) Notably, the only other prokaryotic α-NAGAL reported to date, also used for blood group conversion (A to O), from *C. perfringens*, has been reported to require reducing conditions (DTT) to maintain catalytic activity.(*54*) These contrasting, yet interesting activity requirements compelled us to undertake the detailed mechanistic investigation of GH109 α-NAGAL, the results of which will be reported in Chapter 4.

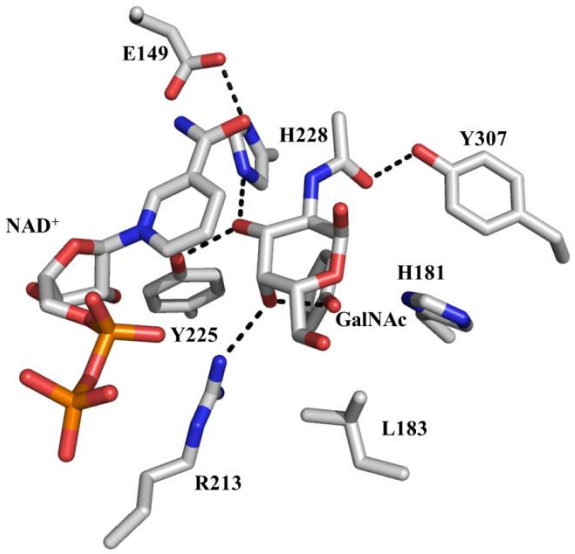
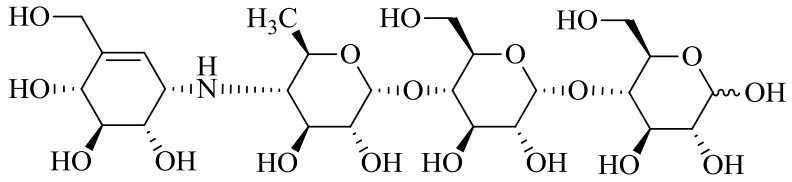


Figure 1-24. Active site architecture of GH109 *E. meningosepticum* α-NAGAL, bound with reaction product GalNAc. Also shown is a molecule of NAD^+ bound in a deeply embedded tunnel like structure.(*11*) (Thanks to Dr. Tracey Gloster for providing the representation of the active site)

1.7. An Overview of Glycosidase Inhibitors

The development of rationally designed glycosyl hydrolase inhibitors has been an advancing field of research because of the biological importance of these enzymes,(*25, 88*) and its usefulness as mechanistic tools.(*89, 90*) Considerable effort has been targeted towards the composition of both potent(*91, 92*) and selective(*93, 94*) inhibitors. The enormous rate enhancements afforded by glycosidases (~10^{17} fold) in comparison to the spontaneous hydrolysis reaction suggests that the design of inhibitors should be based on knowledge of the enzymatic transition states. That is, the goal is to try to design true transition state analogues as inhibitors.(*95*) The inhibitory potency may result from a combination of charge and/or geometrical resemblance to the transition state, as initially proposed by Pauling.(*96*) Transition states, however, are extremely fleeting species with a lifetime of 10^{-13} sec, the time of a single bond vibration. Hence, elucidating the structural aspects of transition states is an extremely arduous task. While tight binding of glycosidase inhibitors is almost always interpreted as reflecting mimicry of the transition state, it is really quite unclear which inhibitors are true mimics, which are merely adventitious binders, and,

equally importantly, which features of these compounds actually give rise to potency.(*95*) Many of the inhibitors developed over the years that researchers claim to be "transition-state analogues" have not been rigorously evaluated in order to support such an assignment. Nature has, as expected, a source of inspiration for the design of inhibitors. For example acarbose (**Figure 1-25**), that was isolated from *Actinoplanes species*, is commonly used for treatment of type 2 diabetes.(*97*) The inhibitory power of acarbose has been shown to be due to the following interactions with active site residues of α-glucosidases: a) formation of a salt bridge between the basic site (N) on the compound with a catalytic proton donor; and b) constraint of the carbocyclic ring to a half-chair $^{2}H_{3}$ conformation that is thought to mimic the TS conformation.(*97*)



Acarbose, **1.25**

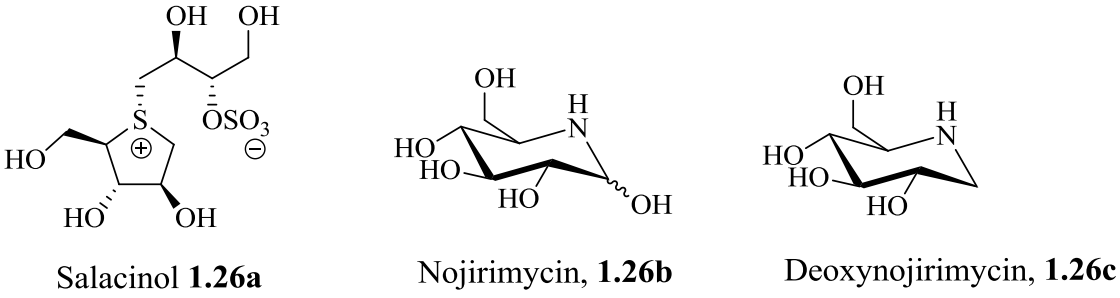
Figure 1-25. Acarbose, isolated from *Actinoplanes species* is a potent inhibitor of α-glucosidase and amylase.(*97*)

Glycosyl hydrolases can be inhibited by a wide variety of structurally different compounds. Nevertheless, no single inhibitor discovered to date, satisfies all the structural requirements of a true TS analogue. Discussing the structural aspects of the countless glycosidase inhibitors is beyond the scope of this thesis. However, a brief listing of the categories into which the glycosyl hydrolase inhibitors are commonly divided based on their structural motif, is as follows, with examples:

a) Oxacarbenium ion-like character of the TS

Many inhibitors possess structural features that mimic certain facets of glycopyranosylium ions **1-20a**, a high-energy species formed during the acid-catalyzed hydrolysis of glycosides. An interesting class of inhibitors which is thought to confer oxacarbenium ion-like properties is that of the sulfonium ion containing inhibitors, for example salacinol **1.26a**, which is also used as a therapeutic for type 2 diabetes.(*98*) It is

postulated that under physiological conditions, the positively charged sulphur atom can act in the same way as the protonated ammonium center of the amine-based glycosidase inhibitors when binding in the active site of glucosidases.(*98*) The inhibitors that contain an endocyclic nitrogen atom, by virtue of protonation at the basic site, are believed to mimic a glycopyranosylium ion. For example, the quintessential natural product with an endocyclic nitrogen atom is nojirimycin, 5-amino-5-deoxyglucose **1.26b**.(*99*) Subsequently, the *manno* and *galacto* analogues were also isolated as natural products. However, due to their inherent instability, the reduced analogue, deoxynojirimycin **1.26c**, which also possesses similar inhibitory potency, has been considered as key structural motif for the design of glycosidase inhibitors.(*100*)

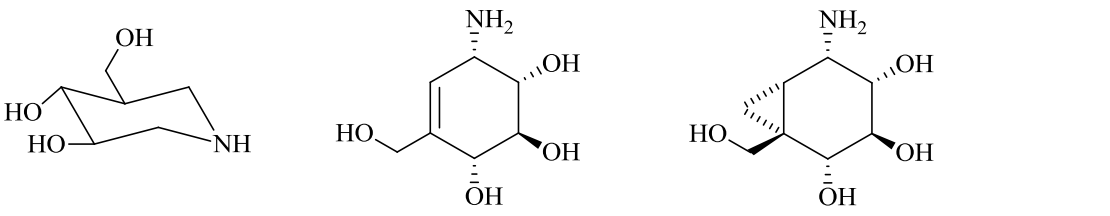


Salacinol **1.26a** Nojirimycin, **1.26b** Deoxynojirimycin, **1.26c**

Figure 1-26. Inhibitors possessing either a permanent positive charge or an endocyclic basic site that upon protonation mimics oxacarbenium ion-like TS.

b) Conformational resemblance to TS

Apart from acarbose,(*97*) yet another α-glucosidase inhibitor valienamine(*101*) **1.27b** is also proposed to bind in a $^{2}H_{3}$ conformation in the enzyme active site, although the proposal remains untested. Modification of the alkene functionality of valienamine (**1.27b**) to that of a fused cyclopropyl ring has augmented inhibitory potency towards GH27 α-galactosidases (for the *galacto* analogue)(*102*) and GH31 yeast α-glucosidase **1.27c.**(*103*)

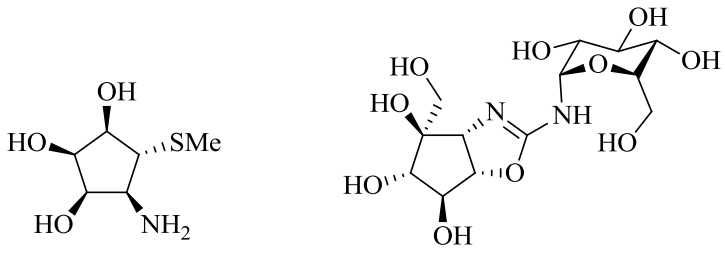


Isofagomine **1.27a** Valienamine **1.27b** Bicyclo [4.1.0]heptane derivatives **1.27c**

Figure 1-27. Possible transition state analogues, mimicking charge **1.27a**, salt bridge interaction between the amine functionality and active site carboxylates and also mimicking conformation, **1.27b** and **1.27c**.(*100, 101, 103*)

c) Salt-bridge/hydrogen bonding interactions

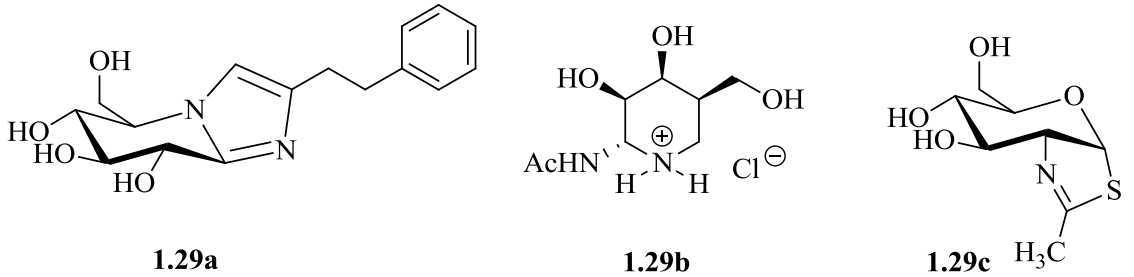
Analogues of glycosidase inhibitors possessing a 2-ammonium instead of the 2-hydroxyl group have been prepared and the anticipated strong inhibition has been shown to involve a salt bridge with the catalytic nucleophile.(*104*) Mannostatin A **1.28a** and trehalozin **1.28b** presumably inhibit on the basis of formation of a salt bridge of the basic group (-NH) on the molecule with the catalytic proton donor on the enzyme. Of note, on the basis of both structural and molecular modelling studies it has been suggested that the potent Golgi α-mannosidase inhibitor mannostatin A **1.28a**, resembles the covalently linked glycosyl enzyme intermediate, and not the expected TS.(*105*) This observation necessitates the consideration of alternative routes to achieve higher inhibitory potency, other than mimicking the structural aspects of the transition state only.



Mannostatin A**1.28a** Trehalozin **1.28b**

Figure 1-28. Structures of glycosidase inhibitors which gains active site stabilization through salt-bridge interactions between the basic site on the molecule and active site carboxylate residues.

Another noteworthy approach towards enhancing the selectivity and strength of inhibitors is by introducing substituents mimicking the aglycone. Introduction of a phenethyl substituent (both hydrophobic and flexible) to a tetrahydroimidazo[1,2a] pyridine (**1.29a**) increased the inhibition potency towards the *C. saccharolyticum* β-glucosidase, leading to the strongest known inhibitor of a β-glucosidase with a dissociation constant of $K_i = 1.1 \times 10^{-10}$ M.(*106*) Presumably, the improved potency is attributed to the binding interactions in the aglycone (+1) subsite; as such Gloster *et al.* speculated the favourable entropy of binding as the reason behind the increased potency.(*106*) Interestingly, strong inhibition ($K_i = 2.7$ μM) of *Streptomyces plicatus* β-hexosaminidase by a GalNAc-isofagomine.HCl **1.29b**(*107*) is in accord with the mechanism involving substrate assisted catalysis, which occurs in GH 18 chitinases(*108*) and GH 20 β-hexosaminidase.(*109*) More recently, Whitworth *et al*. probed the geometric resemblance of NAG-thiazoline **1.29c** to the transition state of GH84 human *O*-GlcNAcase, which also uses a catalytic mechanism involving anchimeric assistance from neighbouring acetamido group.(*110*)



1.29a **1.29b** **1.29c**

Figure 1-29. Phenylethyl-substituted glucoimidazole **1.29a**, a potent inhibitor of β-glucosidase, GalNAc-isofagomine **1.29b** and NAG-thiazoline **1.29c**, are proposed to be a transition state analogues for GH 20 β-hexosaminidases and GH84 human *O*-GlcNAcase, respectively.(*104, 107, 110*)

Further clarification of the finer details of transition state structures for the glycosidases and the subtle differences in the thermodynamic stability of the compounds binding to the GHs, by virtue of charge, shape and electrostatic interactions, should facilitate the generation of more potent and selective inhibitors. Discussion of the structural aspects of

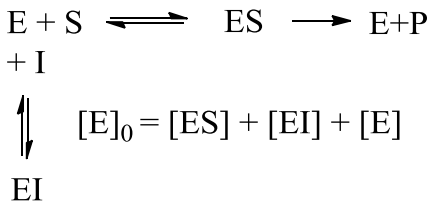
all known glycosidase inhibitors is beyond the scope of this thesis; however, the fundamental basis for different modes of inhibition will be discussed in the next section.

1.8. Modes of Enzyme Inhibition

Conceptually, an inhibitor is a molecule that upon binding to an enzyme reduces the catalytic activity of the protein. There are many different possible modes of inhibition.

1.8.1. Reversible Inhibition

If the inhibitor competes with the substrate for the active site so that in a mixture of inhibitor, substrate and enzyme both EI (enzyme-inhibitor) and ES (enzyme-substrate) complexes are formed but no ESI complexes are present, then the mode of inhibition is termed as competitive inhibition.(*111*) Competitive inhibition, by definition, requires that the inhibitor competes with the substrate for binding to the enzyme, and that this binding is mutually exclusive. In other words, if the substrate binds to the enzyme, the inhibitor will not be able to bind and *vice versa*. Importantly, competitive inhibition also suggests that the degree of inhibition can be lessened in the presence of saturating amounts of substrate, in which case all the enzyme active sites will be occupied by substrate thus preventing inhibitor binding. Typically, a competitive inhibitor is a molecule that resembles the chemical structure and molecular geometry of the substrate. The inhibitor binds to the free enzyme (EI) and hence lowers the concentration of free enzyme, but it does not affect the ES complex once it is formed. Hence, the inhibitor increases the apparent substrate K_m, while leaving the magnitude of maximum reaction rate V_{max} unchanged. Schematically the mechanism of competitive inhibition can be represented as in **Scheme 1-1**:



Scheme 1-1. Schematic representation of competitive inhibition.

The standard Michaelis-Menten equation still holds for competitive inhibition but with the K_m term now including the concentration of inhibitor and its dissociation constant K_i (**Equation 1**):

Equation 1:

$$v = \frac{V_{max}\,[\mathrm{S}]}{[\mathrm{S}] + K_m\left(1 + \frac{[\mathrm{I}]}{K_i}\right)}$$

Competitive reversible inhibitors usually bind non-covalently and are desirable from a pharmaceutical perspective. Two other examples of reversible inhibition are called: i) non-competitive inhibition such inhibitors that bind to both free enzyme (E) and enzyme-substrate complex (ES), with the result that V_{max} decreases and K_m increases; and ii) uncompetitive inhibition – in this case the inhibitor binds to the ES complex only with the overall effect being to lower V_{max} and leave K_m unaltered. Other types of reversible inhibition can occur but will not be discussed further. However, most rationally designed reversible inhibitors are competitive inhibitors. The designing of such inhibitors often incorporates bio-isosteric modifications to the natural enzyme substrate.

1.8.2. Irreversible Inhibition

Irreversible inhibition is commonly associated with the formation of a covalent bond between an inhibitor and the enzyme where the bond cannot be easily broken and a time dependent loss of enzyme activity is observed. A general kinetic scheme for irreversible inhibition is shown in **Scheme 1-2**.

$$\mathrm{E} + \mathrm{I} \overset{K_i}{\rightleftharpoons} \mathrm{EI} \xrightarrow{k_{inact}} \mathrm{EI}^*$$

Scheme 1-2. Schematic representation of irreversible inhibition.

Where, enzyme E binds to the inhibitor I with a dissociation constant of K_i to give a Michaelis-type complex EI which proceeds via a first-order reaction to give the permanently inactive and modified enzyme EI*.

Broadly, irreversible inhibitors can be classified into primarily two types, based on their mechanism of action: a) affinity labels and active site directed irreversible inhibitors; and b) mechanism based irreversible inactivators.

1.8.2.1. Affinity Labelling

$$E + I \underset{}{\overset{K_i}{\rightleftharpoons}} EI \xrightarrow{k_{inact}} E\text{-}I$$

These inactivators are reactive compounds which initially forms a reversible EI complex, then an active-site nucleophilic amino acid residue (-OH, -NH_2, -SH) reacts with the enzyme-bound inactivator to form a covalent adduct, hence 'labelling' or irreversibly modifying the enzyme. The term K_i represents the concentration of the inactivator that gives half-maximal inactivation, and k_{inact} is the maximal rate constant for inactivation. A prominent application for the usage of affinity labels is as a biochemical tool to probe the role of active site residues in the enzyme's catalytic process. Many irreversible inhibitors modify one of the catalytic carboxylate residues of the enzyme active site. In this way, the enzyme is trapped in an inactive form. However, due to their lack of selectivity, affinity labels are undesirable from therapeutics perspective.

1.8.2.2. Mechanism-Based Inactivation

$$E + S \underset{}{\overset{K_i}{\rightleftharpoons}} [ES] \xrightarrow{k_{cat}} [ES^*] \xrightarrow{k_{inact}} [E\text{-}S]$$

Mechanism based inhibitors, often called "suicide inhibitors", are those which are inherently unreactive but when exposed to normal catalytic processing by the enzyme, are activated into highly reactive species. These reactive species, in turn, can either: a) irreversibly modify an active site nucleophilic residue in the enzyme; or b) modify an essential cofactor. Both modes of alkylation leads to the "death" of the enzyme ("suicide"). The kinetic scheme is outlined above. The enzyme (E) binds to the substrate (inhibitor) S, with a dissociation constant of K_i. The ES complex is activated in the enzyme active site to

ES*, which eventually alkylates the enzyme, forming E-S complex. However, in order to be designated as suicide inhibitor, the key requirements would be the following: a) inactivation should be time-dependent; b) the enzyme should show saturation phenomena, hence first order kinetics should be followed; and c) the stoichiometry of active site and inhibitor should be 1:1, in other words a reversible active site inhibitor should be able to protect the enzyme from inactivation.

1.8.3. Transition State Analogues and Fortutious Tight Binding Inhibitors

In retaining glycosidases **Figure 1-15**, both glycosylation and deglycosylation occur via transition states (TSs) that have oxacarbenium ion-like character **1.21**(*46, 69, 112*) and a distorted six-membered ring.(*112*) The high affinity for the transition state makes glycosidases particularly rewarding targets for inhibitor design through mimicry of this transition state.(*95*) That is, if a molecule resembles the transition state of an enzymatic reaction by virtue of its shape and/or charge, then the enzyme should bind more tightly to the transition state analogue in comparison to both the substrates and products. For glycosidase inhibitors, three important transition state features are considered: i) geometry; ii) charge distribution; and iii) salt bridge/hydrogen bonding interactions with the catalytic carboxylates.(*113, 114*) Although no single compound incorporates all the desired TS features, numerous postulated TS analogues have been reported. Those for glucosidase enzymes include: nojirimycin (5-amino-5-deoxyglucose) (**1.26b**),(*100*) valienamine(*101*)(**1.27b**) and the bicyclo[4.1.0]heptyl derivative (**1.27c**).(*103*) However, as the enzyme-substrate complex (Michaelis complex) progresses towards the transition state, the structural changes that occur enable comparatively tighter electrostatic and hydrogen bond interactions to be formed in the active site. The electrostatic binding could further be enhanced by introducing differential charge interactions between the substrate and the possible transition state. An insightful paper by Wolfenden further explains that chemically stable transition state analogues would be anticipated to bind more tightly than substrates by factors resembling the rate enhancement imposed by enzymes.(*113*) Nevertheless, the experimental approach that provides unmatched insights into the enzymatic transition state structure is through analysis of the kinetic isotope effects.(*115, 116*)

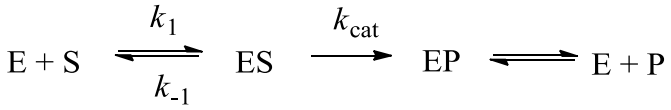
1.9. Experiments and Methods

1.9.1. Cloning and Expression

The currently accepted first and most important step when planning to perform a mechanistic study on a new enzyme is to clone the gene of interest. That is, the important steps in obtaining a clone involve: i) design forward and reverse primers for the gene based on the published genomic sequence; ii) perform PCR amplification of the gene; iii) digest the gene of interest and the chosen vector (pET28 in this thesis) using restriction enzymes to produce "sticky ends"; iv) ligate the desired gene into the vector pET28; v) transform the plasmid into competent cells (BL21DE3 Gold cells in this thesis); and vi) grow bacterial colonies on antibiotic-resistant agar plates. In order to confirm the presence of the gene of interest, a check is usually performed on the growing colonies to ensure that the gene of interest had been incorporated. Subsequently, the protein of interest is expressed in optimized media. Addition of 2-propyl 1-thio-β-D-galactoside (IPTG) induces expression of the protein and finally upon obtaining the crude lysate, the protein of interest is purified using Nickel affinity column. Nickel column purification is commonly used when the protein of interest has a His_6 tag, incorporated on either the C- or N-terminus.

1.9.2. Enzyme Kinetics

The quintessential equation that underlies most enzyme kinetic experiments was developed by Michaelis and Menten in 1913, and is therefore known as the Michaelis–Menten equation.(*111*) For a single substrate enzyme, the substrate (S) and enzyme (E) are in reversible equilibrium with a non-covalent enzyme-substrate complex (ES) that has a dissociation constant of K_S. Following complex formation, the rate-determining step (k_{cat}) converts the substrate to an analogous enzyme-product complex (EP), which then dissociates to give the final product with regeneration of the active catalyst.



Scheme 1-3. Schematic presentation of Michaelis-Menten kinetics

The final equation can be written as,

Equation 2:

$$\nu_0 = \frac{V_{max}[S]}{K_M + [S]}$$

Where ν_0 is the current reaction rate; V_{max} is the maximum reaction rate and it is related to k_{cat} as $V_{max} = k_{cat}$ [E], where E refers to the enzyme concentration; K_m, the Michaelis-constant, which is the substrate concentration where ν_0 is half the value of V_{max}; and [S] stands for the substrate concentration. The terminologies which are important to explain from the context of the thesis content would be: i) k_{cat} is the first-order rate constant for the conversion of the enzyme:substrate complex (ES) to product within the active site of the catalyst, and this is often referred to as turnover number (unit s^{-1}); and ii) k_{cat}/K_m = the apparent second-order rate constant, which is first order with respect to both the free enzyme and substrate, and this constant is also known as the catalytic efficiency (units M^{-1} s^{-1}).(*111*) Additionally, the catalytic proficiency of an enzyme ($k_{cat}/K_m \times 1/k_{uncat}$) needs to be considered when trying to understand by how much an enzyme has stabilized the reaction transition states.(*117*)

1.9.3. Kinetic Isotope Effects

This is a powerful technique for the deduction of structural information about the transition state of a reaction. Substitution of one isotope for another at or near an atom at which bonds are breaking or rehybridizing conventionally leads to a change in rate of reaction. When the bond breaking/making involves those to hydrogen, replacing the

hydrogen (H) for the heavier isotope deuterium (D) leads to a prominent measurable isotope effect. The main origin for isotope effects in solution is the difference in frequencies of various vibrational modes of a molecule that are altered by isotopic substitution. By virtue of the Born–Oppenheimer approximation, the potential energy of the system remains unchanged as a result of the isotopic substitution. At this point, however, discussing the zero-point energy (ZPE) and its relation to vibrational frequency of the bond would be useful. ZPE is usually referred to as the energies measured at the lowest point in the potential energy well (**Figure 1-30**). It is represented as:

Equation 3:

$$e_0 = \frac{1}{2}h\nu$$

where h = Planck's constant and ν = vibration frequency. For a diatomic molecule the vibration frequency (ν), for a simplified harmonic oscillator model, is directly proportional to the square root of the force constant (κ) and inversely proportional to the reduced mass (m_r), and can be represented as:

Equation 4:

$$\nu = \frac{1}{2\pi}\sqrt{\frac{k}{m_r}}$$

The reduced mass for a bond between a heavy atom like O, N, C with that of H is altered on isotopic substitution of H for D. Comparison of a C-D bond with a C-H bond clearly indicates a larger activation energy required to break a C-D bond. Typically, a kinetic isotope effect is defined as the ratio of reaction rates of two isotopologues, for example shown in **Equation 5** is the definition for a deuterium KIE: where, k_H = rate constant of the hydrogen isotopologue, and k_D = rate constant of the deuterium isotopologue. When the isotopic substitution is in the bond that is being broken or formed in the rate-determining step, the KIE is termed a primary isotope effect. When the substitution is at a site in which

only weakening or strengthening of bonds occurs, it is termed a secondary kinetic isotope effect. The magnitude of kinetic isotope effects can help in the determination of bonding changes that are occurring at the reaction transition state.

Equation 5:

$$KIE = \frac{k_H}{k_D}$$

The magnitude of isotope effect provides important information regarding the reaction mechanism. When, $k_H/k_D > 1$, this is termed a normal kinetic isotope effect, whereas when $k_H/k_D < 1$, it is termed as an inverse kinetic isotope effect.

1.9.3.1. Origin of Secondary Kinetic Isotope Effects

Secondary kinetic isotope effect (SDKIE) arises due to isotopic substitution at a bond that is not being broken in the rate-limiting step, and these typically involve a change in hybridisation or hyperconjugation. They are defined as α- or β-secondary isotope effects, depending on the position of the bond undergoing the change due to isotopic substitution. Of note, SDKIE can be either normal or inverse. SDKIE can be divided into three three types based on the origin of effect, where the origin could be due to one of the following: a) hybridization changes; b) hyperconjugation; and c) steric effects. With respect to the thesis content, only SDKIE due to hybridization changes and hyperconjugation will be discussed in detail.

a) Hybridization changes

In order to understand a secondary isotope effect, all the changes in the vibrational modes that occur in an atom (or atoms) associated with a bond undergoing rehybridization, needs to be needs to be taken into consideration. The C-H bond strengths and the associated force constansts (κ) decrease in the order $sp > sp^2 > sp^3$. Therefore, when the hybridization of the C-H bond changes from sp^3 to sp^2, the force constant for the stretching vibrations also changes. However, the energy difference between the out-of-plane bending modes result in the most significant differences in zero point energies (ZPE) between the ground state C-H

and C-D bonds and those in the transition state, which gives rise to the measurable secondary kinetic isotope effect. The out-of-plane bend weakens as the reaction approaches the TS, making ΔZPE_{TS} less than ΔZPE_{GS}, giving a normal KIE ($k_H/k_D > 1$). Since bending modes are relatively low-energy, the associated SDKIE is small, with a theoretical maximum of ~1.4. On the other hand, if the rehybridization is from sp^2 to sp^3, the out-of-plane bend strengthens as the reaction approaches the TS, making ΔZPE_{TS} greater than ΔZPE_{GS}, resulting in an inverse KIE ($k_H/k_D < 1$).

b) Hyperconjugation

The ability of a β-hydrogen/deuterium to stabilize an adjacent carbocation can also effect the SDKIE. In general, these effects are normal and can be just as large as the KIE's arising from rehybridization. Hyperconjugation weakens the C-H/D bond and lowers its associated vibrational frequency. Since the deuterium labelled molecule has a lower ZPE, it participates in a hyperconjugation interaction to a lesser extent than does the protonated molecule, which results in a small normal β-SDKIE being observed.

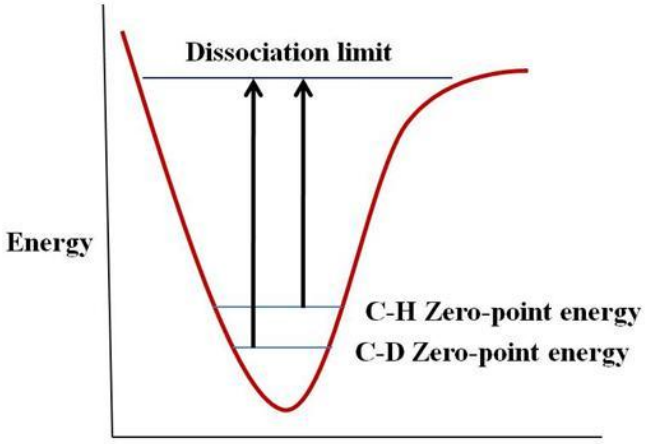


Figure 1-30. A potential energy diagram showing the higher activation energy needed for the dissociation of a C-D bond compared to a C-H bond.

1.9.4. pH Profile for Enzymatic Reactions

It is often the case that the substrate has a required protonation state for binding and/or catalysis. In a similar way, catalytic groups in the enzyme active site usually have required protonation states for catalysis. The determination of the pH-dependence of the

kinetic parameters of the Michaelis-Menten equation is an important consideration while delineating mechanistic details of an enzyme.(*111*) Although enzymes contain a multitude of ionizable groups, it is commonly found that plots of rate of reaction versus pH takes the form of simple single or double ionization curves. A number of conclusions can be drawn from the variation of the catalytic parameters, k_{cat}, K_m and k_{cat}/K_m, with the pH of the reaction media.(*111*) While the pH dependence of K_m follows the ionization of the enzyme-substrate complex, the value of k_{cat} reflects the pK_a of the enzyme-substrate complex whose decomposition is rate-limiting. Plots of k_{cat}/K_m versus pH, on the other hand, yields the pK_a values for free enzyme and/or free substrate. Nevertheless, in some cases in which a bell-shaped curve appears in the pH rate profile, the two apparent pK_a values that characterize it can arise from the same ionization, and not from ionizations of different residues; one of the apparent pK_a values is shifted from the true pK_a value by a factor containing complex rate constants.(*118*) Of note, when analyzing the pH-dependence of an enzyme rate constant one must remember that the enzyme (and cofactors if any) must remain intact throughout the pH range, especially in the pH extremes, for the analysis to be valid.

1.10. Thesis Overview

The work discussed in this thesis includes a survey of the mechanistic aspects of various glycosyl hydrolases, possessing both "classical" and "non-classical" catalytic mechanisms. The goal of the research that forms the basis of this thesis involved building upon a recently unveiled and atypical, oxidation-reduction mechanism adopted by GH4 and GH109 hydrolases. This involved studying previously unexplored or under-investigated enzymes belonging to these two GH families. Of particular interest was the elucidation of the rate-limiting step(s) in the catalytic cycle of these glycosyl hydrolases by measuring multiple deuterium KIEs. Accordingly, two unexplored enzymes, GH4 α-galactosidase from *C. freundii* and GH109 α-*N*-acetylgalactosaminidase from *E. meningosepticum* were chosen as the enzymes of choice to perform the mechanistic studies. These studies have been conducted to provide the knowledge necessary to enable designing of glycosidase inhibitors pertinent to the "unusual" NAD^+ dependent transient oxidation-reduction mechanism followed by GH4 and GH109 hydrolases. With regards to GH36 α-galactosidase from

T. maritima, which follows the classical "retaining" mechanism, we designed and synthesized an active site-directed irreversible inhibitor. Many irreversible inhibitors alkylate one of the two catalytic carboxylate groups of the enzyme, which results in an inactive enzyme. Therefore, we performed irreversible inhibition experiments to aid in the detailed elucidation of the catalytic machinery. The objectives of this thesis entailed charting the flexibility of the mechanistic features of three retaining hydrolases, possessing diverse catalytic mechanisms.

A general introduction of various aspects of the catalytic mechanism of glycosyl hydrolases has been discussed in Chapter 1. In addition, a brief overview of the experimental techniques adopted to perform the mechanistic studies of the aforementioned enzymes has been enumerated in the same chapter.

Chapter 2 focuses on the chemistry and biochemistry of GH4 α-galactosidase from *C. freundii*. This primarily consists of chemical synthesis of substrates for Brønsted plots, cofactor analysis and a detailed pH profile of the enzyme. Additionally, the synthesis of the deuterium labelled isotopologues (1-^{2}H, 2-^{2}H, 3-^{2}H) has been described, combined with the kinetic assays used to measure the detailed KIEs of each isotopologue. The synthesis and the KIE studies with the di-deuterated isotopologue (2-^{2}H, 3-^{2}H) enabled us to propose that the two supposed sequential and partially rate-limiting steps in the catalytic cycle of our GH4 α-galactosidase, were in reality concerted.

In Chapter 3, the focus is on the development and analysis of irreversible inhibitors of the GH36 α-galactosidase from *T. maritima*. Novel carbocyclic α-galactoside substrate analogues were synthesized, possessing intrinsically reactive cyclopropyl and allyl functional groups. One of the three compounds, the *galacto* compound, proved to be an irreversible inhibitor of this enzyme. Consequently, the results of the detailed inhibition experiments and pH profile of inhibition/reactivation of GH36 α-galactosidase by this *galacto* irreversible inhibitor were investigated. We also attempted to identify the important catalytic residue through the alkylation site of the inhibitor using mass spectrometric studies. This elaborate mechanistic elucidation through inhibition of a retaining α-galactosidase, could be used to gain valuable insight into the mechanism and specificities of this family of enzyme, thereby

allowing the generation of more potent inhibitors to treat α-galactosidase-related disorders, like that of lysosomal storage disorder Fabry disease.

Chapter 4 broadens the mechanistic elucidations of another NAD^+ dependent glycosyl hydrolase belonging to GH-109. The chapter begins with the experimental details related to cloning of the gene and expression of the recombinant protein. Subsequently, the synthetic details of several aryl αGalNAc substrates, and the associated Brønsted plots, are detailed. In addition, the results and implications of the diverse substrate specificity and a full pH profile of GH109 α-NAGAL from *E. meningosepticum* have been measured. Preliminary results from deuterium KIE studies conducted for this enzyme using phenyl αGal substrate isotopologues have indicates the mechanistic parallelism between the two NAD^+ dependent hydrolases (GH4 and GH109).

To conclude with, Chapter 5 summarizes the key observations and conclusions from all the individual chapters outlining the advantages and limitations of each undertaking. Additionally, the proposed future work, on the basis of our present results and interpretations, has been briefly described.

1.11. References

1. Lloyd, D. H., Viac, J., Werling, D., Reme, C. A., and Gatto, H. (2007) Role of sugars in surface microbe-host interactions and immune reaction modulation, *Vet Dermatol 18*, 197-204.

2. Frey, P. A. (1996) The Leloir pathway: a mechanistic imperative for three enzymes to change the stereochemical configuration of a single carbon in galactose, *FASEB J 10*, 461-470.

3. Wiegandt, H. (1995) The chemical constitution of gangliosides of the vertebrate nervous system, *Behav Brain Res 66*, 85-97.

4. Hannun, Y. A., and Bell, R. M. (1989) Functions of sphingolipids and sphingolipid breakdown products in cellular regulation, *Science 243*, 500-507.

5. Hakomori, S. (1990) Bifunctional role of glycosphingolipids. Modulators for transmembrane signaling and mediators for cellular interactions, *J Biol Chem 265*, 18713-18716.

6. Harouse, J. M., Bhat, S., Spitalnik, S. L., Laughlin, M., Stefano, K., Silberberg, D. H., and Gonzalez-Scarano, F. (1991) Inhibition of entry of HIV-1 in neural cell lines by antibodies against galactosyl ceramide, *Science 253*, 320-323.

7. Holzel, A., Komrower, G. M., and Schwarz, V. (1957) Galactosemia, *Am J Med 22*, 703-711.

8. Schwarz, V. (1975) Galactosaemia and the problem of galactose toxicity, *Biochem Soc Trans 3*, 234-238.

9. Clausen, H., and Hakomori, S. (1989) ABH and related histo-blood group antigens; immunochemical differences in carrier isotypes and their distribution, *Vox Sang 56*, 1-20.

10. Hoskins, L. C., Larson, G., and Naff, G. B. (1995) Blood group A immunodeterminants on human red cells differ in biologic activity and sensitivity to alpha-N-acetylgalactosaminidase, *Transfusion 35*, 813-821.

11. Liu, Q. P., Sulzenbacher, G., Yuan, H., Bennett, E. P., Pietz, G., Saunders, K., Spence, J., Nudelman, E., Levery, S. B., White, T., Neveu, J. M., Lane, W. S., Bourne, Y., Olsson, M. L., Henrissat, B., and Clausen, H. (2007) Bacterial glycosidases for the production of universal red blood cells, *Nat Biotechnol 25*, 454-464.

12. Olsson, M. L., Hill, C. A., de la Vega, H., Liu, Q. P., Stroud, M. R., Valdinocci, J., Moon, S., Clausen, H., and Kruskall, M. S. (2004) Universal red blood cells-enzymatic conversion of blood group A and B antigens, *Transfus Clin Biol 11*, 33-39.

13. Itzkowitz, S. H., Yuan, M., Montgomery, C. K., Kjeldsen, T., Takahashi, H. K., Bigbee, W. L., and Kim, Y. S. (1989) Expression of Tn, sialosyl-Tn, and T antigens in human colon cancer, *Cancer Res 49*, 197-204.

14. Dahiya, R., Itzkowitz, S. H., Byrd, J. C., and Kim, Y. S. (1989) ABH blood group antigen expression, synthesis, and degradation in human colonic adenocarcinoma cell lines, *Cancer Res 49*, 4550-4556.

15. Brockhausen, I. (2006) Mucin-type *O*-glycans in human colon and breast cancer: glycodynamics and functions, *EMBO Rep 7*, 599-604.

16. Garman, S. C., Hannick, L., Zhu, A., and Garboczi, D. N. (2002) The 1.9 Å structure of alpha-*N*-acetylgalactosaminidase: molecular basis of glycosidase deficiency diseases, *Structure 10*, 425-434.

17. Wang, A. M., Bishop, D. F., and Desnick, R. J. (1990) Human alpha-*N*-acetylgalactosaminidase-molecular cloning, nucleotide sequence, and expression of a full-length cDNA. Homology with human alpha-galactosidase A suggests evolution from a common ancestral gene, *J Biol Chem 265*, 21859-21866.

18. Davies, G. J., Gloster, T. M., and Henrissat, B. (2005) Recent structural insights into the expanding world of carbohydrate-active enzymes, *Curr Opin Struct Biol 15*, 637-645.

19. Henrissat, B., and Bairoch, A. (1996) Updating the sequence-based classification of glycosyl hydrolases, *Biochem J 316 (Pt 2)*, 695-696.

20. Garman, S. C., and Garboczi, D. N. (2002) Structural basis of Fabry disease, *Mol Genet Metab 77*, 3-11.

21. Garman, S. C., and Garboczi, D. N. (2004) The molecular defect leading to Fabry disease: structure of human alpha-galactosidase, *J Mol Biol 337*, 319-335.

22. Brady, R. O., Gal, A. E., Bradley, R. M., Martensson, E., Warshaw, A. L., and Laster, L. (1967) Enzymatic defect in Fabry's disease. Ceramidetrihexosidase deficiency, *N Engl J Med 276*, 1163-1167.

23. Desnick, R. J., Wasserstein, M. P., and Banikazemi, M. (2001) Fabry disease (alpha-galactosidase A deficiency): renal involvement and enzyme replacement therapy, *Contrib Nephrol*, 174-192.

24. Desnick, R. J., and Wasserstein, M. P. (2001) Fabry disease: clinical features and recent advances in enzyme replacement therapy, *Adv Nephrol Necker Hosp 31*, 317-339.

25. Butters, T. D., Dwek, R. A., and Platt, F. M. (2003) New therapeutics for the treatment of glycosphingolipid lysosomal storage diseases, *Adv Exp Med Biol 535*, 219-226.

26. Kanzaki, T., Wang, A. M., and Desnick, R. J. (1991) Lysosomal alpha-*N*-acetylgalactosaminidase deficiency, the enzymatic defect in angiokeratoma corporis diffusum with glycopeptiduria, *J Clin Invest 88*, 707-711.

27. Wang, A. M., Kanzaki, T., and Desnick, R. J. (1994) The molecular lesion in the alpha-*N*-acetylgalactosaminidase gene that causes angiokeratoma corporis diffusum with glycopeptiduria, *J Clin Invest 94*, 839-845.

28. Schindler, D., Bishop, D. F., Wolfe, D. E., Wang, A. M., Egge, H., Lemieux, R. U., and Desnick, R. J. (1989) Neuroaxonal dystrophy due to lysosomal alpha-*N*-acetylgalactosaminidase deficiency, *N Engl J Med 320*, 1735-1740.

29. Linden, H. U., Klein, R. A., Egge, H., Peter-Katalinic, J., Dabrowski, J., and Schindler, D. (1989) Isolation and structural characterization of sialic-acid-containing glycopeptides of the *O*-glycosidic type from the urine of two patients with an hereditary deficiency in alpha-*N*-acetylgalactosaminidase activity, *Biol Chem Hoppe Seyler 370*, 661-672.

30. Desnick, R. J. (2001) Enzyme replacement and beyond, *J Inherit Metab Dis 24*, 251-265.

31. Eng, C. M., Guffon, N., Wilcox, W. R., Germain, D. P., Lee, P., Waldek, S., Caplan, L., Linthorst, G. E., and Desnick, R. J. (2001) Safety and efficacy of recombinant human alpha-galactosidase A-replacement therapy in Fabry's disease, *N Engl J Med 345*, 9-16.

32. Schiffmann, R., Ries, M., Timmons, M., Flaherty, J. T., and Brady, R. O. (2006) Long-term therapy with agalsidase alfa for Fabry disease: safety and effects on renal function in a home infusion setting, *Nephrol Dial Transplant 21*, 345-354.

33. Tomasic, I. B., Metcalf, M. C., Guce, A. I., Clark, N. E., and Garman, S. C. Interconversion of the specificities of human lysosomal enzymes associated with Fabry and Schindler diseases, *J Biol Chem 285*, 21560-21566.

34. Sato, S., Ward, C. L., Krouse, M. E., Wine, J. J., and Kopito, R. R. (1996) Glycerol reverses the misfolding phenotype of the most common cystic fibrosis mutation, *J Biol Chem 271*, 635-638.

35. Fan, J. Q., Ishii, S., Asano, N., and Suzuki, Y. (1999) Accelerated transport and maturation of lysosomal alpha-galactosidase A in Fabry lymphoblasts by an enzyme inhibitor, *Nat Med 5*, 112-115.

36. Matsuda, J., Suzuki, O., Oshima, A., Yamamoto, Y., Noguchi, A., Takimoto, K., Itoh, M., Matsuzaki, Y., Yasuda, Y., Ogawa, S., Sakata, Y., Nanba, E., Higaki, K., Ogawa, Y., Tominaga, L., Ohno, K., Iwasaki, H., Watanabe, H., Brady, R. O., and Suzuki, Y. (2003) Chemical chaperone therapy for brain pathology in G(M1)-gangliosidosis, *Proc Natl Acad Sci U S A 100*, 15912-15917.

37. Goldstein, J., Siviglia, G., Hurst, R., Lenny, L., and Reich, L. (1982) Group B erythrocytes enzymatically converted to group O survive normally in A, B, and O individuals, *Science 215*, 168-170.

38. Liu, Q. P., Yuan, H., Bennett, E. P., Levery, S. B., Nudelman, E., Spence, J., Pietz, G., Saunders, K., White, T., Olsson, M. L., Henrissat, B., Sulzenbacher, G., and Clausen, H. (2008) Identification of a GH110 subfamily of alpha 1,3-galactosidases: novel enzymes for removal of the alpha 3Gal xenotransplantation antigen, *J Biol Chem 283*, 8545-8554.

39. Yip, V. L., Varrot, A., Davies, G. J., Rajan, S. S., Yang, X., Thompson, J., Anderson, W. F., and Withers, S. G. (2004) An unusual mechanism of glycoside hydrolysis involving redox and elimination steps by a family 4 beta-glycosidase from Thermotoga maritima, *J Am Chem Soc 126*, 8354-8355.

40. Fujimoto, Z., Kaneko, S., Kim, W. D., Park, G. G., Momma, M., and Kobayashi, H. (2009) The tetramer structure of the glycoside hydrolase family 27 alpha-galactosidase I from Umbelopsis vinacea, *Biosci Biotechnol Biochem 73*, 2360-2364.

41. Comfort, D. A., Bobrov, K. S., Ivanen, D. R., Shabalin, K. A., Harris, J. M., Kulminskaya, A. A., Brumer, H., and Kelly, R. M. (2007) Biochemical analysis of *Thermotoga maritima* GH36 alpha-galactosidase (*Tm*GalA) confirms the mechanistic commonality of clan GH-D glycoside hydrolases, *Biochemistry 46*, 3319-3330.

42. Imamura, H., Fushinobu, S., Jeon, B. S., Wakagi, T., and Matsuzawa, H. (2001) Identification of the catalytic residue of Thermococcus litoralis 4-alpha-glucanotransferase through mechanism-based labeling, *Biochemistry 40*, 12400-12406.

43. Kitamura, M., Okuyama, M., Tanzawa, F., Mori, H., Kitago, Y., Watanabe, N., Kimura, A., Tanaka, I., and Yao, M. (2008) Structural and functional analysis of a glycoside hydrolase family 97 enzyme from *Bacteroides thetaiotaomicron*, *J Biol Chem 283*, 36328-36337.

44. Guce, A. I., Clark, N. E., Salgado, E. N., Ivanen, D. R., Kulminskaya, A. A., Brumer, H., 3rd, and Garman, S. C. (2010) Catalytic mechanism of human alpha-galactosidase, *J Biol Chem 285*, 3625-3632.

45. Chakladar, S., Cheng, L., Choi, M., Liu, J., and Bennet, A. J. (2011) Mechanistic evaluation of MelA alpha-galactosidase from *Citrobacter freundii*: a family 4 glycosyl hydrolase in which oxidation is rate-limiting, *Biochemistry 50*, 4298-4308.

46. Koshland, D. E., Jr. (1953) Stereochemistry and the mechanism of enzymatic reactions, *Biological Reviews 28*, 416-436.

47. McCarter, J. D., and Withers, S. G. (1994) Mechanisms of enzymatic glycoside hydrolysis, *Curr Opin Struct Biol 4*, 885-892.

48. Coutinho, P. M., Deleury, E., Davies, G. J., and Henrissat, B. (2003) An evolving hierarchical family classification for glycosyltransferases, *J Mol Biol 328*, 307-317.

49. Clark, N. E., and Garman, S. C. (2009) The 1.9 Å structure of human alpha-*N*-acetylgalactosaminidase: The molecular basis of Schindler and Kanzaki diseases, *J Mol Biol 393*, 435-447.

50. Weissmann, B., and Hinrichsen, D. F. (1969) Mammalian alpha-*N*-acetylgalactosaminidase. Occurrence, partial purification, and action on linkages in submaxillary mucins, *Biochemistry 8*, 2034-2043.

51. Sung, S. S., and Sweeley, C. C. (1980) Purification and partial characterization of porcine liver alpha-*N*-acetylgalactosaminidase, *J Biol Chem 255*, 6589-6594.

52. Tuppy, H., and Staudenbauer, W. L. (1966) The action on soluble blood group A substances of an alpha-*N*-acetylgalactosaminidase from *Helix pomatia*, *Biochemistry 5*, 1742-1747.

53. Harun-Or-Rashid, M., Matsuzawa, T., Satoh, Y., Shiraishi, T., Ando, M., Sadik, G., and Uda, Y. Purification and characterization of alpha-*N*-acetylgalactosaminidases I and II from the starfish *Asterina amurensis*, *Biosci Biotechnol Biochem 74*, 256-261.

54. Levy, G. N., and Aminoff, D. (1980) Purification and properties of alpha-*N*-acetylgalactosaminidase from *Clostridium perfringens*, *J Biol Chem 255*, 11737-11742.

55. Rajan, S. S., Yang, X., Collart, F., Yip, V. L., Withers, S. G., Varrot, A., Thompson, J., Davies, G. J., and Anderson, W. F. (2004) Novel catalytic mechanism of glycoside hydrolysis based on the structure of an NAD^+/Mn^{2+} -dependent phospho-alpha-glucosidase from *Bacillus subtilis*, *Structure 12*, 1619-1629.

56. Yip, V. L., Thompson, J., and Withers, S. G. (2007) Mechanism of GlvA from *Bacillus subtilis*: a detailed kinetic analysis of a 6-phospho-alpha-glucosidase from glycoside hydrolase family 4, *Biochemistry 46*, 9840-9852.

57. Yip, V. L., and Withers, S. G. (2006) Mechanistic analysis of the unusual redox-elimination sequence employed by *Thermotoga maritima* BglT: a 6-phospho-beta-glucosidase from glycoside hydrolase family 4, *Biochemistry 45*, 571-580.

58. Yip, V. L., and Withers, S. G. (2004) Nature's many mechanisms for the degradation of oligosaccharides, *Org Biomol Chem 2*, 2707-2713.

59. van Diggelen, O. P., Schindler, D., Willemsen, R., Boer, M., Kleijer, W. J., Huijmans, J. G., Blom, W., and Galjaard, H. (1988) alpha-*N*-acetylgalactosaminidase deficiency, a new lysosomal storage disorder, *J Inherit Metab Dis 11*, 349-357.

60. Kanzaki, T., Yokota, M., Irie, F., Hirabayashi, Y., Wang, A. M., and Desnick, R. J. (1993) Angiokeratoma corporis diffusum with glycopeptiduria due to deficient

lysosomal alpha-*N*-acetylgalactosaminidase activity. Clinical, morphologic, and biochemical studies, *Arch Dermatol 129*, 460-465.

61. Vocadlo, D. J., and Davies, G. J. (2008) Mechanistic insights into glycosidase chemistry, *Curr Opin Chem Biol 12*, 539-555.

62. Tarling, C. A., He, S., Sulzenbacher, G., Bignon, C., Bourne, Y., Henrissat, B., and Withers, S. G. (2003) Identification of the catalytic nucleophile of the family 29 alpha-L-fucosidase from *Thermotoga maritima* through trapping of a covalent glycosyl-enzyme intermediate and mutagenesis, *J Biol Chem 278*, 47394-47399.

63. Lodge, J. A., Maier, T., Liebl, W., Hoffmann, V., and Strater, N. (2003) Crystal structure of *Thermotoga maritima* alpha-glucosidase AglA defines a new clan of NAD^+-dependent glycosidases, *J Biol Chem 278*, 19151-19158.

64. Tews, I., Terwisscha van Scheltinga, A. C., Perrakis, A., Wilson, K. S., and Dijkstra, B. W. (1997) Substrate-Assisted Catalysis Unifies Two Families of Chitinolytic Enzymes, *J Am Chem Soc 119*, 7954-7959.

65. Watson, J. N., Dookhun, V., Borgford, T. J., and Bennet, A. J. (2003) Mutagenesis of the conserved active-site tyrosine changes a retaining sialidase into an inverting sialidase, *Biochemistry 42*, 12682-12690.

66. Watts, A. G., Oppezzo, P., Withers, S. G., Alzari, P. M., and Buschiazzo, A. (2006) Structural and kinetic analysis of two covalent sialosyl-enzyme intermediates on *Trypanosoma rangeli* sialidase, *J Biol Chem 281*, 4149-4155.

67. Amaya, M. F., Watts, A. G., Damager, I., Wehenkel, A., Nguyen, T., Buschiazzo, A., Paris, G., Frasch, A. C., Withers, S. G., and Alzari, P. M. (2004) Structural insights into the catalytic mechanism of *Trypanosoma cruzi* trans-sialidase, *Structure 12*, 775-784.

68. Watts, A. G., Damager, I., Amaya, M. L., Buschiazzo, A., Alzari, P., Frasch, A. C., and Withers, S. G. (2003) *Trypanosoma cruzi* trans-sialidase operates through a covalent sialyl-enzyme intermediate: tyrosine is the catalytic nucleophile, *J Am Chem Soc 125*, 7532-7533.

69. Zechel, D. L., and Withers, S. G. (2000) Glycosidase mechanisms: anatomy of a finely tuned catalyst, *Acc Chem Res 33*, 11-18.

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

71. van Aalten, D. M., Komander, D., Synstad, B., Gaseidnes, S., Peter, M. G., and Eijsink, V. G. (2001) Structural insights into the catalytic mechanism of a family 18 *exo*-chitinase, *Proc Natl Acad Sci U S A 98*, 8979-8984.

72. Tews, I., Perrakis, A., Oppenheim, A., Dauter, Z., Wilson, K. S., and Vorgias, C. E. (1996) Bacterial chitobiase structure provides insight into catalytic mechanism and the basis of Tay-Sachs disease, *Nat Struct Biol 3*, 638-648.

73. Markovic-Housley, Z., Miglierini, G., Soldatova, L., Rizkallah, P. J., Muller, U., and Schirmer, T. (2000) Crystal structure of hyaluronidase, a major allergen of bee venom, *Structure 8*, 1025-1035.

74. Macauley, M. S., Whitworth, G. E., Debowski, A. W., Chin, D., and Vocadlo, D. J. (2005) *O*-GlcNAcase uses substrate-assisted catalysis: kinetic analysis and development of highly selective mechanism-inspired inhibitors, *J Biol Chem 280*, 25313-25322.

75. Kitahata, S., Brewer, C. F., Genghof, D. S., Sawai, T., and Hehre, E. J. (1981) Scope and mechanism of carbohydrase action. Stereocomplementary hydrolytic and glucosyl-transferring actions of glucoamylase and glucodextranase with alpha- and beta-D-glucosyl fluoride, *J Biol Chem 256*, 6017-6026.

76. Hehre, E. J., Brewer, C. F., and Genghof, D. S. (1979) Scope and mechanism of carbohydrase action. Hydrolytic and nonhydrolytic actions of beta-amylase on alpha- and beta-maltosyl fluoride, *J Biol Chem 254*, 5942-5950.

77. Becker, D., Johnson, K. S., Koivula, A., Schulein, M., and Sinnott, M. L. (2000) Hydrolyses of alpha- and beta-cellobiosyl fluorides by Cel6A (cellobiohydrolase II) of *Trichoderma reesei* and *Humicola insolens*, *Biochem J 345 Pt 2*, 315-319.

78. Honda, Y., and Kitaoka, M. (2006) The first glycosynthase derived from an inverting glycoside hydrolase, *J Biol Chem 281*, 1426-1431.

79. Hehre, E. J., Sawai, T., Brewer, C. F., Nakano, M., and Kanda, T. (1982) Trehalase: stereocomplementary hydrolytic and glucosyl transfer reactions with alpha- and beta-D-glucosyl fluoride, *Biochemistry 21*, 3090-3097.

80. Kasumi, T., Tsumuraya, Y., Brewer, C. F., Kersters-Hilderson, H., Claeyssens, M., and Hehre, E. J. (1987) Catalytic versatility of Bacillus pumilus beta-xylosidase: glycosyl transfer and hydrolysis promoted with alpha- and beta-D-xylosyl fluoride, *Biochemistry 26*, 3010-3016.

81. Hehre, E. J., Brewer, C. F., Uchiyama, T., Schlesselmann, P., and Lehmann, J. (1980) Scope and mechanism of carbohydrase action. Stereospecific hydration of 2,6-anhydro-1-deoxy-D-gluco-hept-1-enitol catalyzed by alpha- and beta-glucosidases and an inverting *exo*-alpha-glucanase, *Biochemistry 19*, 3557-3564.

82. Okada, G., Genghof, D. S., and Hehre, E. J. (1979) The predominantly nonhydrolytic action of alpha amylases on alpha-maltosyl fluoride, *Carbohydr Res 71*, 287-298.

83. Tina L. Amyes, and Jencks, W. P. (1989) Lifetimes of oxocarbenium ions in aqueous solution from common ion inhibition of the solvolysis of .alpha.-azido ethers by added azide ion, *J Am Chem Soc 111*, 7888-7900.

84. Huang, X. C., Tanaka, K. S. E., and Bennet, A. J. (1997) Glucosidase-Catalyzed Hydrolysis of R-D-Glucopyranosyl Pyridinium Salts: Kinetic Evidence for Nucleophilic Involvement at the Glucosidation Transition State, *J Am Chem Soc 119*, 11147-11154.

85. Banait, N. S., and Jencks, W. P. (1991) Reactions of anionic nucleophiles with .alpha.-D-glucopyranosyl fluoride in aqueous solution through a concerted, A_ND_N (S_N2) mechanism, *J Am Chem Soc 113*, 7951-7958.

86. Berti, P. J., and Tanaka, K. S. E. (2002) Transition State Analysis Using Multiple Kinetic Isotope Effects: Mechanisms of Enzymatic and Non-enzymatic Glycoside Hydrolysis and Transfer, *Adv Phys Org Chem 37*, 239-314.

87. Sinnott, M. L. (1990) Catalytic Mechanisms of Enzymic Glycosyl Transfer, *Chemical Reviews 90*, 1171-1202.

88. von Itzstein, M., Wu, W. Y., Kok, G. B., Pegg, M. S., Dyason, J. C., Jin, B., Van Phan, T., Smythe, M. L., White, H. F., Oliver, S. W., and et al. (1993) Rational design of potent sialidase-based inhibitors of influenza virus replication, *Nature 363*, 418-423.

89. Tulp, A., Barnhoorn, M., Bause, E., and Ploegh, H. (1986) Inhibition of *N*-linked oligosaccharide trimming mannosidases blocks human B cell development, *EMBO J 5*, 1783-1790.

90. Vosseller, K., Wells, L., Lane, M. D., and Hart, G. W. (2002) Elevated nucleocytoplasmic glycosylation by O-GlcNAc results in insulin resistance associated with defects in Akt activation in 3T3-L1 adipocytes, *Proc Natl Acad Sci U S A 99*, 5313-5318.

91. Vasella, A., Davies, G. J., and Bohm, M. (2002) Glycosidase mechanisms, *Curr Opin Chem Biol 6*, 619-629.

92. Lillelund, V. H., Jensen, H. H., Liang, X., and Bols, M. (2002) Recent developments of transition-state analogue glycosidase inhibitors of non-natural product origin, *Chem Rev 102*, 515-553.

93. Siriwardena, A., Strachan, H., El-Daher, S., Way, G., Winchester, B., Glushka, J., Moremen, K., and Boons, G. J. (2005) Potent and selective inhibition of class II alpha-D-mannosidase activity by a bicyclic sulfonium salt, *Chembiochem 6*, 845-848.

94. Ho, C. W., Lin, Y. N., Chang, C. F., Li, S. T., Wu, Y. T., Wu, C. Y., Liu, S. W., Li, Y. K., and Lin, C. H. (2006) Discovery of different types of inhibition between the

human and *Thermotoga maritima* alpha-fucosidases by fuconojirimycin-based derivatives, *Biochemistry 45*, 5695-5702.

95. Gloster, T. M., and Davies, G. J. Glycosidase inhibition: assessing mimicry of the transition state, *Org Biomol Chem 8*, 305-320.

96. Pauling, L. (1948) Nature of forces between large molecules of biological interest, *Nature 161*, 707-709.

97. Truscheit, E., Frommer, W., Junge, B., Muller, L., Schmidt, D. D., and Wingender, W. (1981) Chemistry and Biochemistry of Microbial α-Glucosidase Inhibitors, *Angew Chem Int Ed Engl 20*, 741-766.

98. Mohan, S., and Pinto, B. M. (2007) Zwitterionic glycosidase inhibitors: salacinol and related analogues, *Carbohydr Res 342*, 1551-1580.

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

100. Hettkamp, H., Bause, E., and Legler, G. (1982) Inhibition by nojirimycin and 1-deoxynojirimycin of microsomal glucosidases from calf liver acting on the glycoprotein oligosaccharides Glc1-3Man9GlcNAc2, *Biosci Rep 2*, 899-906.

101. Kameda, Y., Asano, N., Yoshikawa, M., and Matsui, K. (1980) Valienamine as an alpha-glucosidase inhibitor, *J Antibiot (Tokyo) 33*, 1575-1576.

102. Wang, Y., and Bennet, A. J. (2007) A potent bicyclic inhibitor of a family 27 alpha-galactosidase, *Org Biomol Chem 5*, 1731-1738.

103. Tanaka, K. S., Winters, G. C., Batchelor, R. J., Einstein, F. W., and Bennet, A. J. (2001) A new structural motif for the design of potent glucosidase inhibitors, *J Am Chem Soc 123*, 998-999.

104. Panday, N., Muthuppalaniappan, M., and Vasella, A. (2000) A Comparison of Glucose- and Glucosamine-Related Inhibitors: Probing the Interaction of the 2-Hydroxy Group with Retaining β-Glucosidases, *Helv Chim Acta 83*, 513-538.

105. Tailford, L. E., Offen, W. A., Smith, N. L., Dumon, C., Morland, C., Gratien, J., Heck, M. P., Stick, R. V., Bleriot, Y., Vasella, A., Gilbert, H. J., and Davies, G. J. (2008) Structural and biochemical evidence for a boat-like transition state in beta-mannosidases, *Nat Chem Biol 4*, 306-312.

106. Gloster, T. M., Roberts, S., Perugino, G., Rossi, M., Moracci, M., Panday, N., Terinek, M., Vasella, A., and Davies, G. J. (2006) Structural, kinetic, and thermodynamic analysis of glucoimidazole-derived glycosidase inhibitors, *Biochemistry 45*, 11879-11884.

107. Mark, B. L., Vocadlo, D. J., Knapp, S., Triggs-Raine, B. L., Withers, S. G., and James, M. N. (2001) Crystallographic evidence for substrate-assisted catalysis in a bacterial beta-hexosaminidase, *J Biol Chem 276*, 10330-10337.

108. Funkhouser, J. D., and Aronson, N. N., Jr. (2007) Chitinase family GH18: evolutionary insights from the genomic history of a diverse protein family, *BMC Evol Biol 7*, 96.

109. Jiang, Y. L., Yu, W. L., Zhang, J. W., Frolet, C., Di Guilmi, A. M., Zhou, C. Z., Vernet, T., and Chen, Y. Structural Basis for the Substrate Specificity of a Novel beta-*N*-Acetylhexosaminidase StrH Protein from Streptococcus pneumoniae R6, *J Biol Chem 286*, 43004-43012.

110. Whitworth, G. E., Macauley, M. S., Stubbs, K. A., Dennis, R. J., Taylor, E. J., Davies, G. J., Greig, I. R., and Vocadlo, D. J. (2007) Analysis of PUGNAc and NAG-thiazoline as transition state analogues for human *O*-GlcNAcase: mechanistic and structural insights into inhibitor selectivity and transition state poise, *J Am Chem Soc 129*, 635-644.

111. Fersht, A., (Ed.) (1985) *Enzyme structure and mechanism* 2nd ed., W.H. Freeman, New York.

112. Jencks, W. P. (1975) Binding energy, specificity, and enzymic catalysis: the circe effect, *Adv Enzymol Relat Areas Mol Biol 43*, 219-410.

113. Wolfenden, R. (1969) Transition state analogues for enzyme catalysis, *Nature 223*, 704-705.

114. Gloster, T. M., Meloncelli, P., Stick, R. V., Zechel, D., Vasella, A., and Davies, G. J. (2007) Glycosidase inhibition: an assessment of the binding of 18 putative transition-state mimics, *J Am Chem Soc 129*, 2345-2354.

115. Cleland, W. W. (1982) Use of isotope effects to elucidate enzyme mechanisms, *CRC Crit Rev Biochem 13*, 385-428.

116. Cleland, W. W. (1995) Isotope effects: determination of enzyme transition state structure, *Methods Enzymol 249*, 341-373.

117. Albery, W. J., and Knowles, J. R. (1976) Evolution of enzyme function and the development of catalytic efficiency, *Biochemistry 15*, 5631-5640.

118. Knowles, J. R. (1976) The intrinsic pK_a-values of functional groups in enzymes: improper deductions from the pH-dependence of steady-state parameters, *CRC Crit Rev Biochem 4*, 165-173.

2. Mechanistic Evaluation of MelA α-Galactosidase from Citrobacter Freundii: A Family 4 Glycosyl Hydrolase in which Oxidation is Rate-Limiting

This Chapter comprises the manuscript "**Mechanistic Evaluation of MelAα-Galactosidase from Citrobacter Freundii: A Family 4 Glycosyl Hydrolase in which Oxidation is Rate-Limiting"** which was published in the Journal of *Biochemistry* (**2011**, *50*, 4298-4308).

Saswati Chakladar,[a] Lydia Cheng,[b] Mary Choi,[a] James Liu,[b] and Andrew J. Bennet[a]

[a] Department of Chemistry, Simon Fraser University, Burnaby, British Columbia, Canada V5A 1S6

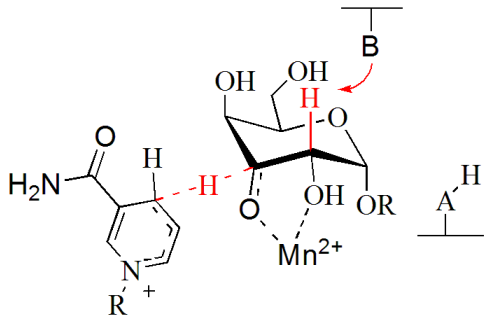
[b] Department of Molecular Biology and Biochemistry, Simon Fraser University, Burnaby, British Columbia, Canada V5A 1S6

The following manuscript is a verbatim copy of the original paper published in *Biochemistry*, **2011**, 50, 4298-4308 and is formatted as per thesis rules and regulations of Simon Fraser University. Permission to reproduce the published material was obtained from the American Chemical Society, and from co-authors Lydia Cheng, Mary Choi and James Liu. These authors made the following contribution to this work: Lydia Cheng performed the cloning of gene, Mary Choi aided in substrate synthesis and James Liu contributed to the measurement of the enzyme kinetics. The manuscript was prepared by myself in collaboration with my thesis supervisor.

2.1. Abstract

The MelA gene from *Citrobacter Freundii*, which encodes for a glycosyl hydrolase family 4 (GH4) α-galactosidase, has been cloned and expressed in *E. coli*. The recombinant enzyme catalyzes the hydrolysis of phenyl α-galactosides via a redox- elimination–addition mechanism involving oxidation of the hydroxyl group at C-3 and elimination of phenol across the C1–C2 bond to give an enzyme-bound glycal intermediate. For optimal activity the MelA enzyme requires two cofactors, NAD^+ and Mn^{2+}, and the addition of a reducing agent, such as mercaptoethanol. In order to delineate the mechanism of action for this GH4 enzyme leaving group effects were measured, and the derived β_{lg} values on V and V/K are indistinguishable from zero (–0.01 ± 0.02 and 0.02 ± 0.04, respectively). Deuterium kinetic isotope effects (KIEs) were measured for the weakly activated substrate phenyl α-D-galactopyranoside in which isotopic substitution was incorporated on C1, C2 or C3. KIEs of 1.06 ± 0.07, 0.91 ± 0.04 and 1.02 ± 0.06 were measured on V for the 1-(^{2}H), 2-(^{2}H) and 3-(^{2}H) isotopic substrates, respectively. The corresponding values on V/K were 1.13 ± 0.07, 1.74 ± 0.06 and 1.74 ± 0.05. To determine if the KIEs report on a single step or on a virtual transition state, KIEs were measured using doubly deuterated substrates. The measured $^{D}V/K$ KIEs for MelA catalyzed hydrolysis of phenyl α-D-galactopyranoside on the dideuterated substrates, $^{D}V/K_{(3\text{-}D)/(2\text{-}D,3\text{-}D)}$ and $^{D}V/K_{(2\text{-}D)/(2\text{-}D,3\text{-}D)}$ are 1.71 ± 0.12 and 1.71 ± 0.13, respectively. In addition, the corresponding values on V, $^{D}V_{(3\text{-}D)/(2\text{-}D,3\text{-}D)}$ and $^{D}V_{(2\text{-}D)/(2\text{-}D,3\text{-}D)}$ are 0.91 ± 0.06 and 1.01 ± 0.06, respectively. These observations are consistent with oxidation

at C-3, which occurs via a hydride-transfer to the on-board NAD^+, being concerted with proton removal at C-2 and that this step is the first irreversible step for the MelA α-galactosidase-catalyzed reactions of aryl substrates. In addition, the rate-limiting step for V_{max} must come after this irreversible step in the reaction mechanism.



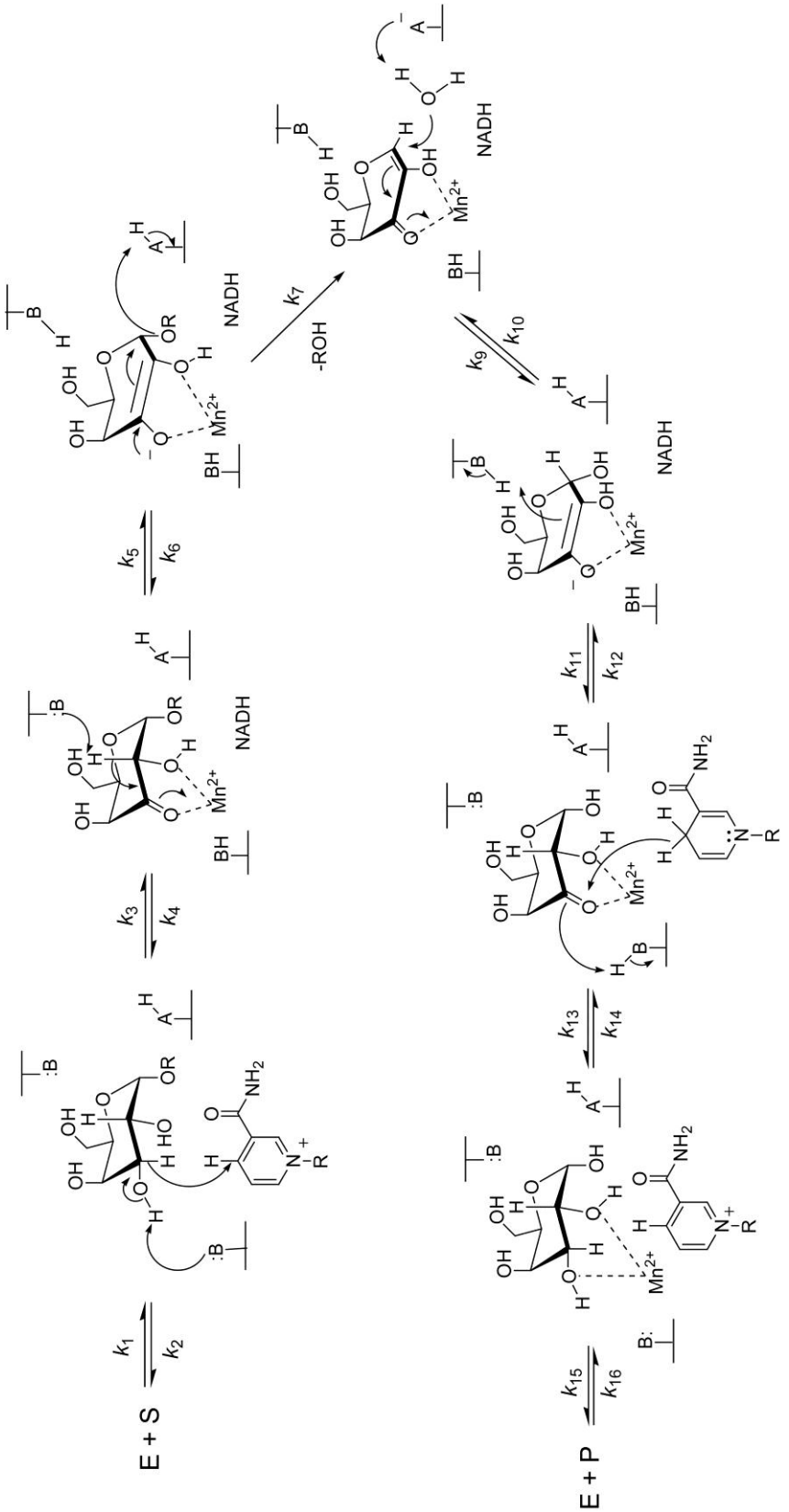
2.2. Abbreviations

1-(^{2}H)PαG, phenyl α-D-(1-^{2}H)galactoside; 2-(^{2}H)PαG, phenyl α-D-(2-^{2}H)galactoside; 2,3-($^2H_2$)PαG, phenyl α-D-(2-^{2}H, 3-^{2}H)galactoside; 3-(^{2}H)PαG, phenyl α-D-(3-^{2}H)galactoside, 4NPαG, 4-nitrophenyl α-D-galactoside; BSA, Bovine Serum Albumin; C. freundii, Citrobacter freundii; DMAP, 4-dimethylaminopyridine; DTT, dithiothreitol; E. coli, Escherichia coli; GH, glycosyl hydrolase; GH4, glycosyl hydrolase family 4; HEPES, (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid); IPTG, isopropyl 1-thio-β-D-galactopyranoside; KIE, kinetic isotope effect; NAD^+, β-nicotinamide adenine dinucleotide; NADH, β-nicotinamide adenine dinucleotide, reduced; PαG, phenyl α-D-galactoside, PNP, p-nitrophenyl; TBDPS, tert-butyl diphenylsilyl; TCEP, tris-(carboxyethyl)phosphine; THF, tetrahydrofuran; TLC, thin layer chromatography; UV-Vis, ultraviolet-visible; TS, transition state; TRIS, tris(hydroxymethyl)aminomethane;

2.3. Introduction

Glycosidases (glycosyl hydrolases) are an important class of carbohydrate processing enzymes, which number over a 110 families that are classified based on their protein sequence (*1*). Glycosyl hydrolases (GH) hydrolyze their respective substrates with either retention or inversion of anomeric configuration, by using a variety of catalytic strategies (*2*). Most GH families use a combination of general-acid catalysis, which assists departure of the aglycone, and either nucleophilic catalysis (retaining families) or general-base catalysis of water attack at the anomeric center (inverting families). In contrast, family 4 (GH4) members, which are solely derived from bacterial sources, contain a characteristic glycine rich sequence that, in combination with the topological arrangement of the secondary structure is consistent with a Rossmann fold, a structural motif that is primarily involved in the binding of NAD(H) (*3*). The crystal structure of α-glucosidase A, AglA, from *Thermotoga maritima* complexed with NAD^+ and maltose, has been solved to a resolution of 1.9 Å where Cys-174 and neighboring histidines have been shown to be key catalytic residues (*4*). Importantly, this research group suggested, based on electron density maps, that Cys-174 is easily oxidized to a sulfinic acid. This observation is consistent with the noted requirement for the presence of a reducing agent such as 2-mercaptoethanol or dithiothreitol for maximal catalytic activity. Recently, another GH family (GH109) has been shown to require NAD^+ for activity (*5*). In addition to the requirement for NAD^+, the catalytic activity of family-4 GHs is also dependent on the presence of a divalent metal, typically Mn^{2+}, which in the single crystal X-ray diffraction based structure is shown to be chelated to the O-2 and O-3 of the sugar ring (*6*). A detailed structural analysis of GlvA 6-phospho-α-glucosidase showed that Mn^{2+} ion promotes association of the protein to give the active tetramer (*7*). Notably, different members of GH4 possess catalytic activity against α- and β-glycosides, a characteristic not seen in most glycosyl hydrolase families that directly promote nucleophilic reactions at the anomeric center. Members of GH4 can hydrolyze both phosphorylated and non-phosphorylated disaccharide substrates suggesting the possible involvement of the phosphoenolpyruvate-dependent sugar phosphotransferase system (PEP-PTS) in the species from which the gene has been isolated (*8*). GH4 members also show approximately 15 % sequence similarity to those of lactate/malate dehydrogenase.

As a result of these unusual features that are present in family-4 glycosyl hydrolases, in-depth kinetic studies have been performed in order to delineate the prototypical mechanism of action for this GH family. Yip *et. al* reported such mechanistic studies on two GH4 family members, BglT from *Thermotoga maritima* and GlvA from *Bacillus subtilis* (*6, 9, 10*). These authors showed that GH4 members catalyze glycoside hydrolysis via an NAD^+ dependent redox reaction which is coupled to an *α,β*-elimination reaction involving the formation of an *α,β*-unsaturated ketone (glycal) intermediate (*10*). Moreover, based on kinetic isotope effect (KIE) measurements they concluded that oxidation at C3 and the subsequent C2 deprotonation steps are both partially rate-limiting (*6, 9*). The proposed mechanistic outline for GH4 enzymes is depicted in Scheme 1 for an α-galactosidase.



Scheme 2-1. Current Proposal for the Mechanism of GH4 Enzymes.

The present study details a mechanistic investigation of a GH4 α-galactosidase that is encoded by the *MelA* gene from the species *C. freundii*. This gene encodes an enzyme that hydrolyzes the terminal α-galactosyl units from glycolipids and glycoproteins. Melibiose (α-D-Galp-(1→6)-D-Glcp) utilization in *E. coli* is dependent on the melibiose locus (*Mel*), which forms an operon consisting of at least 2 structural genes, *MelA* and *MelB*, which code for an α-galactosidase and a melibiose carrier protein, respectively (*11*). Results from identification and purification studies on *MelA* have shown it to be a tetrameric protein with a molecular weight of 200 kDa for the active enzyme (*11*). In addition, Anggraeni *et al* showed that rMel4A α-galactosidase showed the highest k_{cat} values on the natural substrate melibiose, but showed highest affinity for raffinose (*12*).

2.4. Materials and Methods

Materials. All chemicals were of analytical grade or better and were purchased from Sigma-Aldrich unless noted otherwise. Milli-Q water (18.2 MΩ cm^{-1}) was used for all kinetic experiments. All pH values were measured using a standard pH electrode attached to a VWR pH meter. All NMR spectra were acquired on either a Bruker 400, 500 or 600 MHz spectrometer. Chemical shifts are reported in parts per million downfield from signals for TMS. The signal residues from deuterated chloroform and external TMS salts (D_2O) were used for ^1H NMR spectral references; for ^{13}C NMR spectra, natural abundance signals from $CDCl_3$ and external TMS salts (D_2O) were used as references. Coupling constants (*J*) are reported in hertz. Melting points were determined on a Gallenkamp melting point apparatus and are not corrected. Optical rotations were measured on a Perkin-Elmer 341 polarimeter and are reported in units of deg cm^2 g^{-1} (concentrations reported in units of g per 100 mL). All fitting of kinetic data was performed using the appropriate non-linear least squares equation in the computer program Graph Pad Prism (v 4.0).

2.4.1. Synthesis of Aryl α-D-Galactosides.

All aryl α-D-galactopyranosides, except the 3-nitrophenyl derivative, were synthesized from 1,2,3,4,6-penta-*O*-acetyl-D-galactopyranose using stannic chloride as the activator in CH_2Cl_2. Typically, 1,2,3,4,6-penta-*O*-acetyl-*β*-D-galactose (1.6 g, 4.0 mmol) and

the appropriate phenol (1.5 g, 8 mmol) were dissolved in anhydrous CH_2Cl_2 (50 mL) and then $SnCl_4$ (0.8 mL, 8 mmol) was added. The reaction mixture was stirred at ambient temperature under an inert atmosphere for 48 h, at which time TLC analysis showed that the reaction was complete. Following the addition of water (35 mL) the reaction mixture was neutralized by adding saturated $NaHCO_3$ (20 mL). The product was extracted from the aqueous layer using CH_2Cl_2 (3 × 35 mL), and the combined organic layer was washed with brine (2 × 20 mL), dried over Na_2SO_4 and then concentrated under reduced pressure to obtain the crude product. This material was purified by column chromatography using EtOAc-Hexane (1:4) as the eluent to obtain the pure α-anomer. Deprotection was accomplished under Zemplen conditions using catalytic NaOMe in MeOH followed by neutralization of the solution with Amberlite (H^+) resin. Finally, product purification was achieved by crystallization from ethanol. The 3-nitrophenyl α-D-galactopyranoside was synthesized using BF_3.OEt as the activator by following the reported reaction conditions (*13*). The detailed characterization of phenyl α-D-galactoside (**2.1a**) and 4-nitrophenyl α-D-galactoside (**2.2a**) have been previously reported (*14*). The characterization of all other unlabeled substrates used in this study are given below:

*3-Nitrophenyl α-D-galactoside (**2.2b**).* Mpt = 165–168 °C; $[\alpha]^{20}_D$ = +195.7 (c=0.15, H_2O); ^{1}H NMR (600 MHz, D_2O) δ 8.04 (s, 1 H, ArH), 8.00–7.97 (t, 1 H, ArH), 7.61–7.56 (m, 2 H, ArH), 5.79 (d, *J* = 3.8, 1 H, H-1), 4.12–4.01 (m, 4 H, H-2, H-3, H-4, H-5), 3.71 (app dd, *J* = 8.8, 6.2, 2 H, H-6, H-6'). ^{13}CNMR (151 MHz, D_2O) δ 156.42, 130.44, 124.03, 117.80, 112.04, 97.36, 72.02, 69.34, 69.05, 67.91, 60.95. ESI-MS for $C_{12}H_{15}NO_8$ *m*/*z* calcd for (M + Na^+) 324.0695: found 324.0682.

*3,4-Dichlorophenyl α-D-galactoside (**2.2c**).* Mpt = 163–166 °C; ^{1}HNMR (400 MHz, D_2O) δ 7.50 (d, *J* = 8.9, 1 H, ArH), 7.40 (d, *J* = 2.8, 1 H, ArH), 7.11 (dd, *J* = 8.9, 2.8, 1 H, ArH), 5.66 (d, *J* = 3.7, 1 H, H-1), 4.10–3.96 (m, 4 H, H-2, H-3, H-4, H-5), 3.74–3.69 (m, 2 H, H-6, H-6').^{13}C NMR (126 MHz, D_2O) δ 155.32, 130.81, 125.37, 119.03, 117.12, 97.52, 71.83, 69.33, 69.04, 67.85, 60.91. ESI-MS for $C_{12}H_{14}Cl_2O_6$ *m*/*z* calcd for (M + Na^+) 347.0065: found 347.0054.

*3-Chlorophenyl α-D-galactoside (**2.2d**).* Mpt = 139–142 °C; $[\alpha]^{20}_D$ = +195.7 (c=0.499, H_2O); ^{1}HNMR (400 MHz, D_2O) δ 7.34 (t, *J* = 8.2 1 H, ArH), 7.25 (t, *J* = 2.1, 1 H, ArH), 7.18–7.07 (m, 2 H, ArH), 5.67 (d, *J* = 3.9, 1 H, H-1), 4.11–3.93 (m, 4 H, H-2, H-3, H-4, H-5), 3.74–3.65 (m, 2 H, H-6, H-6'). ^{13}C NMR (126 MHz, D_2O) δ 156.78, 134.25, 130.69, 122.91, 117.47, 115.62, 97.28, 71.79, 69.35, 69.05, 67.95, 60.93. ESI-MS for $C_{12}H_{15}ClO_6$ *m*/*z* calcd for (M + Na^+) 313.0455: found 313.0441.

*4-Chlorophenyl α-D-galactoside (**2.2e**).* Mpt = 157–161 °C; $[\alpha]^{20}_D$ = +193.5 (c=0.452, H_2O); ^{1}HNMR (600 MHz, D_2O) δ 7.37 (d, *J* = 9.0, 2 H, ArH), 7.15 (d, *J* = 9.0, 2 H, ArH), 5.64 (d, *J* = 3.8, 1 H, H-1), 4.05–3.99 (m, 4 H, H-2, H-3, H-4, H-5), 3.70 (app d, *J* = 5.9, 2 H, H-6, H-6'), 3.34 (d, *J* = 1.0, OH). ^{13}C NMR (151 MHz, D_2O) δ 154.79, 129.45, 127.27, 118.72, 97.46 (C-1), 71.72, 69.36, 69.08, 67.98, 60.95. ESI-MS for $C_{12}H_{15}ClO_6$ *m*/*z* calcd for (M + Na^+) 313.0455: found 313.0445.

*4-Methylphenyl α-D-galactoside (**2.2f**).* Mpt = 160–163 °C; $[\alpha]^{20}_D$ = +174.3 (c=0.562, H_2O); ^{1}HNMR (500 MHz, D_2O) δ 7.22 (d, *J* = 8.4, 2 H, ArH), 7.08 (d, *J* = 8.5, 2 H, ArH), 5.60 (d, *J* = 3.9, 1 H, H-1), 4.13–4.03 (m, 3 H, H-3, H-4, H-5), 3.98 (dd, 1 H, H-2), 3.70 (app d, *J* = 5.7, 2H, H-6, H-6'), δ 2.29 (s, 3H, Ar CH_3), ^{13}C NMR (126 MHz, D_2O) δ 153.92, 133.03, 130.13, 117.44, 97.65, 71.55, 69.40, 69.11, 68.05, 60.94, 19.55. ESI-MS for $C_{13}H_{18}O_6$ *m*/*z* calcd for (M + Na^+) 293.1001: found 293.0990.

2.4.2. Phenyl 2,3,4,6-tetra-*O*-acetyl-α-D-(1-^{2}H)galactopyranoside (2.3b).

1,2,3,4,6-Penta-*O*-acetyl-D-(1-2H)galactose (supporting information) (1.6 g, 4.0 mmol) and phenol (1.5 g, 8 mmol) were dissolved in anhydrous CH_2Cl_2 (50 mL) to which $SnCl_4$ (0.8 mL, 8 mmol) was added. This solution was stirred at room temperature for 48 h at which time TLC analysis (40% EtOAc-Hexane) showed that the reaction was complete. Following the addition of water (35 mL) and saturated $NaHCO_3$ (20 mL) the product was extracted from the aqueous layer using CH_2Cl_2 (3 × 35 mL). The combined organic layer was washed with brine and dried (Na_2SO_4). Removal of the volatiles under reduced pressure gave the crude product in a quantitative yield. Flash chromatographic purification (25% EtOAc-Hexane) gave 340 mg of the pure α-anomer 2.3a and 600 mg of an anomeric mixture.

$[\alpha]^{20}_{D}$ = +165.8 (c=0.32, $CHCl_3$), ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.34–7.27 (m, 2 H, ArH), 7.09–7.03 (m, 3H, ArH), 5.58 (dd, 1 H, J = 10.8, 3.4, H-3), 5.53 (dd, 1 H, J = 3.4, 1.2, H-4), 5.28 (d, 1 H, J = 10.8, H-2), 4.38–4.33 (t, 1 H, H-5), 4.16–4.03 (m, 2 H, H-6, H-6'), 2.17, 2.07, 2.03, 1.94 (4s, 12 H, 4 × CH_3). ^{13}C NMR: (151 MHz, solvent) δ 170.43, 170.33, 170.20, 170.06 (4 × C=O), 156.26 (ArH), 129.62 (ArH), 122.98 (ArH), 116.75 (ArH), 67.89 (C-4), 67.73 (C-2), 67.54 (C-3), 67.10 (C-5), 61.47 (C-6), 20.74, 20.69, 20.64, 20.57 (4 × CH_3). ESI-MS for $C_{20}H_{23}DO10$ m/z calcd for (M + Na+) 448.1330: found 448.1336.

2.4.3. Phenyl α-D-(1-^{2}H)galactopyranoside (2.1b).

Deprotection of **2.3b** under Zemplen conditions, using catalytic NaOMe in MeOH, was followed by neutralization using Amberlite (H^+) resin. The deprotected compound was further recrystallized from ethanol to obtain 110 mg of **2.1b** (Yield: 60%). Mpt = 132-135 °C; $[\alpha]^{20}_{D}$ =+185.6 (c= 0.0.56, H_2O); ^{1}HNMR (400 MHz, D_2O) δ 7.47–7.36 (m, 2 H, ArH), 7.25–7.11 (m, 3H, ArH), 4.16–4.05 (m, 3 H, H-3, H-4, H-5), 4.00 (d, 1 H, *J* = 10.1, H-2), 3.71 (app d, 2 H, *J* = 6.2, H-6, H-6'). ^{13}C NMR: (151 MHz, D_2O) δ 129.82 (ArH), 123.07 (ArH), 117.34 (ArH), 71.60, 69.41, 69.12, 67.95, 60.95. ESI-MS for $C_{12}H_{15}DO_6$ *m*/*z* calcd for ($M + Na^+$) 280.0907: found 280.0904.

2.4.4. Phenyl 2,3,4,6-tetra-*O*-acetyl-α-D-(2-^{2}H)galactopyranoside (2.3c).

Dimeric 3,4,6-tri-*O*-acetyl-2-deoxy-2-nitroso-α-D-galactopyranosyl chloride **2.5**, made from galactal **2.4** (*15*), was dissolved in anhydrous DMF (100 mL) and to this solution was added phenol (2 g, 2.0 equiv). The resultant mixture was stirred for 70 h at ambient temperature. At this time TLC analysis (40% EtOAc-hexane) indicated complete disappearance of the starting material. The reaction mixture was then diluted with Et_2O (50 mL) and extracted with water (3 × 40 mL), washed with brine (10 mL), dried (Na_2SO_4) and concentrated under reduced pressure to yield a yellow syrup. Purification by flash chromatography (35% EtOAc-hexane) gave phenyl 3,4,6-tri-*O*-acetyl-2-deoxy-2-oximino-α-D-*lyxo*-hexopyranoside **2.6** (1.3 g, 30% over two steps). The glycosylated product **2.6** (1.3 g) was dissolved in CH_3CN (15 mL) that contained 1N HCl (0.5 mL) and CH_3CHO (0.2 mL). The reaction mixture was stirred at ambient temperature for 5 h, at which time the reaction

mixture was extracted with EtOAc (3 × 25 mL). The combined organic layer was then dried (Na_2SO_4) and volatiles were removed under reduced pressure to give a pale yellow oil. The crude product was dissolved in anhydrous THF (30 mL) and the solution was cooled to 0 °C to which a solution of $NaBD_4$ (100 mg) in cold D_2O (2 mL) was added and the mixture was stirred at 0 °C for 1 h and then at room temperature for 2 h. Cautious addition of glacial AcOH (5 mL) destroyed the excess borodeuteride and the resultant solution was washed with MeOH (20 mL). Removal of the volatiles under reduced pressure gave a pale yellow syrup that was directly acetylated under standard conditions, pyridine (5 mL) and Ac_2O (4 mL) at 25 °C. After the addition of water (30 mL), the product was extracted into CH_2Cl_2 (3 × 30 mL), and the organic layer was washed with cold 10% H_2SO_4 (10 mL), brine (2 × 10 mL) and dried (Na_2SO_4). Removal of the volatiles under reduced pressure gave a pale yellow syrup that was purified by flash chromatography to give pure phenyl 2,3,4,6-tetra-*O*-acetyl-α-D-(2-^{2}H)galactopyranoside **2.3c** (350 mg, 30% over 3 steps. $[\alpha]^{20}_D$ = +168.3 (c 0.247, $CHCl_3$) ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.38–7.25 (m, 2 H, ArH), 7.14–7.03 (m, 3 H, ArH), 5.78 (s, 1 H, H-1), 5.58 (d, 1 H, *J* = 3.4, H-3), 5.53 (d, 1 H, *J* = 3.3, H-4), 4.36 (t, 1 H, *J* = 7.1, H-5), 4.09 (m, 1 H, H-6), 4.15 (m, 1 H, H-6'), 2.17, 2.08, 2.04, 1.94 (4s, 12 H, 4 × CH_3); ^{13}C NMR (151 MHz, $CDCl_3$) δ 170.42, 170.33, 170.21, 170.06, 156.26, 129.62 (ArH), 122.99 (ArH), 116.75 (ArH), 94.78 (C-1), 67.90 (C-4), 67.48 (C-3), 67.12 (C-5), 61.47 (C-6), 20.74, 20.68, 20.64, 20.56.(4×CH_3, $COCH_3$). ESI-MS for $C_{20}H_{23}DO_{10}$ *m*/*z* calcd for (M + Na^+) 488.1330: found 448.1321.

Complete deprotection of the per-acetylated compound was accomplished under Zemplen conditions using NaOMe/MeOH followed by neutralization using Amberlite (H^+) resin, to obtain 175 mg (90%) of pure phenyl α-D-(2-^{2}H)galactopyranoside **2.1c**. Mpt = 133-135 °C; $[\alpha]^{20}_D$ = +196.0; ^{1}H NMR: (600 MHz, D_2O) δ 7.41 (dd, *J* = 8.3, 7.5, 2 H, ArH), 7.20 (d, *J* = 7.9, 2 H, ArH), 7.15 (s,1 H, ArH), 5.68 (s, 1 H, H-1), 4.06–4.09 (m, 2 H, H-3, H-4), 4.06 (d, *J* = 3.3, 1 H, H-5), 3.70 (app d, *J* = 6.2, 2 H, H-6, H-6'). ^{13}C NMR (151 MHz, D_2O) δ 129.82 (ArH), 123.08 (ArH), 117.35 (ArH), 97.26 (C-1), 71.63 (C-5), 69.35 (C-3), 69.12 (C-4), 60.95 (C-6). ESI-MS for $C_{12}H_{15}DO_6$ *m*/*z* calcd for (M + Na^+) 280.0907: found 280.0905.

2.4.5. Dimeric-3,4,6-tri-*O*-acetyl-2-deoxy-2-nitroso-α-D-(3-^{2}H)galactopyranosyl chloride (2.8).

NOCl gas, which was prepared by dropwise addition of saturated solution of sodium nitrite (13 g) in water (20 mL) into concentrated HCl (50 mL), was bubbled through a cooled solution (–45 °C) of thoroughly dried deuterated galactal **2.7** (supporting information) in EtOAc (40 mL). After the addition was complete, the temperature was allowed to warm up to 0 °C over a period of 20 min, and the volatiles were then removed under reduced pressure to give a bluish-green mass, which was further dried under vacuum for 2 h. The resultant crude product was used in the subsequent reaction without further purification.

2.4.6. Phenyl 2,3,4,6-tetra-*O*-acetyl α-D-(3-^{2}H)galactopyranoside (2.3d).

The crude dimeric 3,4,6-tri-*O*-acetyl-2-deoxy-2-nitroso-α-D-(3-^{2}H)galactopyranosyl chloride **2.8** was dissolved in anhydrous DMF (50 mL) and to this solution phenol (800 mg, 2.0 equiv) was added. Stirring was continued for 70 h at ambient temperature when TLC analysis (40% EtOAc-Hexane) showed the presence of the desired product (R_f = 0.4). Subsequently, the reaction was diluted with Et_2O (50 mL) and extracted with water (3 × 30 mL). The combined ethereal layers were pooled together, washed with brine (20 mL), dried over Na_2SO_4 and concentrated under vacuum to give a yellowish syrup which was purified by flash chromatography (35% EtOAc-Hexane) to yield 330 mg (30% over 2 steps) of phenyl 3,4,6-tri-*O*-acetyl-2-deoxy-2-oximino α-D-(3-^{2}H)*lyxo*-hexopyranoside **2.9**. After dissolving **2.9** (330 mg) in CH_3CN (5 mL) containing 1N HCl (0.5 mL) and CH_3CHO (0.3 mL) the mixture was stirred at ambient temperature for 5 h, when water (20 mL) was added. The aqueous layer was extracted with EtOAc (3 × 25 mL) and, the combined organic layer was dried (Na_2SO_4) and volatiles were then removed under reduced pressure to give a pale yellow oil. The resultant material was dissolved in anhydrous THF (30 mL) and the solution was cooled to 0 °C. Addition of a solution of $NaBH_4$ (100 mg) in cold H_2O (2 mL) was followed by stirring the reaction at 0 °C for 1 h and then at room temperature for 2 h. Cautious addition of glacial AcOH (5 mL) destroyed the excess borohydride and the resultant solution was washed with MeOH (20 mL). Removal of the volatiles under reduced pressure gave a pale yellow syrup which was directly acetylated under standard conditions, stirring in

pyridine (5 mL) and Ac_2O (4mL) for 15 h at 25 °C. Flash chromatographic purification gave pure **2.3d** (80 mg, 40% yield over 3 steps). $[\alpha]^{20}_D$ = +167.0 (c = 0.15, $CHCl_3$) ^{1}H NMR: (400 MHz, $CDCl_3$) δ 7.34–7.28 (m, 2H, ArH), 7.09–7.03 (m, 3 H, ArH), 5.78 (d, 1 H, *J* = 3.6, H-1), 5.53 (d, 1 H, *J* = 1.1, H-4), 5.29 (d, 1 H, *J* = 3.7, H-2), 4.36 (t, 1 H, *J* = 6.6, H-5), 4.12 (dd, 1 H *J* = 11.3, 6.2, H-6), 4.06 (dd, 1 H, *J* = 11.3, 7.1, H-6'), 2.17, 2.07, 2.03, 1.94 (4 × CH_3); ^{13}C NMR (151 MHz, $CDCl_3$) δ 170.41, 170.32, 170.20, 170.05 (4 × C=O, 4 × $OCOCH_3$), 156.28 (Ar C), 129.62 (Ar C), 122.99 (Ar CH), 116.77 (Ar CH), 94.84 (C-1), 67.86 (C-4), 67.74 (C-2), 67.13 (C-5), 61.47 (C-6), 20.73, 20.67, 20.63, 20.55 (4 × CH_3, 4 × $OCOCH_3$). ESI-MS for $C_{20}H_{23}DO_{10}$ *m*/*z* calcd for (M + Na^+) 488.1330: found 448.1330.

Complete deprotection of the peracetylated compound was accomplished under Zemplen conditions using NaOMe/MeOH followed by neutralization using Amberlite (H^+) resin, to obtain pure phenyl α-D-(3-^{2}H)galactopyranoside **2.1d** in quantitative yield. Mpt = 137–140 °C; $[\alpha]^{20}_D$ = +168.3 (c = 0.25, H_2O); ^{1}H NMR (600 MHz, D_2O) δ 7.46–7.38 (m, 2 H, ArH), 7.20 (d, *J* = 8.2, 2 H, ArH), 7.15 (t, *J* = 7.4, 1 H, ArH), 5.68 (d, *J* = 3.9, 1 H, H-1), 4.09 (t, *J* = 6.2, 1 H, H-5), 4.06 (s, 1 H, H-4), 3.99 (d, *J* = 3.9, 1 H, H-2), 3.70 (app d, *J* = 6.2, 2 H, H-6, 6'). ^{13}C NMR (151 MHz, D_2O) δ 129.82 (ArH), 123.07 (ArH), 117.35 (ArH), 97.29 (C-1), 71.63 (C-5), 69.07 (C-4), 67.99 (C-2), 60.95 (C-6). ESI-MS for $C_{12}H_{15}DO_6$ *m*/*z* calcd for (M + Na^+) 280.0907: found 280.0900.

2.4.7. Phenyl 2,3,4,6–tetra-*O*-acetyl α-D-(2-^{2}H, 3-^{2}H)galactopyranoside (2.3e).

A solution of the ketone made from oxime **2.9** (285 mg, made as above) in anhydrous THF (30 mL) was cooled to 0 °C when a solution of $NaBD_4$ (100 mg) in cold D_2O (2 mL) was added dropwise. After the reaction was stirred at 0 °C for 1 h and then at room temperature for 2 h the cautious addition of glacial AcOH (5 mL) destroyed the excess borodeuteride. The resultant solution was washed with MeOH (20 mL) and removal of the volatiles under reduced pressure gave a pale yellow syrup, which was directly acetylated under standard conditions, pyridine (5 mL) and Ac_2O (5 mL) at 25 °C. The crude product was purified by flash chromatography (25% EtOAc-Hexane) to yield pure peracetylated 2,3-dideuteriated compound **2.3e** (100 mg, 30% over 2 steps). ^{1}H NMR(400 MHz, $CDCl_3$) δ 7.38–7.24 (m, 2 H, ArH), 7.13–7.02 (m, 3 H, ArH), 5.78 (s, 1 H, H-1), 5.53 (d, 1 H,

J = 1.0, H-4), 4.36 (t, 1 H, *J* = 6.5, H-5), 4.18–3.99 (m, 2 H, H-6, H-6'), 2.17, 2.05, 2.03, 1.94 (4 × CH_3); ^{13}C (151 MHz, $CDCl_3$) δ 170.42, 170.33, 170.21, 170.06, 156.26, 129.62 (Ar C), 122.99 (Ar C) 116.75 (Ar, C), 94.78 (C-1), 67.86 (C-4), 67.12 (C-5), 61.47 (C-6), 20.74, 20.68, 20.64, 20.57 (4 × CH_3, $OCOCH_3$). ESI-MS for $C_{20}H_{22}D_2O_{10}$ *m*/*z* calcd for ($M + Na^+$) 449.1393: found: 449.1371.

Final deprotection of the compound was accomplished under Zemplen conditions using NaOMe/MeOH followed by neutralization using Amberlite (H^+) resin gave 40 mg (80%) of pure phenyl α-D-(2-^{2}H,3-^{2}H)galactopyranoside **2.1e**. Mpt = 139–140 °C; $[\alpha]^{20}_D$ = +182.9; ^{1}H NMR (600 MHz, D_2O) δ 7.44–7.38 (m, 2 H, ArH), 7.20 (d, 2 H, *J* = 8.5, ArH), 7.18–7.11 (m, 1 H, ArH), 5.68 (s, 1 H, H-1), 4.08 (t, 1 H, *J* = 6.2, H-5), 4.05 (s, 1 H, H-4), 3.70 (d, 2 H, *J* = 6.3, H-6, H-6'). ^{13}C (151 MHz, D_2O) δ 156.09 (ArH), 129.82 (ArH), 123.07 (ArH), 117.34 (ArH), 97.26 (C-1), 71.63 (C-5), 69.07 (C-4), 60.94 (C-6). ESI-MS for $C_{12}H_{14}D_2O_6$ *m*/*z* calcd for ($M + Na^+$) 281.0970: found 281.0966.

2.4.8. Cloning of α-Galactosidase.

The *MelA* gene encoding α-galactosidase was PCR amplified from *Citrobacter freundii* genomic DNA (ATCC 8090D). The amplification was performed in 5% DMSO using the following primers (forward, 5'-CGCGGCTAGCATGATGTCTGCACCC-3'; reverse, 5'GCGCCTCGAGTTAACGGTGCAGCCAG-3'), which introduced *NheI* and *XhoI* restriction sites (underlined) in the forward and reverse primer respectively. The PCR fragment was purified, digested with *NheI* and *XhoI*, inserted into correspondingly digested pET28a vector (Novagen), and transformed in *Escherichia coli* BL-21(DE3). The plasmid DNA was isolated from a single colony and was verified by restriction digest and DNA sequencing by Macrogen using T7 promoter and T7 terminator primers.

2.4.9. *MelA* Expression and Purification.

For expression of the *MelA* gene, cells were inoculated in Luria broth media and then expressed in YT media supplemented with 1% kanamycin at 37 °C to an OD_{600nm} of 0.5 before induction with 0.5 mM IPTG. The cultures were incubated at 25 °C for 5 h and then centrifuged (10 min at 8400 × g) and the pellet from 1 L of culture was re-suspended in

binding buffer containing 5 mM imidazole. The cells were lysed open using 1% lysozyme (from chicken egg white) and protease inhibitor cocktail tablet followed by sonication (20 s on/40 s off cycle at capacity of 60%) to ensure complete lysis of the cells. The lysate was centrifuged at 4 °C (34,000 × g) in order to remove the cell debris and intact cells. The clear supernatant was collected and filtered through a 0.45 micron filter before loading on to a HiTrap Chelating HP column 5 mL, which had been pre-equilibrated with the binding buffer. The column was washed sequentially with 60 mM, 100 mM and 150 mM imidazole before eluting the protein with 250 mM of imidazole. Fractions containing pure His-*MelA*, as determined by 10% SDS–PAGE, were pooled together and dialyzed, at 4 °C, three times against 4 L of Tris buffer (20 mM, pH 7.0) that contained NaCl (100 mM) and DTT (6 mM). The protein was then concentrated by centrifugation through a 10 kDa filter and its concentration was assessed (Bradford Assay or Nanodrop).

2.4.10. Enzyme Kinetics.

All kinetics assays were conducted in either 0.2 cm or 1.0 cm path length quartz cuvettes, unless stated otherwise, using a Cary 300 UV-vis spectrometer equipped with a Cary temperature controller. For all experiments the *MelA* α-galactosidase was pre-incubated in the assay buffer with cofactors at 37 °C for 15 min prior to the addition of the substrate, which initiated the enzymatic reaction. The following buffer (buffer A) at a pH of 7.5 was used for most kinetic experiments: HEPES (50 mM), $MnCl_2$ (1.0 mM), NAD^+ (100 μM), 2-mercaptoethanol (10 mM) and BSA (0.1% w/v). All cofactor solutions were freshly prepared before each set of kinetic assays. The measurements of all deuterium kinetic isotope effects were performed in buffer B, which was identical to buffer A except that the added reducing agent was TCEP (1 mM) instead of 2-mercaptoethanol (10 mM).

2.4.11. Typical Conditions for the Measurement of Michaelis–Menten Parameters.

The concentration of α-galactosidase was chosen such that less than 10% of the total substrate was consumed during the assay. For each assay, the enzyme was incubated in the appropriate buffer at 37 °C for either 15 or 20 min. After which the reaction was initiated by the addition of substrate. The initial rate of hydrolysis was followed spectrophotometrically

at the wavelength of maximal absorbance change. Typically, the substrate concentration was varied between 40 μM and 1 mM and the measured initial rate versus concentration data were fit to a standard Michaelis-Menten equation.

2.4.12. Determination of the pH-Rate Profile.

To determine the effect of pH on enzymatic activity, kinetics parameters were measured over a pH range of 4.0–9.5. The buffers used were NaOAc-HOAc (20 mM, pH 4.0–4.5), MES (20 mM, MES-NaOH, pH 6.0–6.7), HEPES (20 mM, pH 6.5–8.2), and CHES (20 mM, pH 8.5–9.5). Typical assay conditions were: α-galactosidase (final concentration of 4.1 μg/mL) was incubated at 37 °C with the appropriate buffer containing NaCl (50 mM), $MnCl_2$ (1.0 mM), NAD^+ (100 μM), 2-mercaptoethanol (10 mM) and BSA (0.1 % w/v) for 10 min prior to addition of substrate 4NPαG, the hydrolysis reaction was monitored at either 400 nm (pH 6.5–9.0) or 340 nm (pH 4.0–6.0). The difference in extinction coefficients (Δε) for 4NPαG and the released phenolate/phenol was determined at each pH value and, the initial rate measurements were fit to a standard Michaelis-Menten equation.

2.4.13. Kinetic Investigation of NAD^+ Cofactor Dependence.

A sample of MelA (final concentration of 16.3 μg/mL) in Buffer A was pre-incubated for 15 min at 37 °C for 15 min. 4NPαG was added to the solution to initiate the hydrolysis reaction. In another parallel assay, the enzyme was incubated with all other cofactors at pH 7.5 in the presence of 10 mM $NaBH_4$, for the quantitative reduction of NAD^+ to NADH. The enzyme was completely inactive after treatment with $NaBH_4$, but approximately 60% of its activity was restored upon addition of exogenous NAD^+ (200 μM).

2.4.14. Determination of K_d Values for NAD^+.

Solutions containing MelA (final concentration 15.8 μg/mL) and NAD^+ (concentration varied from 20 μM to 4 mM) were pre-incubated in HEPES buffer (50 mM, pH 7.5) containing $MnCl_2$ (1 mM), 2-mercaptoethanol (10 mM) and BSA (0.1% w/v) at 37 °C for 15 min, then 4NPαG (final concentration 100 μM) was added

to initiate the reaction (final volume of 400 µL). The measured initial reaction rates were fit to the Michaelis-Menten equation that includes a term for substrate inhibition.

2.4.15. Determination of K_d Value for Mn^{2+}.

To remove bound metal ions, MelA was dialyzed three times against 4 L of Tris buffer (20 mM, pH 7.1). Solutions containing MelA (final concentration 15.0 µg/mL) and Mn^{2+} (varied from 0–400 µM) were pre-incubated in HEPES buffer (50 mM, pH 7.5) containing NAD^+ (100 µM) 2-mercaptoethanol (10 mM) and BSA (0.1% w/v) at 37 °C for 15 min, then 4NPαG (final concentration 100 µM) was added to initiate the reaction (final volume of 400 µL). In a separate experiment a solution of MelA (200 µL; 3 mg/mL) was dialyzed overnight against a Tris-HCl buffer (500 mL, 20 mM, pH 7.1) containing EDTA (2mM) and NaCl (100 mM). The resultant enzyme solution was then dialyzed twice, for 4 h, against fresh Tris-HCl buffer (500 mL, 20 mM, pH 7.1) containing NaCl (100 mM). The enzyme activity was then monitored as a function of Mn^{2+} concentration as detailed above. After subtraction of the rate in the absence of Mn^{2+} the resultant initial rates were fit to a standard Michaelis-Menten equation.

2.4.16. Deuterium Kinetic Isotope Effect Measurements.

To determine KIEs (k_H/k_D) on V separate measurements of the Michaelis-Menten parameters were carried out for each of the following: phenyl α-D-galactopyranoside, phenyl α-D-(1-^{2}H)galactopyranoside, phenyl α-D-(2-^{2}H)galactopyranoside, phenyl α-D-(3-^{2}H)galactopyranoside and phenyl α-D-(2-^{2}H, 3-^{2}H)galactopyranoside. Typically, for each assay MelA α-galactosidase (final concentration of 3.1 µg/mL) was incubated in buffer B for 20 min prior to initiation of hydrolysis by the addition of substrate. The initial rates were measured by monitoring the release of phenol at 270 nm. Substrate concentration was varied between 0.04–1.0 mM. The cofactor solution was prepared freshly and all the data sets were obtained using the same stock of cofactors and enzyme. Each set of kinetic experiments was completed within a day to ensure the smallest possible variation in enzyme activity. All kinetic data was fit to a standard Michaelis-Menten equation.

2.4.17. Conditions for the Measurement of KIEs on *V/K*.

Measurement of individual *V*/*K* values were performed by monitoring the first-order consumption of substrate, whose concentration was typically $\leq 0.1 \times K_m$, at the wavelength of maximum absorbance change (277 nm). Normally, MelA α-galactosidase (final concentration in the range of 40–50 μg/mL) was incubated in buffer B for 20 min at 37 °C before addition of the appropriate substrate initiated the reaction. The absorbance was monitored at 277 nm until the hydrolysis was complete. The absorbance versus time data were fit to a standard first-order rate equation, for all cases it was determined that the fitting residuals were randomly scattered around zero. Measurements for the *protio* and *deuterio* substrates were performed in an alternating fashion and each measurement was repeated 3 times. The resulting $^{D}V/K$ values were calculated using a standard weighted average method (*16*).

2.4.18. Product Studies.

^{1}H NMR spectroscopy (500 MHz) was employed to identify the stereochemical course of the enzyme-catalyzed reaction. The reaction conditions involved incubating, at 37 °C, enzyme (4.2 μg/ml) in buffer A containing CD_3OD (5 M). After addition of 4NPαG (2.5 mg) the reaction was allowed to proceed at 37 °C until TLC analysis (1:4 v/v MeOH:EtOAc) showed no remaining starting material. Removal of the enzyme by centrifugal ultrafiltration at 4 °C was followed by stirring the resultant solution with Chelex resin (Fluka) for 60 min. Following filtration, the solution was lyophilized and the resulting solid was dissolved in D_2O and the ^{1}H NMR spectrum was acquired.

2.4.19. Linear Free Energy Relationship – Brønsted Analysis.

A series of substrates with varying leaving groups were synthesized to perform a Brønsted analysis. Full Michaelis-Menten curves were measured, in buffer A, for each substrate using the protocol listed above.

2.5. Results

In order to test whether *C. freundii* MelA galactosidase displays any catalytic reactivity towards phosphorylated substrates, 4-nitrophenyl 6-phospho-α-D-galactoside was synthesized (supplementary material). However, it was shown that MelA hydrolyzed neither 4-nitrophenyl 6-phospho-α-D-galactopyranoside nor 4-nitrophenyl β-D-galactopyranoside while it was catalytically active against 4-nitrophenyl α-D-galactopyranoside (data not shown). The absolute requirement for catalytic activity of this GH4 α-galactosidase for NAD^+ was shown by a standard borohydride reduction experiment (supporting information, **Figure 2S 1**). The binding constants for the two cofactors NAD^+ and Mn^{2+} were determined by monitoring the hydrolysis of 4NPαG as a function of co-factor concentration at a pH of 7.5 and at a temperature of 37 °C. The dependence of enzyme activity on the concentration of NAD^+ exhibited marked substrate inhibition (**Figure 2-1**), and the associated kinetic parameters on the MelA α-galactosidase-catalyzed reaction for K_m and K_{is} are 240 ± 40 μM and 2.0 ± 0.4 mM, respectively. In the absence of Mn^{2+} the enzyme displayed an activity of approximately 10% of its maximal value and after subtraction of this residual activity from the initial rates measured in the presence of Mn^{2+}, an estimate of 220 ± 100 μM was made for the binding constant of Mn^{2+} (K_d) (supporting information, **Figure 2S 2**). All other kinetic parameters were evaluated using buffers that contained NAD^+ and Mn^{2+} at concentrations of 100 μM and 1.0 mM, respectively.

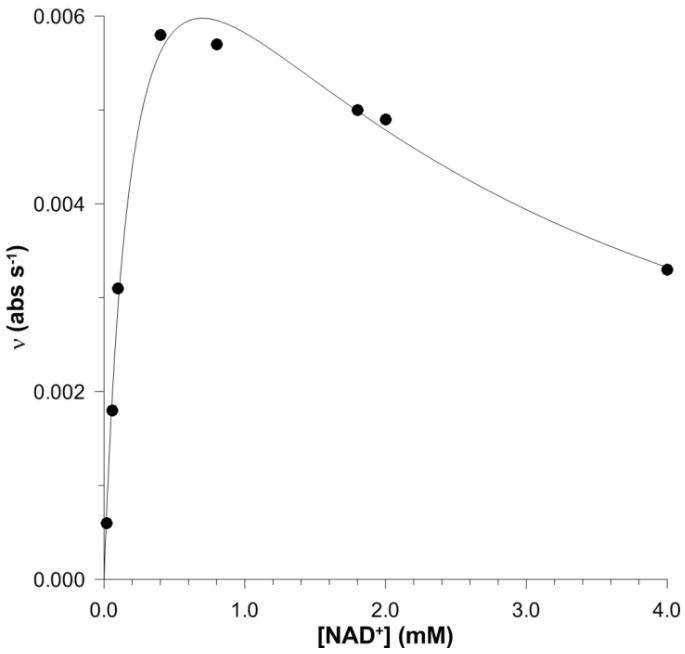


Figure 2-1. Ligand binding curve for NAD^+ using PNPαG as substrate.

Product Studies. In the absence of reducing agent MelA gave no product even after 48 h incubation, and when the reaction was performed with mercaptoethanol in the presence of methanol (5 M) no methyl galactoside was produced. However, an additional new anomeric resonance, in addition to galactose, appeared in the ^{1}H NMR spectrum at a chemical shift of 5.8 ($J_{1,2}$ = 5.5 Hz) that is tentatively assigned to be an α-galactoside.

The variations of the kinetic parameters k_{cat}/K_m and k_{cat} for the enzyme-catalyzed hydrolysis of 4-nitrophenyl α-D-galactopyranoside as a function of pH are listed in **Table 2-1**. The kinetic data for k_{cat} was fit to a classic bell-shaped pH-rate curve (**Figure 2-2**), however, the resultant calculated pK_a values, for the two catalytically important ionizations, were not well resolved with both values being 7.8 ± 0.5. In addition, the gross shape of the pH-rate curve on k_{cat}/K_m is similar with the maximal activity being around a pH value of 8.0 (Figure S3, supporting information), but the fit of the kinetic data to a standard bell-shaped profile is particular ill-defined.

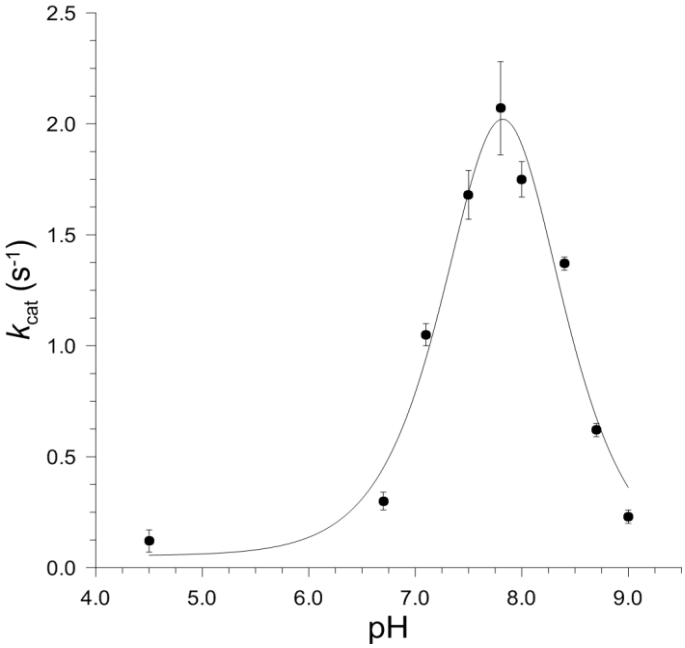


Figure 2-2. pH dependence of k_{cat} for the MelA-catalyzed hydrolysis of 4NPαG.

Table 2-1. Michaelis-Menten kinetic parameters for MelA-catalyzed hydrolysis of 4NPαG as a function of pH at 37 °C.

pH	k_{cat} (s^{-1})	$10^{-3} \times k_{cat}/K_m$ (M^{-1} s^{-1})
4.5	0.12 ± 0.05	0.091 ± 0.070
6.7	0.30 ± 0.04	0.81 ± 0.26
7.1	1.05 ± 0.05	2.7 ± 0.4
7.5	1.68 ± 0.11	5.8 ± 1.1
7.8	2.07 ± 0.21	9.9 ± 1.9
8.0	1.75 ± 0.08	8.9 ± 1.4
8.4	1.37 ± 0.03	15.6 ± 1.5
8.7	0.62 ± 0.03	5.0 ± 1.2
9.0	0.23 ± 0.03	0.96 ± 0.42

A panel of seven aryl α-D-galactosides was synthesized by following literature procedures using peracetylated galactose as the starting material (see experimental section for details). Full Michaelis-Menten curves were measured at a pH of 7.5 for each substrate

(**Table S1**, supporting information) and the associated Brønsted β_{lg} values on k_{cat} and k_{cat}/K_m are –0.01 ± 0.02 and 0.02 ± 0.04, respectively (**Figure 2-3**).

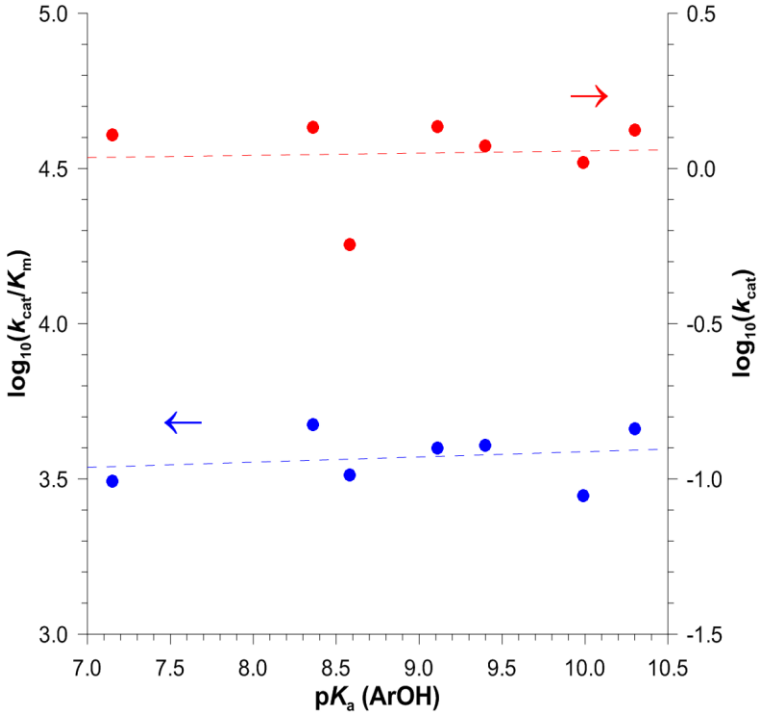
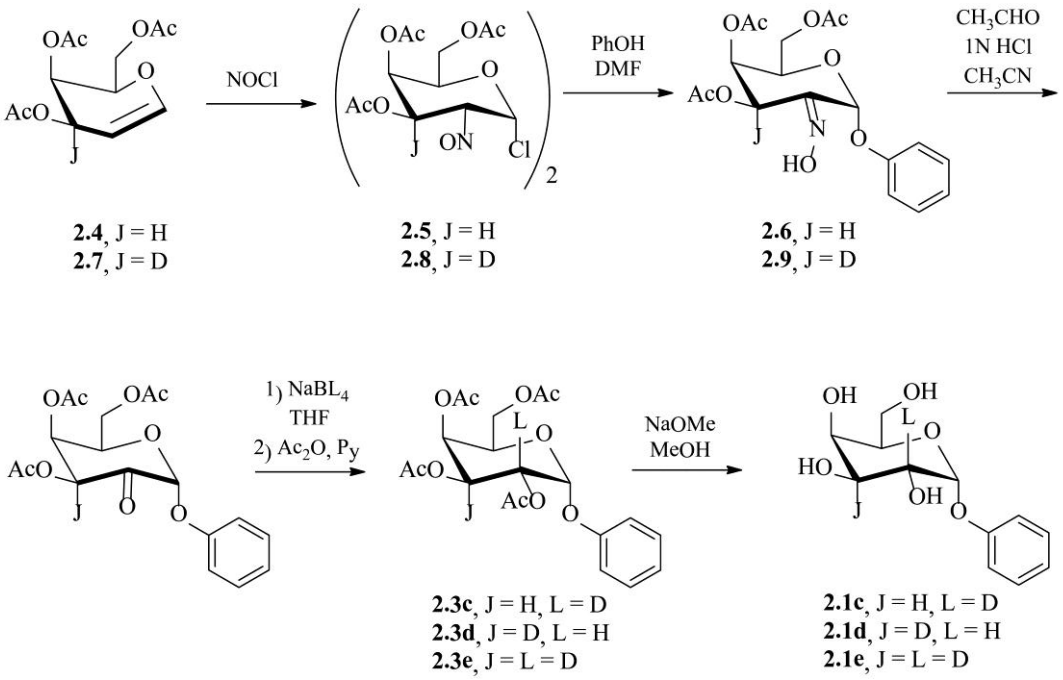


Figure 2-3. Brønsted plots for MelA-catalyzed hydrolysis, measured at 37 °C, of a series of aryl α-D-galactopyranosides and the associated pK_a value for the leaving group phenol. Shown in red and blue are the kcat and the k_{cat}/K_m kinetic data, respectively.

Synthesis of Deuterated Substrates. Phenyl α-D-(1-^{2}H)galactopyranoside was made using peracetylated (1-^{2}H)galactose (supporting information) as the starting material. All other deuterated phenyl α-D-galactopyranosides were made by the appropriate choice of reducing agent and starting tri-*O*-acetal D-galactal (**2.4** or **2.7**, 3,4,6-tri-*O*-acetyl-1,5-anhydro-2-deoxy-D-*lyxo*-hex-1-enitol) using the nitrosyl chloride addition reaction (**Scheme 2-2**) pioneered by Lemieux and co-workers (*17*). Deuterated **2.7** was made by reduction of a known 3-keto compound (*18*) using sodium borodeuteride in place of borohydride (*19*) followed by a standard acetylation reaction (supporting information).



Scheme 2-2. Synthetic reactions used in the synthesis of C-2 and C-3 deuterated PαG substrates.

Deuterium Kinetic Isotope Effects. In order to measure KIEs on *V/K* the substrate depletion method was used (experimental section). Of note, in these experiments TCEP was used as the reducing agent because in the presence of 2-mercaptoethanol the background absorbance at 277 nm displayed a slow increase over time, a phenomenon that was traced to the reducing agent, and this resulted in large errors for the measured rate constants. In addition, to determine if the measured deuterium KIEs for H-2 and H-3 report on two sequential reactions that involve cleavage of a C–H bond, namely: (i) hydride transfer to NAD^+ during oxidation of the 3-OH group; and (ii) base-catalyzed removal of the C2-proton, or if they originate from a single transition state, $^{D}V/K$ effects on deuterated substrates were measured using the dideuterated substrate phenyl α-D-(2-^{2}H,3-^{2}H)galactopyranoside. For consistency, KIEs on V_{max} were also measured with TCEP as the reducing agent. In these experiments, full Michaelis-Menten kinetic profiles were measured and the resultant V_{max} values were used to determine the ^{D}V effects. **Tables 2-2** and **2-3** list the calculated KIEs for the MelA-catalyzed hydrolysis of PαG.

Table 2-2. Kinetic Isotope Measurements on *V*/*K* for the MelA α-Galactosidase-Catalyzed Hydrolysis of Phenyl α-D-Galactopyranoside at a pH of 7.5 and a temperature of 37 °C.

Isotopologue 1	Isotopologue 2	$^{D}V/K^{a}$
PαG	1-(^{2}H)PαG	1.13 ± 0.07
PαG	2-(^{2}H)PαG	1.74 ± 0.06
PαG	3-(^{2}H)PαG	1.74 ± 0.05
3-(^{2}H)PαG	2,3-($^{2}H_2$)PαG	1.71 ± 0.12
2-(^{2}H)PαG	2,3-($^{2}H_2$)PαG	1.71 ± 0.13

[a] Ratio of *V*/*K*(isotopologue 1)/*V*/*K*(isotopologue 2).

Table 2-3. Kinetic Isotope Measurements on *V* for the MelA α-Galactosidase-Catalyzed Hydrolysis of Deuterated Phenyl α-D-Galactopyranosides at a pH of 7.5 and a Temperature of 37 °C.

Isotopologue 1	Isotopologue 2	$^{D}V^{a}$
PαG	1-(^{2}H)PαG	1.06 ± 0.07
PαG	2-(^{2}H)PαG	0.91 ± 0.04[b]
PαG	3-(^{2}H)PαG	1.02 ± 0.06
3-(^{2}H)PαG	2,3-($^{2}H_2$)PαG	0.91 ± 0.06[c]
2-(^{2}H)PαG	2,3-($^{2}H_2$)PαG	1.01 ± 0.06

[a] Ratio of V_{max}(isotopologue 1)/V_{max}(isotopologue 2). [b] Repeat measurement 0.99 ± 0.08. [c] Repeat measurement 0.98 ± 0.08.

2.6. Discussion

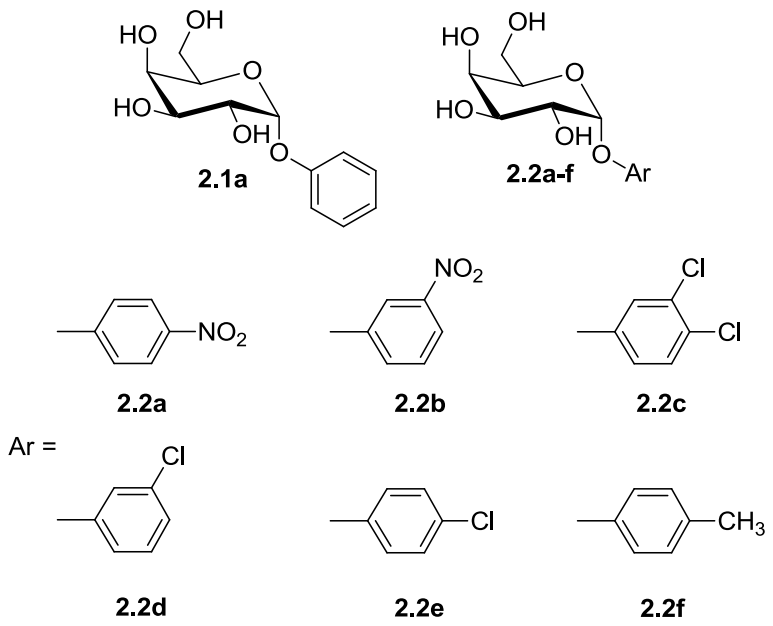
The MelA gene from *E. coli*, which was proposed to encode for a GH4 α-galactosidase, has previously been cloned and sequenced (*11*). The sequence data showed that MelA codes for a 450 amino acid long protein that has a predicted molecular weight of 50.6 kDa (*11*). In the present work, the MelA gene from *Citrobacter freundii* was cloned and recombinantly expressed in *E. Coli* in order to perform a detailed mechanistic study on this GH4 glycosidase. As with other Family-4 glycosyl hydrolases (*6, 9, 10*), the MelA enzyme from *C. freundii* requires two cofactors, NAD^+ and Mn^{2+}, as well as reducing conditions for maximal activity. Moreover, this α-galactosidase is incapable of hydrolyzing 4-nitrophenyl 6-phospho-α-D-galactopyranoside. Anggraeni *et al.* showed that for natural substrates the rMel4A α-galactosidase from *Bacillus halodurans* displays higher turnover and catalytic efficiencies (k_{cat} and k_{cat}/K_m values) for the disaccharide melibiose [α-D-Gal*p*-(1→6)-D-Glc*p*] over the trisaccharide raffinose [α-D-Gal*p*-(1→6)-(α-D-Glc*p*-(1↔2)β-D-Fru*f*] (*12*).

The dependence of the catalytic activity for *C. freundii* MelA α-galactosidase on the presence of NAD^+ was shown by incubating the enzyme with NAD^+ in the presence of $NaBH_4$ (10 mM), conditions under which NAD^+ is reduced to NADH. The *C. freundii* α-galactosidase was completely inactive under these condition, but after addition of a solution of NAD^+ (200 μM) to the reaction medium approximately 60% of the initial enzyme activity was restored (**Figure S1**, supporting information). Furthermore, the catalytic activity of the MelA α-galactosidase varied as a function of NAD^+ concentration and the activity profile is consistent with the occurrence of substrate inhibition ($K_d = 240 \pm 40$ μM, $K_{is} = 2.0 \pm 0.4$ mM; **Figure 2-1**). As a result, it was decided to measure enzyme kinetic parameters at sub-saturating concentrations of NAD^+. The variation in enzyme activity as a function of Mn^{2+} concentration showed no evidence for substrate inhibition and all further kinetic experiments were performed at saturating Mn^{2+} levels (1.0 mM; $K_d \approx 0.2$ mM). Anggraeni *et al.* showed, for the Mel4a GH4 α-galactosidase from *Bacillus halodurans*, that in the presence of the reducing agent DTT both Co^{2+} and Fe^{2+} also gave active enzyme (*12*). In contrast to the reported results with the 6-phospho-glucosidases *GlvA* and *BglT* (*6, 9*), which are also members of glycosyl hydrolase family 4 (GH4), no activity was observed with

MelA in the absence of reducing agents. Also, when MelA-catalyzed hydrolysis of 4NPαG was performed in the presence of methanol (5 M) no trapping by methanol was observed. Consequently, the stereochemical outcome of MelA α-galactosidase-catalyzed reactions was inferred, by analogy to other previously characterized GH4 hydrolases, to be a retaining enzyme. Also of note, when the MelA-catalyzed hydrolysis 4NPαG was performed in D_2O the reaction product, galactose, was completely deuterated at C2 as shown by the ^{1}H NMR spectrum where the anomeric protons for both anomers were singlets.

A detailed pH rate profile for the catalyzed hydrolysis of 4NPαG was measured and the kinetic data was fit to a standard bell-shaped curve where the pK_a values for the two catalytically important groups are approximately 7.8–8.0 (**Figures 2-2** and **2S 3**, supporting information).

In order to probe the effects of leaving group ability on catalytic activity a panel of seven aryl α-D-galactopyranoside substrates were synthesized (structures **2.1a**, **2.2a–f**). The derived Brønsted β_{lg} values, for these substrates in which the pK_a of the conjugate acid of the aglycone varied by over 3 pK_a units, on both k_{cat} and k_{cat}/K_m are, within experimental error, indistinguishable from zero (**Figure 2-3**). Therefore, it can be concluded that cleavage of the anomeric C–O bond is not kinetically significant for the MelA α-galactosidase-catalyzed hydrolysis of aryl α-D-galactopyranosides.



Withers and co-workers have also concluded that glycosidic bond cleavage was not a kinetically important event for either the *T. maritima* BglT 6-phospho-β-glucosidase (*6*) or the *B. subtilis* GlvA 6-phospho-α-glucosidase (*9*) GH4 enzymes. A detailed KIE study was undertaken in order to delineate the important TSs for the MelA enzyme-catalyzed reaction. That is, several different isotopologues that possessed a weakly activated leaving group were synthesized (**Scheme 2-2** and supporting information). Listed in **Tables 2-2** and **2-3** are the measure *V*/*K* and *V* KIE values for deuteration at one of the C-1, C-2 or C-3 carbons. Of note, the tabulated ^{D}V and $^{D}V/K$ values are markedly different. At first glance, the magnitude of the $^{D}V/K$ effects measured at both the C-2 (1.74 ± 0.06) and C-3 (1.74 ± 0.05) centers suggest that both oxidation and proton-abstraction are kinetically significant in accordance with previously reported conclusions (*6*, *9*). In order to verify this analysis it was decided to measure the isotope effects associated with deuterated substrates. That is, if the C-2 and C-3 KIEs arise from a virtual TS resulting from two sequential steps it would be expected that the KIE on C-2 should decrease if both isotopologues (C2-H and C2-D) are deuterated at C-3. In other words, if deuteration at C-3 increases the kinetic barrier for hydride transfer then, for a stepwise mechanism, the magnitude of the ^{2}H-KIE from deuteration at C-2 must decrease for the C-3 deuterated substrates relative to that measured with the C-3 non-deuterated galactosides (*20*), i.e., $^{D}V/K$(C-2) > $^{D}V/K$(C-2,3-^{2}H).

Therefore, phenyl α-D-(2-^{2}H, 3-^{2}H)galactoside was synthesized (**Scheme 2-2**) and both the $^{D}V/K$(C-2,3-^{2}H) and $^{D}V/K$(C-3,2-^{2}H) KIEs were measured (**Table 2-2**). Of note, neither $^{D}V/K$(C-2) nor $^{D}V/K$(C-3) shows any significant decrease in magnitude upon incorporation of a remote deuterium (**Table 2-2**).

An estimate of the expected reduction in $^{D}V/K$ values on incorporation of a second deuterium can be made by deriving the kinetic expressions for **Scheme 2-1** in which oxidation (k_3) and proton abstraction (k_5) are separate kinetic events. Thus, the formula for $^{D}V/K$(C-3) is given in equation 1, where $^{D}k_3$ is intrinsic isotope effect on k_3, $^{D}K_{eq}$ is the equilibrium isotope effect (EIE) on this step. The ratios of individual kinetic terms in equation 1 are normally referred to as commitment factors, either forward or reverse (*21*), i.e., $c_f = k_{3H}/k_2$ and $c_r = k_{4H}/k_{5H}(1 + k_{6H}/k_7)$. However, in order to avoid confusion with different commitment factors for the oxidation and proton removal steps the individual kinetic terms are used in the equations.

$$^{D}\left(\frac{V}{K_{C3}}\right)_{C2-H} = \frac{^{D}k_3 + \frac{k_{3H}}{k_2} + \frac{k_{4H}}{k_{5H}}\left(1+\frac{k_{6H}}{k_7}\right)^{D}K_{eq}}{1+\frac{k_{3H}}{k_2} + \frac{k_{4H}}{k_{5H}}\left(1+\frac{k_{6H}}{k_7}\right)} \qquad 1$$

The corresponding expression for $^{D}V/K$(C-2) is given in equation 2 (**Scheme 2-1**). Given the kinetic insignificance of aglycone departure as shown by the β_{lg} value of zero (k_7, **Scheme 2-1**) the term k_{6H}/k_7 must be small ($k_7 >> k_{6H}$).

$$^{D}\left(\frac{V}{K_{C2}}\right)_{C3-H} = \frac{^{D}k_5 + \frac{k_{5H}}{k_{4H}}\left(1+\frac{k_{3H}}{k_2}\right) + \frac{k_{6H}}{k_7}{}^{D}K_{eq}}{1+\frac{k_{5H}}{k_{4H}}\left(1+\frac{k_{3H}}{k_2}\right) + \frac{k_{6H}}{k_7}} \qquad 2$$

Equation 3 can be derived for the expected V/K KIE on C-3 when C-2 is deuterated, if $k_{6H}/k_7 \approx 0$, (supporting information).

$$ {}^{D}\left(\frac{V}{K_{C3}}\right)_{C2-D} = \frac{{}^{D}K_{eq}\left({}^{D}\left(\frac{V}{K_{C2}}\right)_{C3-H} - 1\right) + {}^{D}\left(\frac{V}{K_{C3}}\right)_{C2-H}}{{}^{D}\left(\frac{V}{K_{C2}}\right)_{C3-H}} $$

Equation 3

Using the measured $^{D}V/K$(C-2) and $^{D}V/K$(C-3) values (both = 1.74) in conjunction with the reported EIE ($^{D}K_{eq}$ = 1.18) for oxidation of cyclohexanol by NAD^{+} (*22*) the estimated value for the magnitude of the $^{D}V/K$(C-3,2-^{2}H) effect is 1.50 (supporting information). Of note, the incorporation of a second deuterium at C-2 also results in a secondary equilibrium isotope effect for the hydride transfer reaction (kinetic terms k_{3H} and k_{4H}), and this is expected to be approximately K_H/K_D = 1.054 (*22*). Incorporation of this value into equation 3, which results in modifying the value of $^{D}K_{eq}$ to be 1.24 (1.18 × 1.054), gives an expected effect for $^{D}V/K$(C-3,2-^{2}H) of 1.53 (supporting information).

A similar formula (equation 4, see supporting information) can be derived for the value of $^{D}V/K$(C-2,3-^{2}H). Coincidentally, the expected value for $^{D}V/K$(C-2,3-^{2}H), based on equation 4, is either 1.50 or 1.53 depending on whether or not the secondary EIE correction noted above is included in the calculations.

$$ {}^{D}\left(\frac{V}{K_{C2}}\right)_{C3-D} = \frac{{}^{D}\left(\left(\frac{V}{K_{C2}}\right)_{C3-H} - 1\right) \times {}^{D}K_{eq}}{{}^{D}\left(\frac{V}{K_{C3}}\right)_{C2-H}} + 1 $$

Equation 4

Of note in the current example, the calculated values for both $^{D}V/K$(C-3,2-^{2}H) and $^{D}V/K$(C-2,3-^{2}H) are independent of the intrinsic KIEs ($^{D}k_3$ and $^{D}k_5$) or of any commitment factors. Of note, given the standard error associated with the measured value for

$^{D}V/K$(C-2,3-^{2}H) of 0.12 (**Table 2-1**) there is either a 4% likelihood, i.e., 100 – 50% (value is greater than 1.74) – 46% (value is between 0 and 1.75 standard deviations below 1.71), that the data is consistent with the stepwise mechanism or a 6.7% probability when the EIE value is included in the calculations. Thus, the measured $^{D}V/K$ values on dideuterated substrates of 1.71 (**Table 2-1**) are more consistent with a concerted reaction for oxidation and proton removal.

A single concerted step requires that abstraction of the C-2 proton by the active site base must start prior to completion of C-3 oxidation by hydride transfer to the onboard NAD^{+} co-factor (**Figure 2-4**). In other words, during oxidation at C-3 to form the ketone the C-2 proton becomes progressively more acidic by virtue of the increasing polarization at C-3 and the proximal Mn^{2+} ion until transfer begins, which in this case results in circumventing formation of the proposed ketone intermediate (**Figure 2-4**).

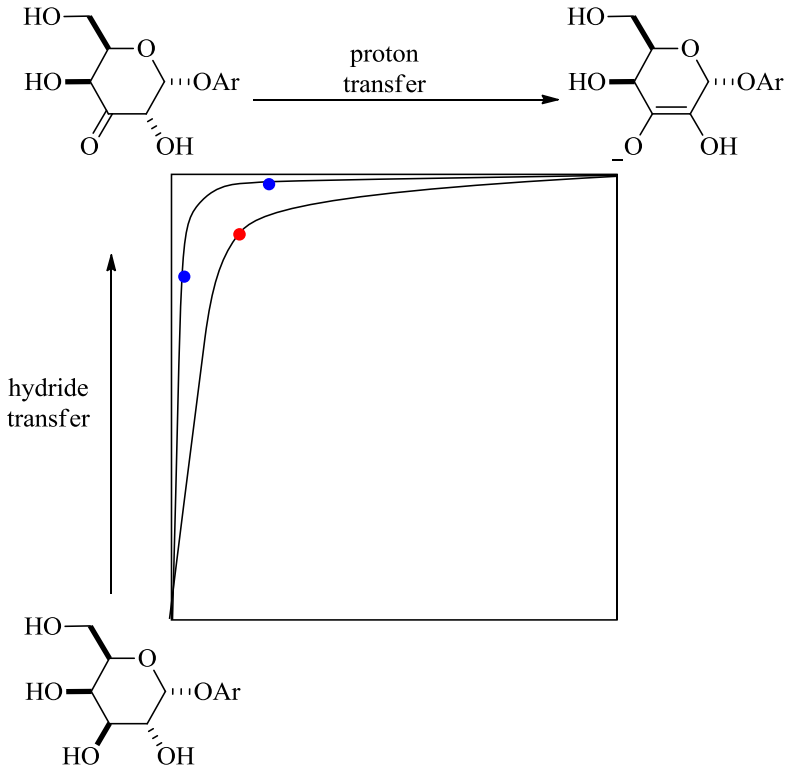


Figure 2-4. More O'Ferrall–Jencks diagram to show possible transition state (red circle) for the concerted reaction from the Michaelis complex to the ene-diol intermediate that circumvents formation of the bound ketone (open circle). Also shown are plausible transition states (blue circle) for the previously proposed stepwise mechanism (references 6 and 9). The enzymatic co-factors, NAD^{+} and Mn^{2+} are omitted from the diagram for clarity reasons.

The magnitude of the $^{D}V/K$ values (on C2 and C3) reported here for the MelA reactions of phenyl α-galactopyranoside are similar to, albeit a little larger than, those reported for the *Bacillus subtilis* GlvA-catalyzed hydrolysis of 4-nitrophenyl 6-phospho-α-glucopyranoside (*9*), a result that implies GH4 enzymatic reactions that involve trans elimination reactions may proceed through very similar transition state structures. In the present study, the negligible β_{lg} values notwithstanding, no information is available concerning the mechanism for the trans-elimination reaction, a situation that contrasts the cis-eliminations encountered with GH4 β-glycosidase where the $^{D}V/K$ on C-2 are significantly higher than those reported here (*6, 10*).

The large relative errors associated with the normal secondary deuterium KIE at the anomeric center, $^{D}V/K$(C-1), make a detailed discussion of the origin of the effect difficult at the present time, although the minimal Brønsted β_{lg} values are inconsistent with the occurrence of any TS rehybridization at this position.

The measured KIEs on *V* for the MelA α-galactosidase-catalyzed hydrolysis of phenyl α-galactoside are markedly different than the $^{D}V/K$ values (cf. **Tables 2-2** and **2-3**). All effects are either equal to 1, within experimental error, or slightly inverse (^{D}V(C-2)). These KIE measurements are consistent with the TS for *V* being associated with a different reaction step. Given that this step must come after those that limit *V*/*K* (*23*) and that the β_{lg} value on *V* is approximately zero the first possible TS for *V* involves the hydration of the enzyme-bound glycal intermediate (k_9, **Scheme 2-1**). However, at the present time given the lack of specific evidence other subsequent steps in the mechanism, such as product release, might be kinetically significant for *V*.

2.7. Conclusions

The MelA α-galactosidase from *Citrobacter freundii* catalyzes hydrolysis of aryl substrates via a mechanism in which the first irreversible step (rate-limiting for *V*/*K*) involves oxidation of the C-3 hydroxyl group and proton abstraction at C-2, a process which starts before completion of the NAD^+-promoted hydride transfer. Subsequent elimination of the

leaving group is kinetically silent, and this event is followed by hydration of the α,β-unsaturated enone intermediate. The catalytic cycle is completed by protonation at C-2 and reduction by the on-board NADH.

2.8. Associated Content

Supporting Information. Synthetic details for 1,2,3,4,6-penta-*O*-acetyl-D-(1-^{2}H)galactose, 4-nitrophenyl 6-phospho-α-D-galactoside (Schemes S1 and S2) and labeled galactal **7**. Derivation of equations for the estimation of expected $^{D}V/K$(C-3,2-^{2}H) and $^{D}V/K$(C-2,3-^{2}H) values for a stepwise mechanism. Michaelis-Menten kinetic parameters for the MelA α-galactosidase-catalyzed hydrolysis of the seven aryl α-D-galactosides are listed in Table S1. Assay of reduced MelA (Figure S1), ligand binding curve for Mn^{2+}. (Figure S2) and pH dependence of k_{cat}/K_m for the MelA-catalyzed hydrolysis of 4NPαG (Figure S3). ^{1}H and ^{13}C NMR spectra for labeled compounds **2.1b–2.1e** and substrates **2.2b–2.2f**. This material is available free of charge via the Internet at http://pubs.acs.org.

2.9. References

1. Henrissat, B. (1991) A classification of glycosyl hydrolases based on amino acid sequence similarities, *Biochem. J. 280*, 309-316.

2. Vocadlo, D. J., and Davies, G. J. (2008) Mechanistic insights into glycosidase chemistry, *Curr. Opin. Chem. Biol. 12*, 539-555.

3. Rao, S. T., and Rossmann, M. G. (1973) Comparison of super-secondary structures in proteins, *J. Mol. Biol. 76*, 241-256.

4. Lodge, J. A., Maier, T., Liebl, W., Hoffmann, V., and Strater, N. (2003) Crystal structure of *Thermotoga maritima* α-glucosidase AglA defines a new clan of NAD^+-dependent glycosidases, *J. Biol. Chem. 278*, 19151-19158.

5. Liu, Q. Y. P., *et al.* (2007) Bacterial glycosidases for the production of universal red blood cells, *Nat. Biotechnol. 25*, 454-464.

6. Yip, V. L. Y., and Withers, S. G. (2006) Mechanistic analysis of the unusual redox-elimination sequence employed by *Thermotoga maritima* BgIT: A 6-phospho-β-glucosidase from glycoside hydrolase family 4, *Biochemistry 45*, 571-580.

7. Thompson, J., Pikis, A., Ruvinov, S. B., Henrissat, B., Yamamoto, H., and Sekiguchi, J. (1998) The gene *glvA* of *Bacillus subtilis* 168 encodes a metal-requiring, NAD(H)-dependent 6-phospho-α-glucosidase: Assignment to family 4 of the glycosylhydrolase superfamily, *J. Biol. Chem. 273*, 27347-27356.

8. Coutinho, P. M., Deleury, E., Davies, G. J., and Henrissat, B. (2003) An evolving hierarchical family classification for glycosyltransferases, *J. Mol. Biol. 328*, 307-317.

9. Yip, V. L. Y., Thompson, J., and Withers, S. G. (2007) Mechanism of GlvA from *Bacillus subtilis*: A detailed kinetic analysis of a 6-phospho-α-glucosidase from glycoside hydrolase family 4, *Biochemistry 46*, 9840-9852.

10. Yip, V. L. Y., *et al.* (2004) An unusual mechanism of glycoside hydrolysis involving redox and elimination steps by a family 4 beta-glycosidase from Thermotoga maritima, *J. Am. Chem. Soc. 126*, 8354-8355.

11. Liljeström, P. L., and Liljeström, P. (1987) Nucleotide-sequence of the *MelA* gene, coding for α-galactosidase in *Escherichia Coli* K-12, *Nucleic Acids Res. 15*, 2213-2220.

12. Anggraeni, A. A., Sakka, M., Kimura, T., Ratanakhaokchai, K., Kitaoka, M., and Sakka, K. (2008) Characterization of *Bacillus halodurans* α-galactosidase Mel4A encoded by the *mel4A* gene (BH2228), *Biosci., Biotechnol., Biochem. 72*, 2459-2462.

13. Sokolov, V. M., Zakharov, V. I., and Studentsov, E. P. (2002) Stereoselectivity of reactions at the glycoside center of carbohydrates: VII. Synthesis of aryl α- and β-D-glucopyranosides by Helferich, catalyzed by boron trifluoride etherate, *Russ. J. Gen. Chem. 72*, 806-811.

14. Comfort, D. A., *et al.* (2007) Biochemical analysis of *thermotoga maritima* GH36 α-galactosidase (*Tm*GalA) confirms the mechanistic commonality of Clan GH-D-glycoside hydrolases, *Biochemistry 46*, 3319-3330.

15. Lemieux, R. U., Nagabhushan, T. L., and O'Neill, I. K. (1964) A new approach to the synthesis of 2-amino-2-deoxyglycosides, *Tetrahedron Lett.* 1909-1916.

16. Taylor, J. R. (1982) *An introduction to error analysis: The study of uncertainties in physical measurements*, University Science Books, Mill Valley, Calif.

17. Lemieux, R. U., Nagabhushan, T. L., and Gunner, S. W. (1968) Synthesis of 3,4,6-tri-*O*-acetyl-2-oximino-α-D-hexopyranosides, *Can. J. Chem. 46*, 405-411.

18. Czernecki, S., Vijayakumaran, K., and Ville, G. (1986) Convenient synthesis of hex-1-enopyran-3-uloses: Selective oxidation of allylic alcohols using pyridinium dichromate, *J. Org. Chem. 51*, 5472-5475.

19. Fujiwara, T., and Hayashi, M. (2008) Efficient synthesis of rare sugar D-allal via reversal of diastereoselection in the reduction of protected 1,5-anhydrohex-1-en-3-uloses: Protecting group dependence of the stereoselection, *J. Org. Chem. 73*, 9161-9163.

20. Hermes, J. D., Roeske, C. A., O'Leary, M. H., and Cleland, W. W. (1982) Use of multiple isotope effects to determine enzyme mechanisms and intrinsic isotope effects. Malic enzyme and glucose-6-phosphate dehydrogenase, *Biochemistry 21*, 5106-5114.

21. Cook, P. F., and Cleland, W. W. (2007) *Enzyme kinetics and mechanism*, Garland Science, London; New York.

22. Cook, P. F., Blanchard, J. S., and Cleland, W. W. (1980) Primary and secondary deuterium isotope effects on equilibrium constants for enzyme-catalyzed reactions, *Biochemistry 19*, 4853-4858.

23. Fersht, A. (1985) *Enzyme structure and mechanism*, 2nd ed., W.H. Freeman, New York.

2.10. Supporting Information

Mechanistic Evaluation of *Mel A* α-Galactosidase from *Citrobacter Freundii*: A Family 4 Glycosyl Hydrolase in which Oxidation is Rate-Limiting

Saswati Chakladar, Lydia Cheng, Mary Choi, James Liu, and Andrew J. Bennet

Department of Chemistry and Department of Molecular Biology and Biochemistry, Simon Fraser University, 8888 University Drive, Burnaby, British Columbia V5A 1S6, Canada

2.10.1. Materials and Methods

2,3,4,6-Tetra-O-acetyl-D-galactono-1,5-1actone **2.11** (**Scheme S1**) 1,2,3,4,6-Penta-*O*-acetyl-D-galactose (3g) was dissolved in THF (50 mL) and hydrazine acetate (2.0 g, 2.0 equiv) was added to the solution, which was stirred at room temperature for 15 h. After quenching by addition of water (30 mL) the reaction mixture was extracted with ethyl acetate (3 × 50 mL). The combined organic layer was washed with water (3 × 30 mL), dried over sodium sulfate and the volatiles were removed under reduced pressure to give a pale yellow syrup of a mixture of α- and β-anomers **2.10** in quantitative yield (3.0 g). This product, which displayed identical spectroscopic properties to those reported previously (*118*), was dried under high vacuum and used without further purification. To a solution of the hemiacetal **2.10** (3 g, 8.6 mmol) in CH_2Cl_2 (70 mL) Dess-Martin periodonane (18.2 g, 40 mmol) was added and the reaction mixture was stirred at room temperature for 48 h. After removal of the white precipitate by filtration through a celite pad the filtrate was partitioned between dichloromethane and water. The combined organic layers were dried over Na_2SO_4 and the solvent was removed under reduced pressure to yield off-white syrup. The product was purified by column chromatography using 40% EtOAc-Hexane as eluent to give lactone **2.11** as a foamy solid (2.5 g, 84 %). ^{1}H NMR (600 MHz,) δ 5.68 (dd, *J* = 2.9, 1.5, 1H, H-4), 5.45 (dd, *J* = 10.3, 2.9, 1H, H-3), 5.26 (d, *J* = 10.3, 1H, H-2), 4.80 (dd, *J* = 6.7, 1.5, 1H, H-5), 4.23 (m, H-6, H-6'), 2.18, 2.17, 2.08, 2.05 (4 × CH_3, 4s, 4 × $OCOCH_3$). ^{13}C NMR (151 MHz) δ 170.17, 170.14, 169.77, 169.43, 164.88, 74.69, 69.20, 66.13, 60.91, 20.58, 20.56, 20.47, 20.44.

1,2,3,4,6-Penta-O-acetyl-D-(1-^{2}H)galactose **2.12**: A solution of the lactone **2.11** (2.5 g) in anhydrous THF (50 mL) was cooled to 0 °C. To this mixture a solution of $NaBD_4$ (100 mg) in ice-cold D_2O (2 mL) was added to the reaction mixture, which was stirred at 0-5 °C under inert conditions for 60 min and for another 2 h at room temperature. Cautious addition of glacial acetic acid (5 mL) destroyed the excess borodeuteride and the solution was then evaporated to give syrup, to which methanol was added and evaporated under vacuum in order to remove the boric acid by-product. The resultant syrup was dissolved in CH_2Cl_2 (20 mL), and to this was added pyridine (5 mL) and acetic anhydride (10 mL). The resulting solution was stirred at ambient temperature for 15 h. This mixture was then poured

into ice-cold water and extracted with CH_2Cl_2 (3 × 40 mL). The combined organic layer was dried over Na_2SO_4 and the volatiles were removed to obtain, after drying under high vacuum for 2 h, the peracetylated anomerically labeled product **2.12** (1.6 g, 55%).

Synthesis of 4-Nitrophenyl 6-Phospho-α-D-galactoside **2.17** (**Scheme S2**). 4-Nitrophenyl α-D-galactoside **2.13** was synthesized in 60 % yield according to a literature procedure (*119*). Selective protection of 6-OH **2.14** was achieved using *tert*-butyl diphenylsilyl chloride and imidazole in 67 % yield according to a literature procedure (*120*). $[\alpha]^{20}_D$ = +35.0 (*c* 0.45, $CHCl_3$); ^{1}HNMR (600 MHz,) δ 8.16 (d, *J* = 9.1, 2H, Ar, 4-NPh), 7.64 (d, *J* = 7.7, 2 H, TBDPS), 7.60 (d, *J* = 7.8, 2 H, TBDPS), 7.48–7.32 (m, 6 H TBDPS), 7.13 (d, *J* = 9.2, 2 H, 4-NPh), 5.70 (d, *J* = 3.5, 1H, H-1), 4.21 (bd, H-3), 4.12 (d, *J* = 7.2, 1 H, H-2), 3.96 (bdd, 1 H, H-4), 3.89 (app dd, *J* = 3.0, 1.9, H-6, H-6'), 3.79 (t, *J* = 4.8, 1H, H-5), 3.11 (d, *J* = 1.6, 1H, OH), 2.57 (d, *J* = 6.6, 1 H, OH), 2.10 (d, *J* = 8.3, 1H, OH), 1.02 (s, 9H, 3 × CH_3). ^{13}CNMR (151 MHz) δ 135.64, 135.46, 132.58, 132.45, 130.07, 130.01, 127.82, 127.80, 125.81, 116.53, 97.39, 70.98, 70.92, 69.58, 69.31, 63.79, 26.73, 19.09.

4-Nitrophenyl 2,3,4,-tri-O-benzoyl-6-tert-butyl-dimethylsilyl α-D-galactoside (**2.15**) Compound **2.14** (413 mg, 7.6 mmol) was dissolved in a solution containing anhydrous CH_2Cl_2 (30 mL), pyridine (0.27 mL, 30.6 mmol) and a catalytic amount of DMAP. This reaction mixture was stirred for 10 min prior cooling to 0 °C using a ice/water bath. Benzoyl chloride (0.7 mL, 61 mmol) was added dropwise to the solution over a period of 10 min. After which the resultant solution was stirred at 0 °C for 15 min and then at 25 °C for 15-20 h. Following the addition of water (30 mL) the aqueous layer was extracted with CH_2Cl_2 (3 × 20 mL), and the organic layer was subsequently washed with ice-cold 10% H_2SO_4 (3 × 20 mL) and saturated $NaHCO_3$ solution (3 × 20 mL). The combined organic layer was dried over Na_2SO_4 and the volatiles were removed to give crude product in a quantitative yield. Flash chromatographic purification using 10% EtOAc-Hexane resulted in a foamy white solid **2.15** (590 mg, 91%). $[\alpha]^{20}_D$ = +150.0 (*c* 0.25, $CHCl_3$); ^{1}H NMR (600 MHz, $CDCl_3$) δ 8.21–8.15 (m, 1H, Ar), 8.08–8.03 (m, 1H, Ar), 7.99–7.95 (m, 1H, Ar), 7.88–7.83 (m, 1H, Ar), 7.67 (t, *J* =7.5, 1H, Ar), 7.60–7.56 (m, 1H, Ar), 7.55–7.44 (m, 3H, Ar), 7.40 (t, *J* = 8.0, 2H, Ar), 7.32 (t, *J* = 7.8, 2H, Ar), 7.25–7.21 (m, 1H, Ar), 7.18 (t,

J = 7.6, 1H, Ar), 6.18 (dd, *J* = 10.7, 3.4, 1H), 6.15 (d, *J* = 3.6, 1H), 6.10 (d, *J* = 3.2, 1H), 5.79 (dd, *J* = 10.7, 3.6, 1H), 4.38 (t, *J* = 6.6, 1H), 3.81 (qd, *J* = 10.5, 6.7, 1H), 0.96 (s, 9H, TBDPS 4 × CH_3). ^{13}C NMR (151 MHz, $CDCl_3$) δ 165.99, 165.60, 165.29, 161.23, 135.43, 135.38, 129.88, 129.83, 129.80, 129.78, 128.64, 128.52, 128.30, 127.72, 127.64, 125.89, 116.78, 95.34, 70.91, 68.83, 68.48, 68.25, 61.64.

4-Nitrophenyl 2,3,4,-tri-O-benzoyl-α-D-galactoside (**2.16**) Compound **2.15** (590 mg, 6.9 mmol) was dissolved in 1:1 v/v Et_2O:MeOH (50 mL) and the reaction mixture was stirred under N_2 at 25 °C. Generation of HCl *in situ* was accomplished by dropwise addition of CH_3COCl (1.4 mL) while vigorously stirring the reaction mixture. Stirring was continued for another 20 h at ambient temperature when TLC analysis (40% EtOAc-Hexane) showed the formation of a more polar compound (R_f = 0.45). After the effervescence caused by the addition of saturated $NaHCO_3$ (40 mL) had ceased, the organic layer was extracted with Et_2O (3 × 30 mL). The combined ethereal layers were pooled together, washed with brine (2 × 20 mL), dried over Na_2SO_4 and concentrated under vacuum to give a yellowish syrup which was purified by flash chromatography (20% EtOAc-Hexane) to obtain a white foamy sticky solid **2.16** (295 mg, 70%). $[\alpha]^{20}_D$ = +122.2 (*c* 0.38, $CHCl_3$) ^{1}H NMR (600 MHz, $CDCl_3$) δ 8.23-8.13 (m, 4H, Ar), 7.97 (d, *J* = 7.2, 2 H, Ar), 7.85 (d, *J* = 8.4, 2 H, Ar), 7.66 (dd, *J* = 11.9, 4.4, 1 H, Ar), 7.56–7.50 (m, 3 H, Ar), 7.47 (t, *J* = 7.4, 1 H, Ar), 7.41–7.36 (m, 2 H, Ar), 7.29-7.23 (m, 5H, Ar), 6.20 (d, *J* = 3.4, 1 H, H-2), 6.18 (t, *J* = 3.2, 1 H, H-3), 5.96 (d, *J* =3.5, 1 H, H-1), 5.91 (dd, *J* = 10.8, 3.6, 1 H, H-4), 4.40 (t, *J* = 6.8, 1 H, H-5), 3.76-3.70 (m, 1H, H-6), 3.67–3.61 (m, 1 H, H-6'). ^{13}C NMR (151 MHz, $CDCl_3$) δ 166.72, 165.51, 161.03, 134.01, 133.76, 133.47, 130.05, 129.82, 129.73, 128.79, 128.65, 128.63, 128.57, 128.41, 125.97, 116.63, 95.24, 70.76, 69.45, 68.71, 68.04, 60.38. ESI-MS for $C_{33}H_{27}NO_{11}$ *m/z* calcd for ($M + H^+$) 614.1662: found 614.1670.

4-Nitrophenyl 6-phospho-α-D-galactoside (**2.17**). Compound **2.15** (295 mg) was dissolved in 1, 4-dioxane (15 mL) and pyridine (3 equiv.) and $POCl_3$ (3 equiv.) were added. The reaction was stirred at 25 °C for 30 min. After the addition of 50% aqueous pyridine the resulting solution was treated as described in a literature procedure (*121*). Flash chromatographic purification of the crude product (eluant 10% MeOH-EtOAc) gave a white solid. This solid was immediately deacetylated using standard Zemplen conditions

(NaOMe/MeOH), which was followed by neutralization using Amberlite (NH_4^+) resin to give the ammonium salt of 4-nitrophenyl 6-*O*-phospho-α-D-galactoside **2.17** in 90% (100 mg). Mpt = 179-180 °C; $[\alpha]^{20}_D$ = -40 (*c* 1.0, H_2O); ^{1}H NMR (400 MHz, D_2O) δ 8.28 (d, *J* = 9.3, 2 H, Ar), 7.32 (d, *J* = 9.3, 2 H, Ar), 5.83 (d, *J* = 3.7, 1 H, H-1), 4.18–4.12 (m, 3 H, H-3, H-4, H-5), 4.06–4.00 (m, 1 H, H-2), 4.00–3.92 (m, 1 H, H-6), 3.92–3.83 (m, 1 H, H-6'); ^{13}C NMR (101 MHz, D_2O) δ 161.73, 142.40, 126.06, 117.00, 97.13, 70.66, 69.10, 68.45, 67.87, 63.04, 54.67, 54.63, 51.87.;^{31}P NMR (162 MHz, D_2O) δ 1.24. ESI-MS for $C_{12}H_{15}NO_{11}P$ *m/z:* calcd for ($M - H^+$) 380.0383: found 380.0390.

3,4,6-Tri-O-acetyl-1,5-anhydro-2-deoxy-D-(3-^{2}H)lyxo-hex-1-enitol **2.7** (**Scheme 2-2**): 1,5-anhydro-2-deoxy-D-*threo-hex-1-en-3-ulose* (2.2 g) (*122*) was dissolved in anhydrous THF (60 mL) and this solution was cooled to 0 °C under an inert atmosphere. Subsequent addition of a solution of $NaBD_4$ (130 mg, 2 equiv) in cold D_2O (2 mL) was followed by stirring the mixture at 0 °C for 1 h and then at ambient temperature for another 2 h. At this point TLC analysis (40% EtOAc-Hexane) showed that the reaction was complete. Cautious addition of glacial AcOH (5 mL) destroyed the excess borodeuteride and the resultant solution was washed with MeOH (20 mL). Removal of volatiles under reduced pressure gave a pale yellow syrup which was directly acetylated under standard conditions, pyridine (5 mL) and Ac_2O (4 mL) with stirring for 15 h at 25 °C. After the addition of water (10 mL) the peracetylated material was extracted from the aqueous layer with CH_2Cl_2 (3 × 40 mL). The combined organic layer was washed with cold 10% H_2SO_4 (10 mL), brine (2 × 10 mL) and dried (Na_2SO_4). Removal of the volatiles under reduced pressure gave a yellow syrup, and this material was purified by flash chromatography (30% EtOAc-Hexane) to give pure **8** (1.3 g, 50 % yield over 2 steps).

Derivation of Equation for Expected Value of ${}^{D}(V/K_{C3})_{C2\text{-}D}$ Based on Experimental Values for ${}^{D}(V/K_{C3})_{C2\text{-}H}$ and ${}^{D}(V/K_{C2})_{C3\text{-}H}$ and Literature Value for ${}^{D}K_{eq(C3)}$.

$$ {}^{D}\left(\frac{V}{K_{C3}}\right)_{C2-H} = \frac{{}^{D}k_3 + \frac{k_{3H}}{k_2} + \frac{k_{4H}}{k_{5H}}\left(1+\frac{k_{6H}}{k_7}\right){}^{D}K_{eq}}{1+\frac{k_{3H}}{k_2}+\frac{k_{4H}}{k_{5H}}\left(1+\frac{k_{6H}}{k_7}\right)} $$

Equation S1 (Equation 1 in paper)

Setting $k_{6H}/k_7 = 0$

$$ \left(1+\frac{k_{3H}}{k_2}+\frac{k_{4H}}{k_{5H}}\right)\times {}^{D}\left(\frac{V}{K_{C3}}\right)_{C2-H} = {}^{D}k_3 + \frac{k_{3H}}{k_2} + \frac{k_{4H}}{k_{5H}}{}^{D}K_{eq} $$

Equation S2

Calculate ${}^{D}k_5$

$$ {}^{D}\left(\frac{V}{K_{C2}}\right)_{C3-H} = \frac{{}^{D}k_5 + \frac{k_{5H}}{k_{4H}}\left(1+\frac{k_{3H}}{k_2}\right) + \frac{k_{6H}}{k_7}{}^{D}K_{eq}}{1+\frac{k_{5H}}{k_{4H}}\left(1+\frac{k_{3H}}{k_2}\right)+\frac{k_{6H}}{k_7}} $$

Equation S3 (Equation 2 in paper)

Setting $k_{6H}/k_7 = 0$

$$ {}^{D}\left(\frac{V}{K_{C2}}\right)_{C3-H} = \frac{{}^{D}k_5 + \frac{k_{5H}}{k_{4H}}\left(1+\frac{k_{3H}}{k_2}\right)}{1+\frac{k_{5H}}{k_{4H}}\left(1+\frac{k_{3H}}{k_2}\right)} $$

Equation S4

$$ {}^{D}k_5 = {}^{D}\left(\frac{V}{K_{C2}}\right)_{C3-H}\left(1+\frac{k_{5H}}{k_{4H}}\left(1+\frac{k_{3H}}{k_2}\right)\right) - \frac{k_{5H}}{k_{4H}}\left(1+\frac{k_{3H}}{k_2}\right) = {}^{D}\left(\frac{V}{K_{C2}}\right)_{C3-H} + \frac{k_{5H}}{k_{4H}}\left(1+\frac{k_{3H}}{k_2}\right)\left({}^{D}\left(\frac{V}{K_{C2}}\right)_{C3-H} - 1\right) $$

Equation S5

Calculate Effect of Deuteration at C-2 by Changing k_{5H} to k_{5D} in Equation S1 (with $k_{6H}/k_7 = 0$) and Solving

$$ {}^{D}\left(\frac{V}{K_{C3}}\right)_{C2-D} = \frac{{}^{D}k_3 + \frac{k_{3H}}{k_2} + \frac{k_{4H}}{k_{5D}}{}^{D}K_{eq}}{1 + \frac{k_{3H}}{k_2} + \frac{k_{4H}}{k_{5D}}} = \frac{{}^{D}k_3 + \frac{k_{3H}}{k_2} + \frac{k_{4H}k_{5H}}{k_{5D}k_{5H}}{}^{D}K_{eq}}{1 + \frac{k_{3H}}{k_2} + \frac{k_{4H}k_{5H}}{k_{5D}k_{5H}}} $$

Equation S6

$$ {}^{D}\left(\frac{V}{K_{C3}}\right)_{C2-D} = \frac{{}^{D}k_3 + \frac{k_{3H}}{k_2} + \frac{k_{4H}{}^{D}k_5}{k_{5H}}{}^{D}K_{eq}}{1 + \frac{k_{3H}}{k_2} + \frac{k_{4H}{}^{D}k_5}{k_{5H}}} $$

Equation S7

Substitute Expression for ${}^{D}k_5$ (Equation S5) into the Denominator of Equation S7

$$ 1 + \frac{k_{3H}}{k_2} + \frac{k_{4H}{}^{D}k_5}{k_{5H}} = 1 + \frac{k_{3H}}{k_2} + \frac{k_{4H}}{k_{5H}}\left({}^{D}\left(\frac{V}{K_{C2}}\right)_{C3-H} + \frac{k_{5H}}{k_{4H}}\left(1 + \frac{k_{3H}}{k_2}\right)\left({}^{D}\left(\frac{V}{K_{C2}}\right)_{C3-H} - 1\right)\right) $$

Expression S1

$$ 1 + \frac{k_{3H}}{k_2} + \frac{k_{4H}}{k_{5H}}{}^{D}\left(\frac{V}{K_{C2}}\right)_{C3-H} + {}^{D}\left(\frac{V}{K_{C2}}\right)_{C3-H} + \frac{k_{3H}}{k_2}{}^{D}\left(\frac{V}{K_{C2}}\right)_{C3-H} - 1 - \frac{k_{3H}}{k_2} = \left(1 + \frac{k_{3H}}{k_2} + \frac{k_{4H}}{k_{5H}}\right) \times {}^{D}\left(\frac{V}{K_{C2}}\right)_{C3-H} $$

Expression S2

Substitute Expression for ${}^{D}k_5$ into tNumerator of Equation S7

$$ {}^{D}k_3 + \frac{k_{3H}}{k_2} + \frac{k_{4H}\,{}^{D}k_5\,{}^{D}K_{eq}}{k_{5H}} = {}^{D}k_3 + \frac{k_{3H}}{k_2} + \frac{k_{4H}\,{}^{D}K_{eq}}{k_{5H}}\left({}^{D}\left(\frac{V}{K_{C2}}\right)_{C3-H} + \frac{k_{5H}}{k_{4H}}\left(1+\frac{k_{3H}}{k_2}\right)\left({}^{D}\left(\frac{V}{K_{C2}}\right)_{C3-H} - 1\right)\right) $$

Expression S3

$$ = {}^{D}k_3 + \frac{k_{3H}}{k_2} + \frac{k_{4H}\,{}^{D}K_{eq}}{k_{5H}}\,{}^{D}\left(\frac{V}{K_{C2}}\right)_{C3-H} + {}^{D}K_{eq}\left(1+\frac{k_{3H}}{k_2}\right)\left({}^{D}\left(\frac{V}{K_{C2}}\right)_{C3-H} - 1\right) $$

Expression S4

Thus Substitution of Expressions S2 and S4 with Rearrangement into Equation S7 gives

$$ \left(1+\frac{k_{3H}}{k_2}+\frac{k_{4H}}{k_{5H}}\right)\times {}^{D}\left(\frac{V}{K_{C2}}\right)_{C3-H} \times {}^{D}\left(\frac{V}{K_{C3}}\right)_{C2-D} = {}^{D}k_3 + \frac{k_{3H}}{k_2} + \frac{k_{4H}\,{}^{D}K_{eq}}{k_{5H}}\,{}^{D}\left(\frac{V}{K_{C2}}\right)_{C3-H} + {}^{D}K_{eq}\left(1+\frac{k_{3H}}{k_2}\right)\left({}^{D}\left(\frac{V}{K_{C2}}\right)_{C3-H} - 1\right) $$

Equation S8

Subtract Equation S2 from Equation S8

$$ \left(1+\frac{k_{3H}}{k_2}+\frac{k_{4H}}{k_{5H}}\right)\left({}^{D}\left(\frac{V}{K_{C2}}\right)_{C3-H} \times {}^{D}\left(\frac{V}{K_{C3}}\right)_{C2-D} - {}^{D}\left(\frac{V}{K_{C3}}\right)_{C2-H}\right) = {}^{D}k_3 + \frac{k_{3H}}{k_2} + \frac{k_{4H}\,{}^{D}K_{eq}}{k_{5H}}\,{}^{D}\left(\frac{V}{K_{C2}}\right)_{C3-H} + {}^{D}K_{eq}\left(1+\frac{k_{3H}}{k_2}\right)\left({}^{D}\left(\frac{V}{K_{C2}}\right)_{C3-H} - 1\right) - \left({}^{D}k_3 + \frac{k_{3H}}{k_2} + \frac{k_{4H}\,{}^{D}K_{eq}}{k_{5H}}\right) $$

Equation S9

$$ = \frac{k_{4H}\,{}^{D}K_{eq}}{k_{5H}}\left({}^{D}\left(\frac{V}{K_{C2}}\right)_{C3-H} - 1\right) + {}^{D}K_{eq}\left(1+\frac{k_{3H}}{k_2}\right)\left({}^{D}\left(\frac{V}{K_{C2}}\right)_{C3-H} - 1\right) = \left(1+\frac{k_{3H}}{k_2}+\frac{k_{4H}}{k_{5H}}\right){}^{D}K_{eq}\left({}^{D}\left(\frac{V}{K_{C2}}\right)_{C3-H} - 1\right) $$

Equation S10

Divide Both Sides by $(1+k_{3H}/k_2+ k_{4H}/k_{5H})$

$$\left({}^{D}\left(\frac{V}{K_{C2}}\right)_{C3-H} \times {}^{D}\left(\frac{V}{K_{C3}}\right)_{C2-D} - {}^{D}\left(\frac{V}{K_{C3}}\right)_{C2-H} \right) = {}^{D}K_{eq}\left({}^{D}\left(\frac{V}{K_{C2}}\right)_{C3-H} - 1 \right)$$

Equation S11

Setting ${}^{D}K_{eq} = 1.18$ (*123*)

$${}^{D}\left(\frac{V}{K_{C3}}\right)_{C2-D} = \frac{{}^{D}K_{eq}\left({}^{D}\left(\frac{V}{K_{C2}}\right)_{C3-H} - 1 \right) + {}^{D}\left(\frac{V}{K_{C3}}\right)_{C2-H}}{{}^{D}\left(\frac{V}{K_{C2}}\right)_{C3-H}} = \frac{1.18 \times (1.74 - 1.00) + 1.74}{1.74} = 1.502$$

Equation S12 (Equation 3 paper)

Inclusion of the secondary EIE of 1.054 (*123*), which for this example changes ${}^{D}K_{eq}$ to 1.24, results in a value of 1.527

Derivation of Equation for Expected Value of ${}^{D}(V/K_{C2})_{C3\text{-}D}$ Based on Experimental Values for ${}^{D}(V/K_{C3})_{C2\text{-}H}$ and ${}^{D}(V/K_{C2})_{C3\text{-}H}$ and Literature Value for ${}^{D}K_{eq(C3)}$.

Calculate ${}^{D}k_3$ from Equation S2

$${}^{D}k_3 = {}^{D}\left(\frac{V}{K_{C3}}\right)_{C2-H}\left(1 + \frac{k_{3H}}{k_2} + \frac{k_{4H}}{k_{5H}}\right) - \frac{k_{3H}}{k_2} - \frac{k_{4H}}{k_{5H}}{}^{D}K_{eq}$$

Equation S13

Calculate Effect of Deuteration at C-3 by Changing k_{3H} to k_{3D} and k_{4H} to k_{4D} in Equation S3 (with $k_{6H}/k_7 = 0$) and Solving

$$^{D}\left(\frac{V}{K_{C2}}\right)_{C3-D} = \frac{^{D}k_5 + \frac{k_{5H}}{k_{4D}}\left(1+\frac{k_{3D}}{k_2}\right)}{1+\frac{k_{5H}}{k_{4D}}\left(1+\frac{k_{3D}}{k_2}\right)} = \frac{^{D}k_5 + \frac{k_{5H}k_{4H}}{k_{4H}k_{4D}}\left(\frac{k_{3H}k_{3D}}{k_2k_{3H}}+1\right)}{1+\frac{k_{5H}k_{4H}}{k_{4H}k_{4D}}\left(\frac{k_{3H}k_{3D}}{k_2k_{3H}}+1\right)} = \frac{^{D}k_5 + \frac{k_{5H}\,^{D}k_3}{k_{4H}\,^{D}K_{eq}} + \frac{k_{5H}k_{3H}}{k_{4H}k_2\,^{D}K_{eq}}}{1+\frac{k_{5H}\,^{D}k_3}{k_{4H}\,^{D}K_{eq}} + \frac{k_{5H}k_{3H}}{k_{4H}k_2\,^{D}K_{eq}}}$$

Equation S14

Substitute Expression for $^{D}k_3$ (Equation S13) into the Denominator of Equation S14

$$1+\frac{k_{5H}}{k_{4H}\,^{D}K_{eq}}\left(^{D}\left(\frac{V}{K_{C3}}\right)_{C2-H}\left(1+\frac{k_{3H}}{k_2}+\frac{k_{4H}}{k_{5H}}\right)-\frac{k_{3H}}{k_2}-\frac{k_{4H}}{k_{5H}}\,^{D}K_{eq}\right)+\frac{k_{5H}k_{3H}}{k_{4H}k_2\,^{D}K_{eq}} = 1+\frac{k_{5H}}{k_{4H}\,^{D}K_{eq}}\,^{D}\left(\frac{V}{K_{C3}}\right)_{C2-H}\left(1+\frac{k_{3H}}{k_2}+\frac{k_{4H}}{k_{5H}}\right)+\frac{k_{5H}k_{3H}}{k_{4H}k_2\,^{D}K_{eq}}-\frac{k_{5H}k_{3H}}{k_{4H}k_2\,^{D}K_{eq}}-1$$

$$=\frac{k_{5H}}{k_{4H}\,^{D}K_{eq}}\,^{D}\left(\frac{V}{K_{C3}}\right)_{C2-H}\left(1+\frac{k_{3H}}{k_2}+\frac{k_{4H}}{k_{5H}}\right)$$

Expression S5

Substitute Expression for $^{D}k_3$ into Numerator of Equation S14

$$^{D}k_5 - 1+\frac{k_{5H}}{k_{4H}\,^{D}K_{eq}}\,^{D}\left(\frac{V}{K_{C3}}\right)_{C2-H}\left(1+\frac{k_{3H}}{k_2}+\frac{k_{4H}}{k_{5H}}\right)$$

Expression S6

Thus Substitution of Expressions S5 and S6 with Rearrangement into Equation S14 gives

$$\left(\frac{k_{5H}}{k_{4H}\,{}^{D}K_{eq}}{}^{D}\left(\frac{V}{K_{C3}}\right)_{C2-H}\left(1+\frac{k_{3H}}{k_2}+\frac{k_{4H}}{k_{5H}}\right)\right){}^{D}\left(\frac{V}{K_{C2}}\right)_{C3-D}={}^{D}k_5-1+\left(\frac{k_{5H}}{k_{4H}\,{}^{D}K_{eq}}{}^{D}\left(\frac{V}{K_{C3}}\right)_{C2-H}\left(1+\frac{k_{3H}}{k_2}+\frac{k_{4H}}{k_{5H}}\right)\right)$$

Equation S15

$$\left(\frac{k_{5H}}{k_{4H}\,{}^{D}K_{eq}}{}^{D}\left(\frac{V}{K_{C3}}\right)_{C2-H}\left(1+\frac{k_{3H}}{k_2}+\frac{k_{4H}}{k_{5H}}\right)\right)\left({}^{D}\left(\frac{V}{K_{C2}}\right)_{C3-D}-1\right)={}^{D}k_5-1$$

Equation S16

$$\left(\frac{1}{{}^{D}K_{eq}}{}^{D}\left(\frac{V}{K_{C3}}\right)_{C2-H}\left(1+\frac{k_{3H}k_{5H}}{k_2k_{4H}}+\frac{k_{5H}}{k_{4H}}\right)\right)\left({}^{D}\left(\frac{V}{K_{C2}}\right)_{C3-D}-1\right)={}^{D}k_5-1$$

Equation S17

Use Equation S5 and subtract 1 for each side

$${}^{D}k_5-1={}^{D}\left(\frac{V}{K_{C2}}\right)_{C3-H}+\frac{k_{5H}}{k_{4H}}\left(1+\frac{k_{3H}}{k_2}\right)\left({}^{D}\left(\frac{V}{K_{C2}}\right)_{C3-H}-1\right)-1=\left(\frac{k_{5H}}{k_{4H}}+\frac{k_{3H}k_{5H}}{k_2k_{4H}}\right)\left({}^{D}\left(\frac{V}{K_{C2}}\right)_{C3-H}-1\right)+{}^{D}\left(\frac{V}{K_{C2}}\right)_{C3-H}-1=\left({}^{D}\left(\frac{V}{K_{C2}}\right)_{C3-H}-1\right)\left(1+\frac{k_{3H}k_{5H}}{k_2k_{4H}}+\frac{k_{5H}}{k_{4H}}\right)$$

Equation S18

Substitute Right-Hand Side Equation S18 with Left-Hand Side Equation S17

$$\left(\frac{1}{{}^{D}K_{eq}}{}^{D}\left(\frac{V}{K_{C3}}\right)_{C2-H}\right)\left({}^{D}\left(\frac{V}{K_{C2}}\right)_{C3-D}-1\right)\left(1+\frac{k_{3H}k_{5H}}{k_2k_{4H}}+\frac{k_{5H}}{k_{4H}}\right)={}^{D}\left(\left(\frac{V}{K_{C2}}\right)_{C3-H}-1\right)\left(1+\frac{k_{3H}k_{5H}}{k_2k_{4H}}+\frac{k_{5H}}{k_{4H}}\right)$$

Equation S19

$$ {}^{D}\left(\frac{V}{K_{C2}}\right)_{C3-D} = \frac{{}^{D}\left(\left(\frac{V}{K_{C2}}\right)_{C3-H} - 1\right)}{\left(\frac{1}{{}^{D}K_{eq}}\,{}^{D}\left(\frac{V}{K_{C3}}\right)_{C2-H}\right)} + 1 = \frac{{}^{D}\left(\left(\frac{V}{K_{C2}}\right)_{C3-H} - 1\right) \times {}^{D}K_{eq}}{\left(\frac{V}{K_{C3}}\right)_{C2-H}} + 1 $$

Equation S20 (Equation 4 paper)

2.10.2. References

1. Sim, M. M., Kondo, H., and Wong, C. H. (1993) Synthesis and use of glycosyl phosphites: An effective route to glycosyl phosphates, sugar nucleotides, and glycosides, *J. Am. Chem. Soc*. 115, 2260-2267.

2. Apparu, M., Blancmuesser, M., Defaye, J., and Driguez, H. (1981) Stereoselective syntheses of O-nitrophenyl and S-nitrophenyl glycosides. Part III. Syntheses in the α-D-galactopyranose and a-maltose series, *Can. J. Chem*. 59, 314-320.

3. Nashed, E. M., and Glaudemans, C. P. J. (1987) Selective silylation of β-D-galactosides - A new approach to the synthesis of (1->6)-β-D-galactopyranooligosaccharides, *J. Org. Chem*. 52, 5255-5260.

4. Mungall, W. S., Greene, G. L., Miller, P. S., and Letsinge.Rl. (1974) Use of phosphorus oxychloride in synthesizing nucleotides and oligonucleotides, Nucleic Acids Res. 1, 615-627.

5. Czernecki, S., Vijayakumaran, K., and Ville, G. (1986) Convenient synthesis of hex-1-enopyran-3-uloses: Selective oxidation of allylic alcohols using pyridinium dichromate, *J. Org. Chem*. 51, 5472-5475.

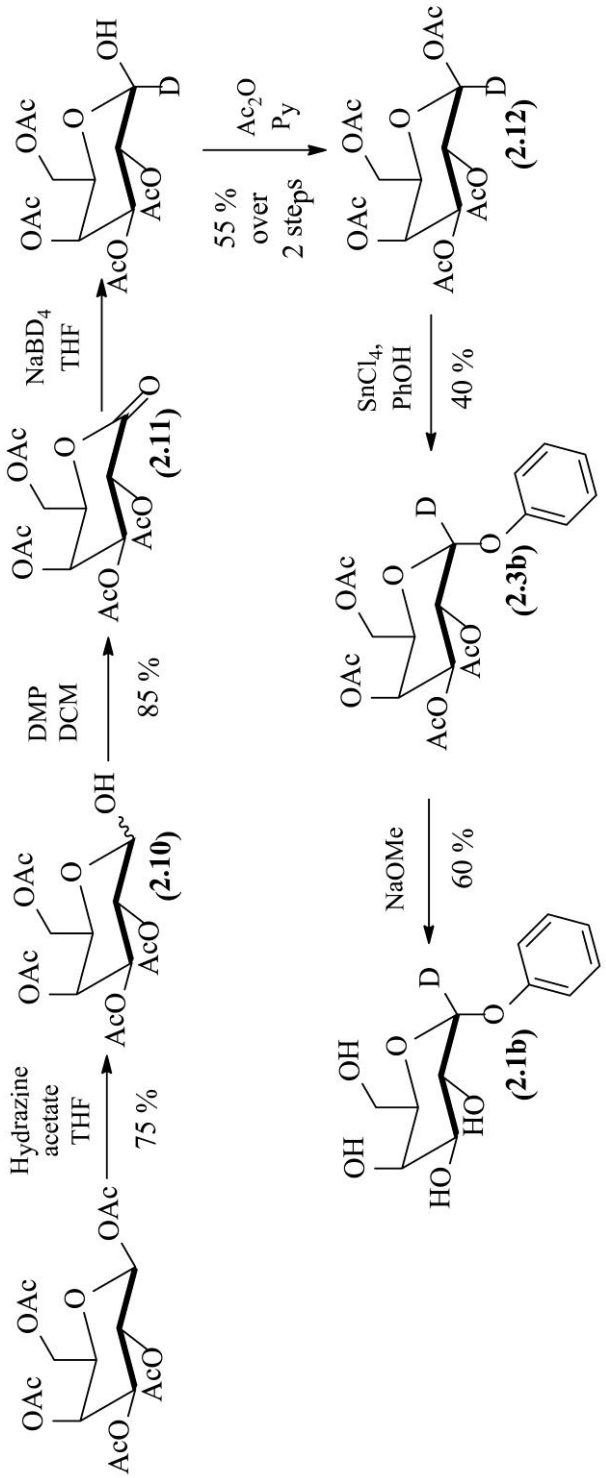
6. Cook, P. F., Blanchard, J. S., and Cleland, W. W. (1980) Primary and secondary deuterium isotope effects on equilibrium constants for enzyme-catalyzed reactions, *Biochemistry* 19, 4853-4858.

7. Serjeant, E. P., and Dempsey, B. (1979) Ionisation constants of organic acids in aqueous solution, Pergamon Press, Oxford ; New York.

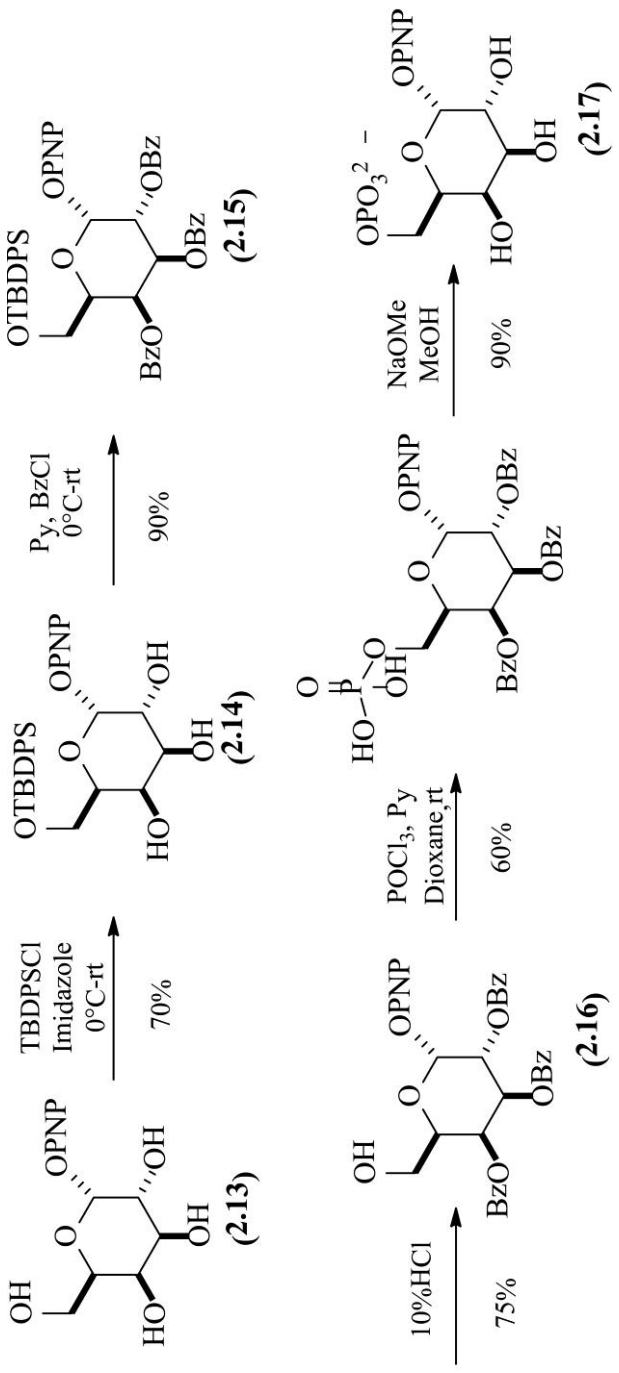
Table 2S 1. Michaelis-Menten Kinetic Parameters of the MelA α-Galactosidase-Catalyzed Hydrolysis of Aryl α-D-Galactopyranosides at 37 °C and pH 7.5.

Aglycone	**pK_a (ArOH)**[a]	**k_{cat} (s^{-1})**	**$10^{-3} \times k_{cat}/K_m$ (M^{-1} s^{-1})**
4-Cresol	10.30	1.33 ± 0.10	4.6 ±1.5
Phenyl	9.99	1.04 ± 0.05	2.8 ± 0.4
4-Chlorophenyl	9.40	1.18 ± 0.14	4.1 ± 1.5
3-Chlorophenyl	9.11	1.36 ± 0.09	4.0 ± 0.7
3,4-Dichlorophenyl	8.58	0.57 ± 0.03	3.3 ± 0.6
3-Nitrophenyl	8.36	1.36 ± 0.11	4.7 ± 1.2
4-Nitrophenyl	7.15	1.28 ± 0.07	3.1 ± 0.4

[a] pK_a values for the conjugate acids of the leaving groups are taken from reference (*124*).



Scheme 2S 1. Synthesis of phenyl α-D-(1-^{2}H)galactopyranoside.



Scheme 2S 2. Synthesis of 4-Nitrophenyl 6-phospho-α-D-galactoside.

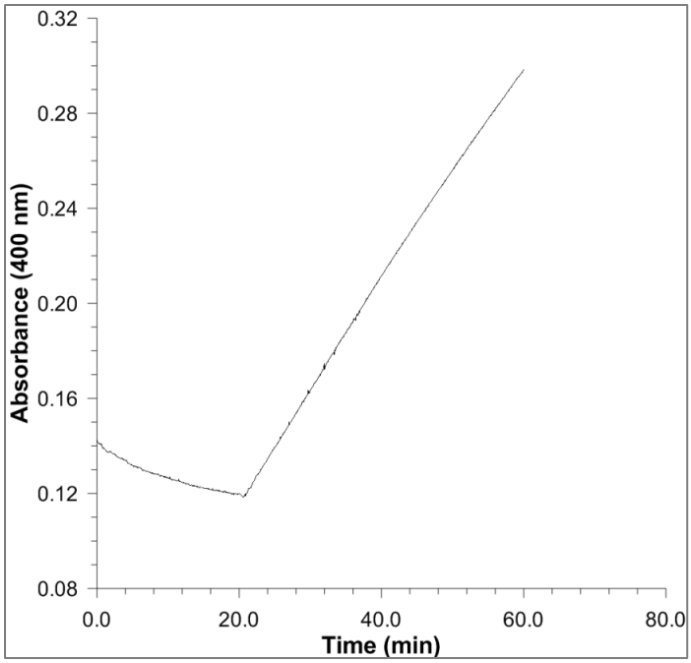


Figure 2S 1. Assay of MelA in the reduced (NADH) state. Observed rates of hydrolysis of 4NPαG by MelA. Assay condition: enzyme preincubated in buffer A with 10 mM $NaBH_4$, at t = 20 min NAD^+ (200 μM) was added to the reaction mixture.

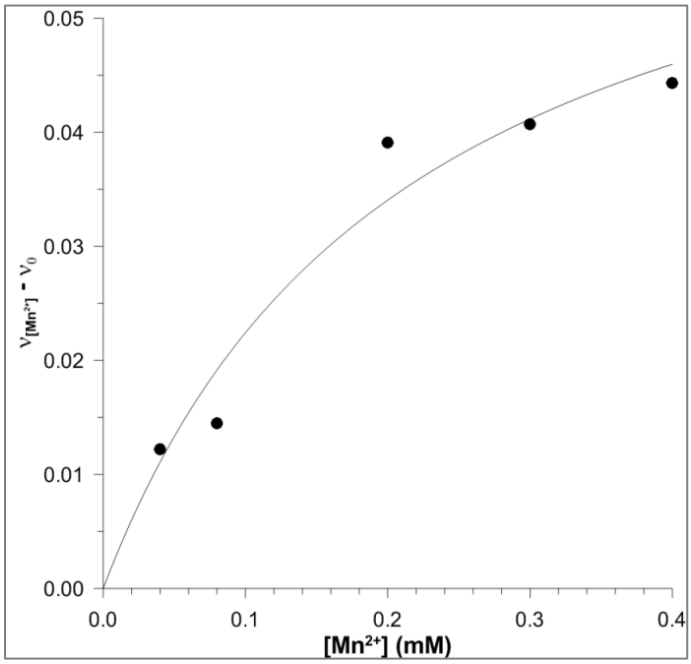


Figure 2S 2. Ligand binding curve for Mn^{2+}.

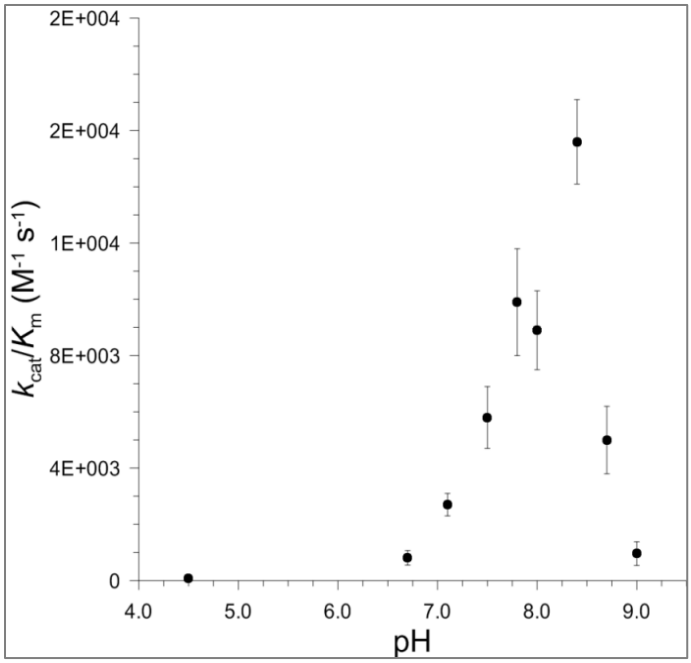


Figure 2S 3. pH dependence of k_{cat}/K_m for the MelA-catalyzed hydrolysis of 4NPαG.

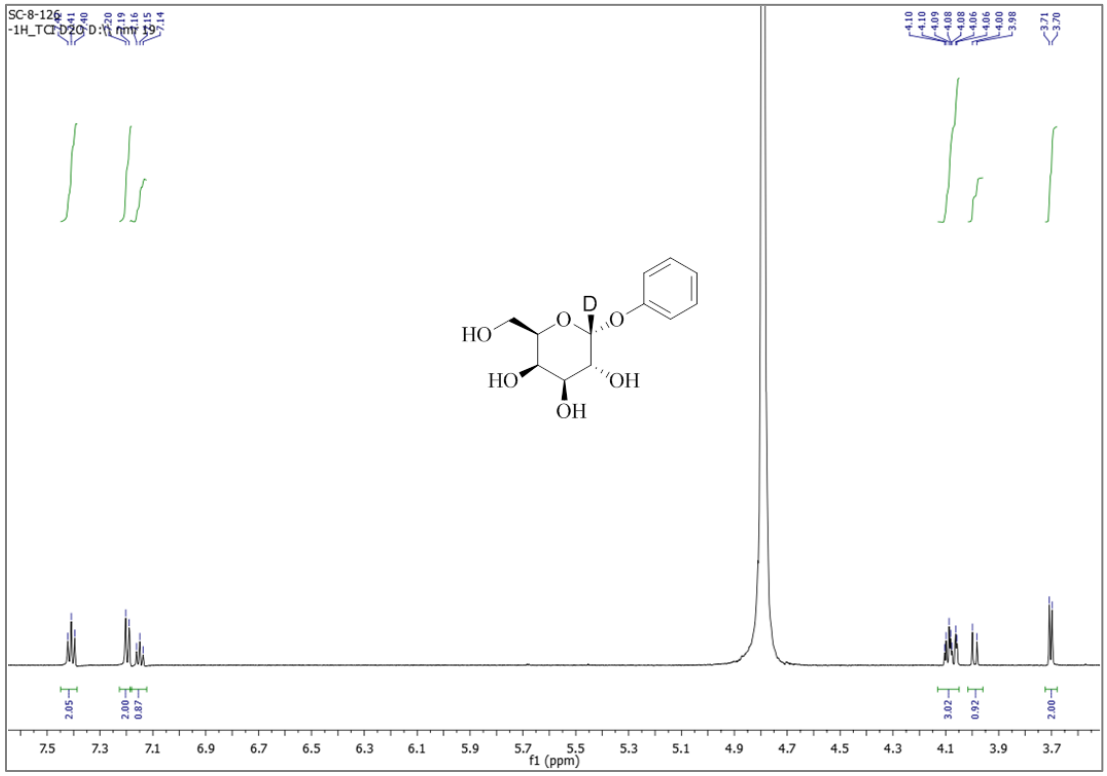


Figure 2S 4. ^{1}H NMR: Phenyl α-D-(1-^{2}H)galactopyranoside (600 MHz, D_2O).

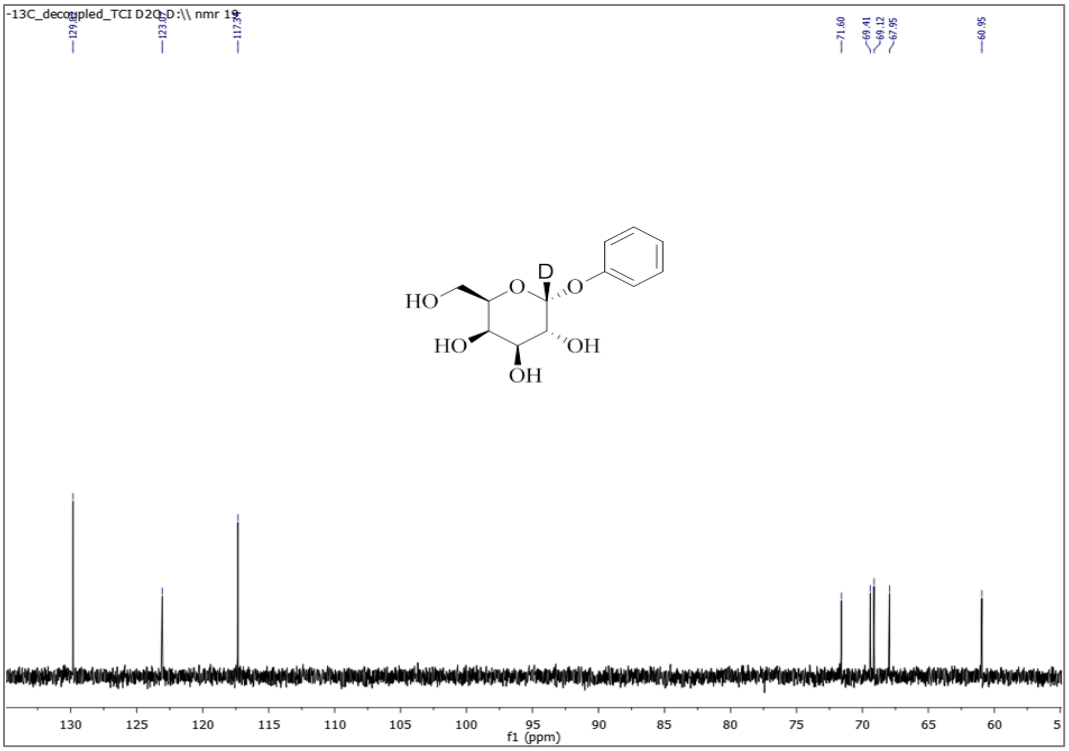


Figure 2S 5. ^{13}C NMR: Phenyl α-D-(1-^{2}H)galactopyranoside (600 MHz, D_2O).

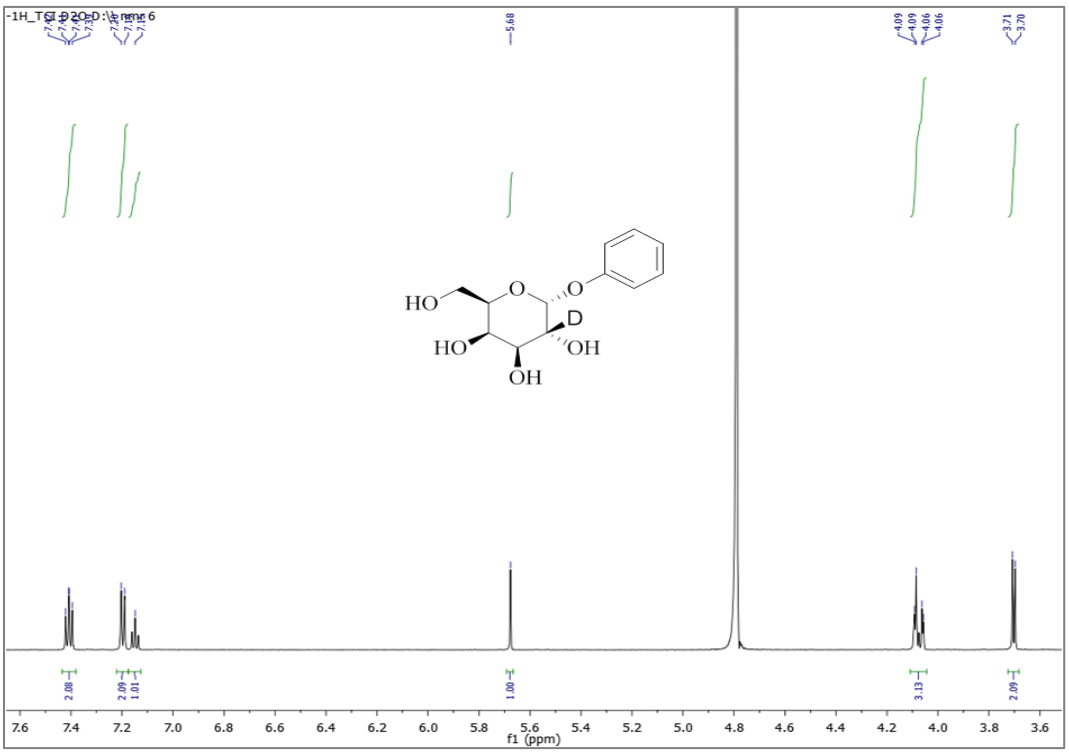


Figure 2S 6. ^{1}H NMR: Phenyl α-D-(2-^{2}H)galactopyranoside (600 MHz, D_2O).

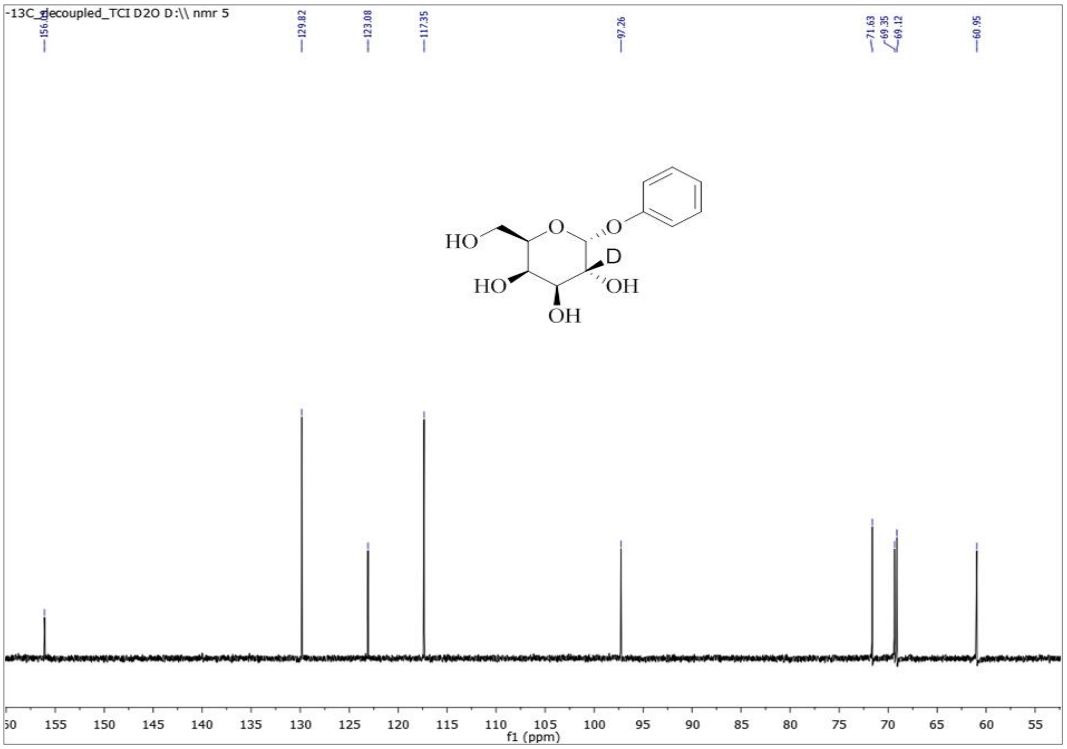


Figure 2S 7. ^{13}C NMR: Phenyl α-D-(2-^{2}H)galactopyranoside (600 MHz, D_2O).

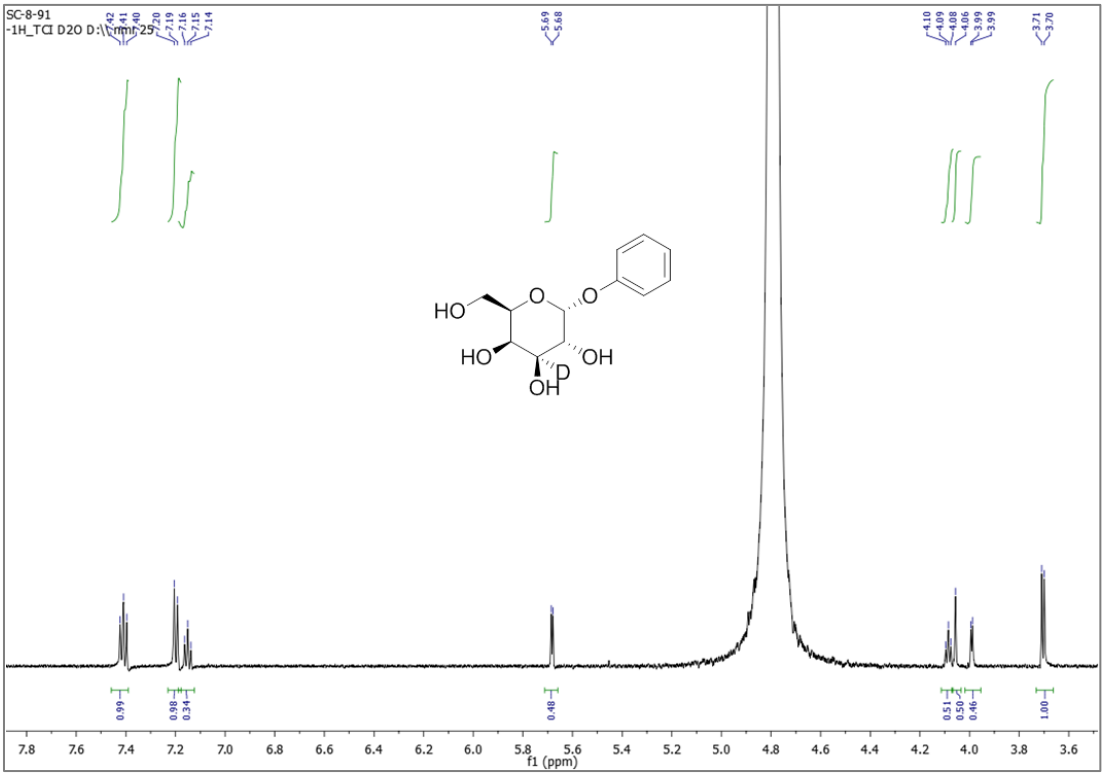


Figure 2S 8. ^{1}H NMR: Phenyl α-D-(3-^{2}H)galactopyranoside (600 MHz, D_2O).

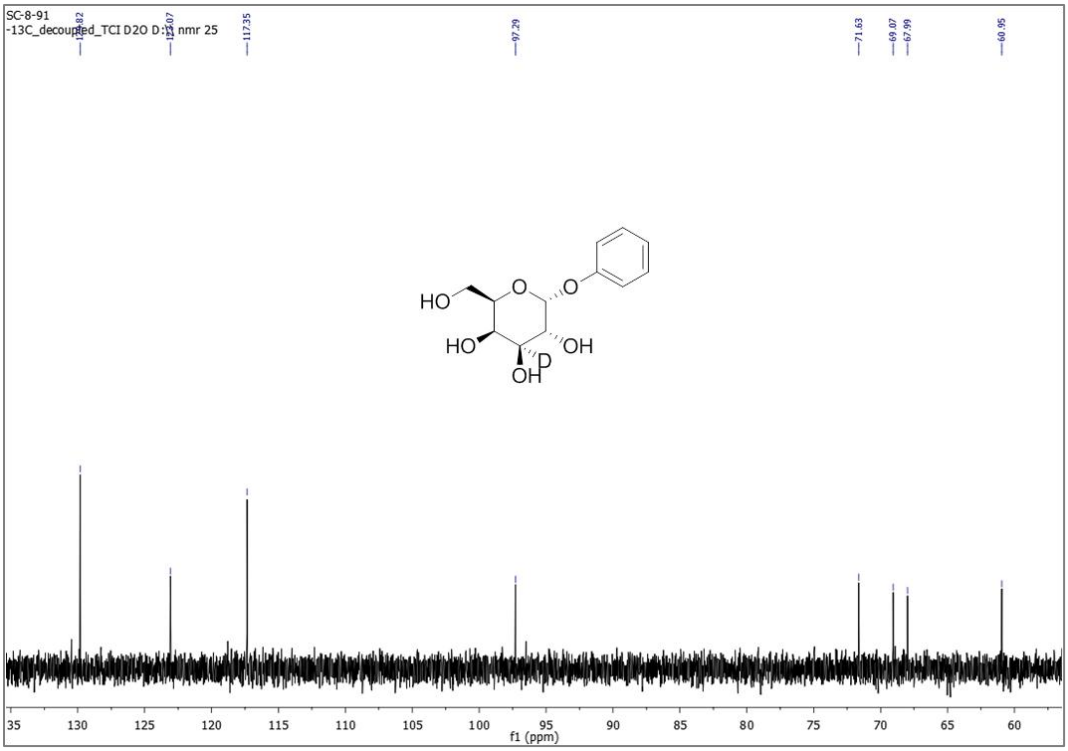


Figure 2S 9. ^{13}C NMR: Phenyl α-D-(3-^{2}H)galactopyranoside (600 MHz, D_2O).

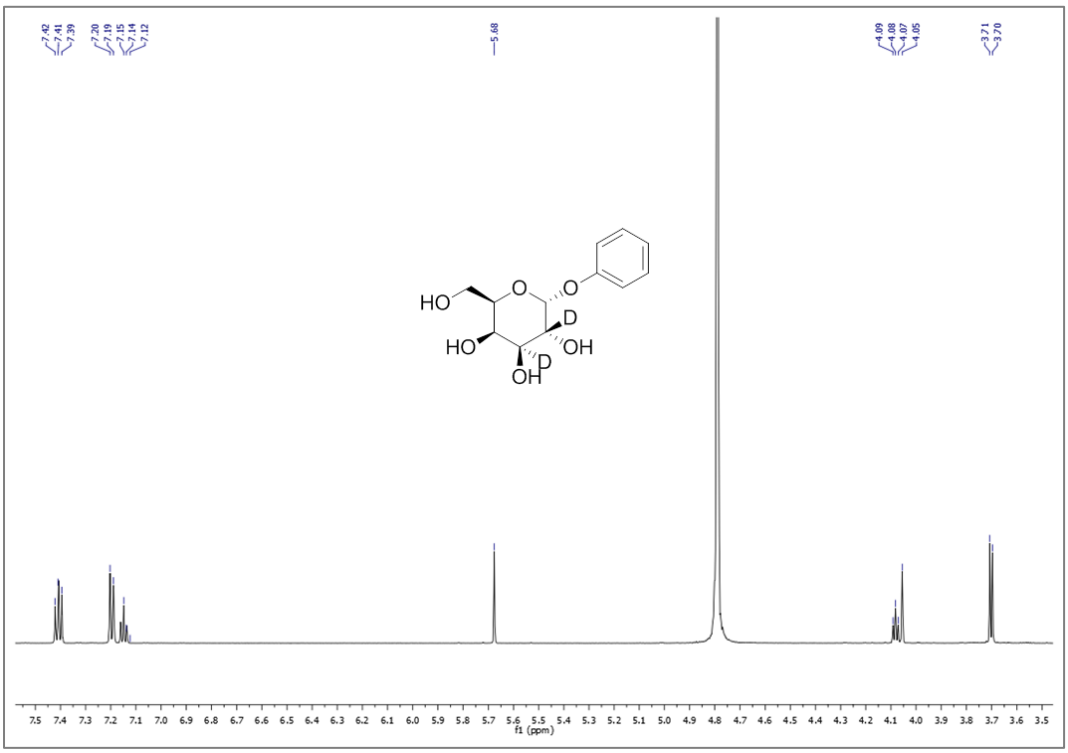


Figure 2S 10. ^{1}H NMR: Phenyl α-D-(2-^{2}H, 3-^{2}H)galactopyranoside (600 MHz, D_2O).

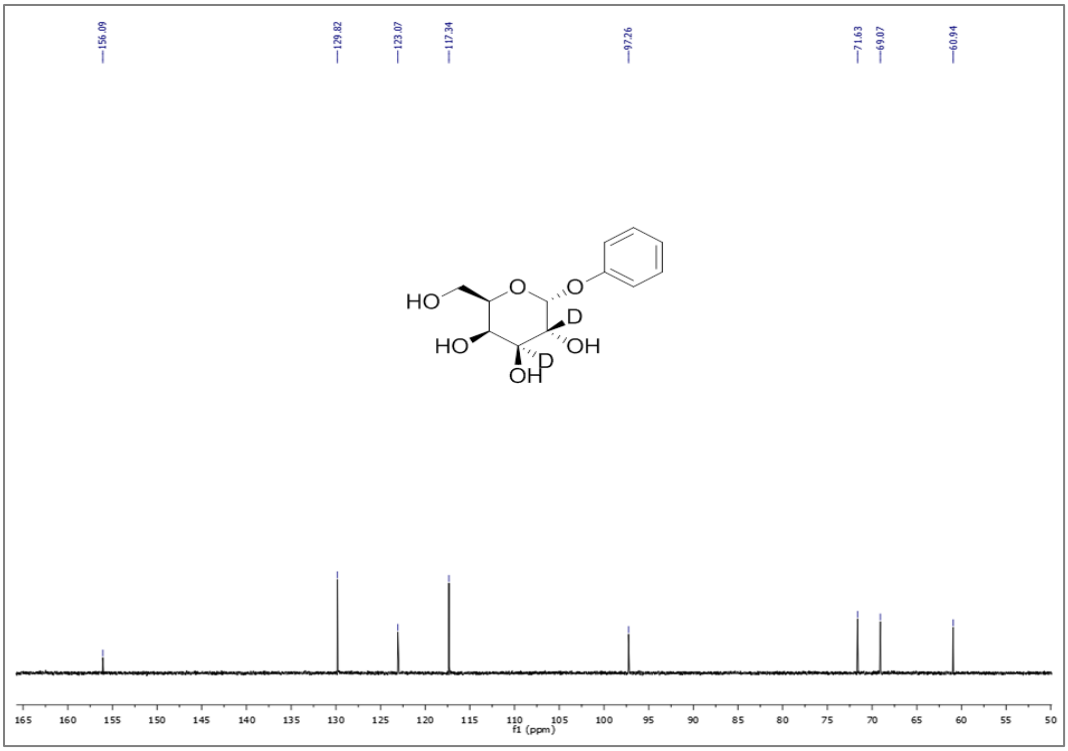


Figure 2S 11. ^{13}C NMR: Phenyl α-D-(2-^{2}H, 3-^{2}H)galactopyranoside (600 MHz, D_2O).

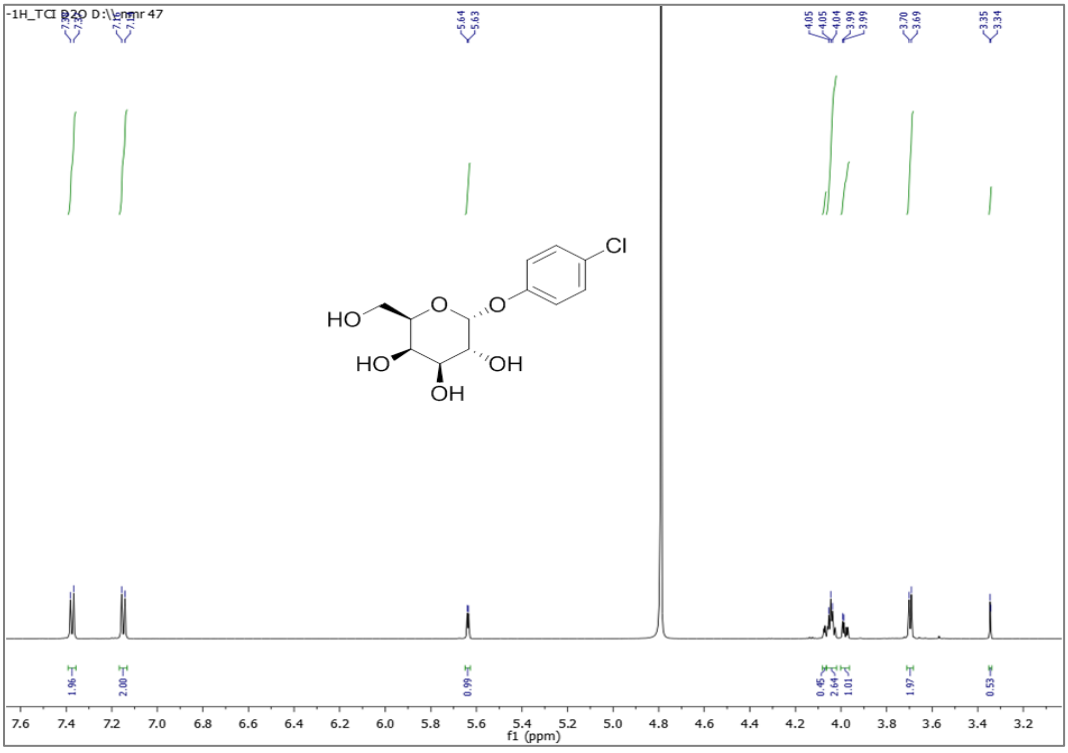


Figure 2S 12. ^{1}H NMR: 4-Chlorophenyl α-D-galactoside (600 MHz, D2O).

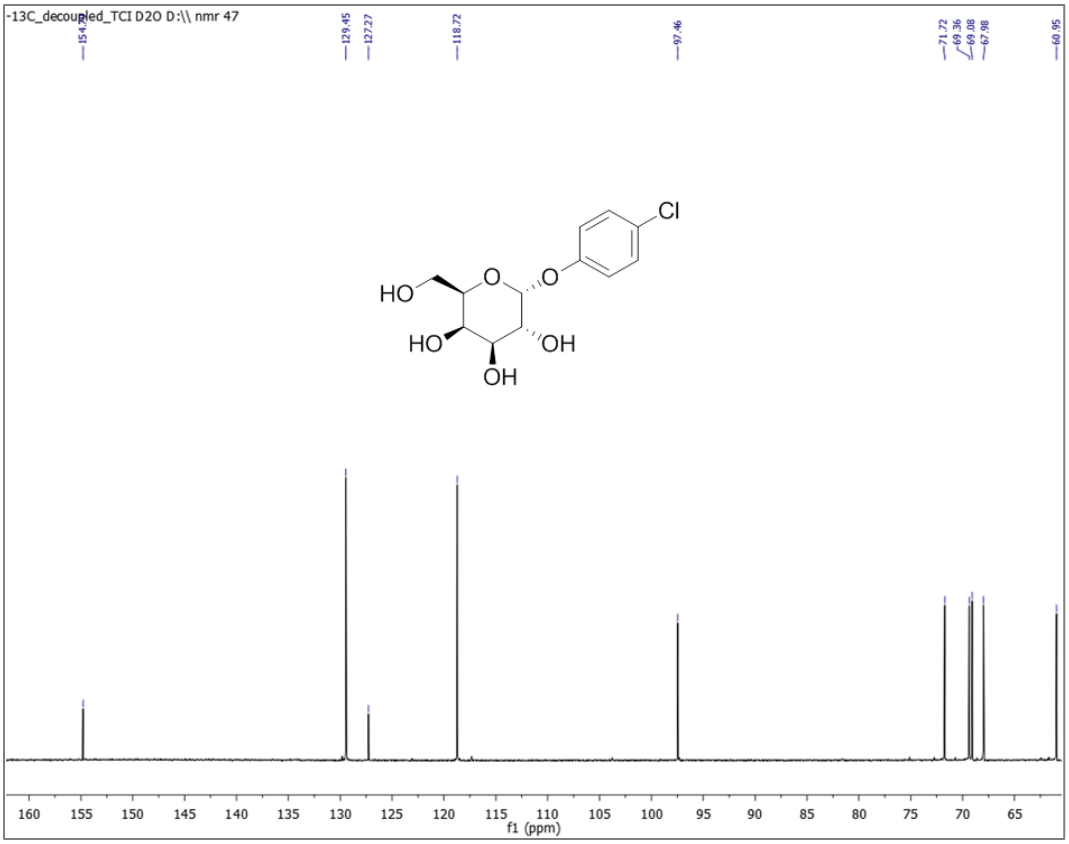


Figure 2S 13. ^{13}C NMR: 4-Chlorophenyl α-D-galactoside (600 MHz, D_2O).

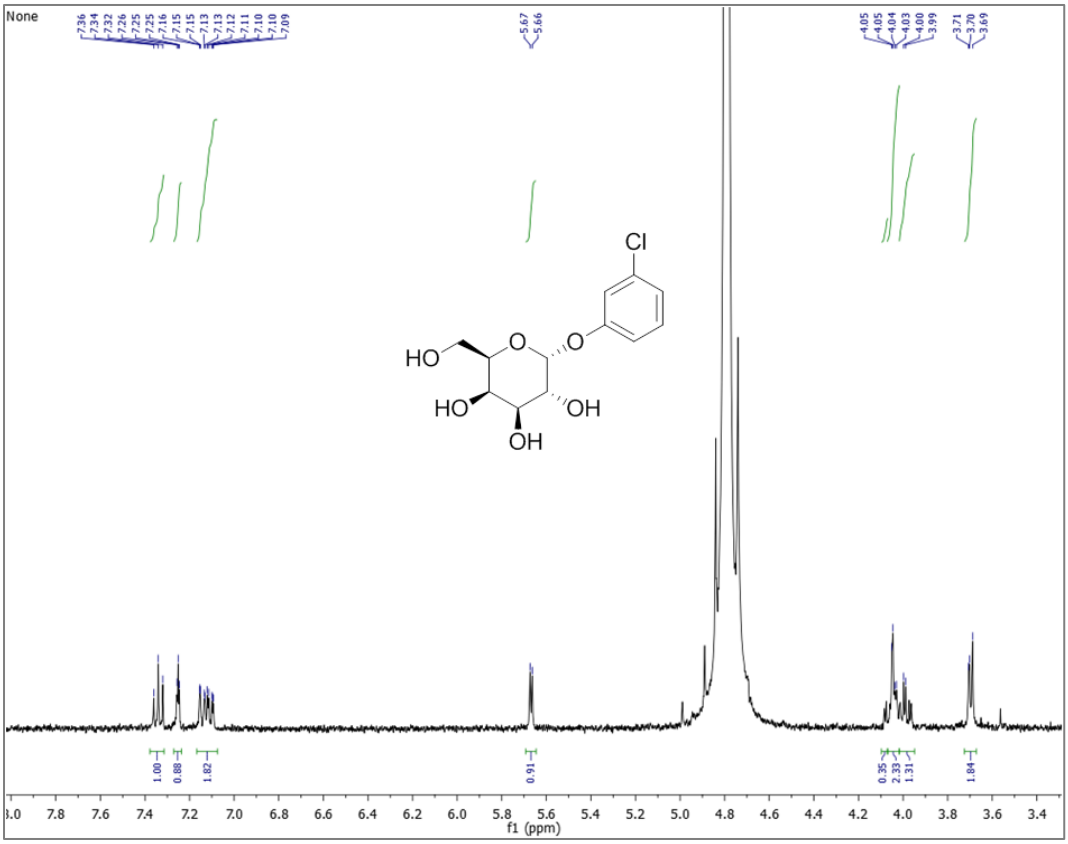


Figure 2S 14. ^{1}H NMR: 3-Chlorophenyl α-D-galactoside (600 MHz, D_2O).

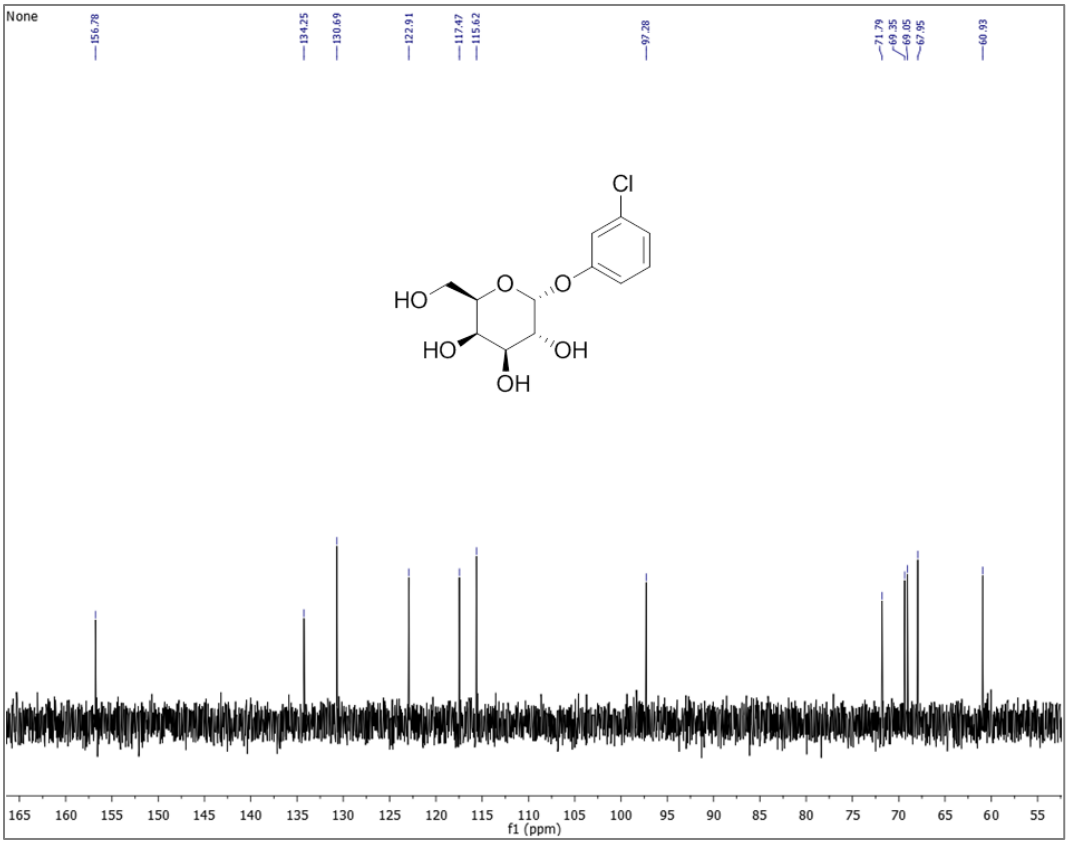


Figure 2S 15. ^{13}C NMR: 3-Chlorophenyl α-D-galactoside (600 MHz, D_2O).

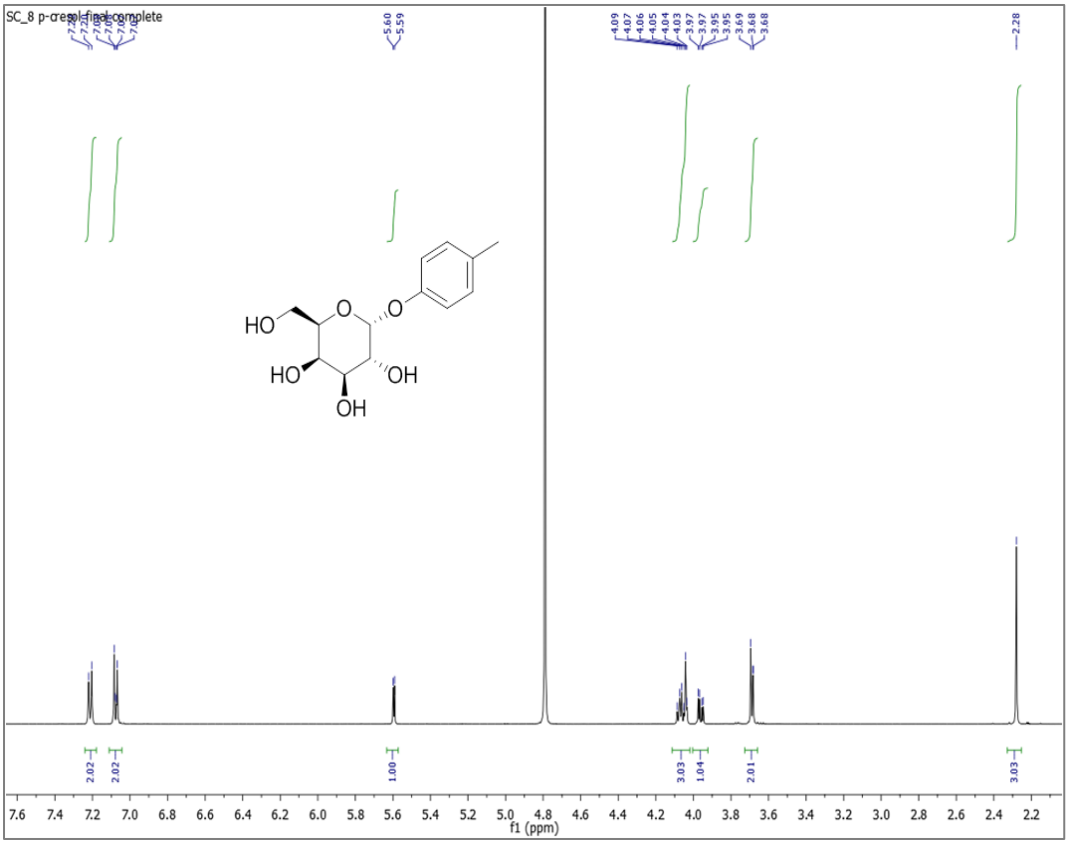


Figure 2S 16. ^{1}H NMR: 4-Methylphenyl α-D-galactoside (600 MHz, D_2O).

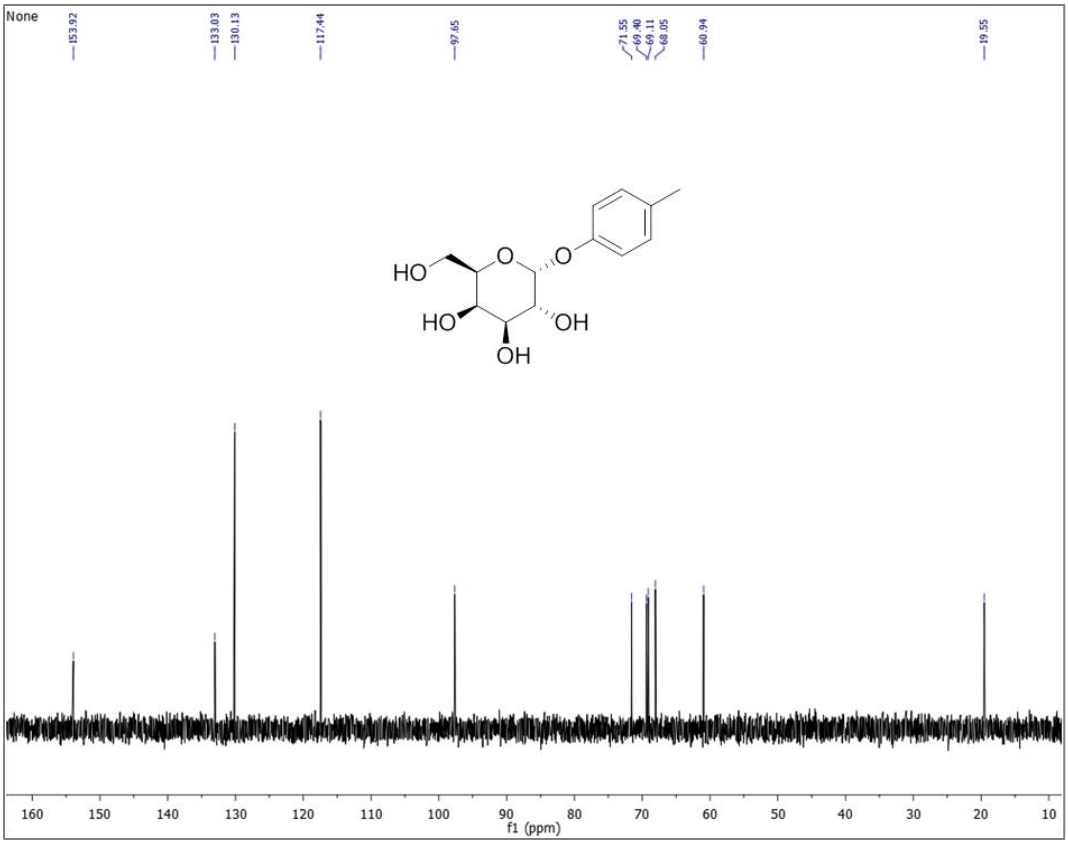


Figure 2S 17. ^{13}C NMR: 4-Methylphenyl α-D-galactoside (600 MHz, D_2O).

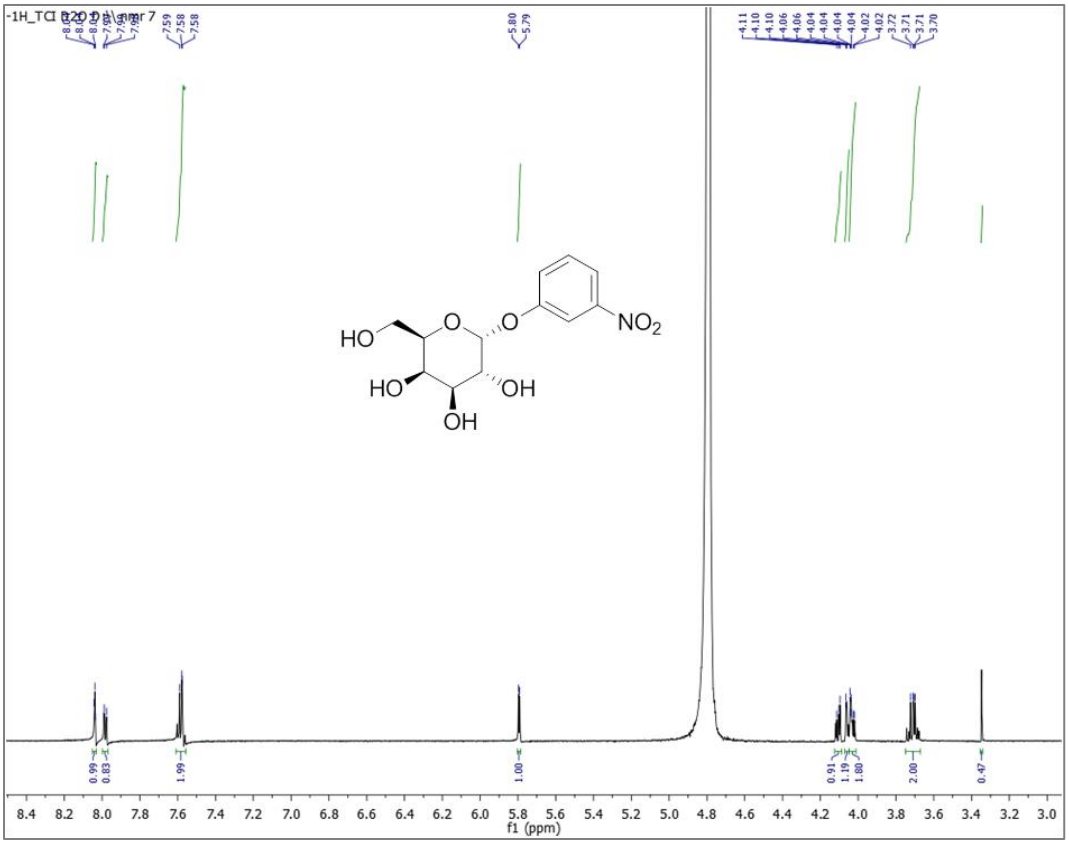


Figure 2S 18. ^{1}H NMR: 3-Nitrophenyl α-D-galactoside (600 MHz, D_2O).

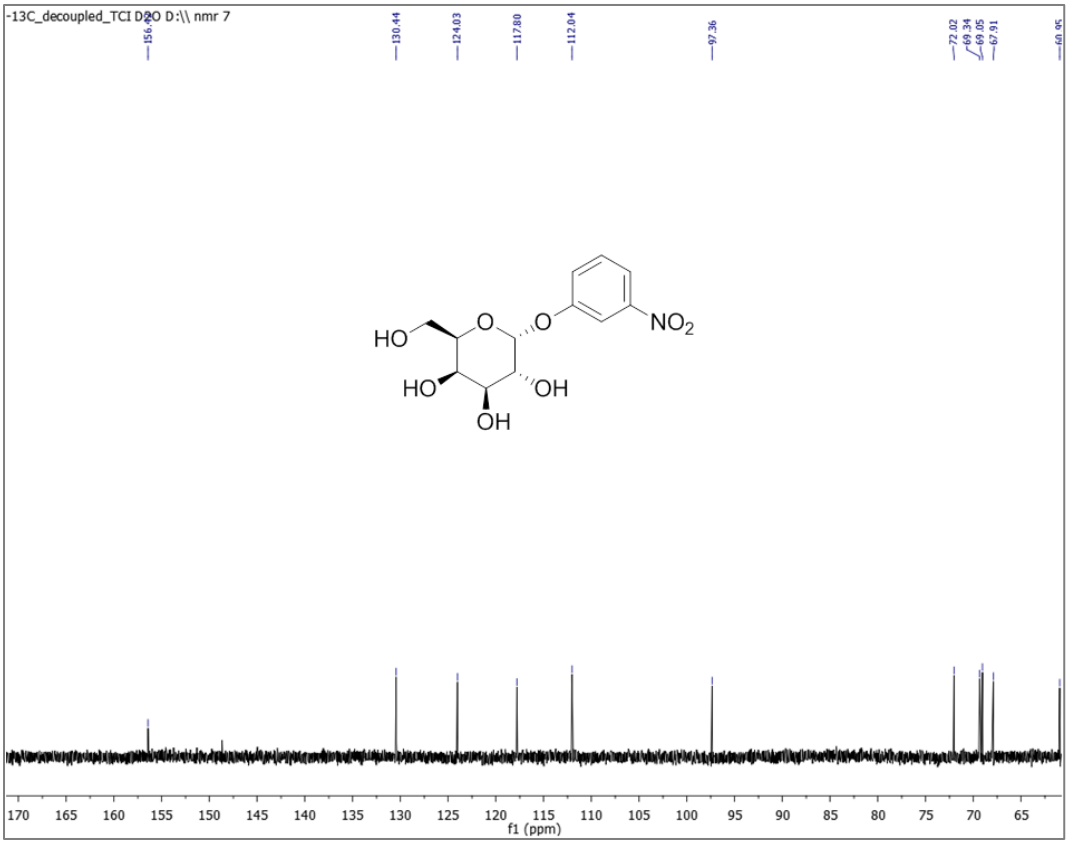


Figure 2S 19. ^{13}C NMR: 3-Nitrophenyl α-D-galactoside (600 MHz, D_2O).

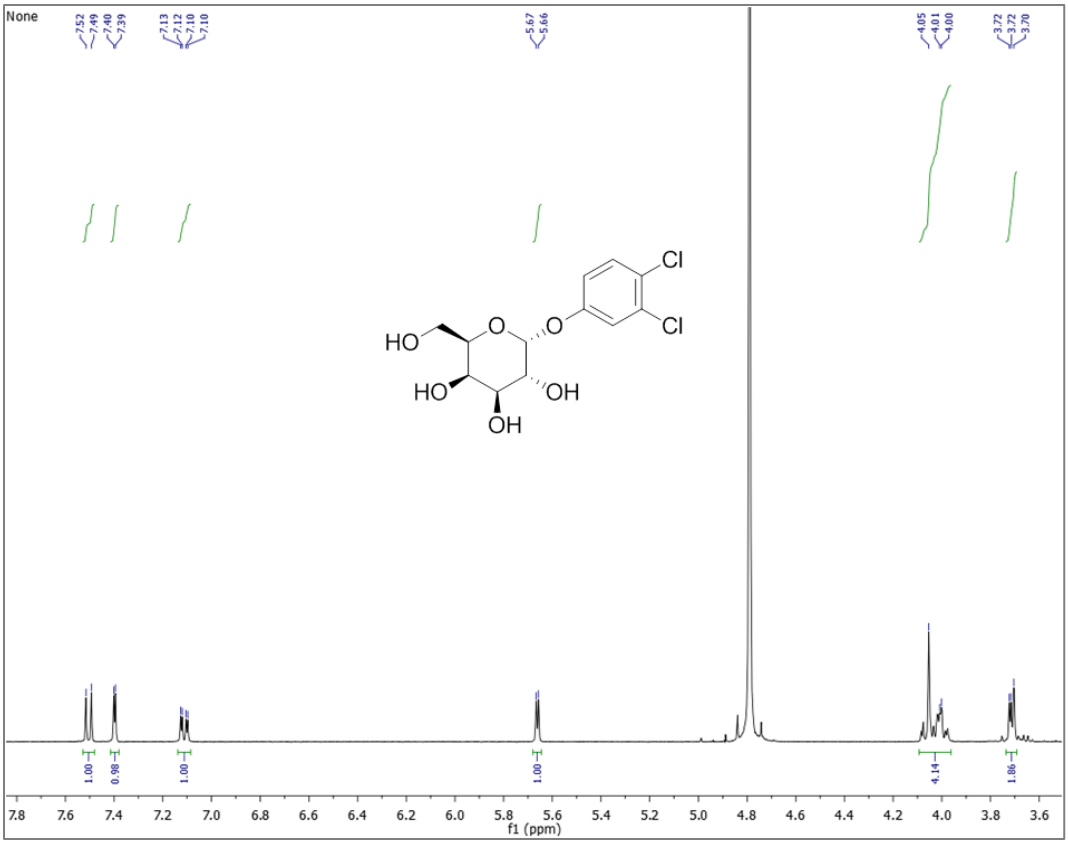


Figure 2S 20. ^{1}H NMR: 3,4-Dichlorophenyl α-D-galactoside (600 MHz, D_2O).

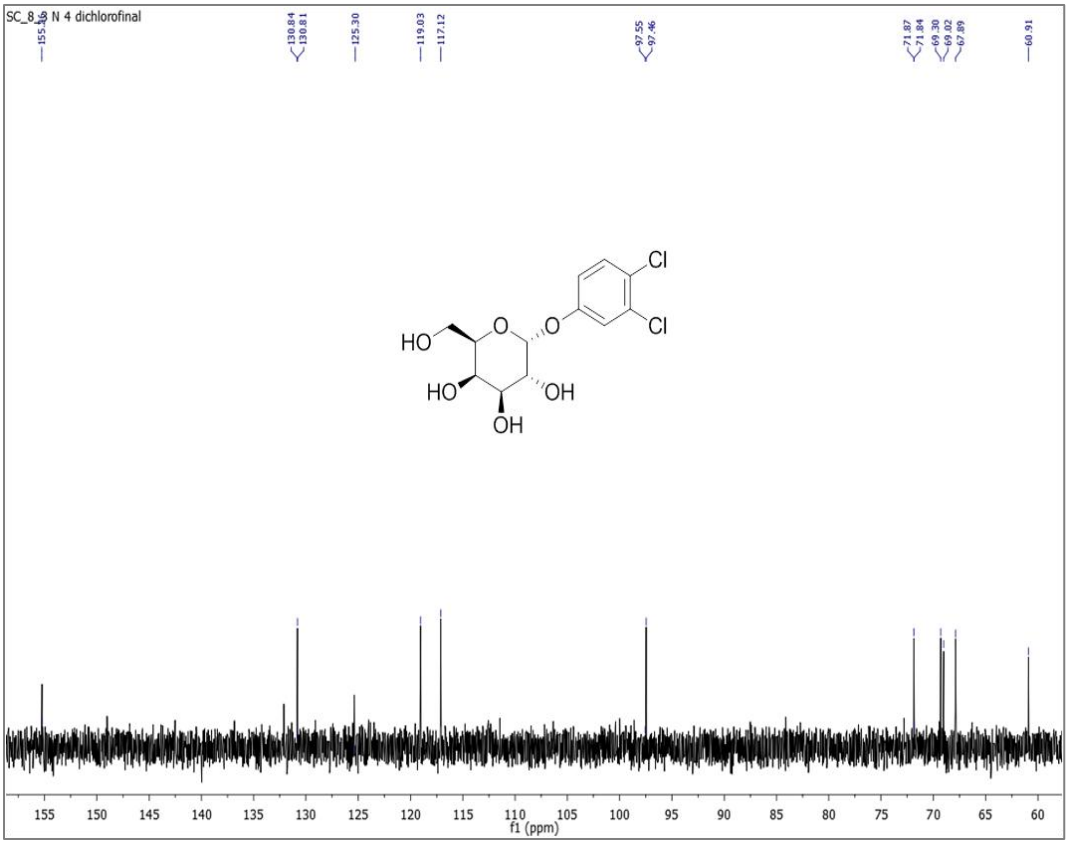


Figure 2S 21. ^{13}C NMR: 3,4-Dichlorophenyl α-D-galactoside (600 MHz, D_2O).

3. A Mechanism-Based Glycosidase Inhibitor with a High Inhibitory Proficiency: Generation of a Cyclopropylmethyl Carbenium Ion in an Enzymatic Active Site

3.1. Introduction

Carbohydrate metabolism is a central process required for life; accordingly enzymes that transfer glycosyl units are ubiquitous in all kingdoms of life. Roughly 1-3% of all proteins encoded by a typical genome possess this transferase activity.(*1*). The importance of these enzymes in the proper functioning of living systems is reflected by over 50 years of research aimed at probing the mechanisms of action of glycosyl transferring enzymes. The state of knowledge is more advanced for understanding the mechanisms utilized by glycosidase enzymes than for glycosyltransferases, in part because it is easier to synthesize substrates for glycosidases than glycosyltransferases. In 1953, Koshland put forward the idea that retaining glycosidase enzymes function via a double inversion mechanism involving an intermediate that is formed by nucleophilic attack of either an enzymatic residue or a neighboring group on the substrate.(*2)* Research findings since that time have refined this generic view, and we now know that, in cases where a covalent glycosyl-enzyme intermediate is formed, the nucleophile can be an aspartate, a glutamate(*3,4*) or a tyrosine(*5,6*) residue. Nucleophilic substitution reactions such as these likely proceed either by weakly associative (S_N2, A_ND_N)(*7,8*) or dissociative (S_N1, $D_N{*}A_N$)(*8,9*) transition states. In any case, it is clear that most glycosidases have evolved to stabilize transition states in which a significant degree of positive charge has developed on the carbohydrate moiety(*10,11*).

As understanding of the mechanistic details of glycosidase-catalyzed reactions continues to improve, researchers use this knowledge to explore strategies for designing covalent inhibitors of retaining glycosidases, which can be used as chemical tools to probe the biological roles of these enzymes.(*12*) In contrast, the development of irreversible

inhibitors for inverting enzymes has lagged behind that for retaining glycosidases; since the inverting enzymes require a water molecule as the nucleophile, potent electrophiles generated in the active site usually react to produce a hydrolysis product rather than an inactivated enzyme.

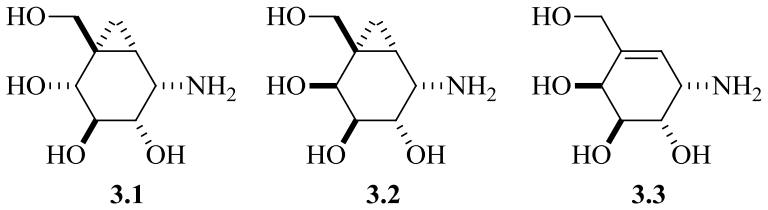
One of the earliest designs for making irreversible labeling agents for retaining enzymes involved appending an intrinsically reactive electrophile, such as a bromoketone(*13*) or bromoamide,(*14*) to a carbohydrate scaffold. Another useful type of reactive functionality, which is also found in natural products such as cyclophellitol,(*15*) is an epoxide, which upon binding to the enzymatic active site, undergoes irreversible nucleophilic ring opening.(*16*) Other strategies for inactivation of glycosidases require a latent electrophile affixed to a carbohydrate structure, which upon protonation generates a potent electrophile; examples of these functionalities include triazenes(*17*) and diazoketones.(*18*)

The most recognizable class of "mechanism-based" irreversible inhibitors is the fluorinated sugars developed by Withers and co-workers.(*19*) In this case, the inhibitor is designed to slow down both glycosylation and deglycosylation (retaining enzymes only) by incorporating a fluorine atom in close proximity to the anomeric center, while ameliorating the effect on the deglycosylation step by incorporating a good leaving group.

As part of an ongoing program of research for the discovery of new structural motifs that inhibit glycosidase, we proposed to test carbohydrate mimics containing known functionalities that are predisposed to form carbocationic intermediates during solvolysis reactions. For more than 50 years,(*20*) scientists have known that cyclopropylmethyl derivatives undergo S_N1 ($D_N + A_N$)(*21*) reactions at greatly accelerated rates in comparison to the corresponding acyclic homologues.(*20, 22, 23)* Accordingly, we decided to test these functional groups.

In the current study we used the bicyclic core skeleton of tight-binding competitive glycosidase inhibitors (**3.1**(*24*) and **3.2**(*25*)) and the carbocyclic structure of the allylic analogue **3.3,**(*26*) and modified the pseudo-anomeric group so that it becomes a potential leaving group that, upon acid-catalyzed departure will generate a cationic intermediate. In

this report we detail the synthesis and biological evaluation, against two retaining α-galactosidases from different glycosyl hydrolase (GH) families, of three analogues (**3.4**, **3.5** and **3.6**; Ar = 3,5-difluorophenyl) of the known galactosidase inhibitors galacto-bicyclo[4.1.0]heptyl amine (**3.2**)(*25*) and galacto-valienamine (**3.3**).(*26*) In addition, three non-hydroxylated analogues of these structural motifs were made to gain insight into how closely the "mechanism-based" inhibitor mimics the mechanism of the GH36 retaining α-galactosidase from *T. maritima* by estimating the enzymatic catalytic proficiency ($k_{cat}/K_m \times 1/k_{uncat}$) and the corresponding inhibitory proficiencies ($k_{inact}/K_i \times 1/k_{uncat}$).



3.2. Methods

Ms Yi Wang contributed significantly in the synthesis and purification of the cyclohexenol intermediates and also performed the initial kinetics experiments reported in this chapter. Ms Lydia Cheng aided this research work by performing the cloning and expression of the recombinant protein (*T. maritima* α-Galactosidase) and helping for the kinetics experiments related to pH profile of inhibition. Ms Shirley Ko carried out the mass spectrometric studies reported in this chapter.

3.2.1. Cloning, Expression, and Purification of *T. maritima* α-Galactosidase.

The *galA* gene encoding *α*-galactosidase from *Thermotoga maritima* MSB8 genomic DNA (ATCC 43589D) was amplified by standard PCR techniques. The amplification was performed in solutions containing 5% DMSO with the following primers (forward,

5’-CGCGGCTAGCATGGAAATATTCGGAAAG-3’; reverse, 5’GCGCGAATTCTACCTG AGTTCCATCATTC-3’), which introduced *NheI* and *EcoRI* restriction sites (underlined) in the forward and reverse primer, respectively. As a result, the first codon of the gene was changed from GTG to ATG. The PCR fragment was purified, digested with *NheI* and *EcoRI*, inserted into correspondingly digested pET28a vector (Novagen), and transformed in *Escherichia coli* BL-21(DE3). The plasmid DNA was isolated from a single colony and was verified by restriction digest and DNA sequencing by Macrogen using T7 promoter and T7 terminator primers.

The protocol used for expression and purification of *Tm* GalA was identical to that reported previously.(*40*) The *galA* gene was expressed in *E. coli* BL-21(DE3) and the recombinant *E. coli* cells were grown in Luria broth media containing 30 mg/L kanamycin. The expression of protein was induced by addition of IPTG (1 mM) when the OD_{600nm} of the cell culture was 0.8. After expression the cells were lysed using 1% lysozyme and 6 cycles of sonication (20 s on 40 s off per cycle with 50 % amplitude) in the presence of PMSF (1 mM). Purification of the protein involved heating at 75 °C for 20 min followed by running a standard Ni-NTA affinity column (Qiagen). The purity of protein was analyzed by activity assay and SDS-PAGE gel. Fractions containing pure protein were combined, concentrated by ultracentrification, and dialyzed against 10 mM Tris pH 7.1 buffer. Quantification of protein concentration was performed by Bradford assay using bovine serine albumin as the standard.

3.2.2. Enzyme Kinetics.

Standard buffer conditions for coffee bean α-galactosidase inhibition were 50 mM sodium phosphate pH 6.52. The kinetic parameters K_i and k_{inact} were determined by using a classic dilution assay method. That is, the inactivation of both coffee bean and *T. maritima* α-galactosidases were monitored by incubating the enzyme with various concentrations of inhibitor in the appropriate buffer. Periodically, the activity of the enzyme was measured by taking an aliquot (5 μL) of the inactivation stock solution and adding to a solution of 4-nitrophenyl α-D-galactopyranoside (0.15 mM in 395 μL of 50 mM NaOAc buffer, pH = 5.2; coffee bean enzyme) or 6,8-difluoro-4-methylumbelliferyl α-D-galactopyranoside

3.7 (0.15 mM in 395 μL of 50 mM NaOAc buffer, pH = 5.2; *T. maritima* enzyme). Pseudo first-order rate constants (k_{obs}) were determined, at each inhibitor concentration, by fitting the absorbance/fluorescence versus time data to a standard first order equation. The first- and second-order rate constants (k_{inact} and k_{inact}/K_i) for the inactivation process were determined by fitting the k_{obs} versus [inhibitor] data to a standard Michaelis-Menten expression.

3.2.3. Initial Reactivation Kinetics.

In a typical assay, *Tm* GalA (40 μL, 0.6 mg/mL) in 20 mM Tris-HCl buffer containing BSA (1% w/v) was added to a mixture containing Tris-HCl buffer (20 μL 20 M, pH 7.1), and **3.4** (20 μL of a 5 mM solution in MeOH). After incubating the solution for various time periods (1 h, 12 h, 48 h) at 25 °C, excess inhibitor was removed and the inactivation buffer was exchanged by filtering the solution using a 5 K molecular weight cutoff centrifugal filter and washing three times with reactivation buffer (3 × 300 μL) at 4 °C to give a final volume of 50 μL. The resultant solution was added to the reactivation buffer (final volume 300 μL), which was incubated at 25 °C. Periodically, α-galactosidase activity was monitored, using a Cary Eclipse Fluorescence Spectrometer at 25 °C, by adding aliquots (5 μL) taken from the reaction mixture to a pre-incubated solution of 6,8-difluoro-4-methylumbelliferyl α-D-galactopyranoside (3 μM) in NaOAc-HOAc buffer (50 mM, pH 5.2) containing BSA (1%). The measured initial rates were fit to a standard first-order rate equation (Prism 4.0) to give estimates for the rate constant of reactivation (k_{react}).

3.2.4. pH-Rate Profile for Inactivation.

A detailed pH profile for inactivation of *Tm* GalA by *galacto*-inhibitor **3.4** was performed over a pH range of 3.0–9.0. The buffers used were Glycine-HCl (20 mM, pH 2.0-4.0), NaOAc-HOAc (20 mM, pH 4.0–4.5), MES (20 mM, MES-NaOH, pH 6.0-6.7), HEPES (20 mM, pH 6.5–8.2), and CHES (20 mM, pH 8.5–9.5) containing NaCl (50 mM). Typically, *Tm* GalA (2.95 μg/mL) was incubated with inhibitor (10 – 500 μM, MeOH) in the appropriate buffer (50 mM) containing BSA (1 % w/v) at 37 °C. Aliquots of the incubation mixture (5 μL) were added to a pre-incubated mixture of fluorogenic substrate 6,8-difluoro-4-methylumbelliferyl α-D-galactopyranoside **3.7** (3 μM), BSA (1 %) and buffer (500 μM) at

37 °C. The initial rate of hydrolysis was determined by monitoring the increase in fluorescence (excitation 345 nm, emission 450 nM). For each concentration of inhibitor 7–8 time points were obtained over a period of 10 min when **3.4** > 50 μM) and 20 min when **3.4** < 50 μM). The measured initial rate data were fit to a standard first-order rate expression to give a pseudo first-order rate constant (k_{obs}) for inactivation. These calculated rate constants (k_{obs}) were fit to a standard saturation kinetic expression (Michaelis-Menten) to obtain the kinetic parameters k_{inact} and k_{inact}/K_i.

3.2.5. pH-Rate Profile for Reactivation.

In a typical assay, *Tm* GalA (40 μL, 1.2 mg/mL) in 20 mM Tris-HCl buffer was added to a mixture containing Tris-HCl buffer (20 μL 20 mM, pH 7.1), BSA (20μL, 1% w/v) and **3.4** (20 μL of a 10.0 mM solution in MeOH). After incubating this inactivation mixture for 1 h at 37 °C, excess inhibitor was removed and the inactivation buffer was exchanged by filtering the solution using a 5 K molecular weight cutoff centrifugal filter and washing three times with reactivation buffer (3 × 300 μL) at 4 °C to give a final volume of 50 μL. The resultant solution was added to the reactivation buffer (final volume 300 μL), which was incubated at 37 °C. Periodically, α-galactosidase activity was monitored, using a Cary Eclipse Fluorescence Spectrometer at 37 °C, by adding aliquots (5 μL) taken from the reaction mixture to a pre-incubated solution of 6,8-difluoro-4-methylumbelliferyl α-D-galactopyranoside **3.7** (3 μM) in NaOAc-HOAc buffer (50 mM, pH 5.2) containing BSA (1%). The measured initial rates were fit to a standard first-order rate equation (Prism 4.0) to give estimates for the rate constant of reactivation (k_{react}).

3.2.6. Inactivation and Labeling of α-Galactosidase.

Incubation of *Tm* GalA α-galactosidase (50 μL of 0.4 mg/mL in 50 mM sodium acetate buffer, pH = 5.2) with *galacto*-isomer **3.4** (1 mM) overnight at room temperature was performed in parallel with a negative control sample. The activity of the samples were checked using a fluorimeter (excitation 345 nM, emission 450 nm) with 6,8-difluoro-4-methylumbelliferyl α-D-galactopyranoside **3.7** (5 μM) as the substrate. After confirming that the inhibited enzyme had negligible activity relative to the control, both samples were

washed (3 × 500 μL) with ammonium bicarbonate pH 6.5 buffer (10 mM) before submitting the samples to mass spectrometric analysis.

3.2.7. Mass Spectroscopic Equipment and Data Acquisition for Analysis of Full Length Enzymes.

A 4000 QTrap LC/MS system was used for analysis and this consisted of a 4000 Qtrap mass spectrometer (Applied Biosystems) and a "Ultimate 3000" high performance liquid chromatography (HPLC) system (Dionex Corporation), which was composed of a binary LC pump, a vacuum degasser, a temperature controlled autosampler and a thermostatted column compartment set to 40 °C. The control software for data acquisition was Analyst version 1.4.2 (Applied Biosystems, Life Technologies Corporation), DCMS-link version 2.0.0.2315 and Chromeleon version 6.80 SP2 (Dionex Corp.). The mobile phase consisted of 0.1 % formic acid (Fluka) + 50 % acetonitrile (HPLC grade: Caledon) + H_2O (LC-MS grade; Mallinckrodt Baker), was pumped at a flow rate of 100 μL/min. A solution of α-galactosidase (5 μL) that had either been untreated or incubated with inhibitor was injected into the autosampler of the HPLC system, and sample was then delivered into the turboionspray ion source by the HPLC pump. The software began the acquisition prior to the injection of the sample, where the resolution of Q1 (first quadruple) was set to unity, and the scanning range was set to be from 300.00 to 2800.00 amu. The dwell time was set to 5.00 s and the step size was set to 0.35 amu. With positive ionization mode set in the acquisition method, the other parameters were curtain gas (10.00 psi), ion spray (5500.00 V), temperature (0.00 °C), nebulizer gas (30.00 psi), turbo ion spray gas (0.00 psi), declustering potential (120.00 V), and entrance potential (10.00 V).

3.2.8. Mass Spectroscopic Equipment and Data Acquisition for Analysis of Tryptic Digestion Experiments.

T. maritima α-galactosidase (20 μL; ~393 pmol) was added to buffered dithiothreitol (16 μL; 15 mg/mL) in ammonium bicarbonate (25 mM) and incubated separately with either methanol (20 μL; control) or inhibitor **5** (20 μL) for 30 min at 37 °C. These two samples were then alkylated by addition of a solution of iodoacetamide (15 mg/mL) in ammonium bicarbonate (38 μL, 25 mM) and incubating the resultant mixture for 30 min at 37 °C.

Sequencing grade porcine trypsin (1 nmol/mL Roche, Fisher Scientific) in ammonium bicarbonate (19.3 µL) was added to each sample, which was then incubated at 37 °C overnight. Before MALDI-TOF/TOF analysis the samples were de-salted by pipetting 5 µL of each sample over a C18 Ziptip (Millipore) and washing with a 3% formic acid/water mixture (3 × 5 µL). The peptides were eluted with a mixture consisting of α-cyano-4-hydroxycinnamic acid (3 mg/mL) and ammonium citrate (1.8 mg/mL) in 70% acetonitrile containing 0.1% trifluoroacetic acid (~ 2 µL) directly onto stainless steel 384 spot MALDI target plate. An Applied Biosystems MDS Sciex 4800 TOF/TOF Mass analyzer was used to collect signal mass range 800-4000m/z for 500 laser shots. The resulting peak list was searched on the public Mascot Server using the following variables: type of search: peptide mass fingerprint; enzyme: trypsin; fixed modifications: carbamidomethyl (C); variable modifications: deamidated (NQ), oxidation (M); mass values: monoisotopic; protein mass: unrestricted; peptide mass tolerance: ± 0.3 Da; peptide charge state: 1+; max missed cleavages:1; and database: NCBInr 20100424 (10924345 sequences; 3720199341 residues).

3.2.9. Measurement of the Rate Constants for the Spontaneous Hydrolysis Reactions.

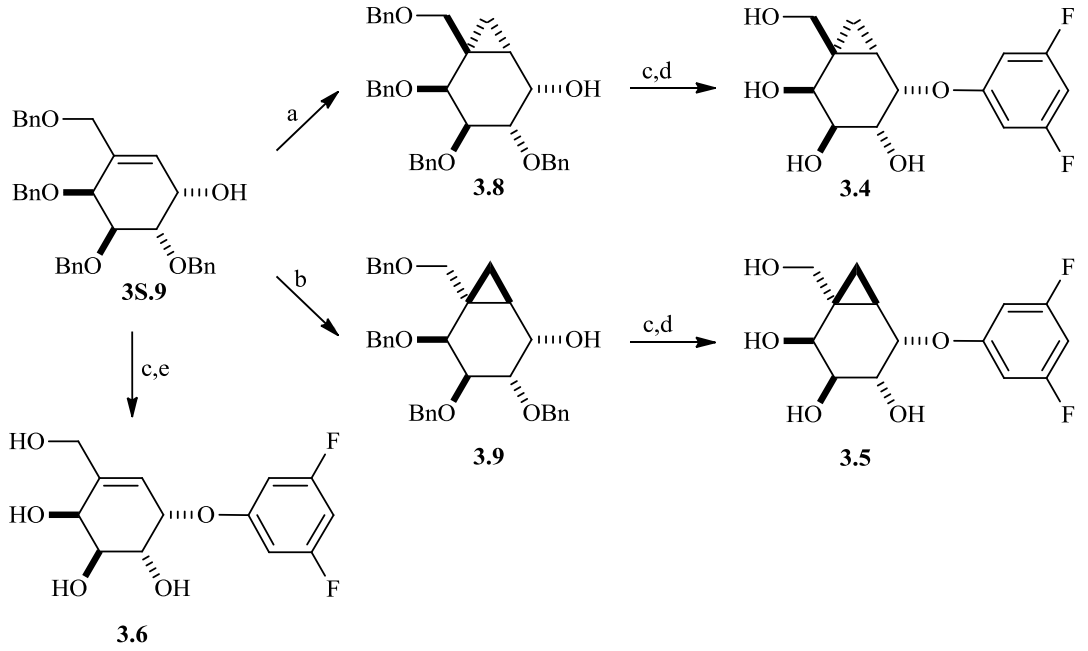
In a typical procedure the substrate was added to a Tris buffer (50 mM, pH 8.5, I = 1.0 KCl) at various temperatures and the rate of hydrolysis was monitored by the release of 4-nitrophenolate anion at 400 nm. First-order rate constants were obtained by following the reaction for at least three half-times and fitting the results to the standard first-order rate equation. In addition, the calculated rate constants were fit to the Eyring equation in order to obtain estimates for the enthalpy and entropy of activation.

3.3. Results

We chose (**Scheme 3-1**) to make the bicyclo[4.1.0]heptyl-based D-*galacto* and L-*altro* analogues **3.4** and **3.5**, which possess pseudo-anomeric phenolic leaving groups, by incorporating minor modifications to the route used for making the *S*-cyclohexenol **3S.9**(*25)* (full details for the synthesis of **3S.9** are given in the supporting information, **Scheme 3S-1**). After several trials, we achieved diastereoselective cyclopropanation of this alkene to give

either **3.8** or **3.9** by inclusion of ~2 equivalents of trifluoroacetic acid in the reaction medium(*28)* and by using toluene or dichloromethane, respectively, as the reaction solvent.

We then subjected each diastereomer to a nucleophilic aromatic substitution reaction (S_NAr, $A_N + D_N$) in order to incorporate, with retention of configuration, a leaving group at the pseudo-anomeric center. Thus, treatment of alcohols **3.8** and **3.9** with sodium hydride in DMSO, followed by the addition of 1,3,5-trifluorobenzene in the presence of potassium benzoate(*29*), gave the corresponding benzylated 3,5-difluorophenyl compounds in yields of 68 and 75%, respectively. We removed the benzyl protecting group by using conventional hydrogenation conditions and obtained the two bicyclo[4.1.0]heptyl carbohydrate analogues **3.4** and **3.5** from methyl α-D-galactopyranoside in a total of 11 steps (**Scheme 3S-1**), with overall yields of 4.2 and 3.7%, respectively. We synthesized the allylic derivative **3.6** by using a similar S_NAr reaction, followed by a BCl_3-promoted debenzylation reaction conducted at –78 °C.



Scheme 3-1. Synthesis of carbocylic analogues **3.4**, **3.5** and **3.6** of an aryl α-D-galactopyranoside. Reagents and conditions: (a) $ZnEt_2$, CH_2I_2, CF_3CO_2H in toluene; (b) $ZnEt_2$, CH_2I_2, CF_3CO_2H in dichloromethane; (c) NaH, 1,3,5-trifluorobenzene in DMF; (d) H_2/Pd-C (10%, w/w) in MeOH; (e) BCl_3, CH_2Cl_2, –78 °C.

We initially assigned the relative stereochemistry of the two cyclopropyl compounds using ^{1}H NMR nOe difference experiments on **3.4**. That is, **Figure 3-1** displays the ^{1}H NMR nOe difference spectra for **3.4**, which show the expected correlation between H-3 and one of the two H-7 cyclopropyl protons.

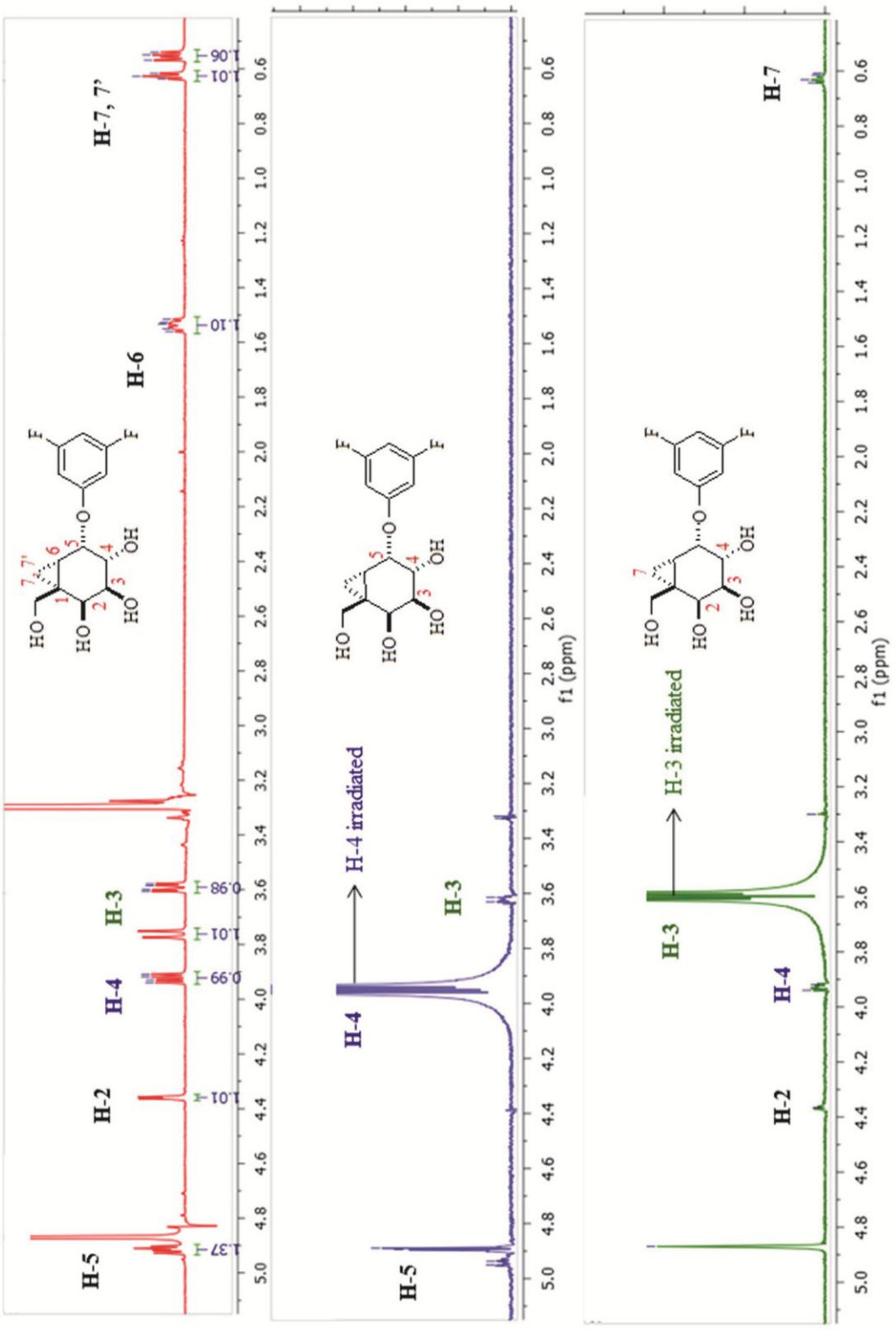


Figure 3-1. [1]H NMR nOe difference spectra used in the analysis of the absolute configuration for **3.4**.

Subsequently, all three 3,5-difluorophenyl compounds were crystallized, and the solved single crystal X-ray structures confirm our assignments (Structures shown in **Figure 3S.4**).

3.3.1. Activity of the Ring Constrained Sugar Analogues 3.4, 3.5 and 3.6.

Incubation of the bicyclo[4.1.0]heptyl carbohydrate analogue **3.4** with the α-galactosidase from coffee bean (GH family 27) results in a time-dependent decrease in enzymatic activity (**Figure 3-2**). Moreover, the rate of inactivation decreases in the presence of the competitive inhibitor **3.2**(*25*) indicating that inhibition occurs at the enzymatic active site (**Figure 3-3**). **Figure 3-4** shows that the inactivation kinetics are consistent with the simplest type of inactivation, one in which the inhibitor binds to the enzyme to give a Michaelis complex (E:I) with a dissociation constant of K_i that undergoes a first-order inactivation event (k_{inact}, **Scheme 3-2**). **Scheme 3-2** also presents the first-order rate constant (k_{react}) for the reactivation of the enzyme, a kinetic term that is included because, at low inhibitor concentration (2 mM, red circles **Figure 3-2**), extrapolation of enzyme activity to infinity is non-zero. In other words, the active enzyme is regenerated by an apparent first-order process (k_{react}), which is slower than k_{inact}.

In comparison to the GH27 coffee bean enzyme, compound **3.4** inactivated the GH36 α-galactosidase from *T. maritima* (*Tm* GalA) at a rate that required the development of a faster and more sensitive assay for enzyme activity. Accordingly, we synthesized 6,8-difluoro-4-methylumbelliferyl α-D-galactopyranoside **3.7**, which gives large changes in fluorescence when it is hydrolyzed (full synthetic details are given in supporting information). Even when using this more sensitive substrate, we observed that the kinetic data for *Tm* GalA inhibition is associated with larger experimental errors because of the rapidity with which it is inactivated; the half-life for reaction of the *Tm*GalA E:I complex is 49 s, while that for the GH27 coffee bean enzyme is 20.6 min (**Table 3-1** lists the measured kinetic parameters).

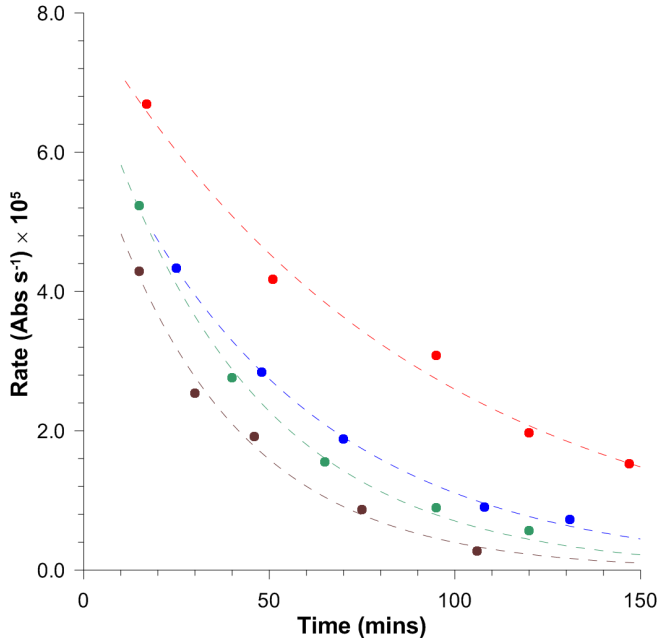


Figure 3-2. Time dependent inactivation of coffee bean α-galactosidase activity during incubation with **3.4**. Inhibitor concentrations are: i) 20 mM (brown circles); ii) 10 mM (green circles); iii) 5 mM (blue circles); and iv) 2 mM (red circles) in 50 mM phosphate buffer at a pH of 6.52 and T = 37 °C. The colored dashed lines are the best non-linear least squares fits to a standard first-order rate equation for the correspondingly colored data points.

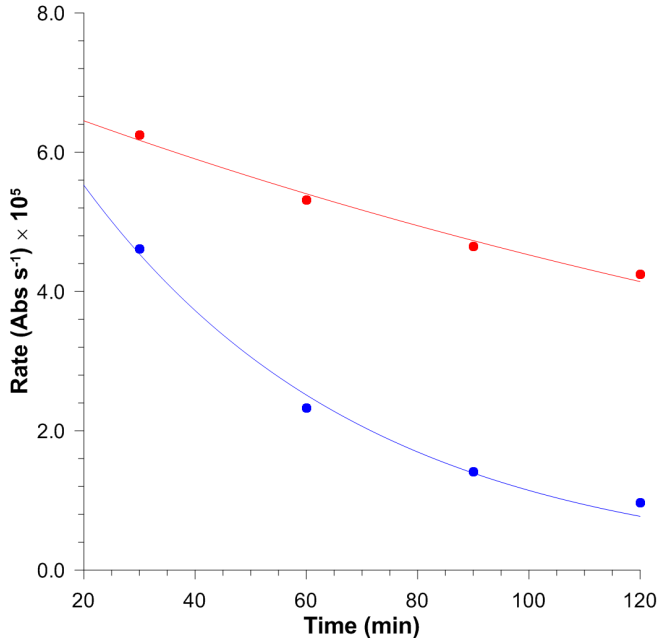


Figure 3-3. Profile for the decrease in coffee bean α-galactosidase activity during incubation with **3.4** (4 mM) in the presence (red circles) and absence (blue circles) of the competitive inhibitor **3.2** (1 μM) in 50 mM phosphate buffer at a pH of 6.52 and T = 37 °C.

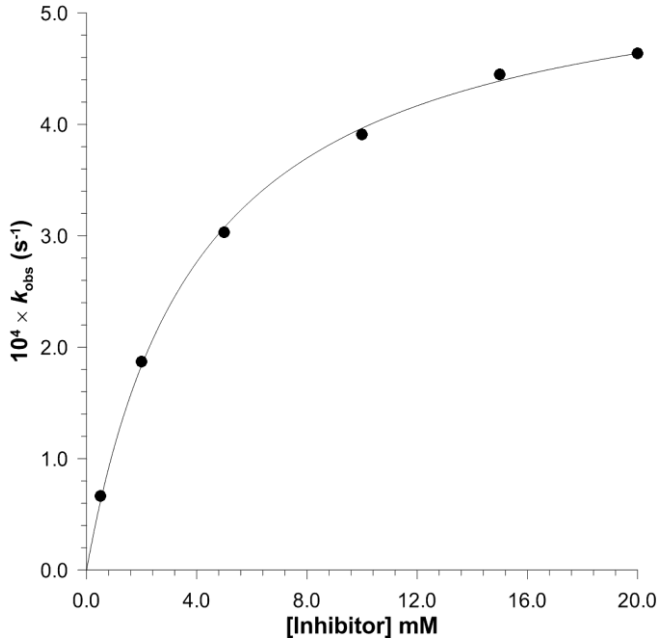


Figure 3-4. Saturation kinetics for the inactivation of coffee bean α-galactosidase activity by **3.4** in 50 mM phosphate buffer at a pH of 6.52 and T = 37 °C. Also shown is the best fit of the data to a standard saturation kinetic expression.

$$E + I \underset{K_i}{\rightleftharpoons} E{:}I \xrightarrow{k_{inact}} E\text{-}I \xrightarrow{k_{react}} E + P$$

Scheme 3-2. Minimal kinetic scheme for the inactivation of α-galactosidases by **3.4**.

Table 3-1. Kinetic parameters for the inactivation of α-galactosidase enzymes by the bicyclo[4.1.0]heptyl compound **3.4** measured at pH values that are close to those for maximal enzyme activity.30,31

Enzyme	Conditions	$10^3 \times k_{inact}$ (s^{-1})	K_i (mM)	k_{inact}/K_i (M^{-1} s^{-1})
Coffee Bean	pH 6.52, *T* = 37 °C	0.56 ± 0.01	4.1 ± 0.2	0.14 ± 0.01
T. maritima	pH 5.20, *T* = 37 °C	14.2 ± 1.1	0.088 ± 0.022	160 ± 40

In contrast to the notable activity of the D-*galacto* diastereomer **3.4**, the L-*altro* isomer **3.5** showed no inhibitory potency whatsoever against the coffee bean and *T. maritima*

α-galactosidases at concentration > 10 mM (data not shown). Whereas, the allylic analogue, of *galacto*-valienamine, **3.6** is a tight-binding competitive inhibitor of the *T. maritima* α-galactosidase with a K_i value of 15.2 ± 1.5 μM (**Figure 3S-1**, supporting information).

3.3.2. Dependence of the Inactivation and Reactivation Rate Constants on pH.

In order to probe whether the protonation states of the two active aspartic acid residues affected the rate of galactosidase inactivation, as they do for the rate of galactoside hydrolysis,(*31)* the rate of *T. maritima* inactivation by **3.4** was monitored as a function of pH (**Figures 3-5** and **3S-3** supporting information). Subsequent reactivation of the enzyme upon inhibitor removal was also followed (**Figure 3-5**).

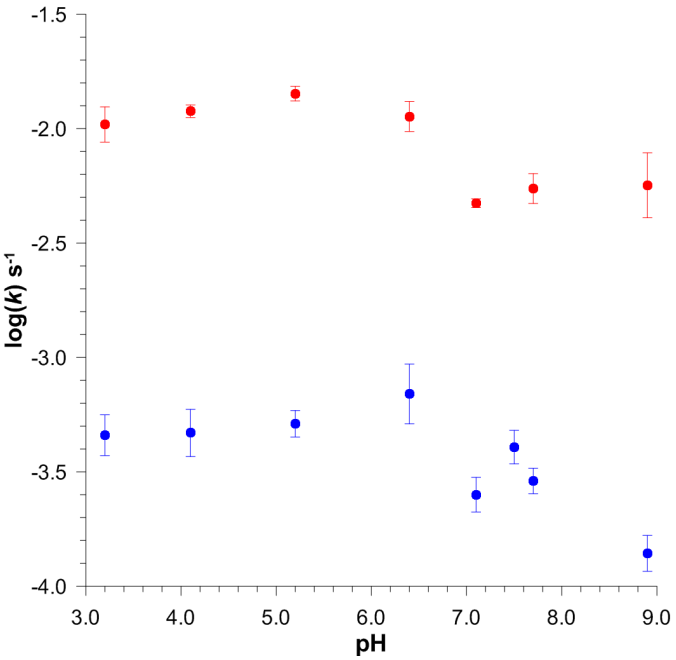


Figure 3-5. pH-Rate profile of the first-order rate constants for inactivation (k_{inact}, red circles) of *Tm* GalA by **3.4** and reactivation (k_{react}, blue circles) of the inhibited enzyme at T = 37 °C.

3.3.3. *Galacto*-Isomer 3-4 Alkylates *T. Maritima* α-Galactosidase.

We probed the mechanism of enzyme inactivation by **3.4** by incubating the α-galactosidase from *T. Maritima* with an excess of inhibitor (see experimental section for full details). After we removed the excess inhibitor, the inactivated enzyme and untreated control native form were analyzed by mass spectrometry. The reconstituted mass spectra of

the full-length protein and the inhibited enzyme are shown in **Figure 3-6**. Importantly, the calculated mass difference between the two samples (66154-65978 = 176) is close to the expected value (172) for alkylation of the enzyme by the pseudo-carbohydrate portion of cyclopropylcarbinyl **3.4**.

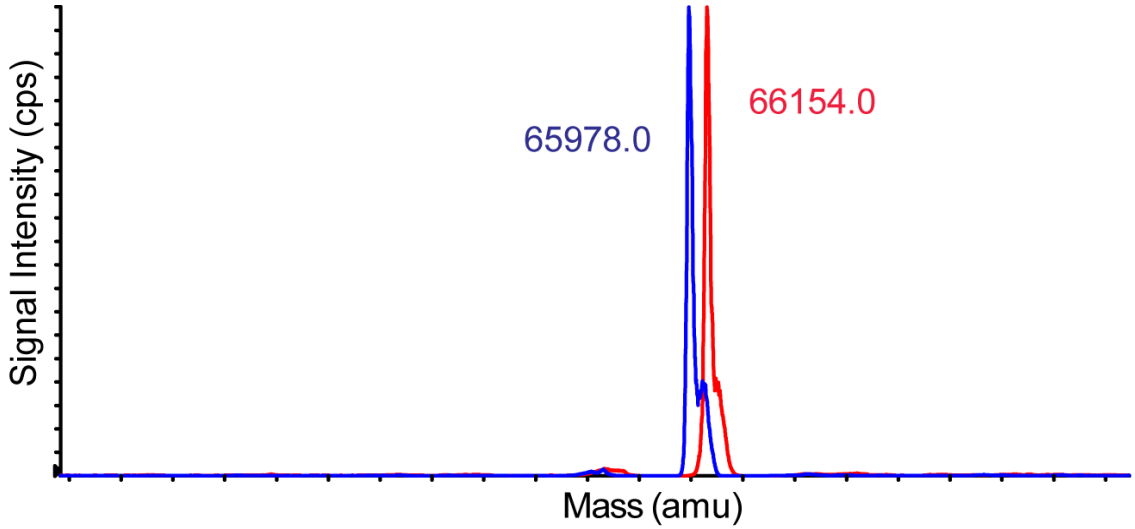


Figure 3-6. Reconstituted ESI-MS for *Thermotoga maritima* α-galactosidase treated (red spectrum) or untreated (blue spectrum) with **3.4** (1 mM) for 12 hrs at RT.

With regard to the reactivation of the enzyme, we have shown that both the rate constant and the final activity of the enzyme are independent of the incubation time of the *T. maritima* α-galactosidase with **3.4** (see experimental section). Specifically, in Tris-HCl buffer (20 mM, pH 7.1) at 25 °C, the values for final enzyme activity and the rate constants for reactivation (k_{react}) after 1 h, 12 h, and 48 h incubations are, respectively, 24.4, 21.8 and 24.3 arbitrary units of fluorescence and 3.27 ± 0.33, 3.56 ± 0.26 and $3.59 \pm 0.28 \times 10^{-4}$ s^{-1}, respectively. Based on these results, we conclude that there is no evidence for a "maturation" process involving two, or more, sites of modification in which one site forms and then reactivates faster.

We tried to identify the site of modification by using standard MS sequencing techniques that include digestion with trypsin or pepsin, however, we have been unable to

obtain sufficient sequence coverage. We show in **Figure 3-7** the protein sequence coverage obtained during the MS analysis of the *T. maritima* α-galactosidase.

1 MEIFGKTFRE GRFVLKEKNF TVEFAVEKIHLGWKISGRVK GSPGRLEVLR
51 TK**APEKVLVN NWQSWGPCRV VDAFSFKPPEIDPNWRYTAS VVPDVLERNL**
101 **QSDYFVAEEG KVYGFLSSK**I AHPFFAVEDG ELVAYLEYFD VEFDDFVPLE
151 PLVVLEDPNT PLLLEK**YAEL VGMENNARVP K**HTPTGWCSW YHYFLDLTWE
201 ETLKNLKLAK **NFPFEVFQID DAYEKDIGDW LVTRGDFPSV EEMAK**VIAEN
251 GFIPGIWTAP FSVSETSDVF NEHPDWVVK**E NGEPKMAYRN WNKKIYALDL**
301 **SKDEVLNWLF DLFSSLR**KMG YRYFKI*D*FLF AGAVPGERKK **NITPIQAFRK**
351 GIETIRK**AVG EDSFILGCGS PLLPAVGCVD GMRIGP*D*TAP FWGEHIEDNG**
401 **APAARWALRN AITRYFMHDR FWLNDPDCLI LR**EEKTDLTQ KEK**ELYSYTC**
451 **GVLDNMIIES DDLSLVR**DHG KK **VLKETLEL LGGRPRVQNI MSEDLR**YEIV
501 SSGTLSGNVK **IVVDLNSREY HLEK** EGKSSL KKRVVKREDG RNFYFYEEGE
551 RE

Figure 3-7. Peptide sequence coverage determined from peptide mass fingerprinting. The two conserved active site aspartic acid residues are underlined (D327 - nucleophile and D387 general-acid catalyst).

3.3.4. Catalytic and Inhibitory Proficiencies.

Our plan was to estimate the enzymatic free energy stabilizations for the catalytic and the inhibitory TSs for this GH36 α-galactosidase; this was not possible, however, due to the extremely slow spontaneous reactions of carbohydrates ($t_{1/2}$ for methyl α-D-glucopyranoside at 37 °C is estimated to be about 1.2×10^6 years).(*32*) We therefore truncated our model for the spontaneous hydrolysis of pyranosyl, bicyclo[4.1.0]heptyl and cyclohex-2-enyl derivatives to the parent structures themselves, i.e., **3.11–3.13**. Synthesis of the 4-nitrophenyl bicyclo[4.1.0]heptane and 4-nitrophenyl cyclohex-2-ene followed standard procedures (full details can be found in the supporting information section). The spontaneous hydrolyses (uncatalyzed hydrolysis reactions) of **3.11–3.13** were monitored at pH values of 7.5 and 8.3 and we observed that the rate constants for the separate pH values were identical. As the reactions of **3.12** and **3.13** are too slow to measure at 37 °C we extrapolated the rate constants for these hydrolyses from kinetic data acquired at high temperatures (**Table 3S-1** lists the measure rate constant values at various temperatures at a pH of 8.3, $I = 1.0$, KCl). In addition, we calculated the activation parameters for the

spontaneous hydrolyses of **3.11**, **3.12** and **3.13** from the Eyring plots **(Figure 3S-2)** and these are $\Delta H^{\ddagger} = 101 \pm 1$, 102 ± 1 and 107 ± 1 kJ mol^{-1}, respectively, while the associated values for $\Delta S^{\ddagger}$ are $+11 \pm 4$, -14 ± 1 and -28 ± 4 J mol^{-1} K^{-1}, respectively.

3.4. Discussion

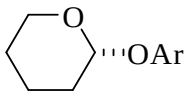
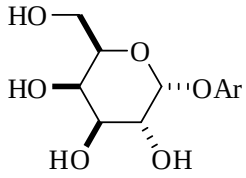
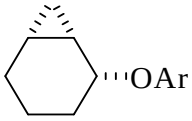
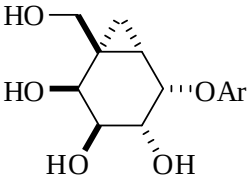
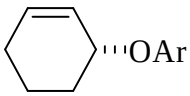
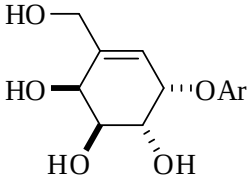
Clearly, the mechanism of inactivation of *Tm* GalA by *galacto* analogue **3.4** needs to be considered in the context of the natural reaction, hydrolysis of α-D-galactopyranosides. The mechanism of action of *Tm* GalA was studied in detail by Comfort *et al.(31)* Briefly, at the pH optimum for activity, the enzyme displayed negligible change in the catalytic constants k_{cat} and k_{cat}/K_m as the leaving group ability of the aglycone was changed (the pK_a of the conjugate acid of the aglycone was varied between 3.96 and 10.32).(*31)* Our analysis of their data gives β_{lg} values on k_{cat} and k_{cat}/K_m of -0.03 ± 0.02 and -0.04 ± 0.04, respectively. Given that k_{cat}/K_m provides information on steps up to and including the first irreversible step, glycosylation, the flat Brønsted plot suggests that a conformational change is kinetically significant, while k_{cat} is likely limited by deglycosylation. In addition, the catalytic activity of *Tm*GalA decreases at both high and low pH values with apparent pK_a values associated with these charges of 3.3 and 6.3, respectively.(*31*). A critical difference for the kinetic analysis of our mechanism-based inhibitors and the standard enzyme-catalyzed hydrolysis reaction is that both of the kinetic parameters for inhibition (k_{inact}/K_i and k_{inact}) report on steps up to and including the "pseudo-glycosylation" step, while k_{react} is the first-order rate constant for the "pseudo-deglycosylation" reaction. Given the above information, we suggest that the pH independence of the inactivation parameter k_{inact} (**Figure 3-5**) is consistent with a kinetically significant non-chemical step because one would expect acid catalysis to be more important for reaction of the intrinsically less reactive cyclopropyl system than for an α-galactoside.

In the model systems, the spontaneous reactions of **3.12** and **3.13** occur with identical enthalpies of activations and their reactivity differences result from differences in their respective entropies of activation. Clearly, the negative entropy of activation for **3.11** results from a more ordered transition state structure that is likely the consequence of stricter

conformational requirement for σ-bond participation than is the case for delocalization of electrons from the n_p orbital on the pyranosyl oxygen atom at the TS. **Table 3-2** presents the rate constants (k_{cat}/K_m or k_{inact}/K_i) for the reactions of the *T. maritima* α-galactosidase with 3,5-difluorophenyl α-D-galactopyranoside and the two carbocycles **3.4** and **3.6**. Also, listed in **Table 3-2** are the spontaneous rates constants for hydrolysis of **3.11**–**3.13** at 37 °C and their corresponding proficiencies (second-order enzymatic rate constant divided by the spontaneous hydrolysis rate constant).

Although the *T. maritima* α-galactosidase has evolved to hydrolyze α-galactosides it is clear that the conformational requires for the formation of a cyclopropylmethyl carbenium ion from **3.4** are closely matched to the glycosylation TS structure as the inhibitory proficiency of **3.4** is only a factor of 20 less that the catalytic proficiency of the enzyme with the corresponding α-galactoside substrate (**Table 3-2**).

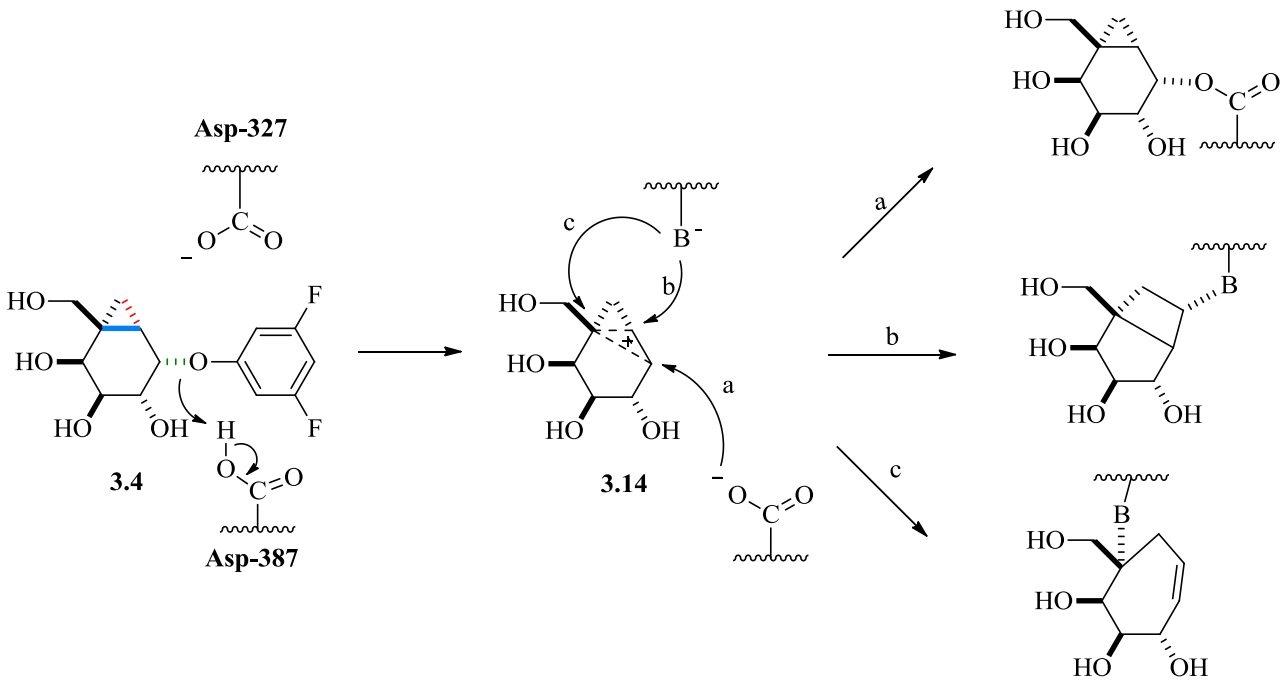
Table 3-2. Kinetic parameters for the spontaneous hydrolysis of compounds **3.11**–**3.13** and the second-order rate constants for the *T. maritima* α-galactosidase-catalyzed hydrolysis of 3,5-diflurophenyl α-D-galactoside and inhibition by **3.4** and **3.6** at 37 °C.

k_{hyd}	k_{cat}/K_m or (k_{inact}/K_i)	$(k_{cat}/K_m \times 1/k_{uncat})$ or $(k_{inact}/K_i \times 1/k_{uncat})$
		$4.17 \times 10^8\ M^{-1}$
$2.40 \times 10^{-4}\ s^{-1}$	$\sim 10^5\ M^{-1}\ s^{-1}$ [a]	
		$2.11 \times 10^7\ M^{-1}$
$7.60 \times 10^{-6}\ s^{-1}$	$160 \pm 40\ M^{-1}\ s^{-1}$	
		
$2.31 \times 10^{-7}\ s^{-1}$	N.I.D.[b]	

Note. [a] reference 31, [b] N.I.D. No inactivation detected; $K_i = 15.2 \pm 1.5\ \mu M$.

3.4.1. Possible Mechanisms of Inactivation.

For more than 50 years, scientists have known that formation of a cyclopropylcarbinyl carbocation is favored when the cation is formed in a bisected geometry rather than in a perpendicular conformation,(*33, 34*) and that rearrangements in the carbon skeleton are generally fast.(*22, 35*) In the case of the chiral molecule **3.4** the two possible C-C σ-bonds that can participate during heterolytic cleavage of the pseudo-glycosidic bond have distinctly different orientations relative to the $\sigma^*_{(C-O)}$ orbital into which electron density must flow in order to form the cationic intermediate. Of the two possibilities (red and blue bonds highlighted in **Scheme 3-3**) with the alignment of the blue bond being more conducive to sigma-bond participation to give cation **3.14** (**Scheme 3-3**).



Scheme 3-3. Postulated mechanism for inhibition of the *T. maritima* α-galactosidase by the mechanism-based inactivator **3.4**.

Within this scheme, we have shown a delocalized cation it is possible that this carbenium ion intermediate (**3.14**) might be a set of rapidly equilibrating classical cations. Of note, reaction of cyclopropylcarbinyl derivatives commonly give three products, namely: (i) cyclopropylcarbinyl; (ii) cyclobutyl; and (iii) homoallylic (reaction arrows a, b and c shown in Scheme 3).(*20*) We show the capture of cationic intermediate **3.14** occurring via pathway a to give a retained product because by microscopic reversibility if σ-bond participation occurs from the backside of the 3,5-difluorophenoxy leaving group then nucleophilic attack, if it occurs prior to any skeletal rearrangement, must occur from the same face from which the leaving group departed. A similar stereochemical outcome has been reported for homoallylic participation that proceeds via such a cationic intermediate.(*36-38*)

Initially, we chose the enzyme from *T. maritima* because a structure of this enzyme had been deposited in the protein crystal database, albeit at a resolution of 0.234 nm (PDB code 1ZY9).(*31,39*) However, despite numerous attempts we were unable to crystallize our protein. As a result, we tried to locate the point(s) of alkylation by tryptic digestion and protein fragment sequencing using mass spectrometry. Unfortunately, the *T. maritima*

α-galactosidase is a remarkably robust enzyme and we need to use very harsh conditions to effect proteolytic cleavage and under these conditions the presumed covalent bond between the enzymatic residue and the carbocyclic component of **3.4** was not stable.

3.4.2. Possible Modes of Reactivation.

Given that the first-order rate constant for reactivation of the inhibited *T. maritima* α-galactosidase (k_{react}) is less than a factor of 100–fold slower that the first-order rate constant for inactivation of the Michaelis complex (k_{inact}) it is likely that the carbocyclic moiety alkylates the enzyme as a cyclopropylmethyl conjugate (pathway a, **Scheme 3-3**) because cyclobutyl and homoallylic isomers are generally much less reactive.(*20*)

3.5. Acknowledgment

This work was supported by a Natural Sciences and Engineering Research Council of Canada.

3.6. References

1 Cantarel, B. L. *et al.* The Carbohydrate-Active enZymes database (CAZy): an expert resource for glycogenomics. *Nucleic Acids Res.* **37**, D233-D238 (2009).

2 Koshland Jr, D. E. Stereochemistry and the mechanism of enzymic reactions. *Biol. Rev. Cambridge Philos. Soc.* **28**, 416-436 (1953).

3 Davies, G., Sinnott, M. L. & Withers, S. G. in *Comprehensive Biological Catalysis* (ed M. L. Sinnott) 119-209 (Academic Press, 1998).

4 Zechel, D. L. & Withers, S. G. Glycosidase Mechanisms: Anatomy of a Finely Tuned Catalyst. *Acc. Chem. Res.* **33**, 11-18 (2000).

5 Watson, J. N., Dookhun, V., Borgford, T. J. & Bennet, A. J. Mutagenesis of the conserved active-site tyrosine changes a retaining sialidase into an inverting sialidase. *Biochemistry* **42**, 12682-12690 (2003).

6 Watts, A. G., Oppezzo, P., Withers, S. G., Alzari, P. M. & Buschiazzo, A. Structural and kinetic analysis of two covalent sialosyl-enzyme intermediates on *Trypanosoma rangeli* sialidase. *J. Biol. Chem.* **281**, 4149-4155, doi:10.1074/jbc.M510677200 (2006).

7 Huang, X. C., Tanaka, K. S. E. & Bennet, A. J. Glucosidase-catalyzed hydrolysis of α-D-glucopyranosyl pyridinium salts: Kinetic evidence for nucleophilic involvement at the glucosidation transition state. *J. Am. Chem. Soc.* **119**, 11147-11154 (1997).

8 Lee, J. K., Bain, A. D. & Berti, P. J. Probing the transition states of four glucoside hydrolyses with C-13 kinetic isotope effects measured at natural abundance by NMR spectroscopy. *J. Am. Chem. Soc.* **126**, 3769-3776 (2004).

9 Tanaka, Y., Tao, W., Blanchard, J. S. & Hehre, E. J. Transition state structures for the hydrolysis of a-D-glucopyranosyl fluoride by retaining and inverting reactions of glycosidases. *J. Biol. Chem.* **269**, 32306-32312 (1994).

10 Chan, J., Lewis, A. R., Gilbert, M., Karwaski, M. F. & Bennet, A. J. A direct NMR method for the measurement of competitive kinetic isotope effects. *Nat. Chem. Biol.* **6**, 405-407 (2010).

11 Chan, J., Lu, A. & Bennet, A. J. Turnover is rate-limited by deglycosylation for *Micromonospora viridifaciens* sialidase-catalyzed hydrolyses: Conformational implications for the Michaelis complex. *J. Am. Chem. Soc.* **133**, 1877-1884 (2011).

12 Witte, M. D., van der Marel, G. A., Aerts, J. M. & Overkleeft, H. S. Irreversible inhibitors and activity-based probes as research tools in chemical glycobiology. *Org Biomol Chem* **9**, 5908-5926, doi:10.1039/c1ob05531c (2011).

13 Howard, S. & Withers, S. G. Bromoketone C-glycosides, a new class of β-glucanase inactivators. *J. Am. Chem. Soc.* **120**, 10326-10331 (1998).

14 Black, T. S., Kiss, L., Tull, D. & Withers, S. G. N-Bromoacetyl-glycopyranosylamines as affinity labels for a β-glucosidase and a cellulase. *Carbohydr. Res.* **250**, 195-202 (1993).

15 Atsumi, S., Iinuma, H., Nosaka, C. & Umezawa, K. Biological activities of cyclophellitol. *J. Antibiot.* **43**, 1579-1585 (1990).

16 Withers, S. G. & Umezawa, K. Cyclophellitol: a naturally occurring mechanism-based inactivator of beta-glucosidases. *Biochem Biophys Res Commun* **177**, 532-537, doi:0006-291X(91)92016-D [pii] (1991).

17 Marshall, P. J., Sinnott, M. L., Smith, P. J. & Widdows, D. Active-site-directed Irreversible inhibition of Glycosidases by the Corresponding Glycosyl methyl-(p-nitrophenyl)triazenes. *J. Chem. Soc., Perkin Trans. 1*, 366-376 (1981).

18 Bemiller, J. N., Gilson, R. J., Myers, R. W. & Santoro, M. M. Suicide-substrate inactivation of β-galactosidase by diazomethyl β-D-galactopyranosyl ketone. *Carbohydr. Res.* **250**, 101-112 (1993).

19 Withers, S. G., Street, I. P., Bird, P. & Dolphin, D. H. 2-Deoxy-2-fluoroglucosides: A novel class of mechanism-based glucosidase inhibitors. *J. Am. Chem. Soc.* **109**, 7530-7531 (1987).

20 Roberts, J. D. & Mazur, R. H. Small ring compounds 4. Interconversion reactions of cyclobutyl, cyclopropylcarbinyl and allylcarbinyl derivatives. *J. Am. Chem. Soc.* **73**, 2509-2520 (1951).

21 Guthrie, R. & Jencks, W. P. IUPAC recommendations for the representation of reaction mechanisms. *Acc. Chem. Res.* **22**, 343-349 (1989).

22 Friedrich, E. C. & Jassawalla, J. D. C. Methyl substituent effects upon the chemistry of 2-bicyclo[4.1.0]heptyl 3,5-dinitrobenzoates. *J. Org. Chem.* **44**, 4224-4229 (1979).

23 Moss, R. A., Zheng, F., Johnson, L. A. & Sauers, R. R. Fragmentation of cyclobutoxychlorocarbene: The cyclopropylcarbinyl/cyclobutyl cations revisited. *J. Phys. Org. Chem.* **14**, 400-406 (2001).

24 Tanaka, K. S. E., Winters, G. C., Batchelor, R. J., Einstein, F. W. B. & Bennet, A. J. A new structural motif for the design of potent glucosidase inhibitors. *J. Am. Chem. Soc.* **123**, 998-999 (2001).

25 Wang, Y. & Bennet, A. J. A potent bicyclic inhibitor of a family 27 α-galactosidase. *Org. Biomol. Chem.* **5**, 1731-1738 (2007).

26 Ogawa, S. *et al.* Convenient synthesis and evaluation of glycosidase inhibitory activity of α- and β-galactose-type valienamines, and some *N*-alkyl derivatives. *Bioorg. Med. Chem.* **12**, 995-1002 (2004).

27 Henrissat, B. & Bairoch, A. Updating the sequence-based classification of glycosyl hydrolases. *Biochem. J.* **316**, 695-696 (1996).

28 Davies, S. G., Ling, K. B., Roberts, P. M., Russell, A. J. & Thomson, J. E. Diastereoselective Simmons-Smith cyclopropanations of allylic amines and carbamates. *Chem. Commun.*, 4029-4031 (2007).

29 Orjales, A. *et al.* Syntheses and binding studies of new [(aryl)(aryloxy)methyl]piperidine derivatives and related compounds as potential antidepressant drugs with high affinity for serotonin (5-HT) and norepinephrine (NE) transporters. *J. Med. Chem.* **46**, 5512-5532 (2003).

30 Zhu, A., Wang, Z. K. & Goldstein, J. Identification of tyrosine 108 in coffee bean a-galactosidase as an essential residue for the enzyme activity. *Biochim. Biophys. Acta* **1247**, 260-264 (1995).

31 Comfort, D. A. *et al.* Biochemical analysis of *Thermotoga maritima* GH36 α-galactosidase (*Tm*GalA) confirms the mechanistic commonality of Clan GH-D glycoside hydrolases. *Biochemistry* **46**, 3319-3330 (2007).

32 Wolfenden, R., Lu, X. & Young, G. Spontaneous hydrolysis of glycosides. *J. Am. Chem. Soc.* **120**, 6814-6815 (1998).

33 Ree, B. R. & Martin, J. C. Perpendicularly twisted cyclopropylcarbinyl and allyl cations. Acetolysis of some 2-substituted derivatives of 1-adamantyl tosylate. *J. Am. Chem. Soc.* **92**, 1660-1666 (1970).

34 de Meijere, A. Bonding properties of cyclopropane and their chemical consequences. *Angew. Chem., Int. Ed. Engl.* **18**, 809-826 (1979).

35 Friedrich, E. C. & Jassawalla, J. D. C. Methyl substituent effects on cyclopropylcarbinyl- cyclopropylcarbinyl cation rearrangements in the 2-bicyclo[4.1.0]heptyl system. *Tetrahedron Lett.*, 953-956 (1978).

36 Huang, X. & Bennet, A. J. Observation of an unusually large rate acceleration caused by a homoallylic double-bond in the solvolyses of an unstrained secondary adamantyl tosylate. *J. Chem. Soc., Perkin Trans. 2*, 1279-1284 (1994).

37 Huang, X. & Bennet, A. J. An unstrained homoallylic secondary adamantyl toluene-p-sulfonate solvolyses with sigma-bond migration to give a protoadamantyl substituted allylic carbenium ion. *J. Chem. Soc., Perkin Trans. 2*, 1027-1033 (1997).

38 Huang, X., Tanaka, K. S. E. & Bennet, A. J. Aqueous Ethanolysis of Unstrained Sterically Congested Homoallylic Halides. *J. Am. Chem. Soc.* **120**, 1405-1409 (1998).

39 Lesley, S. A. *et al.* Structural genomics of the *Thermotoga maritima* proteome implemented in a high-throughput structure determination pipeline. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 11664-11669 (2002).

40 Miller, E. S. *et al.* α-D-Galactosidases from *Thermotoga* species. *Methods Enzymol.* **330**, 246-260 (2001).

3.7. Supporting Information

3.7.1. Materials and Methods

All chemicals were of analytical grade or better and were purchased from Sigma-Aldrich unless noted otherwise. Coffee bean α-galactosidase was purchased from Sigma. Thin-layer chromatography (TLC) was performed on aluminum-backed TLC plates pre-coated with Merck silica gel 60 F_{254}. Compounds were visualized with UV light and/or staining with phosphomolybdic acid (5% solution in EtOH). Flash chromatography was performed using Avanco silica gel 60 (230-400 mesh). Melting points were recorded on a Gallenkamp melting point apparatus and are uncorrected. Solvents used for anhydrous reactions were dried and distilled immediately prior to use. Methanol was dried and distilled over magnesium methoxide. Dichloromethane was dried and distilled over calcium hydride. Glassware for anhydrous reactions was flame-dried and cooled under a nitrogen atmosphere immediately prior to use. NMR spectra were recorded on either a 400 or 500MHz spectrometer. Chemical shifts (δ) are listed in ppm downfield from TMS using the residual solvent peak as an internal reference. ^{1}H and ^{13}C NMR peak assignments are made based on ^{1}H-^{1}H COSY and ^{1}H-^{13}C HMQC experiments. Coupling constants are reported in Hz. IR spectras were recorded on a Bomem IR spectrometer and samples were prepared as cast evaporative films on NaCl plates from CH_2Cl_2. Optical rotations were measured using a Perkin-Elmer 341 polarimeter and are reported in units of deg cm^2 g^{-1} (concentrations reported in units of g/100 mL). 2-(4-Nitrophenoxy)tetrahydro-2H-pyran was made as reported in the literature.(*125*) The synthesis of **3S.4** has been reported in Yi Wang's master's thesis.

3.7.1.1. 2,3,4,6-Tetra-*O*-benzyl-5-methylene-L-*arabino*-hexose diethyl dithioacetal (3S.5).

To a solution of methyltriphenylphosphonium iodide (1.74 g, 4.3 mmol) in THF (20 mL) at –78 °C, n-BuLi (2.5 M in hexane, 1.6 mL, 4.0 mmol) was added slowly. After 1 h, the reaction mixture was warmed up to 0 °C, and a solution of the compound **3S.4** (0.92

g, 1.43 mmol) in THF (10 mL) was added dropwise. The reaction mixture was then allowed to warm slowly to rt over 16 h. The reaction was quenched by the addition of saturated NH_4Cl (50 mL), the solution was then diluted with water (50 mL), and extracted with CH_2Cl_2 (2 × 50 mL). The combined organic layer was washed with brine, dried ($MgSO_4$), and concentrated at reduced pressure. The resultant crude compound was purified by flash column chromatography (Hexane:EtOAc, 9:1) to give the product as a light yellow syrup (0.68 g, 75%): $[\alpha]^{20}_D$ = +13.71° (*c* 1.31, CH_2Cl_2), ^{1}H NMR ($CDCl_3$) δ 1.17 (t, 3 H, $J_{9,8}$ = 7.4, H-9), 1.18 (t, 3 H, $J_{9',8'}$ = 7.5, H-9'), 2.53-2.62 (m, 2 H, H-8), 2.66 (dd, 2 H, H-8'), 3.92–3.95 (m, 1 H, H-5), 4.08 (d, 1 H, $J_{6,5}$ = 5.8, H-6), 4.12–4.21 (m, 4 H, H-3, H-4, H-7, H-7'), 4.26 (d, 1 H, $J_{A,B}$ = 11.6, CH_AH$_B$C$_6$H$_5$), 4.49-4.60 (m, 3 H, CH_2C$_6$H$_5$), 4.66–4.73 (m, 3 H, CH_2C$_6$H$_5$), 4.83 (d, 1 H, $J_{C,D}$ = 11.1, CH$_C$H$_D$C$_6$H$_5$), 5.43 (s, 1 H, H-1), 5.54 (s, 1 H, H-1'), 7.20–7.37 (m, 20 H, ArH); ^{13}C NMR ($CDCl_3$) δ 14.76, 14.81 (C-9, C-9'), 25.2, 25.8 (C-8, C-8'), 54.2 (C-6), 70.6 (*C*H$_2$C$_6$H$_5$), 70.8 (C-7), 73.0, 74.8, 75.6 (3 × *C*H$_2$C$_6$H$_5$), 81.1, 81.6 (C-3, C-4), 83.1 (C-5), 116.2 (C-1), 127.6, 127.6$_2$, 127.8, 127.9, 128.0$_2$, 128.0$_6$, 128.3, 128.4, 128.6, 128.6$_3$, 128.7, 128.8, 129.0, 138.4, 138.6, 139.0, 139.1 (Ar-C). Anal. Calcd For $C_{39}H_{46}O_4S_2$: C, 72.86; H, 7.21. Found; C, 72.79; H, 7.12.

3.7.1.2. (3*R**,4*S*,5*S*,6*S*) 4,5,6-Tribenzyloxy-7-((benzyloxy)methyl)-1,7-octadien-3-ol (3S.6).

Mercury (II) oxide (4.97 g, 22.9 mmol) was added to a solution of **3S.5** (3.7 g, 5.7 mmol) in CH_3CN (150 mL) and H_2O (15 mL). After the solution was stirred for 5 min at room temperature, mercury (II) chloride (2.92 g, 10.7 mmol) was added. The reaction mixture was stirred for another 1.5 h, quenched by saturated $NaHCO_3$ (50 mL), and filtered through celite pad. The filtrate was added into water (50 mL), and extracted with Et_2O (3 × 100 mL). The combined organic layer was washed with brine, dried ($MgSO_4$), and concentrated under reduced pressure to give the crude aldehyde **3S.6**, which was carried on to next step without further purification.

To a solution of **3S.6** in dry THF (50 mL) at –78 °C under nitrogen was slowly added vinyl magnesium bromide (30 mL, 1.0 M in THF, 30 mmol). The reaction mixture was stirred for 1 h at room temperature following which saturated aqueous NH_4Cl (20 mL) was

added and the resultant solution was extracted with diethyl ether (50 mL). The combined organic layer was washed with brine, dried ($MgSO_4$), and concentrated under reduced pressure to afford an orange colored syrup. This residue was purified by flash column chromatography (Hexane: EtOAc, 4:1) to give an almost colorless syrup (1.9 g, compound **3S.7a**/ compound **3S.7b** = 1:2, 60%).

3.7.1.3. (1*R**,4*S*,5*S*,6*S*)-4,5,6-tribenzyloxy-3-((benzyloxy)methyl)cyclohex-2-enol (3S.8 and 3S.9).

To a solution of the diastereomeric mixture of **3S.7a** and **3S.7b** (2.3 g, 4.1 mmol) in dry CH_2Cl_2 (1.1 L) was added the 2nd generation Grubbs catalyst (174.3 mg, 0.21 mmol). This reaction mixture was heated to reflux for 3 h under N_2, and then the volatiles were removed under reduced pressure. The resultant residue was purified by flash column chromatography (Hexane: EtOAc, 2:1 v/v) to afford a colorless syrup. The diastereomers were separated by radial chromatography (CH_2Cl_2: EtOAc, 25:1 v/v) to give **3S.8** (0.7 g, 32%) and **3S.9** (1.2 g 54%) as colorless syrups:

Compound **3S.9**: $[\alpha]^{20}_D$ = +72.6° (*c* 1.22, CH_2Cl_2); IR 3445 cm^{-1} (br, O-H); ^{1}H NMR ($CDCl_3$) δ 2.70 (d, 1 H, $J_{OH,1}$ = 4.4, OH), 3.87 (d, 1 H, $J_{7',7}$ = 12.5, H-7), 3.92 (dd, 1 H, $J_{5,4}$ = 3.5, $J_{5,6}$ = 8.9, H-5), 4.03 (dd, 1 H, $J_{6,1}$ = 4.3, H-6), 4.10 (d, 1 H, H-7'), 4.24 (d, 1 H, H-4), 4.36 (m, 1 H, H-1), 4.38 (d, 1 H, $J_{A,B}$ = 12.0, CH_AH$_B$C$_6$H$_5$), 4.45 (d, 1 H, CH$_A$$H_BC_6H_5$), 4.54 (d, 1 H, $J_{C,D}$ = 11.3, CH_CH$_D$C$_6$H$_5$), 4.67 (d, 1 H, $J_{E,F}$ = 11.7, CH_EH$_F$C$_6$H$_5$), 4.72–4.81 (m, 3 H, CH_2C$_6$H$_5$), 4.84 (d, 1 H, CH$_C$$H_DC_6H_5$), 5.81 (d, 1 H, $J_{2,1}$ = 4.1, H-2), 7.24–7.39 (m, 20 H, H-Ar); ^{13}C NMR ($CDCl_3$) δ 65.7 (C-1), 70.5 (C-7), 72.1, 73.4, 73.4$_3$ (3 × *C*H$_2$C$_6$H$_5$), 73.6 (C-4), 74.3 (*C*H$_2$C$_6$H), 76.6 (C-5, C-6), 126.3 (C-2), 127.5$_4$, 127.5$_8$, 127.6$_2$, 127.6$_5$, 127.7, 127.9, 127.9$_3$, 128.2, 128.2$_7$, 128.3$_1$, 128.3$_5$, 128.5 (Ar-C), 137.2, 138.1, 138.2, 138.6, 138.6$_3$ (Ar-C, C-3). Anal. Calcd For $C_{35}H_{36}O_5$: C, 78.33; H, 6.76. Found; C, 78.15; H, 6.83.

3.7.1.4. (1*S*,2*S*,3*S*,4*S*,5*S*,6*R*)-3,4,5-tribenzyloxy-6-((benzyloxy)methyl)bicyclo[4.1.0]heptan-2-ol (3.8).

Under a nitrogen atmosphere, a diethyl zinc solution in hexane (1.0 M, 3 mL) was added to cooled dry toluene (10 mL; –10 °C). This mixture was stirred at –10 °C for 10 min and then CH_2I_2 (0.2 mL, 1.9 mmol) was added dropwise to the reaction mixture. After

10 min, the reaction mixture under cooling conditions, was treated with trifluoroacetic acid (30 μL, 0.38 mmol), following which the cooling bath was removed and the reaction mixture was stirred at room temperature for 5 min. To the resultant mixture, a solution of **3S.9** (100 mg, 0.19 mmol) in dry toluene (5 mL) was added and the reaction was stirred for 15 h at room temperature. The reaction was quenched by the addition of aqueous H_2SO_4 (10%, 5 mL) and then diluted by adding EtOAc (30 mL). After separation of the two layers, the organic layer was washed with saturated aqueous $NaHCO_3$ (25 mL), brine (20 mL), dried over Na_2SO_4, and the resulting solution was concentrated under reduced pressure. The resultant solid residue was purified via flash column chromatography (eluent 20% EtOAc-Hexane) to afford the pure *galacto*-isomer **3.8** as a colorless oil (81 mg, 79%, >98% de): $[\alpha]^{20}_D$ = +55.9° (*c* 2.40, CH_2Cl_2), IR 3530 cm^{-1} (br, O-H); ^1H NMR ($CDCl_3$) δ 0.72 (dd, 1 H, $J_{7,7'}$ = 9.4, $J_{7,1}$ = 5.7, H-7), 0.88 (m, 1 H, H-7'), 1.21 (m, 1 H, H-1), 2.61 (d, 1 H, $J_{8,8'}$ = 9.3, H-8), 2.87 (d, 1 H, $J_{OH,2}$ = 1.5, OH), 3.53 (dd, 1 H, $J_{3,4}$ = 10.0, $J_{4,5}$ = 2.5, H-4), 3.99 (dd, 1 H, $J_{2,3}$ = 5.5, H-3), 4.14 (dd, 1 H, H-8'), 4.40 (m, 1 H, H-2), 4.42–4.49 (m, 3 H, H-5, C*H*$_2$C$_6$H$_5$), 4.58 (d, 1 H, $J_{A,B}$ = 10.7, C*H*$_A$H$_B$C$_6$H$_5$), 4.66 (d, 1 H, $J_{C,D}$ = 11.6, C*H*$_C$H$_D$C$_6$H$_5$), 4.70 (d, 1 H, $J_{E,F}$ = 11.8, C*H*$_E$H$_F$C$_6$H$_5$), 4.75 (d, 1 H, CH$_E$*H*$_F$C$_6$H$_5$), 4.78 (d, 1 H, CH$_C$*H*$_D$C$_6$H$_5$), 4.99 (d, 1 H, CH$_A$*H*$_B$C$_6$H$_5$), 7.25–7.38 (m, 20 H, H-Ar); ^{13}C NMR ($CDCl_3$) δ 11.6 (C-7), 20.1 (C-1), 27.5 (C-6), 64.5 (C-2), 72.9, 73.0, 73.5, (3 × *C*H$_2$C$_6$H$_5$), 74.3 (C-8), 75.2 (*C*H$_2$C$_6$H), 76.4 (C-5), 77.2 (C-3), 77.3 (C-4), 127.3, 127.4, 127.4, 127.6, 127.7, 127.8, 127.9, 128.0, 128.2, 128.3, 128.4, 138.2$_2$, 138.2$_5$, 138.8, 139.2 (Ar-C), Anal. Calcd For $C_{36}H_{38}O_5$: C, 78.52; H, 6.96. Found; C, 78.38; H, 7.13.

3.7.1.5. (1*R*,2*S*,3*S*,4*S*,5*S*,6*S*)-3,4,5-tribenzyloxy-6-((benzyloxy)methyl)bicyclo[4.1.0]heptan-2-ol (3.9).

Under a nitrogen atmosphere, a diethyl zinc solution in hexane (1 M, 1.5 mL) was added to cooled dry CH_2Cl_2 (5 mL; –78 °C). The resultant solution was stirred at –78 °C for 10 min and then CH_2I_2 (0.10 mL, 0.93 mmol) was added dropwise to the mixture. After 15 min of stirring, the dry-ice acetone cooling bath was replaced by an ice/water bath and the reaction mixture was treated with trifluoroacetic acid (15 μL, 0.19 mmol). After stirring at 0 °C for 15 min, a solution of **3S.9** (50 mg, 0.093 mmol) in CH_2Cl_2 (2.5 mL) was added to the reaction mixture. Subsequently, the cooling bath was removed and the reaction

temperature was allowed to warm up to room temperature over a period of 1.5 h. The reaction was quenched by the addition of saturated NH_4Cl solution (5 mL) and the product was extracted from the aqueous layer with CH_2Cl_2 (2 × 15 mL). The combined organic layers were then washed with saturated aqueous $NaHCO_3$ (30 mL), brine (25 mL) and dried over Na_2SO_4, and the resulting solution was concentrated under reduced pressure. The resultant crude product was purified via flash column chromatography (eluent 20% EtOAc-Hexane) afforded the pure *altro*-isomer **3.9** as a colourless oil (40 mg, 78%, >98% de): $[\alpha]^{20}_D$ = +42.9° (*c* 2.37, CH_2Cl_2), IR 3560 cm^{-1} (br, O-H); ^1H NMR ($CDCl_3$) δ 0.65 (dd, 1 H, $J_{7,7'}$ = 10.2, $J_{7,1}$ = 5.3, H-7), 0.86 (t, 1 H, $J_{7,7'} + J_{7',1}$ = 11.9, H-7'), 1.24 (m, 1 H, H-1), 2.91 (d, 1 H, $J_{OH,5}$ = 1.6, OH), 3.17 (d, 1 H, $J_{8,8'}$ = 10.2, H-8), 3.30 (d, 1 H, H-8'), 3.67 (dd, 1 H, $J_{3,4}$ = 10.0, $J_{2,3}$ = 3.6, H-3), 3.84 (dd, 1 H, $J_{4,5}$ = 4.1, H-4), 4.22 (d, 1 H, H-5), 4.36 (m, 1 H, H-2), 4.44 (d, 1 H, $J_{A,B}$ = 12.1, CH_AH$_B$C$_6$H$_5$), 4.52 (d, 1 H, CH$_A$$H_BC_6H_5$), 4.58 (d, 1 H, $J_{C,D}$ = 12.6, CH_CH$_D$C$_6$H$_5$), 4.61–4.66 (m, 3 H, CH_2C$_6$H$_5$), 4.71 (d, 1 H, CH$_C$$H_DC_6H_5$), 4.84 (d, 1 H, $J_{E,F}$ = 11.5, CH_EH$_F$C$_6$H$_5$), 7.22-7.37 (m, 20 H, H-Ar), ^{13}C NMR ($CDCl_3$) δ 9.2 (C-7), 24.0 (C-1), 25.1 (C-6), 68.2 (C-2), 70.5, 72.3, 72.4 (3 × *C*H$_2$C$_6$H$_5$), 73.3 (C-5), 73.3$_2$ (*C*H$_2$C$_6$H$_5$), 74.5(C-3), 76.3 (C-4), 76.6 (C-8), 127.2, 127.3, 127.5, 127.6, 127.6$_3$, 127.7, 127.8, 128.1, 128.2, 128.3, 128.4, 138.3$_6$, 138.4, 138.8, 138.9 (Ar-C), Anal. Calcd For $C_{36}H_{38}O_5$: C, 78.52; H, 6.96. Found; C, 78.35; H, 7.13.

3.7.1.6. (3*S*,4*S*,5*S*,6*S*)-4,5,6-tribenyzloxy-1-((benzyloxy)methyl)-3-(3,5-difluorophenoxy)cyclohex-1-ene (3S.10).

A suspension of NaH in mineral oil (60%, 99.4 mg, 2.48 mmol) was washed with hexane (2 × 5 mL) before being transferred in dry DMSO (50 mL) to a 100 mL flask maintained at 18 °C. To this mixture a solution of **3S.9** (200 mg, 0.37 mmol) in dry DMSO (25 mL) was added dropwise. This mixture was left for 30 min at 18 °C before treating with potassium benzoate (100 mg, 0.62 mmol). After another 30 min of stirring, 1,3,5-trifluorobenzene (700 μL, 6.7 mmol) was added slowly, continuing the stirring for another 30 min. Upon completion, the reaction mixture was quenched by the addition of a saturated NH_4Cl solution (20 mL). This was followed by addition of brine (50 mL) and extraction of the product from the aqueous layer with diethyl ether (3 × 50 mL). The combined organic layer was dried ($MgSO_4$), and concentrated under reduced pressure. The

resultant residue was purified by flash column chromatography (20% EtOAc-Hexane) to afford the product as colourless syrup **3S.10** (170 mg, 70%) in > 98% purity as determined by ^{1}H NMR. This material was taken forward for the next step without further purification.

3.7.1.7. (1*R*,2*S*,3*S*,6*S*)-4-(hydroxymethyl)-6-(3,5-difluorophenoxy)-cyclohex-4-ene-1,2,3-triol (3.6).

To a solution of the **3S.10** (170 mg) in dry CH_2Cl_2 (50 mL) under an argon atmosphere, boron trichloride (5.0 equiv., 1 M soln in CH_2Cl_2) was added slowly *via* a syringe at –78 °C, and the reaction was maintained at this temperature whilst being stirred for 30 min. Subsequently, the resultant mixture was allowed to warm up to 0 °C over a period of 30 min. Upon completion of reaction, as evident from TLC (2:8, EtOAc:Hexane), a solution of 1:1 MeOH–CH_2Cl_2 (5 mL) was added in order to quench the reaction. The volatiles were removed under diminished pressure and the resultant residue was washed with CH_2Cl_2 (5 × 10 mL). A white solid was obtained upon complete removal of solvent, that was further recrystallized from MeOH to afford the final compound **3.6** (50 mg, 66 %); $[\alpha]^{20}_D$ = +101.5° (*c* 0.2, MeOH), Mpt = 170–171 °C; ^{1}H NMR (500 MHz, CD_3OD) δ 6.55 (dd, *J* = 9.3, 2.0, 2H, H-2′, H-6′, ArH), 6.40 (m, 1H, H-4′, ArH), 5.90 (d, $J_{5,6}$ = 3.6, 1H, H-5), 5.01 (app.br s, 1H, H-6), 4.32 (d, 1H, $J_{3,2}$ = 4.0, H-3), 4.24-4.13 (m, 3H, H-2, H-7a, H-7b), 3.93 (dd, 1H, $J_{1,2}$ = 8.8, $J_{1,6}$ = 4.1, H-1), 3.26 (s, 1H, OH); ^{13}C NMR (151 MHz, CD_3OD) δ 165.13 (dd, $^1J_{C,F}$ = 244.8, $^3J_{C,F}$ = 16.0, C-3', C-5'), 162.14 (t, $^3J_{C,F}$ = 13.9, C-1′), 144.46 (C-4, Alkene), 120.18 (C-5, Alkene), 100.6 (m, ArC, C-2', C-6'), 97.00 (t, $^3J_{C,F}$ = 26.4, C-4′), 75.21 (C-6), 70.92 (C-1), 69.94 (C-2), 67.88 (C-3), 63.64 (C-7); Anal. Calcd for $C_{13}H_{14}F_2O_5$: C, 54.17; H, 4.90. Found; C, 54.10; H, 4.82.

3.7.1.8. (1*R*,2*S*,3*S*,4*R*,5*S*,6*S*)-5-(3,5-difluorophenoxy)-1-(hydroxymethyl)bicyclo[4.1.0]heptan-2,3,4-triol (3.4, Scheme 3-1).

After a suspension of NaH in mineral oil (60%, 100 mg, 2.48 mmol) was washed with hexane (2 × 5 mL), it was transferred in dry DMSO (20 mL) into 100 mL flask maintained at 18 °C. To this mixture a solution of **3.8** (176 mg, 0.32 mmol) in dry DMSO (25 mL) was added dropwise. The mixture was stirred for 30 min at 18 °C before treating with potassium benzoate (97 mg, 0.61 mmol), following which, stirring was continued for another 30 min.

At this time, 1,3,5-trifluorobenzene (624 μL, 6.0 mmol) was added slowly. The reaction was quenched after 30 min of stirring by the addition of a saturated NH_4Cl solution (10 mL). Following the addition of brine (25 mL), the product was extracted from the aqueous layer with diethyl ether (3 × 30 mL). The combined organic layer was dried ($MgSO_4$) and concentrated under reduced pressure. The resultant crude residue was purified via flash column chromatography (EtOAc:Hexane, 1:12) to afford a colourless syrup (145 mg, 0.22 mmol, 68%): $[\alpha]^{20}_D$ = +93.4° (*c* 3.34, CH_2Cl_2); ^{1}H NMR ($CDCl_3$) δ 0.67 (dd, 1 H, $J_{7a,7b}$ = 5.7, $J_{7a,6}$ = 9.3, H-7a), 0.76 (m, 1 H, H-7b), 1.25 (m, 1 H, H-6), 2.64 (d, 1 H, $J_{8a,8b}$ = 9.3, H-8a), 3.72 (dd, 1 H, $J_{3,2}$ = 2.7, $J_{3,4}$ = 9.7, H-3), 4.05 (d, 1 H, $J_{8b,7b}$ = 1.5, H-8b), 4.15 (dd, 1 H, $J_{4,5}$ = 5.1, H-4), 4.35–4.40 (m, 2 H, H-2, CH_AH$_B$C$_6$H$_5$), 4.44 (d, 1 H, $J_{A,B}$ = 12.0, CH$_A$$H_BC_6H_5$), 4.61 (d, 1 H, $J_{C,D}$ = 10.7, CH_CH$_D$C$_6$H$_5$), 4.61 (d, 1 H, $J_{E,F}$ = 12.1, CH_EH$_F$C$_6$H$_5$), 4.69 (d, 1 H, CH$_E$$H_FC_6H_5$), 4.73 (d, 1 H, $J_{G,H}$ = 11.8, CH_GH$_H$C$_6$H$_5$), 4.78 (dd, 1 H, $J_{5,6}$ = 7.3, H-5), 4.86 (d, 1 H, CH$_G$$H_HC_6H_5$), 4.99 (d, 1 H, CH$_C$$H_DC_6H_5$), 6.34-6.45 (m, 3 H, H-4', H-2', H-6'), 7.20–7.39 (m, 20 H, ArH); ^{13}C NMR ($CDCl_3$) δ 12.1 (C-7), 19.1 (C-6), 27.7 (C-1), 72.0 (C-5), 72.9, 73.0, 73.3 (*C*H$_2$C$_6$H$_5$), 73.6 (C-8), 75.2 (*C*H$_2$C$_6$H$_5$), 75.7 (C-4), 76.6 (C-2), 76.7 (C-3), 96.4 (C-4'), 99.6–99.9 (m, C-2', C-6'), 127.4, 127.4$_5$, 127.5, 127.6$_5$, 127.7, 127.8, 128.0, 128.2, 128.2$_2$, 128.3, 128.4, 138.1, 138.4, 138.9, 139.1, 160.4 (t, $^3J_{C,F}$ = 13.7, C-1'), 163.8 (dd, $^1J_{C,F}$ = 245.8, $^3J_{C,F}$ = 15.9, C-3', C-5'). ^{19}F NMR ($CDCl_3$) δ –110.1 (t, $^3J_{H,F}$ = 8.6). Anal. Calcd For $C_{42}H_{40}F_2O_5$: C, 76.11; H, 6.08, Found; C, 76.15; H, 5.88.

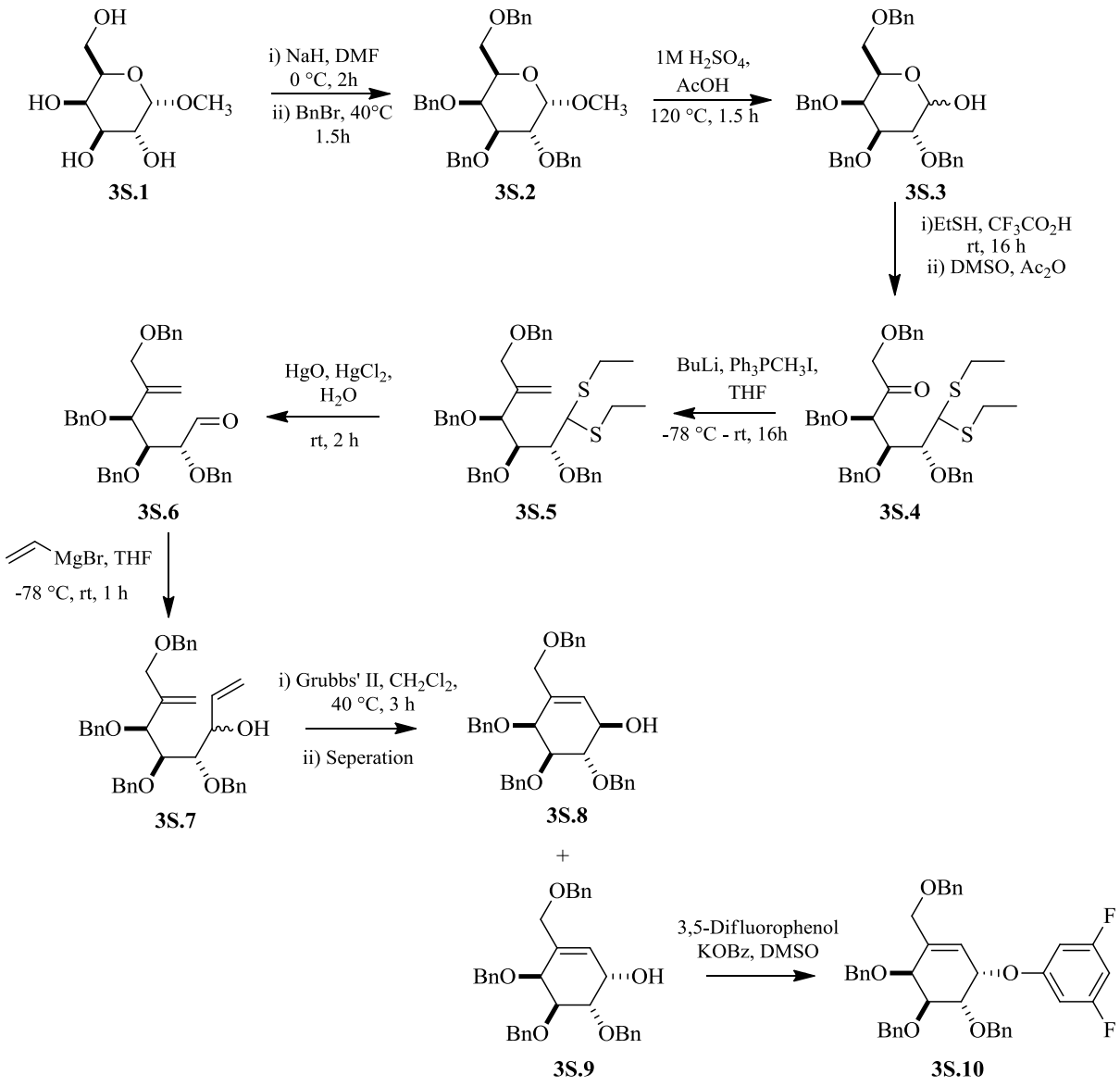
The obtained perbenzylated difluorophenyl protected *galacto* product (149 mg, 0.22 mmol) and 10% Pd-C (60 mg, 10%, w/w) were dissolved in MeOH (30 ml), and stirred at room temperature under a H_2 atmosphere for 6 h **(Scheme 3-1)**. The mixture was filtered through a celite pad, which was subsequently washed thoroughly with MeOH (100 ml). The filtrate and washings were combined and concentrated under reduced pressure to a solid. The obtained *galacto* product **3.4** was further recrystallized from MeOH (62 mg, 91%): $[\alpha]^{20}_D$ = +181.8° (*c* 0.06, MeOH); Mpt = 157–158 °C; ^{1}H NMR (D_2O) δ 0.58 (dd, 1 H, $J_{7a,7b}$ = 5.7, $J_{7a,6}$ = 9.5, H-7a), 0.72 (td, 1 H, $J_{7a,7b}$ + $J_{7b,6}$ = 11.0, $J_{7b,8b}$ = 1.5, H-7b), 1.58 (m, 1 H, H-6), 2.99 (d, 1 H, $J_{8a,8b}$ = 11.5, H-8a), 3.71 (dd, 1 H, $J_{3,2}$ = 3.3, $J_{3,4}$ = 11.1, H-3), 3.93 (dd, 1 H, $J_{4,5}$ = 5.2, H-4), 4.03 (d, 1 H, H-8b), 4.43 (d, 1 H, H-2), 5.00 (dd, 1 H, H-5), 6.58 (m, 1 H, H-4'), 6.63–6.70 (m, 2 H, H-2', H-6'); ^{13}C NMR (D_2O) δ 10.4 (C-7), 19.3 (C-6), 29.2

(C-1), 66.0 (C-8), 66.6 (C-3), 67.2 (C-4), 69.2 (C-2), 72.7 (C-5), 96.6 (t, $^{3}J_{C,F}$ = 26.2, C-4'), 99.7–99.9 (m, C-2', C-6'), 159.1 (t, $^{3}J_{C\text{-}F}$ = 14.1, C-1'), 163.6 (dd, $^{1}J_{C,F}$ = 243.7, $^{3}J_{C,F}$ = 16.1, C-3', C-5'). ^{19}F NMR ($CDCl_3$) δ –112.1 (t, $^{3}J_{H,F}$ = 9.1). Anal. Calcd For $C_{14}H_{16}F_2O_5$: C, 55.63; H, 5.34. Found; C, 55.65; H, 5.50.

3.7.1.9. (1*S*,2*S*,3*S*,4*R*,5*S*,6*R*)-5-(3,5-difluorophenoxy)-1-(hydroxymethyl)-bicyclo[4.1.0]heptan-2,3,4-triol (3.5, Scheme 3-1).

After a suspension of NaH in mineral oil (60%, 100 mg, 2.5 mmol) was washed with hexane (2 × 5 mL), it was transferred in dry DMSO (10 mL) into 100 mL flask held at 18 °C. To this mixture a solution of **3.9** (180 mg, 0.33 mmol) in dry DMSO (10 mL) was added dropwise. The mixture was stirred for 30 min at 18 °C before treating with potassium benzoate (100 mg, 0.66 mmol), and stirring was continued for a further 30 min. At which time, 1,3,5-trifluorobenzene (620 μL, 6 mmol) was added slowly. The reaction was quenched, after another 30 min of stirring, by the addition of a saturated NH_4Cl solution (10 mL). Following the addition of brine (25 mL), the product was extracted from the aqueous layer with diethyl ether (3 × 20 mL). The combined organic layer was dried ($MgSO_4$) and volatiles were removed under reduced pressure. The resultant residue was further purified via flash column chromatography (EtOAc:Hexane, 1:12) to afford a colourless syrup (164 mg, 0.25 mmol, 75%): $[\alpha]^{20}_{D}$ = +57.2° (*c* 5.30, CH_2Cl_2); IR 1115 cm^{-1} (C-F); ^{1}H NMR ($CDCl_3$) δ 0.69 (dd, 1 H, $J_{7a,7b}$ = 5.1, $J_{7a,6}$ = 9.5, H-7a), 1.10 (t, 1 H, $J_{7a,7b}$ + $J_{7b,6}$ = 11.2, H-7b), 1.19 (m, 1 H, H-6), 2.98 (d, 1 H, $J_{8a,8b}$ = 10.5, H-8a), 3.61 (d, 1 H, H-8b), 3.85 (dd, 1 H, $J_{4,3}$ = 7.7, $J_{4,5}$ = 3.4, H-4), 3.95 (dd, 1 H, $J_{3,2}$ = 4.4, H-3), 4.37 (d, 1 H, H-2), 4.39 (d, 1 H, $J_{A,B}$ = 12.0, CH_AH$_B$C$_6$H$_5$), 4.54 (d, 1 H, CH$_A$$H_BC_6H_5$), 4.56–4.67 (m, 5 H, H-5, 4 H-CH_2C$_6$H$_5$), 4.69 (br s, 2 H, CH_2C$_6$H$_5$), 6.42 (m, 1 H, H-4'), 6.46–6.53 (m, 2 H, H-2', H-6'), 7.19–7.39 (m, 20 H, H-Ar); ^{13}C NMR ($CDCl_3$) δ 11.1 (C-7), 22.2 (C-6), 25.4 (C-1), 71.6, 72.3, 73.1 (*C*H$_2$C$_6$H$_5$), 73.1$_4$ (C-2), 73.7 (*C*H$_2$C$_6$H$_5$), 75.0 (C-40, 75.7 (C-8), 76.1, 76.2 (C-3, C-5), 96.5 (C-4'), 99.7–99.9 (m, C-2', C-6'), 127.6, 127.6$_4$, 127.6$_8$, 127.0, 127.8, 127.8$_3$, 128.1, 128.4, 128.4$_2$, 128.4$_3$, 128.4$_4$, 138.6, 138.7, 138.7$_1$, 138.9, 160.4 (t, $^{3}J_{C,F}$ = 13.5, C-1'), 163.8 (dd, $^{1}J_{C,F}$ = 245.9, $^{3}J_{C,F}$ = 15.8, C-3', C-5'); ^{19}F NMR ($CDCl_3$) δ -109.9 (t, $^{3}J_{H,F}$ = 8.8). Anal. Calcd For $C_{42}H_{40}F_2O_5$: C, 76.11; H, 6.08, Found; C, 75.86; H, 5.82.

The obtained perbenzylated difluorophenyl protected *altro* product (164 mg, 0.25 mmol) dissolved in MeOH (35 ml) was treated with 10% Pd-C (67 mg, 10%, w/w), and stirring was continued at room temperature under a H_2 atmosphere for 9 h (**Scheme 3-1**). The mixture was filtered through a celite pad, which was subsequently washed thoroughly with MeOH (100 ml). The filtrate and washings were combined and concentrated under reduced pressure to afford the completely deprotected compound. The resultant residue was purified via flash column chromatography (CH_2Cl_2:MeOH, 10:1) to afford a white solid, which was further recrystallized from MeOH to obtain the final *altro* compound **3.5** (54.5 mg, 73%): $[\alpha]^{20}_D$ = +97.7° (*c* 0.62, MeOH); Mpt = 154-155 °C; ^{1}H NMR (CD_3OD) δ 0.65 (dd, 1 H, $J_{7a,7b}$ = 5.1, $J_{7a,6}$ = 9.6, H-7a), 0.97 (t, 1 H, $J_{7a,7b}$ + $J_{7b,6}$ = 11.1, H-7b), 1.21 (ddd, 1 H, $J_{6,7a}$ = 9.5, $J_{7b,6}$ = 6.0, $J_{6,5}$ = 1.6, H-6), 3.13 (d, 1 H, $J_{8a,8b}$ = 11.4, H-8a), 3.68 (d, $J_{8a,8b}$ = 11.3, 1 H, H-8b), 3.81 (dd, 1 H, $J_{4,5}$ = 3.3, $J_{4,3}$ = 8.5, H-4), 3.94 (dd, 1 H, $J_{3,2}$ = 5.0, $J_{3,4}$ = 8.5, H-3), 4.39 (d, 1 H, $J_{2,3}$ = 5.0, H-2), 4.73 (dd, 1 H, H-5), 6.50 (m, 1 H, H-4'), 6.63-6.68 (m, 2 H, H-2', H-6'); ^{13}C NMR (CD_3OD) δ 10.8 (C-7), 22.1 (C-6), 29.7 (C-1), 68.0 (C-2), 68.9 (C-4), 69.0 (C-8), 70.8 (C-3), 77.6 (C-5), 96.9 (t, $^3J_{C,F}$ = 26.4, C-4'), 100.5–100.7 (m, 2 C, C-2', C-6'), 162.1 (t, $^3J_{C,F}$ = 13.9, C-1'), 165.2 (dd, $^1J_{C,F}$ = 244.6, $^3J_{C,F}$ = 15.9, C-3', C-5'); ^{19}F NMR ($CDCl_3$) δ –112.0 (t, $^3J_{H,F}$ = 9.1); Anal. Calcd For $C_{14}H_{16}F_2O_5$: C, 55.63; H, 5.34. Found; C, 55.53; H, 5.26.



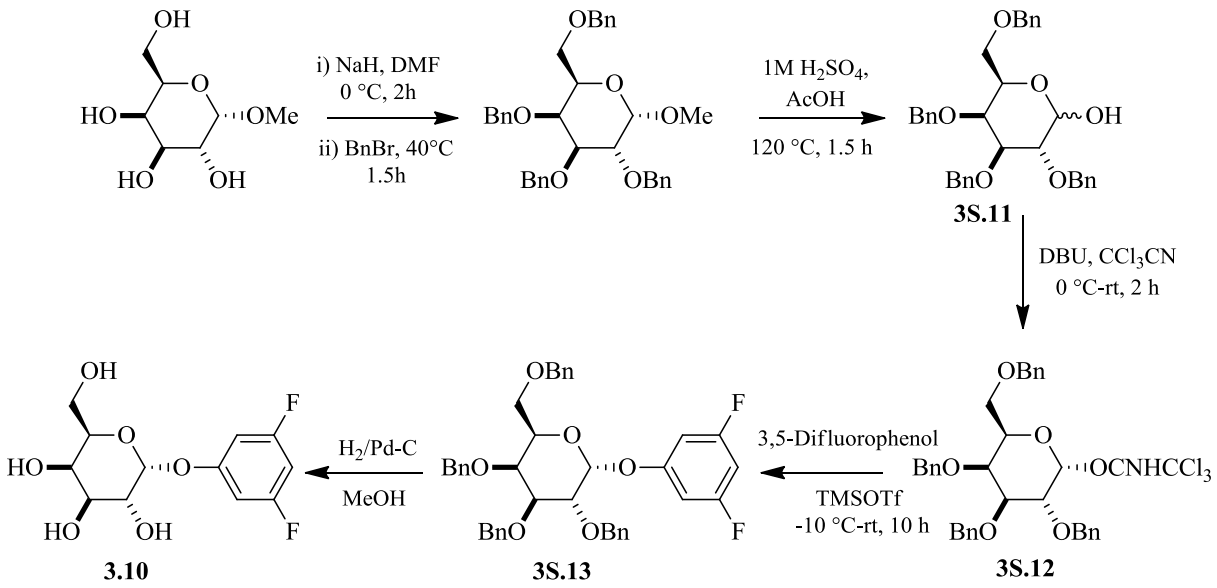
Scheme 3S 1. Synthetic route to the diastereomeric cyclohexenol intermediates towards the synthesis of the precursors to the final inhibitors **3.4**, **3.5** and **3.6**.

3.7.1.10. 3,5-Difluorophenyl α-D-galactopyranoside (3.10) (Scheme 3S-2).

A solution of 2,3,4,6-tetra-*O*-benzyl-galactopyranose **3S.11** (1 g, 1.8 mmol) in anhydrous CH_2Cl_2 (25 mL) was cooled to 0 °C under an inert atmosphere. To this solution, DBU (0.17 mL, 1.1 mmol) and CCl_3CN (2.2 mL, 2 mmol) were added and the resultant reaction mixture was stirred at 0 °C for 20 min. The reaction was warmed up to 25° C over a period of 2 h, which led to the completion of reaction as indicated by TLC. Subsequent

removal of the volatiles under reduced pressure afforded the desired trichloroacetimidate intermediate **3S.12** (1.3 g). In the next step, dissolved the trichloroacetimidate in diethyl ether (30 mL) and treated with 3,5-difluorophenol (0.5 g, 3.8 mmol). Upon cooling the resultant mixture to –10 °C, under an inert atmosphere, TMSOTf (30 μL, 0.47 mmol) was added and stirring was further continued for another 10 h while the reaction mixture was allowed to warm to room temperature. The reaction was quenched by addition of water (10 mL) and the product was extracted from the aqueous layer with EtOAc (3 × 40 mL). The combined organic layers were washed subsequently with saturated $NaHCO_3$ (3 × 50 mL) and brine (60 mL). The organic fractions were dried ($NaSO_4$), and concentrated under reduced pressure. The resultant crude glycosylated product was purified by column chromatography (EtOAc: Hexane, 3:7) to afford pure 3,5-difluorophenyl 2,3,4,6-tetra-*O*-benzyl-α-D-galactopyranoside **3S.13** (450 mg, 36% yield).

A solution of **3S.13** (450 mg, 0.69 mmol) in MeOH (33 mL) was treated with 10% Pd-C (80 mg, 10%, w/v) and stirring was continued under a H_2 atmosphere for 10 h at 25° C. The resulting mixture was filtered through a celite pad, which was subsequently washed thoroughly with MeOH (200 mL). The filtrate and washings were combined and concentrated under pressure to give the product as a white solid, which was recrystallized from MeOH to afford the final substrate **3.10** (100 mg, 50% yield); Mpt = 150-151 °C; $[\alpha]^{20}_D$ = +136.3° (*c* 0.04, MeOH); ^{1}H NMR (600 MHz, D_2O) δ 6.84 (m, 2H, ArH, H-2′, H-6′), 6.71 (m, 1H, ArH, H-4′), 5.71 (d, *J* = 3.6, 1H, H-1), 4.11–4.01 (m, 4H, H-2, H-3, H-4, H-5), 3.77–3.69 (m, 2H, H-6, H-6’); ^{13}C NMR (151 MHz, D_2O) δ 163.23 (dd, *J* = 244.5, 15.9, C-3′, C-5′), 157.82 (t, *J* = 14.0, C-1′), 100.84-100.69 (m, 2 C, C-2', C-6'), 98.12 (t, $^3J_{C,F}$ = 26.4, C-4'), 97.26 (C-1), 71.98, 69.34, 69.04, 67.87 (C-2, C-3, C-4, C-5), 60.96 (C-6). Anal. Calcd for $C_{12}H_{14}F_2O_6$: C, 49.32; H, 4.83, Found; C, 49.30; H, 4.75.



Scheme 3S 2. Synthetic route to 3,5-difluorophenyl substrate **3.10**.

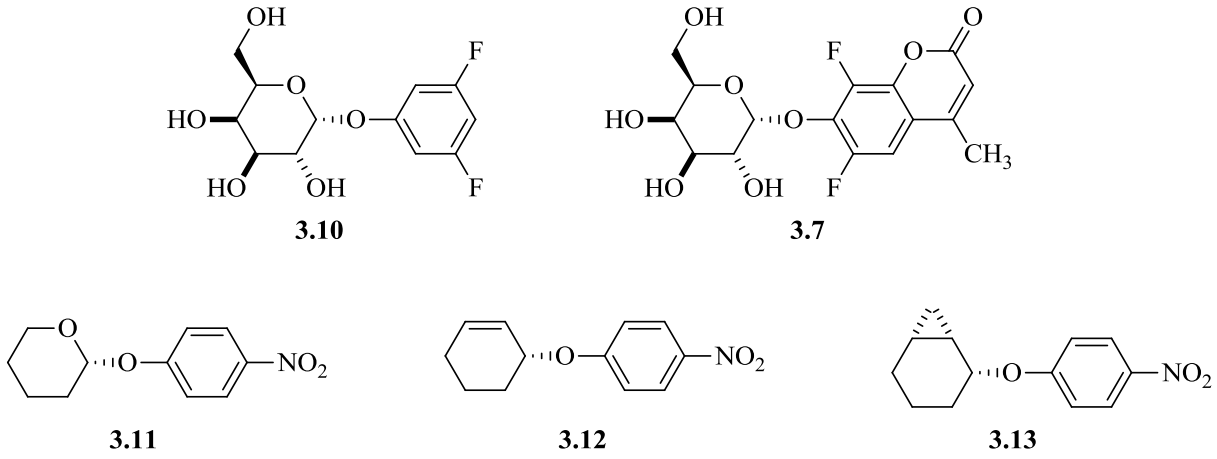


Figure 3S 1. Substrates (**3.10** and **3.7**) and the model compounds (**3.11**, **3.12** and **3.13**) synthesized.

3.7.1.11. 3-(4-Nitrophenoxy)cyclohexene (3.12).

In a round-bottom flask charged with cyclohex-2-ene-1-one (2 g, 20 mmol) and $CeCl_3.7H_2O$ (2 g, 5 mmol) were dissolved in methanol (100 mL). To this stirred solution, $NaBH_4$ (700 mg, 18 mmol) was added in small portions. After the vigorous evolution of gas had subsided, the reaction mixture was carefully neutralized with dilute aqueous HCl (15 mL). The resultant solution was extracted with diethyl ether (3 × 50 mL) and the

combined organic layers were washed with brine (2 × 20 mL) and dried (Na_2SO_4). The volatiles were removed under reduced pressure to afford the crude allylic alcohol (1.6 g, 82% yield) in >98% purity as determined by ^{1}H NMR. This material was carried forward to the next step without further purification. The resultant product, allylic alcohol, was dissolved in dimethoxyethane (100 mL), charged with 4-nitrofluorobenzene (2.3 g, 16 mmol) and cooled the resultant solution to 0 °C. At this point, NaH (700 mg, 29 mmol, 60% dispersion in oil) was added portion wise with continuous stirring. Following which, the reaction mixture was warmed to 60 °C and stirring was continued at this temperature for another 12 h. Upon quenching the reaction by addition of water (40 mL), the product was extracted from the aqueous layer with $CHCl_3$ (3 × 50 mL). The organic layers were combined and washed with brine (2 × 20 mL) and dried (Na_2SO_4). This was followed by removal of the volatiles under reduced pressure to afford a yellow syrup. The crude residue was purified via flash column chromatography (5% EtOAC-Hexane, 50% Hexane-Toluene) to obtain pure product **3.12** as a yellow syrup (2.2 g, 62% yield); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.27–8.17 (m, 2H), 7.04-6.92 (m, 2H), 6.06-6.01 (m, 1H), 5.85-5.81 (m, 1H), 4.93 (dt, J = 7.2, 3.4, 1H), 2.20-2.01 (m, 2H), 2.0-1.93 (m, 1H), 1.92-1.79 (m, 2H), 1.72-1.62 (m, 1H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 163.14, 141.09, 133.44, 125.92, 124.72, 115.22, 71.65, 28.05, 24.94, 18.68; HRMS expected for $C_{12}H_{13}NO_3$ is 242.0793 ($M + Na^+$); Found 242.0780 ($M + Na^+$).

3.7.1.12. rel-(1R,2S,6S)-2-(4-nitrophenoxy)bicyclo[4.1.0]heptane (3.13).

Cyclohex-2-enol (1.5 g, 15 mmol) was dissolved in 1, 2-dichloroethane (50 mL) and the resultant solution was cooled down to 0 °C under an inert atmosphere. To this solution, Et_2Zn (20 mL, 1 M solution in hexane) was added with continuing stirring at 0 °C for 15 min. Subsequently, CH_2I_2 (3.5 mL, 45 mmol) was added to the reaction mixture at 0° C. The resultant solution was gradually warmed to 25 °C and stirring was continued for another 2 h. The reaction was quenched by addition of water (40 mL) and saturated NH_4Cl (30 mL). The product was extracted from the aqueous layer with CH_2Cl_2 (3 × 50 mL), washed with brine (2 × 20 mL) and dried (Na_2SO_4). The volatiles were removed under reduced pressure to afford the cyclopropanated alcohol as a colourless syrup (1.4 g, 83% yield). Upon drying under high vacuum (30 min), this cyclopropanated alcohol (1.4 g, 12.5 mmol) was dissolved

in dimethoxyethane (100 mL) and treated with 4-nitrofluorobenzene (1.8 g, 13 mmol). The resultant solution was cooled to 0 °C, when NaH (620 mg, 26 mmol, 60% dispersion in oil) was added portion wise. After the addition was complete the mixture was warmed to 60 °C and the stirring was continued for another 12 h. The reaction was quenched by the addition of water (25 mL), and the product was extracted into $CHCl_3$ (3 × 50 mL), washed with brine (2 × 20 mL) and dried (Na_2SO_4). Removal of the volatiles under reduced pressure afforded a bright yellow syrup which was purified via flash column chromatography (5% EtOAC-Hexane, 50% Hexane-Toluene) to obtain the pure compound **3.13** (2.2 g, 76% yield). Mpt = 54–55 °C; ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.24–8.14 (m, 2H), 7.05–6.92 (m, 2H), 4.93 (dd, *J* = 11.9, 6.0, 1H), 1.89-1.80(m, 1H), 1.68-1.61 (m, 2H), 1.49-1.44 (m, 2H), 1.38-1.32 (m, 2H), 1.22-1.14 (m, 1H), 0.58-0.53 (m, 1H), 0.48-0.44 (m, 1H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 163.58, 140.88, 125.88, 115.52, 73.48, 27.63, 22.70, 18.16, 13.97, 12.70, 6.95; HRMS expected for $C_{13}H_{15}NO_3$ is 256.0950 ($M + Na^+$); Found 256.0947 ($M + Na^+$).

Table 3S 1. Rate constants and thermodynamic parameters obtained for the spontaneous hydrolysis of model compounds **3.11**, **3.12** and **3.13**.

Temperature (°C)	**3.11** (s^{-1})	**3.13** (s^{-1})	**3.12** (s^{-1})
15.0	$(1.07 \pm 0.05) \times 10^{-5}$		
37.0	$(2.40 \pm 0.04) \times 10^{-4}$ [a]	7.60×10^{-6} [b]	2.31×10^{-7} [b]
50.0	$(1.15 \pm 0.02) \times 10^{-3}$	$(3.88 \pm 0.09) \times 10^{-5}$	
60.0		$(1.26 \pm 0.02) \times 10^{-4}$	$(4.29 \pm 0.14) \times 10^{-6}$
70.0		$(3.76 \pm 0.08) \times 10^{-4}$	$(1.39 \pm 0.04) \times 10^{-5}$
80.0			$(4.03 \pm 0.15) \times 10^{-5}$
$\Delta H^{\ddagger}$ (kJ mol^{-1})	101 ± 1	102 ± 1	107± 1
$\Delta S^{\ddagger}$ (J mol^{-1} K^{-1})	+11 ± 4	–14 ± 1	–28 ± 4

[a] Value from reference 1 at 39 °C (*I* = 1.0 KCl) is 2.19×10^{-2} min^{-1}. [b] Extrapolated value calculated from the Eyring plot.

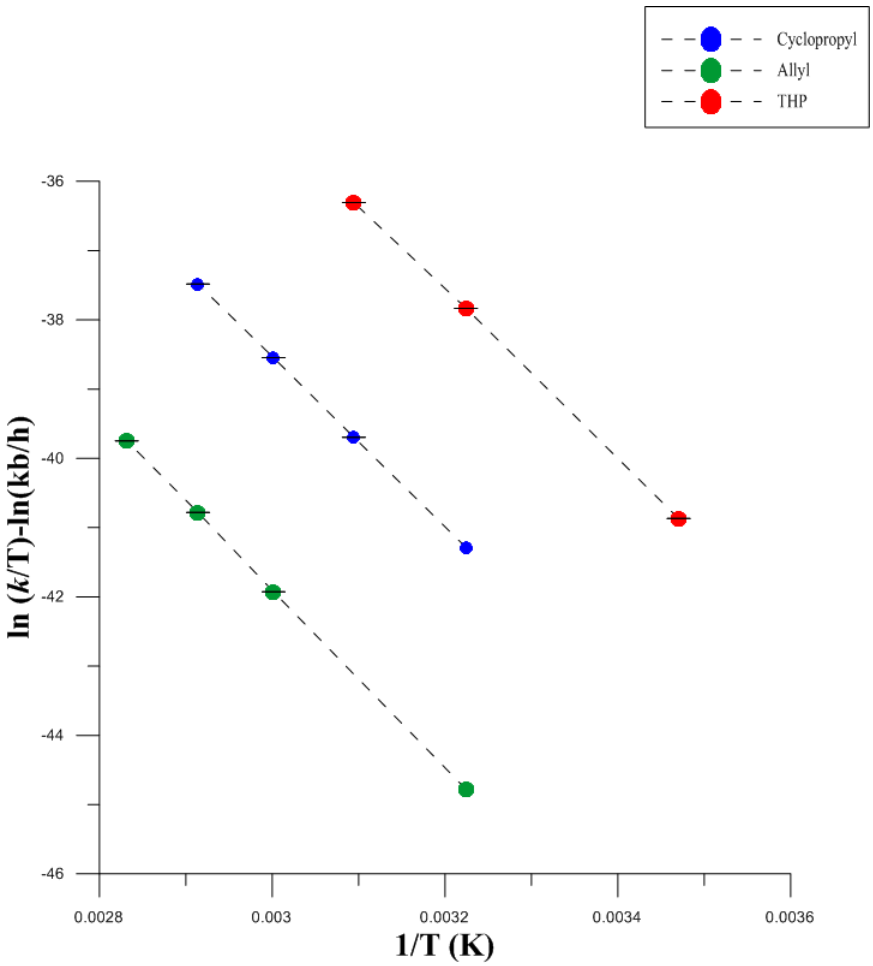


Figure 3S 2. Eyring plots for the spontaneous hydrolysis of model compounds **3.11**(THP analogue), **3.12** (allyl analogue) and **3.13** (cyclopropyl analogue).

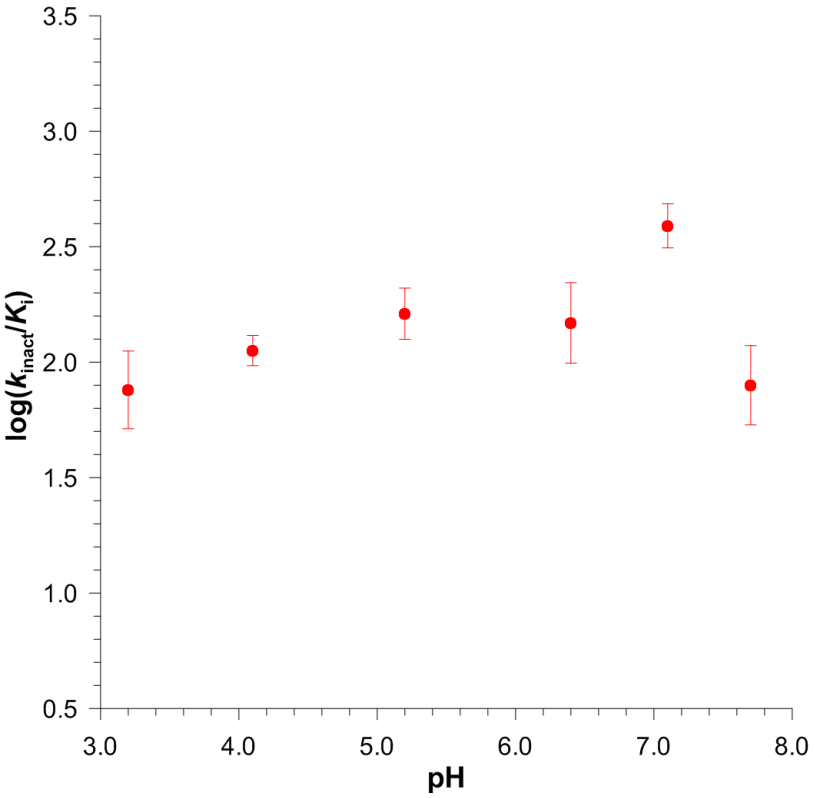
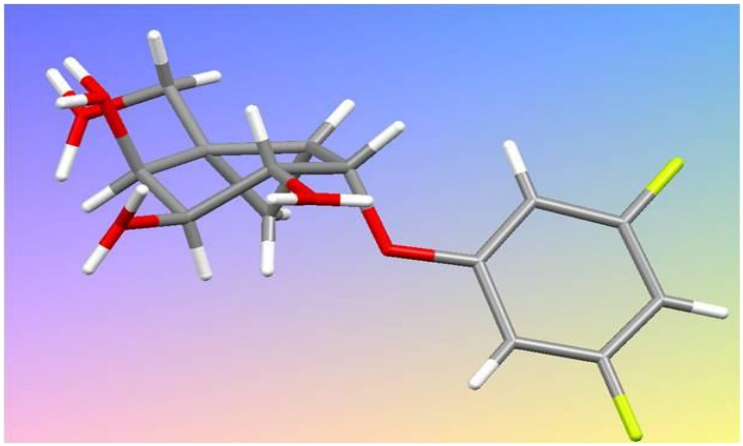
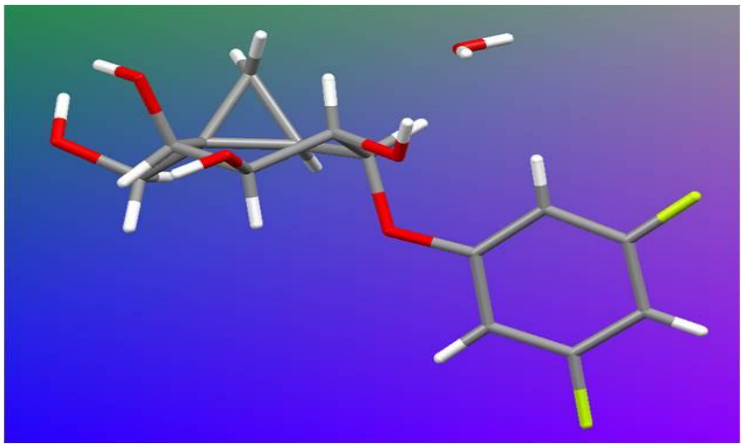


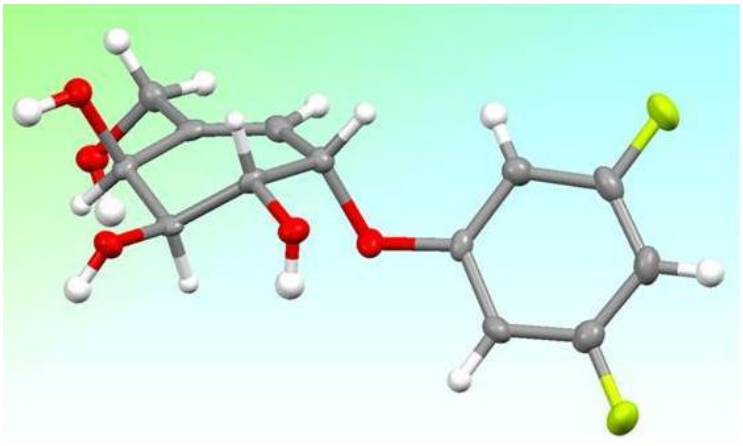
Figure 3S 3. Inactivation of *T. maritima* by *galacto* inhibitor **3.4** as a function of pH.



galacto compound (**3.4**)



altro compound (**3.5**)



Allyl compounds (**3.6**)

Figure 3S 4. X-ray crystal structures of the three inhibitors, **3.4**, **3.5** and **3.6** (These X-ray diffraction studies were conducted by Dr. Mike Katz, to whom I extent my gratitude).

3.7.2. References

1. Craze, G. A. & Kirby, A. J. The spontaneous hydrolysis of acetals: Sensitivity to the leaving group. *J. Chem. Soc., Perkin Trans*. 2, 354-356 (1978).

2 Sun, W. C., Gee, K. R. & Haugland, R. P. Synthesis of novel fluorinated coumarins: Excellent UV-light excitable fluorescent dyes. *Bioorg. Med. Chem. Lett*. 8, 3107-3110 (1998).

3 Miller, E. S. et al. α-D-Galactosidases from Thermotoga species. *Methods Enzymol*. 330, 246-260 (2001).

4. Mechanistic Analysis of Family 109 α-*N*-acetylgalactosaminidase from *Elizabethkingia meningosepticm*: A multi-substrate substrate specific hydrolase possessing redox mechanism.

4.1. Introduction

α-*N*-Acetylgalactosaminidases (α-NAGAL, E.C 3.2.1.49) are the carbohydrate processing enzymes that specifically cleave terminal non-reducing *N*-acetyl-α-D-galactosamine (αGalNAc) residues from glycoconjugates. Out of the approximately 130 sequence-based glycosyl hydrolase families (GH), α-NAGAL has been classified into families 27, 36 and 109.(*1-3*) Of these, families GH27 and GH36 are thought to share a common ancestral gene, forming the glycoside hydrolase clan (GHD) with a common triosephosphate isomerase (TIM) barrel structure.(*4, 5*) α-NAGAL can be found in both prokaryotic pathogens (GH109)(*1*) and humans (GH27).(*6*) The *N*-acetylated derivative of α-D-galactosamine (αGalNAc) can be commonly found in various biological systems, for example, mucin glycoproteins,(*7, 8*) terminal carbohydrate of the blood group A antigen(*9*) and the sensory nerve structures of both humans and animals.(*10*) The α-NAGAL found in human tissue was initially thought to be an isoenzyme of α-galactosidase, due to their similarity in: a) protein folds, both are homodimeric with each monomer consisting of an N-terminal $(\beta/\alpha)_8$ barrel containing the active site and C-terminal antiparallel β-domain; and b) amino acid sequence, α-NAGAL (EC 3.2.1.49) and α-GAL (EC 3.2.1.22) possesses 46% amino acid sequence identity.(*6, 11*) In *O*-linked glycoproteins (**Figure 4-1**), the addition of a core GalNAc residue to either a serine or threonine residue involves an *O*-glycosidic bond that is added post-translationally in the Golgi apparatus. In mucins, such post translational modifications with αGalNAc are further elongated to give an oligosaccharide chain by addition of sugars like galactose, *N*-acetylglucosamine, fucose or sialic acid (**Figure 4-2**).(*7, 8*) These heavily *O*-glycosylated proteins, mucins, are present as a protective layer at or on

the epithelial surfaces of the body, including the gastrointestinal, genital and respiratory tracts, in order to provide lubrication and water retention at these surface that are exposed to the environment.(*7*) However, the widespread importance of α-NAGAL as an *exo*-glycosidase acting upon pharmaceutically significant substrates with terminal αGalNAc residues are in; a) blood group A antigen **Figure 4-2**;(*12*) and, b) lysosomal glycolipids.(*13, 14*)

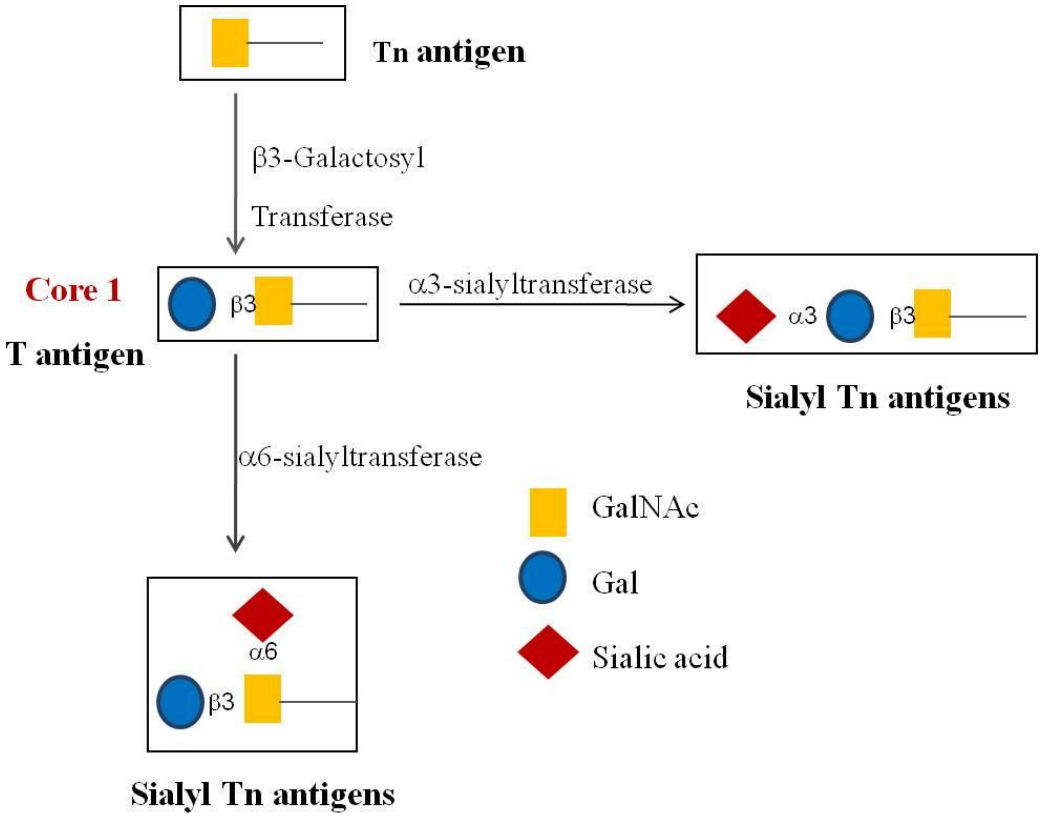


Figure 4-1. *O*-linked oligosaccharides underlying the core structures of mucin glycoproteins.

Blood group (ABO) specificity is determined by the nature and linkage of the monosaccharides found at the termini of the oligosaccharides chains on glycoconjugates on the surface of erythrocytes.(*15, 16*) The antigenic determinants of blood group A and B are the terminally positioned monosaccharides, αGalNAc and αGal respectively.(*16, 17*) Structurally, the blood group A determinant is characterized by the trisaccharide structure 2-acetamido-2-deoxy-α-D-galactopyranosyl-(1→3)-[α-L-fucopyranosyl-(1→2)]-D-galactopyranosyl and blood group B by α-D-galactopyranosyl-(1→3)-[α-L-fucopyranosyl-

(1→2)]-D-galactopyranosyl, attached to the glycolipids or glycoproteins that are present on the surface of erythrocytes and endothelial and on most epithelial cells (**Figure 4-2**).(*1, 16*)

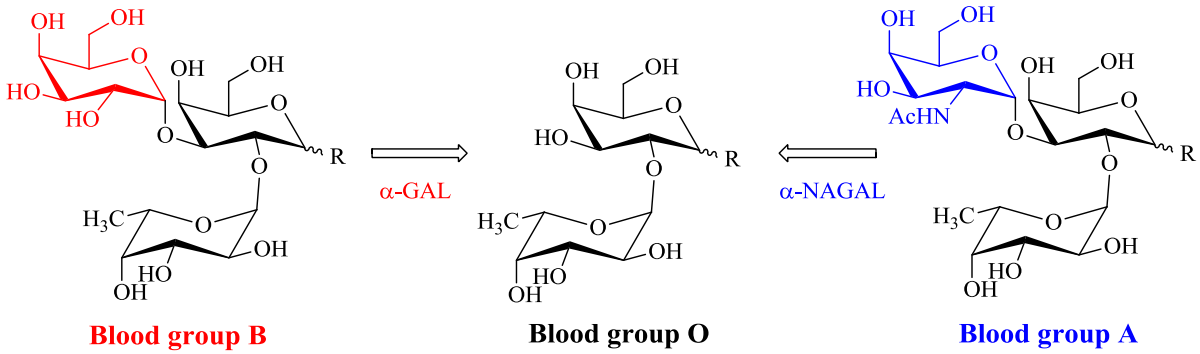


Figure 4-2. Enzymatic blood group conversion of blood group A and B to universal donor O, showing the immunodominant epitopes present at the termini of the oligosaccharide chains linked to the glycolipid and glycoprotein, represented as R.(*11, 12, 126*)

Plasma of blood group A individuals contains naturally occurring antibodies to the B antigen and *vice versa*. Blood group AB plasma has neither antibodies, and blood group O plasma has both antibodies. Group O erythrocytes are considered "universal donors".(*15*) Enzymatic removal of ABO antigens producing universal blood red blood cells (RBCs) using substrate specific *exo*-glycosidases, to hydrolyze off the α-linked *N*-acetylgalactosamine (blood group A) or galactose (blood group B) releasing the reducing end, have attracted considerable attention with regards to blood transfusion (**Figure 4-2**).(*12, 15*) Originally pioneered almost 25 years ago, Goldstein first showed that by using an α-galactosidase (*coffee bean* α-galactosidase (E.C 3.2.1.22)) B erythrocytes could be converted to blood group O cells without either compromising their viability *in vitro* or affecting their ability to survive normally post transfusion.(*17*) The disadvantage of this protocol being that the *coffee bean* α-galactosidase employed in the *in vitro* experiments possesses a very low pH optimum (pH 3.5), a situation that necessitated the use of as much as 1–2 g enzyme per unit of group B RBCs for efficient conversion to universal blood group O.(*17*) Several reports have been published regarding the isolation and purification of substrate (αGalNAc) specific α-NAGAL for efficient blood group conversion (A to O).(*8, 18-20*) In contrast to blood group B epitope (terminal 1→3 linked αGal), blood group A (terminal 1→3 linked α-GalNAc) immuno-determinants on human RBCs differ in susceptibility and biological

activity to α-NAGAL.(*1*) As a result, complete analysis of blood group conversion A to O, has proven to be more complicated than is the blood group B to O conversion.

To date, α-NAGAL has been purified from multiple animal sources; which include from bovine(*8*) and porcine liver(*18*) and snail.(*19*) Many isolated α-NAGAL enzymes also possess α-GAL activity. The α-NAGAL enzyme isolated from bacteria *Clostridium perfringens*, showed optimum catalytic activity at physiological pH and temperature.(*21*) Of note, this α-NAGAL showed two striking features: a) a strict substrate specificity for α-GalNAc; and b) an absolute requirement for reducing conditions (DTT inclusion) to maintain its activity.(*21*) A follow up report by Hsin-Yeh *et al.* showed that successful blood group conversion (A to O) was possible using the recombinant *Clostridium perfringens* α-NAGAL enzyme.(*22*) Recently, in 2007, Liu *et al.* reported the X-ray crystal structure of an α-NAGAL enzyme from *E. meningosepticum*, which is a GH109 enzyme and this family is found exclusively in prokaryotes.(*1*) Moreover, the 3D structure revealed several similarities between GH4 glycosidases and GH109 α-*N*-acetylgalactosaminidases.(*1, 23*) Although, both GH4 and GH109 hydrolases share a remote sequence based relationship with oxidoreductases,(*1, 24*) the cofactor dependence of GH109 α-*N*-acetylgalactosaminidase is markedly different from that of GH4 hydrolases.(*1*) Glycosidases (both α and β) from GH4 have been reported to be dependent on divalent metal ion (Mn^{2+}), NAD^+ and reducing conditions (DTT or TCEP) for catalytic activity (**Figure 1-17,** Introduction).(*25-27*)

In contrast, supported by the conclusions from the X-ray crystallographic structure of *E. meningosepticum* α-*N*-acetylgalactosaminidase with the reaction product GalNAc, the deeply embedded NAD^+ (**Figure 4-3**) is the only necessary cofactor for GH109 catalytic activity.(*1*) The observation that upon addition of NAD^+ to GH109 α-NAGAL from *E. meningosepticum*, results in no rate enhancement is consistent with the enzyme being saturated with NAD^+ in its purified recombinant form.(*1*) Absolute dependence of the catalytic activity of GH109 α-NAGAL on NAD^+ was substantiated by precipitating the protein with ammonium sulfate, prior to reactivation using exogenous NAD^+.(*1*) Tyr307, Tyr225, His228 and Glu149 have been reported to be the important carbohydrate (GalNAc) binding residues, typically invariant within the GH109 family.(*1*) Additionally, unlike GH4

hydrolases, the active site of *E. meningosepticum* GH109 α-*N*-acetylgalactosaminidase does not contain a metal binding site (**Figure 4-3**).(*1*) These findings have some similarities with the GH4 family, yet the differing cofactor requirements for GH109 catalysis, motivated us to undertake an in-depth mechanistic study on a GH109 family member using similar experimental protocols that are listed in Chapter 2.

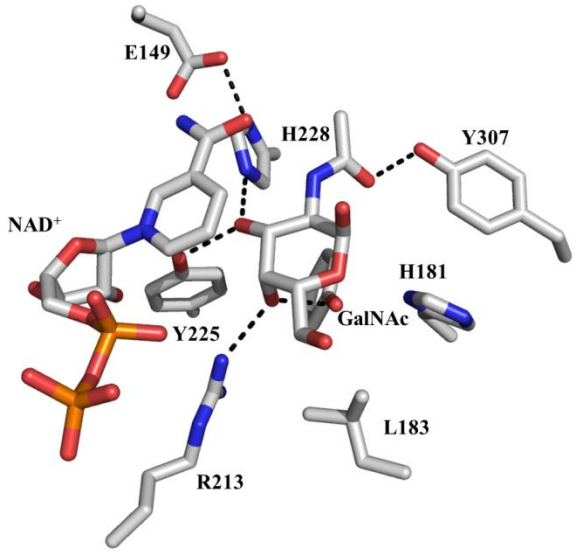


Figure 4-3. Active site architecture of GH109 α-*N*-acetylgalactosaminidase from *E. meningosepticum*, complexed with its reaction product, GalNAc. Also shown is the deeply embedded NAD^+ molecule. (*1*) (Thanks to Dr. Tracey Gloster for providing the representation of the active site)

Consequently, this chapter details the cloning of the Flv109 gene from *E. meningosepticum* (ATCC 51720D) and the expression of the encoded protein (α-NAGAL), which hydrolyzes the terminal α-GalNAc units from glycolipids and glycoproteins. In agreement with literature precedents,(*1, 21, 22*) the expressed α-NAGAL demonstrated multi-substrate specificity (α/βGalNAc and α/βGal). The optimum pH for catalytic activity was determined to be 7.5. The cofactor dependence (Mn^{2+}, NAD^+ and DTT) and the stereochemical outcome (inversion or retention of anomeric configuration) of the α-NAGAL catalyzed hydrolyzed product have been further examined. In order to measure the KIEs required to delineate the mechanism of action, it is necessary to synthesize deuterium isotopologues of a substrate for the enzyme. It was decided to make such labelled

compounds as a 2-naphthyl αGalNAc series. The synthesis of 2-naphthyl 2-acetamido-2-deoxy-α-D-(2-^{2}H)galactopyranoside has been accomplished, unfortunately, owing to synthetic difficulties the other deuterium isotopologues were not made. However, the previously made isotopologues of phenyl α-D-galactosides could be employed to measure kinetic isotope effects on this enzyme. The insights from these studies on the α-NAGAL-catalyzed hydrolysis of glycosides will further facilitate the development of mechanism-based inhibitor for α-galactosaminidases from GH109 enzymes, which are only found in prokaryotes.

4.2. Materials and Methods

Materials. All chemicals were of analytical grade or better and were purchased from Sigma-Aldrich unless noted otherwise. Milli-Q water (18.2 MΩ cm^{-1}) was used for all kinetic experiments. All pH values were measured using a standard pH electrode attached to a VWR pH meter. All NMR spectra were acquired on either a Bruker 400, 500 or 600 MHz spectrometer. Chemical shifts are reported in parts per million downfield from signals for TMS. The signal residues from deuterated chloroform and external TMS salts (D_2O) were used for ^{1}H NMR spectral references; for ^{13}C NMR spectra, natural abundance signals from $CDCl_3$ and external TMS salts (D_2O) were used as references. Coupling constants (*J*) are reported in hertz. Melting points were determined on a Gallenkamp melting point apparatus and are not corrected. Optical rotations were measured on a Perkin-Elmer 341 polarimeter and are reported in units of deg cm^2 g^{-1} (concentrations reported in units of g per 100 mL).

4.2.1. Synthesis of aryl 2-acetamido-2-deoxy-α-D-galactosaminides.

All aryl 2-acetamido-2-deoxy-α-D-galactosaminides, were synthesized from 1,3,4,6-tetra-*O*-acetyl-2-deoxy-(2-methoxycarbonylamino)-β-D-galactopyranoside using stannic chloride as the activator in CH_2Cl_2 (**Scheme 4-2**).(*28*) In a typical procedure, 1,3,4,6-tetra-*O*-acetyl-2-deoxy-(2-methoxycarbonylamino)-β-D-galactopyranoside (1 g, 2.4 mmol) and the appropriate phenol (4.9 mmol) were dissolved in anhydrous CH_2Cl_2 (50 mL) and then $SnCl_4$ (0.3 mL, 2.4 mmol) was added. The reaction mixture was stirred at

ambient temperature under an inert atmosphere for 48 h, at which time TLC analysis showed that the reaction was complete. Following the addition of water (40 mL) the reaction mixture was neutralized by adding saturated $NaHCO_3$ (30 mL). The product was extracted from the aqueous layer using CH_2Cl_2 (3 × 40 mL), and the combined organic layer was washed with brine (2 × 25 mL), dried over Na_2SO_4 and then concentrated under reduced pressure to obtain the crude product. This material was purified by column chromatography using EtOAc-Hexane (35:65, EtOAc: Hexane) as the eluent to obtain the pure α-anomer. Deprotection of the *N*-carbamate was accomplished using tetra-n-butylammonium fluoride (TBAF) under refluxing conditions.(*29*) Typically, the glycosylated product (1.8 mmol) was dissolved in anhydrous THF (70 mL) and after addition of TBAF (18 mmol) the resultant mixture was heated to reflux for 3 h under anhydrous conditions. After the reaction had been quenched by the addition of methanol (10 mL) the volatiles were removed under reduced pressure to yield crude deprotected aryl α-D-galactosamine. Without further purification, the resultant crude product was per-acetylated under standard conditions by dissolving in a pyridine (10 mL) acetic anhydride (10 mL) mixture. Stirring was continued for another 15 h at ambient temperature. Following which the reaction was quenched by the addition of water (30 mL) and the product was extracted from the aqueous layer using CH_2Cl_2 (3 × 35 mL). The combined organic layer was washed with brine (2 × 20 mL), dried over Na_2SO_4 and then concentrated under reduced pressure to obtain the crude product. Excess pyridine was removed azeotropically using toluene. Purification of the final product was achieved by using column chromatography in 5% MeOH-CH_2Cl_2, to obtain the pure per-acetylated product. Complete deprotection of the obtained peracetylated compound was accomplished under Zemplen conditions using NaOMe/MeOH followed by neutralization using Amberlite (H^+) resin to obtain the pure substrates in the series of aryl 2-acetamido-2-deoxy-α-D-galactopyranosides. The detailed characterization of the series of substrates used in this study is given as follows.

4.2.1.1. Aryl 3,4,6-tri-*O*-acetyl-2-deoxy-(2-methoxycarbonylamino)-α-D-galactopyranosides (Compounds 4.1a-4.1f).

3-Nitrophenyl 3,4,6-tri-O-acetyl-2-deoxy-(*2-methoxycarbonylamino*)*-α-D-galactopyranoside* (**4.1a**) $[\alpha]^{20}_D$ = +29.3 (*c* 1.0, $CHCl_3$); ^{1}H NMR (600 MHz, $CDCl_3$) δ 7.95

(s, 1H, ArH), 7.92 (d, *J* = 8.1, 1H, ArH), 7.47 (t, 1H, ArH), 7.39 (d, *J* = 8.2, 1H, ArH), 5.67 (d, *J* = 3.1, 1H, H-1), 5.45 (app.d, 1H, H-4), 5.34 (dd, *J* = 11.4, 2.9, 1H, H-3), 4.97 (d, *J* = 9.5, 1H, NH), 4.48 - 4.52 (m, 1H, H-2), 4.26 (t, *J* = 6.5, 1H, H-5), 4.11 (m, 2H, H-6, 6′), 3.68 (s, 3H, NHCOO*C*H_3), 2.17 (s, 3H), 2.03 (s, 3H), 1.91 (s, 3H) (3× OAc); ^{13}C NMR (151 MHz, $CDCl_3$) δ 170.70, 170.25, 170.08 (3 × CO), 156.52 (NH*C*OOCH_3), 149.10, 130.25 (Ar C), 122.89 (Ar C), 117.91(Ar C), 111.87 (Ar C), 97.23 (C-1), 67.98 (C-5), 67.81 (C-3), 66.98 (C-4), 61.60 (C-6), 52.48 (NHCOO*C*H_3), 49.55 (C-2), 20.63, 20.58, 20.38 (3 × OAc); ESI-MS for $C_{20}H_{24}N_2O_{12}$ *m*/*z* calcd for (M + Na^+) 507.12: Found 507.1 (M + Na^+).

Phenyl 3,4,6-tri-O-acetyl-2-deoxy-(*2-methoxycarbonylamino*)*-α-D-galactopyranoside* (**4.1b**) $[\alpha]^{20}_D$ = +19.5 (*c* 1.12, $CHCl_3$); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.29 - 7.33 (m, 2H, ArH), 7.05 - 7.08 (m, 3H, ArH) 5.60 (d, *J* = 3.4, 1H, H-1), 5.45 (app.d, 1H, H-4), 5.35 (dd, *J* = 11.4, 3.0, 1H, H-3), 5.0 (d, *J* = 9.8, 1H, NH), 4.48-4.43 (m, 1H, H-2), 4.28 (t, *J* = 6.5, 1H, H-5), 4.03-4.13 (m, 2H, H-6, 6′), 3.65 (s, 3H, NHCOO*C*H_3), 2.18 (s, 3H) , 2.03 (s, 3H), 1.92 (s, 3H) (3 × OAc); ^{13}C NMR (126 MHz, $CDCl_3$) δ 170.75, 170.30, 170.23 (3 × CO, 3 × OAc), 156.56 (NH*C*OOCH_3), 129.65 (ArC), 116.57 (ArC), 96.64 (C-1), 68.29 (C-3), 67.50 (C-5), 67.23 (C-4), 61.57 (C-6), 52.44 (NHCOO*C*H_3), 49.64(C-2), 20.74, 20.68, 20.53 (3 × OAc); ESI-MS for $C_{20}H_{25}NO_{10}$ *m*/*z* calcd for (M + Na^+) 462.15: Found 462.1 (M + Na^+).

4-Methoxyphenyl 3,4,6-tri-O-acetyl-2-deoxy-(*2-methoxycarbonylamino)-α-D-galacto pyranoside* (**4.1c**) $[\alpha]^{20}_D$ = +20.5 (*c* 0.58, $CHCl_3$); ^{1}H NMR (600 MHz, $CDCl_3$) δ 6.99 (d, *J* = 9.2, 2H, ArH), 6.83 (d, *J* =9.1, 2H, ArH), 5.47 (d, *J* = 3.3, 1H,H-1), 5.45 (app.d, *J* = 2.8, 1H, H-4), 5.32 (dd, *J* = 11.3, 3.0, 1H, H-3), 5.01 (d,*J* = 9.9, 1H, NH), 4.43-4.42 (m, 1H, H-2), 4.32 (t, *J* = 6.6, 1H, H-5), 4.13-4.04 (m, 2H, H-6, 6′), 3.77 (s, 3H, OCH_3), 3.66 (s, 3H, $NHCOOCH_3$), 2.17 (s, 3H), 2.03 (s, 3H), 1.96 (s, 3H) (3 × OAc); ^{13}C NMR (151 MHz, $CDCl_3$) δ 170.74, 170.30, 170.24 (3 × CO, 3 × OAc), 156.58 (CO, NH*C*OOCH_3), 118.00 (ArC), 114.67 (ArC), 97.63 (C-1), 68.34 (C-3), 67.40 (C-5), 67.31 (C-4), 61.69 (C-6), 55.64 (O*C*H_3), 52.42 (NHCOO*C*H_3), 49.69 (C-2), 20.72, 20.67, 20.57 (3 × OAc); ESI-MS for $C_{21}H_{27}NO_{11}$ *m*/*z* calcd for (M + Na^+) 492.15: Found 492.1 (M + Na^+).

3-Chlorophenyl 3,4,6-tri-O-acetyl-2-deoxy-(*2-methoxycarbonylamino*)-*α-D-galactopyranoside* (**4.1d**) $[\alpha]^{20}_{D}$ = +22.9 (*c* 0.8, $CHCl_3$); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.23 (d, *J* = 8.2, 1H, ArH), 7.13 (s, 1H, ArH), 7.06 (d, *J* = 8.0, 1H, ArH), 6.96 (dd, *J* = 8.1, 1.7, 1H, ArH), 5.58 (d, *J* = 3.0, 1H, H-1), 5.45 (app.d, 1H, H-4), 5.32 (dd, *J* = 11.3, 2.8, 1H, H-3), 4.94 (d, *J* = 9.6, 1H, NH), 4.48-4.44 (m, 1H, H-2), 4.26 (t, *J* = 6.3, 1H, H-5), 4.14-4.02 (m, 2H, H-6, 6′), 3.67 (s, 3H, $NHCOOCH_3$), 2.19 (s, 3H), 2.04 (s, 3H), 1.97 (s, 3H) (3 × OAc); ^{13}C NMR (126 MHz, $CDCl_3$) δ 170.77, 170.38, 170.21(3 × CO, 3 × OAc), 156.58 ($NHCOOCH_3$), 130.45 (ArC), 123.34 (ArC), 117.23 (ArC), 115.06 (ArC), 96.76 (C-1), 68.11 (C-3), 67.80 (C-5), 67.18 (C-4), 61.70 (C-6), 52.53 ($NHCOOCH_3$), 49.63 (C-2), 20.74, 20.68, 20.55 (3 × OAc); ESI-MS for $C_{20}H_{24}ClNO_{10}$ *m*/*z* calcd for ($M + Na^+$) 496.11: Found 496.1 ($M + Na^+$).

4-Chlorophenyl 3,4,6-tri-O-acetyl-2-deoxy-(*2-methoxycarbonylamino*)-*α-D-galactopyranoside* (**4.1e**) $[\alpha]^{20}_{D}$ = +20.9 (*c* 0.32, $CHCl_3$); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.26 (d, *J* = 8.9, 2H, ArH), 7.00 (d, *J* = 9.0, 2H, ArH), 5.56 (d, *J* = 3.4, 1H, H-1), 5.44 (d, 1H, H-4), 5.32 (dd, *J* = 11.3, 2.9, 1H, H-3), 4.98 (d, *J* = 9.8, 1H, NH), 4.45 (m, 1H, H-2), 4.24 (t, *J* = 6.6, 1H, H-5), 4.07-4.01 (m, 2H, H-6, 6′), 3.66 (s, 3H, $NHCOOCH_3$), 2.18 (s, 3H), 2.04 (s, 3H), 1.94 (s, 3H) (3 × OAc); ^{13}C NMR (126 MHz, $CDCl_3$) δ 170.76, 170.26, 170.18 (3 × CO, 3 × OAc), 156.53 ($NHCOOCH_3$), 154.57, 129.57 (ArC), 117.97 (ArC), 96.92 (C-1), 68.10 (C-3), 67.67 (C-5), 67.14 (C-4), 61.59 (C-6), 52.49 ($NHCOOCH_3$), 49.61 (C-2), 20.72, 20.66, 20.53 (3 × OAc); ESI-MS for $C_{20}H_{24}ClNO_{10}$ *m*/*z* calcd for ($M + Na^+$) 496.11: Found 496.1 ($M + Na^+$).

2-Naphthyl 3,4,6-tri-O-acetyl-2-deoxy-(*2-methoxycarbonylamino*)-*α-D-galactopyranoside* (**4.1f**) $[\alpha]^{20}_{D}$ = +30.1 (*c* 1.0, $CHCl_3$); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.79 (d, *J* = 8.9, 2H, ArH), 7.74 (d, *J* = 8.1, 1H, ArH), 7.49-7.43 (m, 1H, ArH), 7.42-7.36 (m, 1H, ArH), 7.22 (dd, *J* = 8.9, 2.5, 1H, ArH), 5.75 (d, *J* = 3.2, 1H, H-1), 5.48 (app.d, 1H, H-4), 5.41 (dd, *J* = 11.3, 3.0, 1H, H-3), 5.08 (d, *J* = 9.9, 1H, NH), 4.48-4.54 (m, 1H, H-2), 4.33 (t, *J* = 6.5, 1H, H-5), 4.17-4.04 (m, 2H, H-6, 6′), 3.66 (s, 3H, $NHCOOCH_3$), 2.20 (s, 3H), 2.05 (s, 3H), 2.04 (s, 3H) (3 × OAc); ^{13}C NMR (101 MHz, $CDCl_3$) δ 171.10, 170.28, 170.22 (3 × CO, 3 × OAc), 156.59 ($NHCOOCH_3$), 134.15, 129.91, 129.70 (ArC), 127.64 (ArC), 127.11

(ArC), 126.67 (ArC), 124.67 (ArC), 118.37 (ArC), 111.21 (ArC), 96.78 (C-1), 68.34 (C-3), 67.66 (C-5), 67.29 (C-4), 61.65 (C-6), 52.45 (NHCOO*C*H$_3$), 49.73 (C-2), 20.73, 20.67, 20.42 (3 × OAc); ESI-MS for $C_{24}H_{27}NO_{10}$ *m/z* calcd for (M + Na^+) 512.15: Found 512.2 (M + Na^+).

4.2.1.2. Aryl 2-acetamido-2-deoxy-α-D-galactopyranosides (Compounds 4.2a-4.2f)

3-Nitrophenyl 2-acetamido-2-deoxy-α-D-galactopyranoside (**4.2a**)

Mpt = 159-161 °C; $[\alpha]^{20}_D$ = +271.4 (c 0.29, H_2O); 1H NMR (600 MHz, D_2O) δ 7.97 (t, *J* = 2.2, 1H, ArH), 7.95 (d, *J* = 7.8, 1H, ArH), 7.56 (t, *J* = 8.2, 1H, ArH), 7.51 (dd, *J* = 8.3, 2.3, 1H, ArH), 5.73 (d, *J* = 3.6, 1H, H-1), 4.37 (dd, *J* = 11.0, 3.6, 1H, H-2), 4.15 (dd, *J* = 11.1, 3.2, 1H, H-3), 4.05-4.02 (m, 2H, H-4, H-5), 3.77-3.71 (m, 2H, H-6, 6′), 2.02 (s, 3H, NHAc); ^{13}C NMR (151 MHz, D_2O) δ 174.77 (CO, NHAc), 156.35 (ArC), 148.67 (ArC), 130.51 (ArC), 123.84 (ArC), 117.85 (ArC), 111.93 (ArC), 96.26 (C-1), 72.16 (C5), 68.38 (C-4), 67.58 (C-3), 61.06 (C-6), 49.59 (C-2), 21.91 (NHCO*C*H$_3$); HRMS expected for $C_{14}H_{18}N_2O_8$ is 365.0961 (M + Na^+) : Found 365.0953 (M + Na^+).

Phenyl 2-acetamido-2-deoxy-α-D-galactopyranoside (**4.2b**)

Mpt = 249-250 °C (decomp.); $[\alpha]^{20}_D$ = +220.0 (*c* 0.56, H_2O); ^{1}H NMR (600 MHz, D_2O) δ 7.43 (t, *J* = 8.0, 2H), 7.17 (dd, *J* = 12.6, 7.7, 3H), 5.66 (d, *J* = 3.6, 1H, H-1), 4.37 (dd, *J* = 11.1, 3.5, 1H, H-2), 4.19 (dd, *J* = 11.2, 3.2, 1H, H-3), 4.13 (t, *J* = 6.1, 1H, H-5), 4.10 (app d, 1H, H-4), 3.75-3.77 (m, 2H, H-6, 6′), 2.07 (s, 3H, NHAc); ^{13}C NMR (151 MHz, D_2O) δ 174.27 (CO, NH*C*OCH$_3$), 155.69 (ArC), 129.39 (ArC), 122.64 (ArC), 116.70 (ArC), 95.73 (C-1), 71.26 (C-5), 67.94 (C-4), 67.15 (C-3), 60.56 (C-6), 49.28 (C-2), 21.40 (NHCO*C*H$_3$); HRMS expected for $C_{14}H_{19}NO_6$ is, 320.1110 (M + Na^+): Found 320.1103 (M + Na^+).

4-Methoxyphenyl 2-acetamido-2-deoxy-α-D-galactopyranoside (**4.2c**)

Mpt = 175-177 °C; $[\alpha]^{20}_D$ = +200.8 (*c* 0.35, H_2O); ^{1}H NMR (600 MHz, D_2O) δ 7.05 (d, *J* = 9.1, 2H, ArH), 6.92 (d, *J* = 9.1, 2H, ArH), 5.43 (d, *J* = 3.7, 1H, H-1), 4.26 (dd, *J* = 10.9, 3.5, 1H, H-2), 4.09-4.05 (m, 2H, H-3, H-5), 4.01 (d, *J* = 2.5, 1H, H-4), 3.74 (s, 3H, OCH$_3$), 3.68 (app. d, 2H, H-6, 6′), 2.00 (s, 3H, NHAc); ^{13}C NMR (151 MHz, D_2O) δ

174.26 (CO, NH*C*OCH$_3$), 154.16 (ArC), 150.03 (ArC), 118.40 (ArC), 114.56 (ArC), 96.80 (C-1), 71.21 (C-5), 67.96 (C-4), 67.13 (C-3), 60.59 (C-6), 55.31 (OCH$_3$), 49.33 (C-2), 21.41 (NHCO*C*H$_3$); HRMS expected for $C_{15}H_{21}NO_7$ is 350.1216 (M + Na$^+$): Found 350.1210 (M + Na$^+$).

3-Chlorophenyl 2-acetamido-2-deoxy-α-D-galactopyranoside (**4.2d**)

Mpt = 180-181 °C (decomp.); $[\alpha]^{20}_D$ = +210.5 (*c* 0.63, H_2O); ^{1}H NMR (600 MHz, D_2O) δ 7.32 (t, *J* = 8.0, 1H, ArH), 7.20 (t, *J* = 2.2, 1H, ArH), 7.14-7.10 (m, 1H, ArH), 7.06 (dd, *J* = 8.4, 2.4, 1H, ArH), 5.61 (d, *J* = 3.7, 1H, H-1), 4.33 (dd, *J* = 11.1, 3.7, 1H, H-2), 4.12 (dd, *J* = 11.0, 3.2, 1H, H-3), 4.05-4.02 (m, 2H, H-4, H-5), 3.74-3.68 (m, 2H, H-6, 6′), 2.02 (s, 3H, NHAc); ^{13}C NMR (151 MHz, D_2O) δ 174.77 (CO, NH*C*OCH$_3$), 156.83 (ArC), 134.30 (ArC), 130.78 (ArC), 122.99 (ArC), 117.35 (ArC), 115.49 (ArC), 96.22 (C-1), 71.94 (C-5), 68.42 (C-4), 67.60 (C-3), 61.06 (C-6), 49.67 (C-2), 21.89 (NHCO*C*H$_3$); HRMS expected for $C_{14}H_{18}ClNO_6$ is 354.0720 (M + Na$^+$): Found 354.0714 (M + Na$^+$).

4-Chlorophenyl 2-acetamido-2-deoxy-α-D-galactopyranoside (**4.2e**)

Mpt = 218-220 °C; $[\alpha]^{20}_D$ = +190.5 (*c* 0.8, H_2O); ^{1}H NMR (600 MHz, D_2O) δ 7.40 (d, *J* = 9.0, 2H), 7.14 (d, *J* = 9.0, 2H), 5.63 (d, *J* = 3.6, 1H, H-1), 4.37 (dd, *J* = 11.1, 3.7, 1H, H-2), 4.16 (dd, *J* = 11.1, 3.2 Hz, H-3), 4.09 (m, 2H, H-4, H-5), 3.79-3.72 (m, 2H, H-6, 6′), 2.07 (s, 3H, NHAc); ^{13}C NMR (151 MHz, D_2O) δ 174.25 (CO, NH*C*OCH$_3$), 154.35 (ArC), 129.02 (ArC), 126.82 (ArC), 118.08 (ArC), 95.85 (C-1), 71.37 (C-5), 67.90 (C-4), 67.11 (C-3), 60.56 (C-6), 49.20 (C-2), 21.40 (NHCO*C*H$_3$); HRMS expected for $C_{14}H_{18}ClNO_6$ is 354.0720 (M + Na$^+$): Found 354.0717 (M + Na$^+$).

2-Naphthyl 2-acetamido-2-deoxy-α-D-galactopyranoside (**4.2f**) Mpt = 240-243 °C; $[\alpha]^{20}_D$ = +205.7 (*c* 0.36, H_2O); ^{1}H NMR (500 MHz, D_2O) δ 7.95-7.89 (m, 2H, ArH), 7.85 (d, *J* = 8.2, 1H, ArH), 7.59-7.52 (m, 2H, ArH), 7.46 (t, *J* = 7.5 1H, ArH), 7.33 (dd, *J* = 8.9, 2.2, 1H, ArH), 5.76 (d, *J* = 3.5, 1H, H-1), 4.40 (dd, *J* = 11.1, 3.5, 1H, H-2), 4.20 (dd, *J* = 11.1, 3.0, 1H, H-3), 4.15-4.05 (m, 2H, H-4, H-5), 3.79–3.64 (m, 2H, H-6, 6′), 2.04 (s, 3H, NHAc); ^{13}C NMR (126 MHz, D_2O) δ 170.42 (CO, NH*C*OCH$_3$), 153.52 (ArC), 134.17 (ArC), 129.90 (ArC), 129.79 (ArC), 127.69 (ArC), 127.17 (ArC), 126.78 (ArC), 124.76 (ArC), 118.31 (ArC), 111.24 (ArC), 96.34 (C-1), 68.36 (C-5), 68.23 (C-4), 67.65 (C-3), 61.73 (C-6),

47.92 (C-2), 20.80 (NHCO*C*H_3); HRMS expected for $C_{18}H_{21}NO_6$ is 370.1267 ($M + Na^+$): Found 370.1210 ($M + Na^+$).

4.2.2. Dimeric-3,4,6-tri-O-acetyl-2-deoxy-2-nitroso-α-D-galactopyranosyl chloride (4.4).

NOCl gas, which was prepared by dropwise addition of saturated solution of sodium nitrite (13 g) in water (20 mL) into concentrated HCl (50 mL), was bubbled through a cooled solution (–45 °C) of thoroughly dried D-galactal(*30*) **4.3** in EtOAc (40 mL). After the addition was complete, the temperature was allowed to warm up to 0 °C over a period of 20 min, and the volatiles were then removed under reduced pressure to give a bluish-green mass, which was further dried under vacuum for 2 h. The resultant crude product was used in the subsequent glycosylation without further purification.

4.2.3. 2-Naphthyl 2-acetamido-2-deoxy-α-D-(2-^{2}H)galactopyranoside (4.8, Scheme 4-3).

Dimeric 3,4,6-tri-*O*-acetyl-2-deoxy-2-nitroso-α-D-galactopyranosyl chloride **4.4**, made from galactal **4.3,**(*30*) was dissolved in anhydrous DMF (100 mL) and to this solution was added 2-naphthol (2 g, 2.0 equiv). The resultant mixture was stirred for 70 h at ambient temperature. At this time TLC analysis (40% EtOAc-Hexane) indicated complete disappearance of the starting material. The reaction mixture was then diluted with Et_2O (50 mL) and extracted with water (3 × 40 mL), washed with brine (2 × 20 mL), dried (Na_2SO_4) and concentrated under reduced pressure to yield a yellow syrup. Purification by flash chromatography (40% EtOAc-Hexane) afforded the pure product 2-naphthyl 3,4,6-tri-*O*-acetyl-2-deoxy-2-oximino-α-D-*lyxo*-hexopyranoside **4.5** (1.0 g, 30% over two steps), which was thoroughly dried under high vacuum (10 h). To the solution of **4.5** (1.0 g, 2.2 mmol) in anhydrous THF (100 mL), under refluxing conditions, was added lithium aluminum deuteride (940 mg, 22.4 mmol) in portions during a period of 10 min.(*31*) After the reaction mixture was kept under reflux for 40 min, excess lithium aluminum deuteride was destroyed by careful addition of methanol (20 mL). The resultant flocculent precipitate was separated by filtration through celite and washed several times with methanol (3 × 25 mL). Subsequently, the combined filtrate was evaporated under reduced pressure to

yield 1.2 g of the crude completely deprotected compound **4.7** as yellowish syrup. This crude material was immediately acetylated under standard conditions using a mixture of acetic anhydride (10 mL) and pyridine (10 mL), for which stirring was continued for 15 h at ambient temperature. The reaction was then quenched by addition of water (30 mL). The per-acetylated product was extracted into CH_2Cl_2 (3 × 40 mL), and the organic layer was washed with cold 10% H_2SO_4 (2 × 10 mL), brine (2 × 10 mL) and dried (Na_2SO_4). After removing the volatiles under reduced pressure, excess pyridine was removed azeotropically using toluene to afford the desired product as a pale yellow syrup. This was further column purified via flash chromatography using 5% CH_3OH-CH_2Cl_2 to obtain the 2D labelled product **4.8** in > 98% purity as determined by ^{1}H NMR. Mpt = 239–240 °C; $[\alpha]^{20}_D$ = +200.7 (*c* 0.3, H_2O); ^{1}H NMR (500 MHz, D_2O) δ 7.84 (t, *J* = 7.6, 1H, ArH), 7.79 (d, *J* = 8.1, 1H, ArH), 7.52 (s, 1H, ArH), 7.47 (t, *J* = 7.5, 1H, ArH), 7.40 (t, *J* = 7.6, 1H, ArH), 7.27 (d, *J* = 8.9, 1H, ArH), 5.71 (s, 1H, H-1), 4.14 (d, *J* = 2.3 Hz, 1H, H-3), 4.07-4.02 (m, 2H, H-4, H-5), 3.72-3.62 (m, 2H, H-6, 6′), 1.97 (s, 3H, NHAc); ^{13}C NMR (151 MHz, D_2O) δ 174.76 (CO, NH*C*OCH_3), 154.01 (ArC), 133.93 (ArC), 129.78 (ArCH), 127.67 (ArCH), 127.07 (ArCH), 126.87 (ArCH), 124.83 (ArCH), 119.02 (ArCH), 111.37 (ArCH), 96.14 (C-1), 71.85 (C-5), 68.43 (C-4), 67.62 (C-3), 61.05 (C-6), 21.90 (NHCO*C*H_3); HRMS expected for $C_{18}H_{20}DNO_6$ is 371.1329 ($M + Na^+$): Found 371.1325 ($M + Na^+$).

4.2.4. Cloning of α-*N*-acetylgalactosaminidase (α-NAGAL).

The *FlvA* gene encoding α-*N*-acetylgalactosaminidase was PCR amplified from *Elizabethkingia meningosepticum* genomic DNA (ATCC 51720D). The amplification was performed in 5% DMSO using the following primers forward 5'-GCGGGATCCGAATTCTAAAAAGGTAAGAATAGCTTTTATAGC-3'/*EcoR*I and reverse 5'-CGGCGGCCGCTTAGTAGTCGTCATTTATTGCAAATG-3'/*Not*I which introduced *EcoRI* and *NotI* restriction sites (underlined) in the forward and reverse primer, respectively. The PCR fragment was purified, digested with *EcoRI* and *NotI*, inserted into correspondingly digested pET28a vector (Novagen), and transformed in *Escherichia coli* BL-21(DE3 Gold) competent cells. The plasmid DNA was isolated from a single colony and

was verified by restriction digest and DNA sequencing by Macrogen using T7 promoter and T7 terminator primers.

4.2.5. Expression and Purification of the α-*N*-Acetylgalactosaminidase.

For expression of α-NAGAL from *FlvA* gene, cells were inoculated in Luria broth media and then expressed in "terrific broth" (TB media) supplemented with 1% kanamycin at 37 °C to an OD_{600nm} of 0.5 before induction with 0.5 mM IPTG. The cultures were incubated at 25 °C for 5 h and then centrifuged (10 min at 8400 × g) and the pellet from 1 L of culture was re-suspended in binding buffer containing 5 mM imidazole. The cells were lysed open using 1% lysozyme (from chicken egg white) and protease inhibitor cocktail tablet followed by sonication (20 s on/40 s off cycle at capacity of 60%) to ensure complete lysis of the cells. The lysate was centrifuged at 4 °C (34,000 × g) in order to remove the cell debris and intact cells. The clear supernatant was collected and filtered through a 0.45 micron filter before loading on to a HiTrap Chelating HP column 5 mL, which had been pre-equilibrated with the binding buffer. The column was washed sequentially with 60 mM, 100 mM and 150 mM imidazole before eluting the protein with 250 mM of imidazole. Fractions containing pure His-α-NAGAL, as determined by 10% SDS-PAGE, were pooled together and dialyzed, at 4 °C, three times against 4 L of Tris buffer (20 mM, pH 7.0) that contained NaCl (100 mM). The protein was then concentrated by centrifugation through a 10 kDa filter and its concentration was assessed (Bradford Assay).

4.2.6. Typical Conditions for the Measurement of Michaelis–Menten Parameters.

The concentration of α-*N*-acetylgalactosaminidase (α-NAGAL) was chosen such that less than 10% of the total substrate was consumed during the assay. For each assay, the enzyme was incubated in the appropriate buffer at 37 °C for 5 min. After which the reaction was initiated by the addition of substrate. The initial rate of hydrolysis was followed spectrophotometrically at the wavelength of maximal absorbance change. Typically, the substrate concentration was varied between 40 μM and 500 mM. The initial rates versus substrate concentration kinetic data were fitted to the Michaelis-Menten equation using a standard nonlinear least-squares program Prism 4.0.

4.2.7. Determination of the pH-Rate Profile.

To determine the effect of pH on enzymatic activity, kinetics parameters were measured over a pH range of 6.3-8.8. The buffers used were MES (20 mM, MES-NaOH, pH 6.0-6.7), HEPES (20 mM, pH 6.5-8.2), and CHES (20 mM, pH 8.5-9.5). Typical assay conditions were: α-NAGAL (final concentration of 0.23 μg/mL) was incubated at 37 °C with the appropriate buffer containing NaCl (50 mM) and BSA (0.1 % w/v) for 5 min prior to addition of substrate PNPαGalNAc, the hydrolysis reaction was monitored at 400 nm (pH 6.3-8.8). The difference in extinction coefficients (Δε) for PNPαGalNAc and the released 4-nitro phenolate was determined at each pH value and the initial rate measurements were fit to a standard Michaelis-Menten equation using a standard nonlinear least-squares program Prism 4.0.

4.2.8. Kinetic Investigation of Cofactor Dependence of GH109 α-NAGAL.

A sample of α-NAGAL (final concentration of 0.23 μg/mL) in buffer (50 mM HEPES, pH 7.5) was pre-incubated with NAD^+ (5-100 μM) at 37 °C for 5 min. PNPαGalNAc was added to the solution to initiate the hydrolysis reaction, which was monitered at 400 nm. Similar experiments were implemented to evaluate the dependence of catalytic activity on reducing agent TCEP (1-100 μM) and divalent metal ion Mn^{2+} (10-100 μM).

4.2.9. Product Studies.

^1H NMR spectroscopy (500 MHz) was employed to identify the stereochemical course of the enzyme-catalyzed reaction. The reaction conditions involved incubating a mixture of enzyme GH109 α-NAGAL (3.0 μg/ml) in buffer (pH 7.5, 50 mM HEPES) containing methanol-D_4 (5 M) at 37 °C. After addition of PNPαGalNAc (1.2 mg) the reaction was allowed to proceed at 37 °C for ~10 h. Removal of the enzyme by centrifugal ultra-filtration (10 KDa filtration unit) was accomplished at 4 °C, followed by lyophilization of the resultant solution to afford a white solid which was re-dissolved in D_2O to aquire the 1H NMR spectrum. Additionally, in order to evaluate the acidity of the C-2 proton under enzymatic conditions, GH109 α-NAGAL catalyzed hydrolysis of both the activated

substrates, PNPαGalNAc and PNPαGal, was performed in D_2O, at pD 7.5 (50 mM HEPES, D_2O) and at 37 °C.

4.2.10. Linear Free Energy Relationship-Brønsted Analysis.

A series of substrates with varying leaving groups were synthesized to perform a Brønsted analysis (**Scheme 4-1** and **Figure 4-7 and 4-8**). Full Michaelis-Menten curves were measured, in buffer (50 mM HEPES, pH 7.5), for each substrate using the above listed protocol.

4.3. Results

4.3.1. Cloning and Expression of α-NAGAL.

In the present work, the *FlvA* gene from from *Elizabethkingia Meningosepticum* was amplified by PCR from the genomic DNA (ATCC 51720D) and recombinantly expressed in *E. coli* in order to perform a detailed mechanistic study on the resulting GH109 α-NAGAL enzyme. The genomic sequence results shows that the *FlvA* gene encodes for a 418 amino acid protein, and that this material has a predicted molecular weight of 47.4 kDa. Protein expression was carried out in "terrific broth" (TB) media to give a final purified enzyme concentration, as quantified by Bradford assay, of 0.2 mg/mL protein (total volume 3 mL) that was obtained from 6 L of expression media.

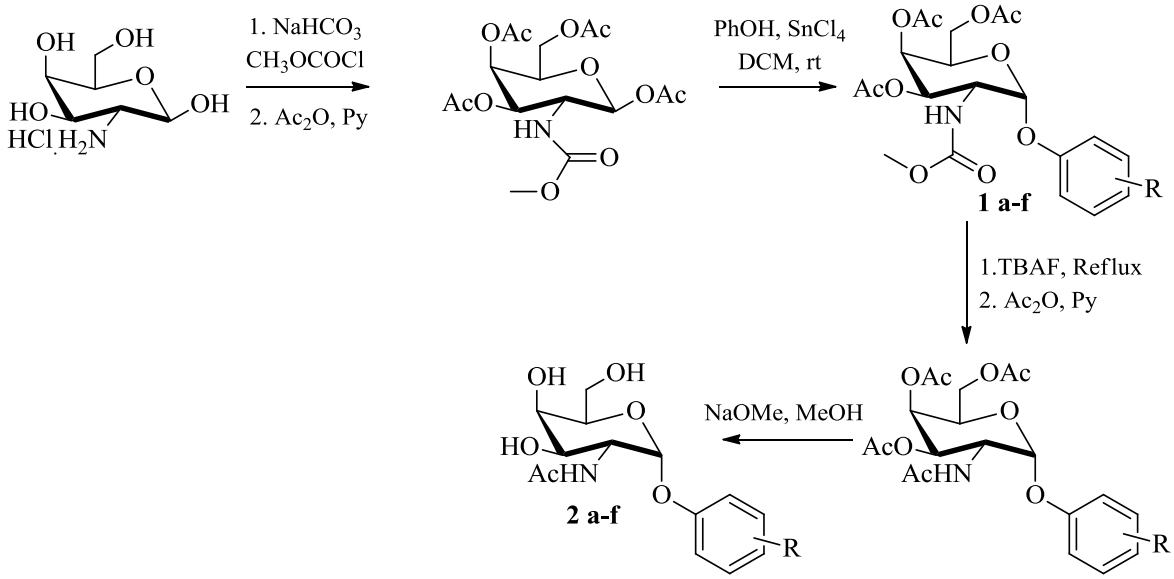
4.3.2. Cofactor Analysis.

As the active site architectures of GH109 α-NAGAL from *E. meningosepticum* and GH4 glycosidases are similar it is possible that a comparable dependence exists for catalytic activity on cofactor concentrations.(*1*) Therefore, we tested the dependence of catalytic activity for our recombinant GH109 α-NAGAL on NAD^+ (25-300 µM), reducing agent TCEP (1-100 µM) and divalent metal ion Mn^{2+} (25-300 µM) concentrations at 37 °C in buffer (50 mM, HEPES, pH 7.5) by employing PNPαGalNAc as the substrate. The measured rate of hydrolysis was unaltered upon addition of any of these cofactors to the

assay media. As a result, it was decided to measure all kinetic parameters in buffers that were devoid of added NAD^+, Mn^{2+} and reducing agent.

4.3.3. Substrate Synthesis.

A panel of seven aryl 2-acetamido-2-deoxy-α-D-galactopyranosides were synthesized for the Brønsted analysis by following literature procedures using 1,3,4,6-tetra-*O*-acetyl-2-deoxy–(2-methoxycarbonylamino)-α-D-galactopyranoside as the starting material (**Scheme 4-1**).(*28, 29*)



Scheme 4-1. Synthetic route to aryl 2-acetamido-2-deoxy-α-D-galactopyranosides.

4.3.4. Product Studies.

Hydrolysis of PNPαGalNAc in the presence of methanol-D_4 and GH109 α-NAGAL (3.0 µg/ml) was allowed to proceed for 15 h at 37 °C. After which the enzyme was removed by filtration (10K filtering unit) and the resultant supernatant was lyophilized. The ^{1}H-NMR spectrum of the solid residue, after dissolving in D_2O, showed an anomeric proton at a chemical shift δ 5.22 ($J_{1,2}$ = 3.8 Hz), in addition to the peak corresponding to the unreacted substrate, indicating the formation of methyl 2-acetamido-2-deoxy-α-D-galactopyranoside. **Figure 4-4** shows the ^{1}H-NMR spectrum, containing PNPαGalNAc and the methyl glycoside, when the reaction was stopped before going to

completion. As a result, we can conclude that the hydrolysis reaction catalyzed by GH109 α-NAGAL proceeds with a retention of anomeric configuration. Of note, when GH109 α-NAGAL-catalyzed hydrolysis of PNPαGalNAc was performed in D_2O, the reaction product, 2-acetamido-2-deoxy-α-D-galactose was completely deuterated at C-2 as shown by the ^1H NMR spectrum (**Figure 4-5)** where the anomeric protons for both anomers were singlets. Similar proton deuterium exchange was observed on C-2 when PNPαGal was employed as the substrate (data not shown).

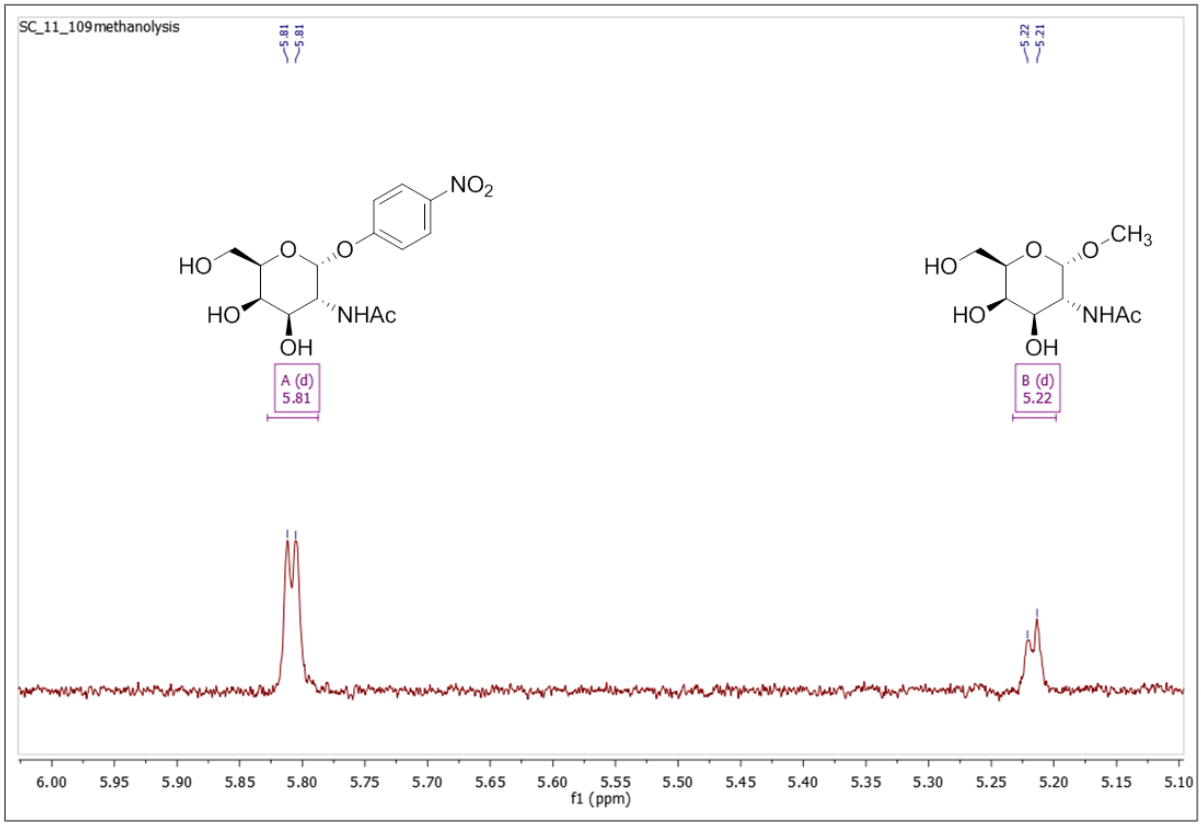


Figure 4-4. GH109 α-NAGAL catalyzed hydrolysis of PNPαGalNAc in the presence of methanol-D4, showing the appearance of a doublet at δ 5.22 ($J_{1,2}$ = 3.8 Hz) corresponding to the anomeric proton of methyl 2-acetamido-2-deoxy-α-D-galactopyranoside, the doublet at δ 5.81 corresponds to unreacted starting material.

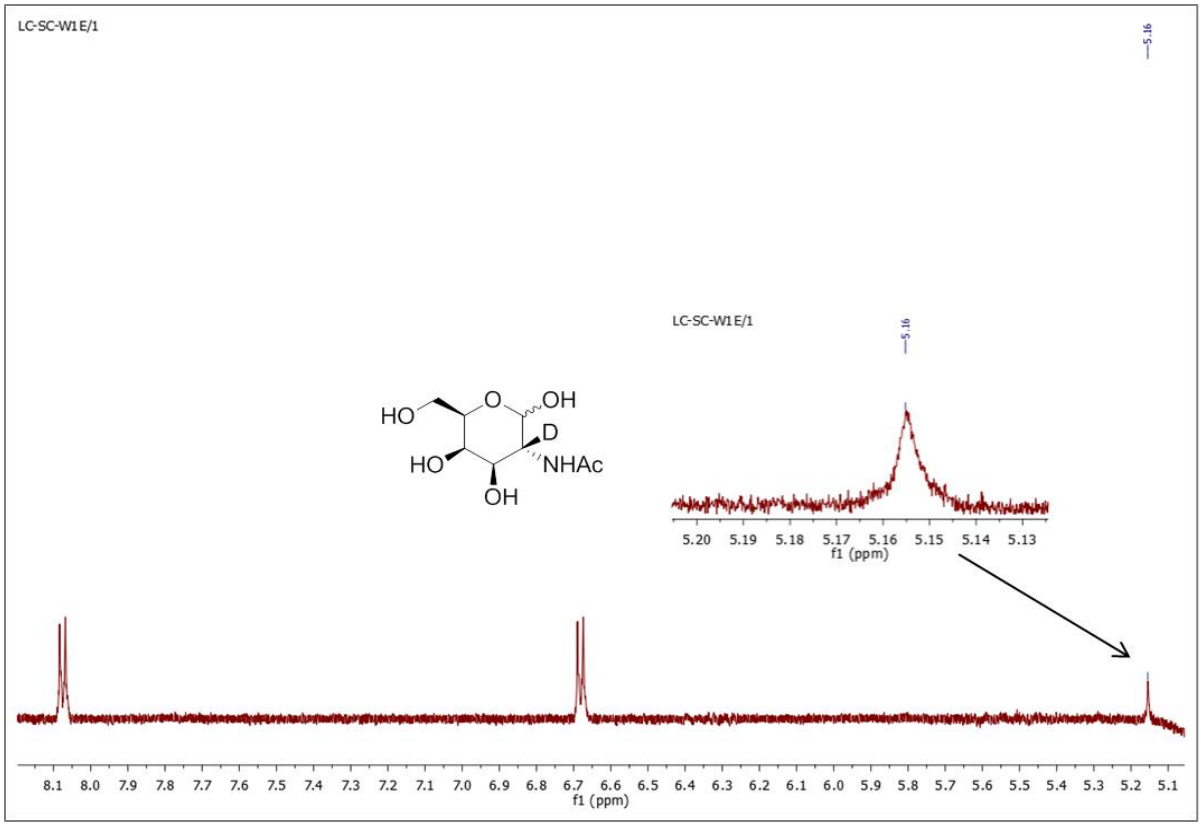


Figure 4-5. Change in the multiplicity of the anomeric proton upon performing the hydrolysis of PNPαGalNAc in D_2O in the presence of GH109 α-NAGAL. The proton on C-2 position exchanges with a solvent deuterium, a process that results in a singlet (δ 5.22 ppm) for the anomeric proton of the α- and β-product.

4.3.5. pH Profile of α-NAGAL Catalyzed Hydrolysis.

The variation in the kinetic parameters, k_{cat}/K_m and k_{cat} for the enzyme-catalyzed hydrolysis of PNPαGalNAc as a function of pH are listed in **Table 4-1**. A fit of the kinetic data for kcat to a bell-shaped pH-rate curve was ill-defined. Instead, obtained data showed a good fitting (R^2 > 99%) to a bell-shaped curve incorporating Hill coefficients of 3.4 and 2.8. Although, catalytic parameter kcat varied with solution pH, K_m values over the entire pH range (6.3-8.8) remained invariant, except at the pH extremes (**Figure 4-6**). However, it is apparent that for both kinetic parameters, k_{cat} and k_{cat}/K_m the maximal activity for GH109 α-NAGAL (from *E. meningosepticum*) occurred around a pH value of 7.2-7.5 (**Figure 4-6**).

Table 4-1. Michaelis-Menten kinetic parameters for GH109 α-NAGAL catalyzed hydrolysis of PNPαGalNAc as a function of pH at 37 °C.

pH	k_{cat} (s^{-1})	$10^5 \times k_{cat}/K_m$ (M^{-1} s^{-1})
6.3	0.95 ± 0.07	0.05 ± 0.01
6.6	1.50 ± 0.10	0.29 ± 0.11
6.8	3.91 ± 0.22	1.44 ± 0.29
7.0	10.5 ± 0.7	1.68 ± 0.58
7.2	17.1 ± 0.42	6.11 ± 1.08
7.5	19.7 ± 0.32	12.6 ± 1.84
7.8	14.2 ± 0.31	8.38 ± 0.74
8.1	5.01 ± 0.03	2.91 ± 0.13
8.4	0.58 ± 0.07	0.04 ± 0.02
8.8	0.17 ± 0.03	0.008 ± 0.004

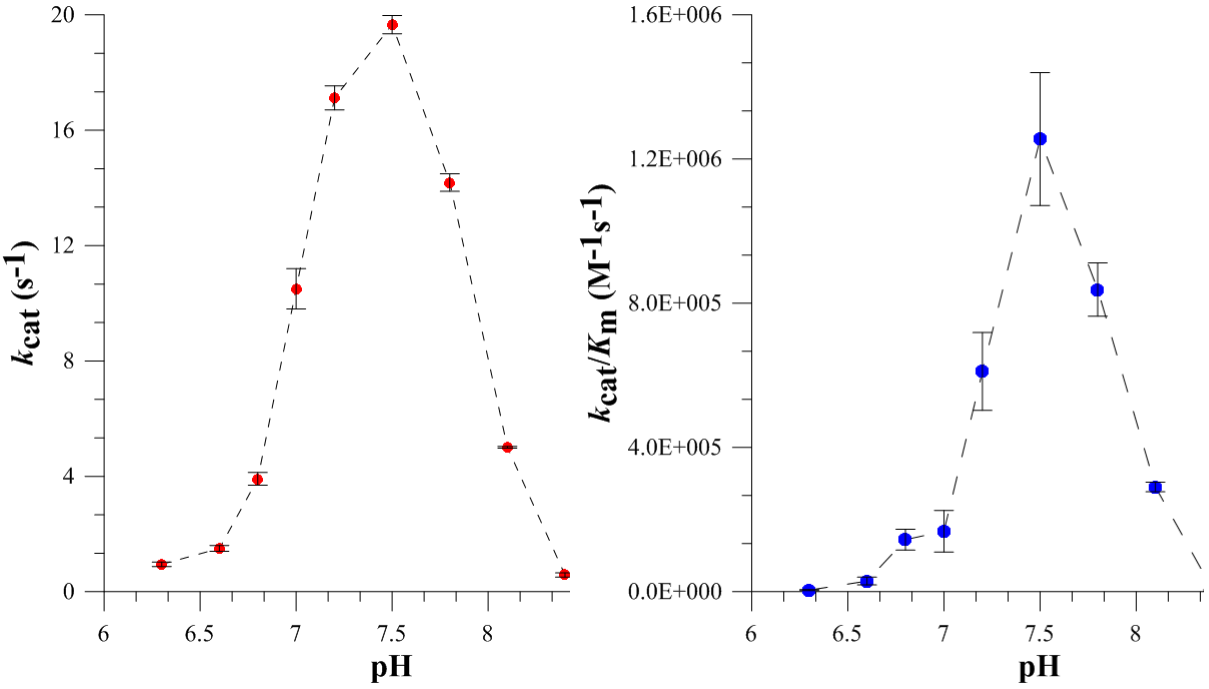


Figure 4-6. pH dependence of k_{cat} for the α-NAGAL-catalyzed hydrolysis of PNPαGalNAc (red circles) and the corresponding variation in k_{cat}/K_m for the same (blue circles).

4.3.6. Substrate Specificity.

The anomeric specificity for the recombinant GH109 α-NAGAL was determined by performing scanning kinetic experiments using both PNPαGalNAc and PNPβGalNAc. For each assay, the enzyme (GH109 α-NAGAL from *E. meningosepticum*, 0.46 mg/mL) was incubated in the buffer (50 mM, pH 7.5, HEPES) at 37 °C for 5 min, when the reaction was initiated by the addition of substrate (PNPαGalNAc/ PNPβGalNAc, 100 mM). These scanning experiments were performed using a wavelength range of 200-500 nm over a time-period of 120 min. Additionally, the rate of hydrolysis of the 2-OH derivatives, PNPαGal (α and β) as substrates were also evaluated under these assay conditions. Michaelis-Menten kinetic parameters for GH109 α-NAGAL-catalyzed hydrolysis of both of the activated substrates (PNPαGal and PNPαGalNAc), were determined at pH 7.5 and 37 °C, within a substrate concentration range of $K_m/4$ to $4 \times K_m$. The rate versus substrate concentration kinetic data were fitted to the Michaelis-Menten equation using a standard

nonlinear least-squares program Prism 4.0 to give the catalytic parameters for PNPαGal (k_{cat} = 2.43 s^{-1}, $k_{cat}/K_M = 6.75 \times 10^3$ $M^{-1}s^{-1}$) and PNPαGalNAc (k_{cat} = 3.08 s^{-1}, $k_{cat}/K_m = 9.13 \times 10^4$ M^{-1} s^{-1}).

To obtain the complete Brønsted plots, the pK_a range of the conjugate acid of the leaving group was varied between 7.15-10.19. Full Michaelis-Menten curves were measured at a pH of 7.5 for each substrate (**Table 4-2**) and the associated Brønsted β_{lg} values on k_{cat} and k_{cat}/K_m are -0.08 ± 0.06 and -0.31 ± 0.12, respectively. Additionally, a complete Brønsted analysis was performed for a panel of six aryl α-D-galactopyranosides (synthesized previously, Chapter 2), with a pK_a range of the conjugate acid of the leaving group varied from 7.15 to 10.21, at pH 7.5 and at 37 °C. The catalytic parameters obtained are listed in **Table 4-3**. The derived β_{lg} values for the α-Gal series on k_{cat} and k_{cat}/K_m are 0.01 ± 0.10 and 0.15 ± 0.09, respectively. A comparison of the k_{cat} values for both the substrates are shown in **Figure 4-7** and the corresponding k_{cat}/K_m values are compared in **Figure 4-8**.

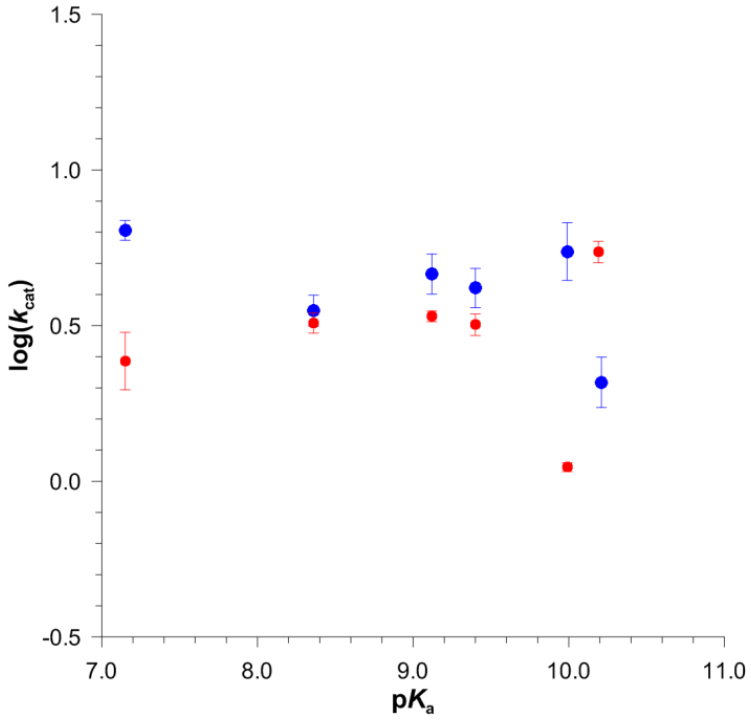


Figure 4-7. Effect of leaving-group ability on k_{cat} for GH109 α-NAGAL- catalyzed hydrolysis α-GalNAc (blue circles) and α-Gal (red circles) series of substrates measured at 37 °C and at pH 7.5.

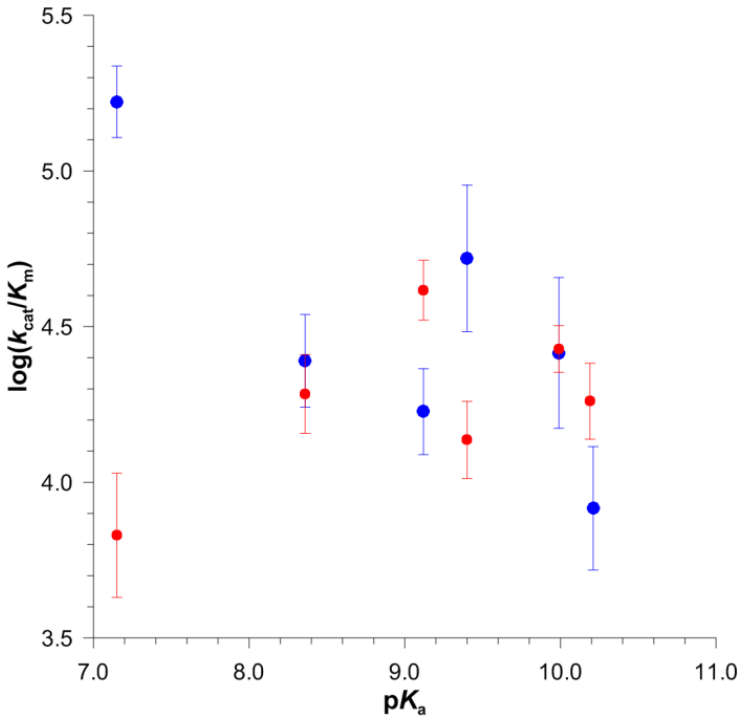


Figure 4-8. Effect of leaving-group ability on k_{cat}/K_m for GH109 α-NAGAL catalyzed hydrolysis α-GalNAc (blue circles) and α-Gal (red circles) series of substrates measured at 37 °C and at pH 7.5.

Table 4-2. Kinetic parameters for the hydrolysis of a series of aryl 2-acetamido-2-deoxy-α-D-galactopyranosides catalyzed by GH109 α-NAGAL at 37 °C.

p*K*a[a] (ArOH)	Log (k_{cat})	Log (k_{cat}/K_m)
7.15	0.81 ± 0.03	5.22 ± 0.11
8.36	0.55 ± 0.05	4.39 ± 0.15
9.12	0.67 ± 0.06	4.23 ± 0.14
9.40	0.62 ± 0.06	4.72 ± 0.23
9.99	0.74 ± 0.09	4.42 ± 0.24
10.21	0.32 ± 0.08	3.91 ± 0.21

[a] pK_a values used for phenols were taken from refs 32

Table 4-3. Kinetic parameters for the hydrolysis of a series of aryl α-D-galactopyranosides catalyzed by GH109 α-NAGAL at 37 °C.

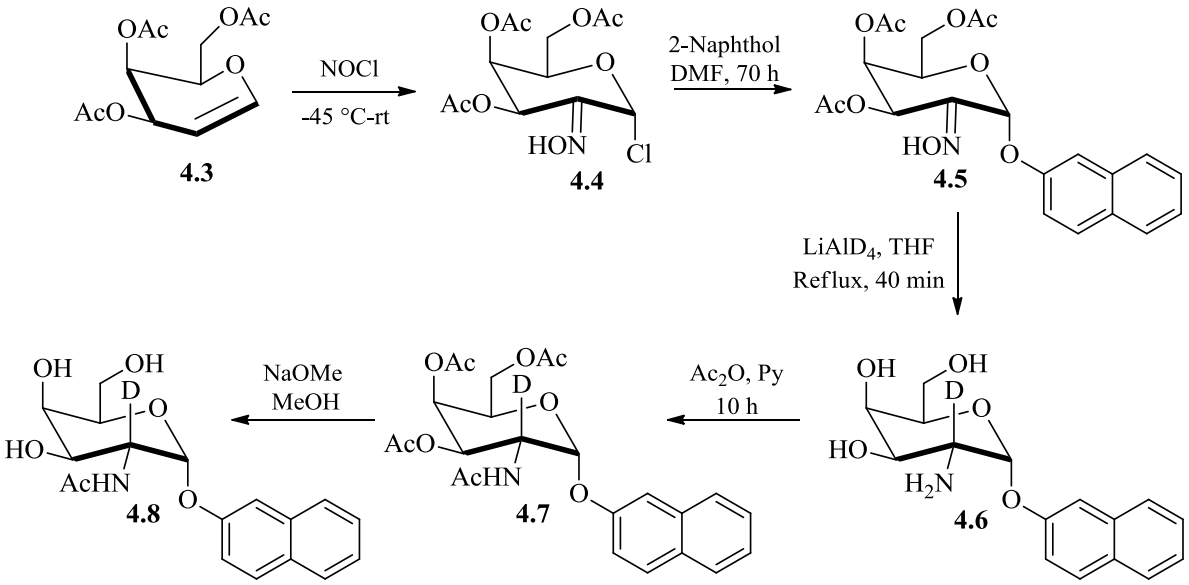
p*K*a[a] (ArOH)	Log (k_{cat})	Log (k_{cat}/K_m)
7.15	0.39 ± 0.09	3.83 ± 0.20
8.36	0.51 ± 0.03	4.28 ± 0.15
9.12	0.53 ± 0.02	4.62 ± 0.09
9.40	0.50 ± 0.03	4.14 ± 0.12
9.99	0.05 ± 0.01	4.42 ± 0.07
10.19	0.74 ± 0.03	4.26 ± 0.12

[a] pK_a values used for phenols were taken from refs 32

4.3.7. Synthesis of Deuterated α-GalNAc Series of Substrates.

2-naphthyl 2-acetamido-2-deoxy-α-D-(2-^{2}H)galactopyranoside **4.8** was synthesized by the appropriate choice of reducing agent by starting with tri-*O*-acetal-D-galactal (3,4,6-tri-*O*-acetyl-1,5-anhydro-2-deoxy-D-*lyxo*-hex-1-enitol) **4.3** using the nitrosyl chloride addition reaction (**Scheme 4-3**) pioneered by Lemieux and co-workers.(*33*) Reduction of the oxime **4.5** to give the 2-^{2}H isotopologue was accomplished using lithium aluminium deuteride ($LiAlD_4$)(*31*) and followed by standard peracetylation, purification and deprotection series of reactions on the product mixture. A comparison of the ^{1}H NMR of the deuterium labelled compound, 2-naphthyl 2-acetamido-2-deoxy-α-D-(2-^{2}H) galactopyranoside **4.8**, and the unlabelled compound, 2-naphthyl 2-acetamido-2-deoxy-α-D-galactopyranoside **4.2f** is shown in **Figure 4-9**. However, the yield obtained for the reduction of the oxime using $LiAlD_4$ (10 equivalents) was extremely low (< 15%). Hence, we anticipated that synthesis of the two other necessary isotopologues, 2-naphthyl 2-acetamido-2-deoxy-α-D-(3-^{2}H)galactopyranoside and 2-naphthyl 2-acetamido-2-deoxy-α-D-(2-^{2}H, 3-^{2}H)galactopyranoside both of which would involve several extra steps, would require the reactions to be run on a much larger scale in order to isolate sufficient quantities of pure substrates for the kinetic experiments. As a result, we decided to take advantage of the

remarkable efficiency of GH109 α-NAGAL in hydrolyzing α-Gal substrates so we could use the deuterium isotopologues of phenyl α-D-galactosides that we had already made, to measure KIEs on GH109 α-NAGAL-catalyzed hydrolysis reactions.



Scheme 4-2. Synthetic route to 2-naphthyl 2-acetamido-2-deoxy-α-D-(2-^{2}H) galactopyranoside starting with peracetylated D-galactal.

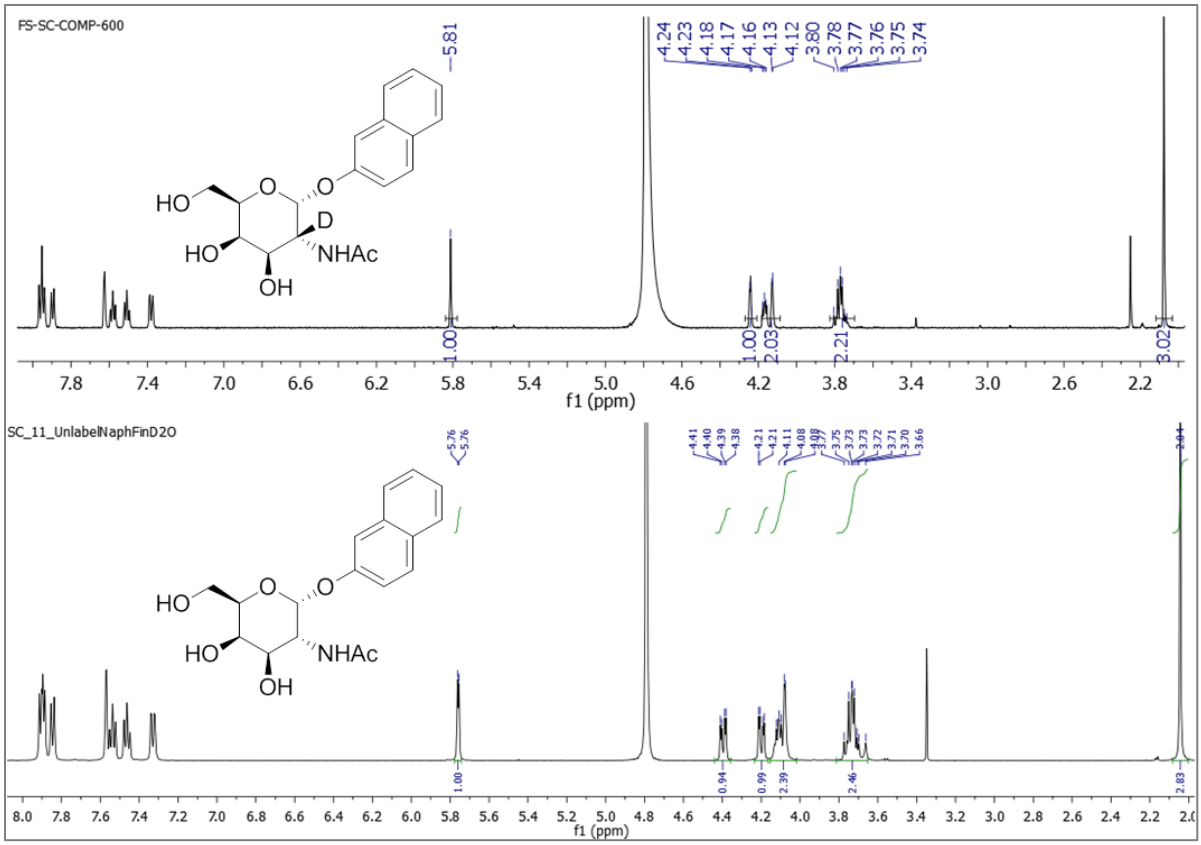


Figure 4-9. Comparison of the ^{1}H NMR of 2-naphthyl 2-acetamido-2-deoxy-α-D-(2-^{2}H)galactopyranoside and 2-naphthyl 2-acetamido-2-deoxy-α-D-galactopyranoside, showing the anomeric proton H-1 as a singlet and H-3 as a doublet in the deuterium isotopologue.

4.3.8. Synthesis of Deuterated α-Gal Series of Substrates.

To determine the KIEs (k_H/k_D) for the GH109 α-NAGAL-catalyzed hydrolysis of α-Gal substrates the following labelled substrates were required: phenyl α-D-galactopyranoside, phenyl α-D-(1-^{2}H)galactopyranoside, phenyl α-D-(2-^{2}H)galactopyranoside, phenyl α-D-(3-^{2}H)galactopyranoside. The synthesis of these labelled compounds is reported in Chapter 2 of this thesis.

4.3.9. Deuterium Kinetic Isotope Effects.

Measuring the kinetic isotope effects on $^{D}V/K$ using the substrate depletion method necessitates employing substrate concentration significantly below the K_m value.(*34*) The observed K_m for GH109 α-NAGAL-catalyzed hydrolysis of phenyl α-D-galactopyranoside

was as low as 42 μM, hence it would have been unfeasible to employ the desired substrate concentration ($1/5 \times K_m$) for such measurements. Therefore, we decided to measure the KIE on ^{D}V instead. In order to measure isotope effects (KIE) on the maximal enzymatic velocity V (i.e ^{D}V), direct comparison of individually measured rates of unlabeled and labeled substrates at saturating levels ($5 \times K_m$) were performed. The substrates employed for the KIE studies were deuterium labeled at C-1, C-2 and C-3 of phenyl α-D-galactopyranoside.(*25*) **Table 4-4** lists the calculated KIEs on ^{D}V for the GH109 α-NAGAL-catalyzed hydrolysis of isotopologues of PαG.

Table 4-4. Kinetic Isotope Measurements on V for the GH109 α-NAGAL-catalyzed hydrolysis of phenyl α-D-Galactopyranoside at a pH of 7.5 and a temperature of 37 °C.

Isotopologue 1	Isotopologue 2	$^{D}V^{a}$
PαG	1-(^{2}H)PαG	1.47 ± 0.10
PαG	2-(^{2}H)PαG	2.36 ± 0.18
PαG	3-(^{2}H)PαG	2.12 ± 0.18

[a] Ratio of V(isotopologue 1)/V(isotopologue 2) at $5 \times K_m$.

4.4. Discussion

Recently, we described a detailed mechanistic evaluation of the GH4 α-galactosidase from *Citrobactor freundii* that incorporated the measurement of deuterium kinetic isotope effects (KIEs) and a Brønsted analysis.(*25*) The single crystal X-ray diffraction structure of GH109 α-NAGAL from *E. meningosepticum*, at a resolution of 2.3 Å, was shown to possess a striking similarity to the active site architecture of several GH4 glycosidases.(*1, 23*) Moreover, both GH families have been shown, using the technique of X-ray diffraction, to bind NAD^+ in their respective structures. When the structure for the GH109 α-NAGAL from *E. meningosepticum(1)* was compared with those of *Bacillus subtilis* phospho-α-glucosidase, GlvA(*27*) and *Thermotoga maritima* 6-phospho-β-glycosidase, BglT,(*24*) a perfect overlap was observed for the dinucleotide cofactors, the sugar rings and an enzymatic tyrosine

residue. Given that the catalytic machinery of GH4 glycosidases (**Figure 1-17**, Introduction) is dependent on three cofactors,(*24*, *25*) namely: a) NAD^+; b) divalent metal ion, Mn^{2+}; and c) reducing conditions (DTT, TCEP), we were interested in performing a detailed mechanistic analysis of a GH109 α-NAGAL enzyme. Therefore, GH109 α-NAGAL from *Elizabethkingia meningosepticum* (ATCC 51720D) was cloned and recombinantly expressed in *E.Coli*. To confirm the reported cofactor requirements for activity of GH109 α-NAGAL,(*1*) we measured the rate of α-NAGAL-catalyzed hydrolysis as a function of NAD^+ (25-300 μM), Mn^{2+} (25-300 μM) and reducing agent TCEP (1-100 μM) concentrations. Of note, GH109 α-NAGAL did not show any changes in its catalyzed rate of reaction on addition of exogenous NAD^+, TCEP or Mn^{2+} to the assay solution. These observations are similar to those reported by Liu *et al*.(*1*) The lack of rate enhancement on addition of Mn^{2+} is not surprising as no metal binding site was apparent in the active site structure of GH109 α-NAGAL from *E. meningosepticum* (ATCC 13253). In 1980, Aminoff and coworkers reported the isolation and purification of α-NAGAL from *Clostridium perfringens*, the only other bacterial α-NAGAL characterization reported to date.(*21*) Interestingly, the α-NAGAL from *C. perfringens* showed activity only in the presence of reducing agent (DTT, dithiothreitol in this case).(*21*) The authors supported their conclusion that a thiol group (-SH) is important for enzymatic activity by showing that addition of mercury resulted in complete abolition of α-NAGAL activity.(*21*) In contrast, our observation is that the activity of GH109 α-NAGAL from *E. meningosepticum*, is invariant upon addition of a reducing agent, such as TCEP (1-100 μM). As a result, we decided to measure all of the kinetic parameters in buffered solutions that are devoid of either Mn^{2+} or TCEP.

4.4.1. Product Studies.

Most glycosyl hydrolases are highly specific with regard to the identity of the sugar residue that binds into the catalytic site. When α-NAGAL-catalyzed hydrolysis of PNPαGalNAc was performed in the presence of methanol (5 M) the formation of methyl 2-acetamido-2-deoxy-α-D-galactopyranoside was inferred based on the appearance of a new doublet, at a chemical shift δ 5.22 ($J_{1,2}$ = 3.8 Hz), in the ^1H NMR spectrum, a

resonance that was assigned to the anomeric proton (**Figure 4-4**). Consequently, the stereochemical outcome of GH109 α-NAGAL catalyzed reactions was deduced to be retention of configuration. Also of note, when the α-NAGAL-catalyzed hydrolysis PNPαGalNAc was performed in D_2O, complete deuteration at C-2 of the reaction product, 2-acetamido-2-deoxy-D-galactopyranose, was observed (**Figure 4-5**). That is, exchange of the C-2 proton occurs concurrently with the enzymatic hydrolysis of the substrate PNPαGalNAc, an observation that requires C2-H2 bond cleavage to be a part of the catalytic cycle. Therefore, it can be concluded that these results are inconsistent with a standard retaining glycosidase double-displacement mechanism. Of note, the X-ray crystal structure of α-NAGAL from *E. meningosepticum* when co-crystallized with a GalNAc substrate revealed that the C-3 of the GalNAc moiety is located about 3 Å from the C-4 atom of the nicotinamide ring of NAD^+.(*1*) The other hydrolase family whose X-ray crystal structure possesses a bound NAD^+ in the active site belong to GH4.(*23*) Further, detailed mechanistic analysis for GH4 glycosidases(*24, 25*) have shown that this hydrolase family functions via an oxidation-reduction mechanism, which is unlike classical glycosyl hydrolases that function by either a double-displacement (retaining) or a single displacement (inverting) mechanism.(*35*) Furthermore, GH4 glycosidases utilize the NAD^+ cofactor to oxidize the C-3 hydroxyl group of the substrate to give a ketone, and as a result the C-2 proton is activated for deprotonation, possibly by a suitably positioned tyrosine residue (Figure 1-17, Introduction).(*26, 27*) In summary, the observations of C-2 proton-deuterium exchange in the hydrolysis of PNPαGalNAc by α-NAGAL from *E.meningosepticum* and the presence of a bound NAD^+ in the enzyme active site(*1*) are consistent with mechanistic parallelism occurring in GH4 glycosidases and GH109 α-NAGAL families.

4.4.2. pH Profile of α-NAGAL-Catalyzed Hydrolysis.

The dependence of enzymatic activity as a function of the buffer pH can provide useful information about active site ionizable groups. Although a neutral pH optimum for α-NAGAL from *E. meningosepticum* has been reported,(*1*) to our knowledge, a complete pH-rate profile for the GH109 α-NAGAL-catalyzed reactions has not been published. Therefore, a detailed pH rate profile (pH range 6.3–8.8) for the GH109 α-NAGAL-catalyzed

hydrolysis of PNPαGalNAc was measured. In these experiments, the maximal enzyme activity was noted around a pH value of 7.2–7.5, therefore all the enzymatic assays were performed at pH 7.5. The pH–rate profile (**Figure 4-6**) for the *E. meningosepticum* GH109 α-NAGAL-catalyzed hydrolysis of PNPαGalNAc exhibited a bell-shape and could be fit to a modified pH-rate equation that incorporated Hill coefficients for the ionization events. In theory, a bell-shaped curve for pH profile of k_{cat} is commonly interpreted to be indicative of two ionisable groups on the enzyme, of which one is protonated and the other deprotonated, in the catalytic cycle.(*36*) However cooperative behaviour as that observed can be due to several complicating factors such as, enzyme and or cofactor stability. At the present time, it is therefore prudent not to over-interpret the measured pH profiles (**Figure 4-6**).(*36*) Several possible conclusions has been stated in literature regarding the variation of k_{cat} and of k_{cat}/K_m with pH.(*37*) According to which, plots of k_{cat}/K_m vs pH yields the pK_a values for free enzyme/ or free substrate and the corresponding plots for variation of k_{cat} with pH provide the pK_a values of the enzyme-substrate complex.(*38*) The two apparent pK_a values that characterize the enzyme (GH109 α-NAGAL) are 7.0 and 7.9. In summary, incorporating the conclusions from the present data, the pH-rate profile for GH109 α-NAGAL (*E. meningosepticum*, ATCC 51720D) shows that the enzyme has evolved to exhibit a peak of activity in a narrow pH range of 7.2-7.5 (**Figure 4-6**).

4.4.3. Substrate Specificity.

Liu *et al.* showed that α-NAGAL from *E. meningosepticum* (ATCC 13253) is maximally active at a pH value of 7.5, and that it exhibits a relative high specificity for the blood group A tetrasaccharide as substrate (0.25 U/mg) labelled with AMC (7-amino-4-methylcoumarin) although the enzyme is more efficient towards the activated substrate PNPαGalNAc (> 10 U/mg).(*1*) These authors also reported some activity with PNPβGalNAc and the 2-OH analogues (α and β PNPGal) as substrates. Of note, the apparent K_m value for PNPαGalNAc has been reported to be similar to that of PNPβGalNAc, although k_{cat} values differ more than 1,000-fold, the former being hydrolyzed more readily.(*1*) In contrast, a stringent substrate specificity for PNPαGalNAc was reported by Levy *et al.* for α-NAGAL from *Clostridium perfringens*.(*21*) Our enzyme GH109

α-NAGAL from *E.meningosepticum* (ATCC 51720D) showed high activity with PNPαGalNAc and low but measureable activity with PNPβGalNAc. As for the 2-OH substrates, both PNPαGal and PNPβGal (data not shown) was hydrolyzed by our α-NAGAL, the former being hydrolysed more efficiently. Notably, the catalytic parameters obtained for PNPαGal (k_{cat} = 2.43 s^{-1}, k_{cat}/K_M = 6.75 × 10^3 M^{-1} s^{-1}) with GH109 α-NAGAL were comparable to those measured using PNPαGalNAc as substrate (k_{cat} = 3.08 s^{-1}, k_{cat}/K_m = 9.13 × 10^4 M^{-1} s^{-1}). These observations are consistent with substrates that possess a hydroxyl group at C-2 in place of the natural 2-acetamido substituent fitting into the enzymatic pocket of α-NAGAL and as a result they undergo α-NAGAL catalyzed hydrolysis.

4.4.4. Brønsted Plots.

Kinetic analysis of enzymatic reactions using unnatural substrates that bear different leaving groups is one of the more common strategies for mechanistic characterization of glycosyl hydrolases. In typical retaining glycosyl hydrolases, the aglycone is cleaved during the first chemical step of the reaction, so the rate of this step (glycosylation) is influenced by leaving group ability, such correlations of rate and leaving group ability are known as Brønsted plots. In the case of GH4 glycosidases it has been demonstrated, on the basis of Brønsted plots and α-secondary deuterium kinetic isotope effects, that the rate of the glycosidic bond cleavage is not kinetically significant for either k_{cat} or k_{cat}/K_m (**Figure 1-17**, Introduction, Chapter 1) for the enzymes from *T. maritima* BglT 6-phospho-β-glucosidase, *B. subtilis* GlvA 6-phospho-α-glucosidase and *C. freundii* α-galactosidase.(*24*, *25*, *27*) We decided to evaluate whether the catalytic parameters (k_{cat} and k_{cat}/K_m) for GH109 α-NAGAL from *E. meningosepticum* depend on the pK_a of the conjugate acid of the leaving group phenol in order probe the mechanistic similarities between GH4 and GH109 glycosyl hydrolases. Usually, such experiments involve making a series of aryl glycosides and plotting the logarithm of the measured rate constant against the pK_a of the conjugate acid of the leaving group; the slope of such a graph (Brønsted parameter β_{lg}) can give mechanistic information on the enzymatic mechanism. Taking into consideration the comparable efficiency demonstrated by GH109 α-NAGAL from *E. meningosepticum* (ATCC 51720D) in

hydrolyzing PNPαGal (k_{cat} = 2.43 s^{-1}, k_{cat}/K_M = 6.75 × 10^3 $M^{-1}s^{-1}$), we decided to perform a full Brønsted analysis on a series of α-GAL substrates as well. To this end, we synthesized a series of aryl αGalNAc substrates (**Scheme 4-1**) with varying leaving groups(*28, 29*) (pK_a range of the conjugate acid 7.15-10.21) and an aryl αGal substrate series (Chapter 2), the pK_a values of the conjugate acid of leaving groups of which range from 7.15-10.19. The GH109 α-NAGAL (*E. meningosepticum*) catalyzed-rate of hydrolyses of both of these activated substrates (aryl αGal and aryl αGalNAc) were measured at a pH of 7.5 and at 37 °C. The kinetic parameters (k_{cat} and k_{cat}/K_m) determined for each substrate are summarized in **Table 4-2** and **Table 4-3**. The kinetic data for both αGalNAc and αGal substrates are remarkably similar for the complete series of aryl substrates where pK_a of the conjugate acid of the aglycone varied by approximately 3 pK_a units (**Figure 4-7** and **Figure 4-8**). The derived β_{lg} values on *V* and *V*/*K* for GH109 α-NAGAL with αGalNAc substrates are -0.08 ± 0.06 and -0.31 ± 0.12, respectively. The corresponding values for αGal substrates are 0.01 ± 0.10 and 0.15 ± 0.09, for *V* and *V*/*K*, respectively. However, if not for the 4-nitrophenyl substrates the β_{lg} values for both series (αGalNAc and αGal), on k_{cat} and k_{cat}/K_m would be indistinguishable from zero (**Figure 4-7** and **Figure 4-8**). Therefore, it is likely that cleavage of the anomeric C–O bond is not a kinetically significant process for the GH109 α-NAGAL catalyzed hydrolysis of either aryl α-D-galactopyranosides or aryl 2-acetamido-2-deoxy-α-D-galactopyranosides with the less activated leaving groups. With regard to the possible closeness in catalytic mechanism of GH109 and GH4, the simplest and most likely interpretation of the Brønsted analysis would be that the elimination of aglycone is relatively fast and non-rate limiting as compared to other catalytic steps.

4.4.5. Deuterium Kinetic Isotope Effects.

In order to delineate the important transition states and the rate-limiting steps for GH109 α-NAGAL catalyzed reaction, it was necessary to measure the kinetic isotope effects (KIE) at positions C-1, C-2 and C-3 of aryl αGalNAc substrates. In order to improve the precision of our KIE measurements from those reported in chapter 2 using labelled phenyl α-D-galactosides as substrates, we decided to incorporate one significant change in the synthesis of deuterium isotopologues of αGalNAc series. Specifically, we decided to change

the aglycone from phenyl to 2-naphthyl because we would then be able to use fluorescence rather than absorbance spectroscopy for the kinetic measurements, a situation that should result in more sensitive assays for the reaction of labelled substrates. However, the syntheses of the deuterium labelled isotopologues of aryl αGalNAc substrates are more complicated because of the added synthetic complication of incorporation of an *N*-acetamido group on C-2. The synthesis of 2-naphthyl 2-acetamido-2-deoxy-α-D-(2-^{2}H)galactopyranoside **4.8** was accomplished, albeit in a low yield, by the appropriate choice of reducing agent for the reaction of oxime **4.5** that was made in 2 steps from tri-*O*-acetal D-galactal (3,4,6-tri-*O*-acetyl-1,5-anhydro-2-deoxy-D-*lyxo*-hex-1-enitol **4.3** (**Scheme 4-2**). Indeed, reduction of 2-naphthyl 3,4,6-tri-*O*-acetyl-2-deoxy-2-oximino-α-D-*lyxo*-hexopyranoside **4.5** proved to be the most challenging step in the synthesis. A wide range of reducing reagents including: a) BD_3-THF;(*39*) and b) $NaBD_4$,(*40*) were examined prior to selection of lithium aluminum deuteride ($LiAlD_4$)(*31*) as the best suitable reducing reagent. The duration of the $LiAlD_4$ (10 equiv.) reduction reaction in boiling THF was critical, that is a prolonged reaction times (3 h), led to complete decomposition of the starting material and product.

As the GH109 α-NAGAL (*E. meningosepticum*) enzyme showed comparable catalytic efficiencies with the αGal substrates we decided to take advantage of this diverse substrate tolerance by measuring KIEs using our previously synthesized phenyl α-D-galactopyranoside isotopologues as a convenient tool to gain mechanistic insight into the kinetically important steps for this enzyme. As such, the various isotopologues of αGal substrates, which possess the weakly activated leaving group (phenyl, pK_a of the conjugate acid phenol = 9.99) were employed for the KIE studies (Chapter 2).

For all reactions, kinetic isotope effects can be primarily divided into two types: a) primary kinetic isotope effects (1° KIE); and 2) secondary kinetic isotope effects (2° KIE). Primary isotope effects are associated with bond making or breaking to the isotopic label, and they are therefore often large in magnitude. Whereas, secondary kinetic isotope effects involve changes in the strength of bonding to the isotopic atom, between the ground and transition states in the reaction. As a consequence, these effects are generally smaller in magnitude.(*34, 41*) The rate-limiting step(s) for the enzymatic efficiency (*V*/*K*) can be any

step up to and including the first irreversible step, which for glycosyl hydrolases under initial rate conditions, is often cleavage of the glycosidic bond. In contrast, for the enzymatic turnover (V) the kinetically significant step(s) is the highest on the reaction free energy profile from the Michaelis complex, which has the highest concentration at saturating concentration of substrate, to the separated products.(*34*) The measurement of isotope effects on the maximal enzymatic velocity V (i.e ^{D}V) can only be accomplished by a direct comparison of individually measured rates of unlabelled and labelled substrates at saturating levels.(*34*) However, purity of both the labelled and unlabelled substrates (devoid of any competitive inhibitor) is extremely important in order to determine ^{D}V accurately. Since V/K is the first order rate-constant for reactions at low substrate levels, the precision in measuring $^{D}V/K$ by following individual reactions can be maximized by keeping the substrate concentration significantly below the K_m, i.e $1/5 \times K_m$ or $1/10 \times K_m$. As such, typically fixed levels of substrates, below saturation, is usually employed for measuring $^{D}V/K$ for enzymatic hydrolysis. However, this approach also requires rigorously purified substrates. The measurement of ^{D}V can be more complex than analysis of those on V/K (i.e $^{D}V/K$), but it can yield useful mechanistic information not obtainable by V/K measurements.

In order to determine ^{D}V, with regard to the apparent dissociation constant obtained for the hydrolysis of phenyl α-D-galactoside by GH109 α-NAGAL (K_m = 42 μM), the substrate concentration for the KIE measurements (^{D}V) was chosen to be 200 μM ($5 \times K_m$). The measured ^{D}V values for deuteration at one of the C-1, C-2 and C-3 carbons are listed in **Table 4-4**. The magnitude of the ^{D}V effects measured at both the C-2 ($^{2D}V = 2.36 \pm 0.18$) and C-3 ($^{3D}V = 2.12 \pm 0.18$) centres, suggest that both oxidation at C-3 and proton-abstraction at C-2 are kinetically significant for GH109 α-NAGAL catalyzed hydrolysis. The measured KIEs on ^{D}V for GH109 α-NAGAL from *E. meningosepticum* is larger than the KIE reported for GH4 α-galactosidase (*C. freundii*) catalyzed hydrolysis ($^{2D}V = 0.91 \pm 0.04$, $^{3D}V = 1.02 \pm 0.06$).(*25*) That is, the large primary KIEs on enzymatic turnover (^{2D}V and ^{3D}V) for GH109 α-NAGAL enzyme contrasts the values reported for the GH4 α-galactosidase, which require that Michael addition of water is likely highest energy step in the catalytic cycle for the α-galactosidase enzyme from *C. freundii*. Notably, we also reported that for GH4 α-galactosidase catalyzed hydrolysis of labelled phenyl αGal

substrates, both of these steps: i) oxidation of C-3 by hydride transfer to the onboard NAD^+ co-factor; and ii) abstraction of the acidic C-2 proton by the active site base, occur in a concerted fashion.(*25*) This conclusion was made on the basis of a comparison of KIE on $^{D}V/K$ for the mono- and di-deuterated isotopologues at positions C-2 and C-3.(*25*) During the hydride transfer from C-3, the C-2 proton becomes progressively more acidic by virtue of the increasing polarization at C-3, according to the general mechanism shown in **Figure 1-17** (Introduction).(*26, 27*) Although both these steps have been demonstrated to be kinetically significant for the activity of GH4 6-phospho-α-glucosidase from *B. subtilis* and 6-phospho-β-glucosidase from *T. maritima*, these authors assumed that the steps occurred stepwise rather than in a single concerted step.(*24, 27*)

Nevertheless, the observed large primary KIE on both C-3 and C-2 raises an important question regarding GH109 α-NAGAL-catalyzed reaction, do these effects result from one concerted step or two separate kinetically significant steps.

To summarize the key obervations for this chapter, our enzyme GH109 α-NAGAL from *E. meningosepticum* showed comparable efficiencies in hydrolysing αGalNAc and αGal substrates, an observation which could result from a lack of stabilization of the enolate occurring directly by coordination of an electropositive metal to the NHAc group of the αGalNAc substrates. Of note, GH4 hydrolases, possesses a divalent metal ion (Mn^{2+}) binding site in the active site which results in specific stabilization of the enolate by chelation.(*23*) The inferences drawn from the deuterium KIE studies performed on both GH109 α-NAGAL (*E. meningosepticum*) and GH4 α-galactosidase (*C. freundii*) manifests an important disparity between the mechanistic features of the two families.(*25*) The high KIE on ^{D}V obtained for C-2 and C-3 for GH109 α-NAGAL catalysed hydrolysis demonstrates that the rate-determining step for k_{cat} for this enzyme is different to that of GH4 α-galactosidase. Moreover, the flat Brønsteds obtained for both αGalNAc and αGal substrates suggest that neither V/K nor V is influenced by the rate of departure of the leaving group. If C1-O1 bond cleavage is not rate-limiting, then that step would not be associated with significant rehybridization at the anomeric centre. In other words, there should not be a large ^{D}V KIE observed on C-1. However, the measured KIE on C-1 ($^{1D}V = 1.47 \pm 0.10$)

contradicts the expectations that were based on Brønsted plots. Nevertheless, the large errors and other KIEs associated with this ^{1D}V value are a result of the small absorbance changes observed for the leaving group phenol and also the possibility of minor impurities in the substrate, require that all conclusions at the present time must be considered as preliminary.

The positive derivation of PNPαGalNAc in the Brønsted plot could indicate the possibility of a change in mechanism from $E1_{CB}$ (followed commonly by GH4 hydrolases) to E2 for GH109 α-NAGAL catalysed reactions of activated substrates. However, in order to comment on the mechanism confidently, we would need to synthesize substrates that possess even better leaving groups, for example, 2,4-dinitrophenolate (pK_a of the conjugate acid 2,4-dinitrophenol is 4.1). Furthermore, to be able to comment on the likelihood of the two rate-limiting steps being stepwise or concerted, we need to measure the KIE on ^{D}V for the (2-^{2}H, 3-^{2}H) di-deuterated isotopologue ($^{2D.\ 3D}V$) in comparison to the mono-deuterated ones (^{2D}V and ^{3D}V). Nonetheless, a more sensitive assay needs to be developed before definitive mechanistic conclusions can be made for this family of hydrolase.

4.5. References

1. Liu, Q. P., Sulzenbacher, G., Yuan, H., Bennett, E. P., Pietz, G., Saunders, K., Spence, J., Nudelman, E., Levery, S. B., White, T., Neveu, J. M., Lane, W. S., Bourne, Y., Olsson, M. L., Henrissat, B., and Clausen, H. (2007) Bacterial glycosidases for the production of universal red blood cells, *Nat Biotechnol 25*, 454-464.

2. Comfort, D. A., Bobrov, K. S., Ivanen, D. R., Shabalin, K. A., Harris, J. M., Kulminskaya, A. A., Brumer, H., and Kelly, R. M. (2007) Biochemical analysis of Thermotoga maritima GH36 alpha-galactosidase (*TmGalA*) confirms the mechanistic commonality of clan GH-D glycoside hydrolases, *Biochemistry 46*, 3319-3330.

3. Garman, S. C., Hannick, L., Zhu, A., and Garboczi, D. N. (2002) The 1.9 Å structure of alpha-*N*-acetylgalactosaminidase: molecular basis of glycosidase deficiency diseases, *Structure 10*, 425-434.

4. Coutinho, P. M., Deleury, E., Davies, G. J., and Henrissat, B. (2003) An evolving hierarchical family classification for glycosyltransferases, *J Mol Biol 328*, 307-317.

5. Garman, S. C., and Garboczi, D. N. (2002) Structural basis of Fabry disease, *Mol Genet Metab 77*, 3-11.

6. Clark, N. E., and Garman, S. C. (2009) The 1.9 Å structure of human alpha-*N*-acetylgalactosaminidase: The molecular basis of Schindler and Kanzaki diseases, *J Mol Biol 393*, 435-447.

7. Brockhausen, I. (2006) Mucin-type *O*-glycans in human colon and breast cancer: glycodynamics and functions, *EMBO Rep 7*, 599-604.

8. Weissmann, B., and Hinrichsen, D. F. (1969) Mammalian alpha-acetylgalactosaminidase. Occurrence, partial purification, and action on linkages in submaxillary mucins, *Biochemistry 8*, 2034-2043.

9. Hoskins, L. C., and Boulding, E. T. (2001) Changes in immunologic properties of group A RBCs during treatment with an A-degrading *exo*-alpha-*N*-acetylgalactosaminidase, *Transfusion 41*, 908-916.

10. Wiegandt, H. (1995) The chemical constitution of gangliosides of the vertebrate nervous system, *Behav Brain Res 66*, 85-97.

11. Garman, S. C., and Garboczi, D. N. (2004) The molecular defect leading to Fabry disease: structure of human alpha-galactosidase, *J Mol Biol 337*, 319-335.

12. Hoskins, L. C., Larson, G., and Naff, G. B. (1995) Blood group A immunodeterminants on human red cells differ in biologic activity and sensitivity to alpha-*N*-acetylgalactosaminidase, *Transfusion 35*, 813-821.

13. van Diggelen, O. P., Schindler, D., Willemsen, R., Boer, M., Kleijer, W. J., Huijmans, J. G., Blom, W., and Galjaard, H. (1988) alpha-*N*-acetylgalactosaminidase deficiency, a new lysosomal storage disorder, *J Inherit Metab Dis 11*, 349-357.

14. Kanzaki, T., Wang, A. M., and Desnick, R. J. (1991) Lysosomal alpha-*N*-acetylgalactosaminidase deficiency, the enzymatic defect in angiokeratoma corporis diffusum with glycopeptiduria, *J Clin Invest 88*, 707-711.

15. Olsson, M. L., Hill, C. A., de la Vega, H., Liu, Q. P., Stroud, M. R., Valdinocci, J., Moon, S., Clausen, H., and Kruskall, M. S. (2004) Universal red blood cells--enzymatic conversion of blood group A and B antigens, *Transfus Clin Biol 11*, 33-39.

16. Watkins, W. M. (1980) Biochemistry and Genetics of the ABO, Lewis, and P blood group systems, *Adv Hum Genet 10*, 1-136, 379-185.

17. Goldstein, J., Siviglia, G., Hurst, R., Lenny, L., and Reich, L. (1982) Group B erythrocytes enzymatically converted to group O survive normally in A, B, and O individuals, *Science*, *215*, 168-170.

18. Sung, S. S., and Sweeley, C. C. (1980) Purification and partial characterization of porcine liver alpha-*N*-acetylgalactosaminidase, *J Biol Chem 255*, 6589-6594.

19. Tuppy, H., and Staudenbauer, W. L. (1966) The action on soluble blood group A substances of an alpha-*N*-acetylgalactosaminidase from Helix pomatia, *Biochemistry 5*, 1742-1747.

20. Harun-Or-Rashid, M., Matsuzawa, T., Satoh, Y., Shiraishi, T., Ando, M., Sadik, G., and Uda, Y. Purification and characterization of alpha-*N*-acetylgalactosaminidases I and II from the starfish Asterina amurensis, *Biosci Biotechnol Biochem 74*, 256-261.

21. Levy, G. N., and Aminoff, D. (1980) Purification and properties of alpha-*N*-acetylgalactosaminidase from *Clostridium perfringens*, *J Biol Chem 255*, 11737-11742.

22. Hsieh, H. Y., Chapman, L. F., Calcutt, M. J., Mitra, M., and Smith, D. S. (2005) Recombinant *Clostridium perfringens* alpha-*N*-acetylgalactosaminidase blood group A2 degrading activity, *Artif Cells Blood Substit Immobil Biotechnol 33*, 187-199.

23. Rajan, S. S., Yang, X., Collart, F., Yip, V. L., Withers, S. G., Varrot, A., Thompson, J., Davies, G. J., and Anderson, W. F. (2004) Novel catalytic mechanism of glycoside hydrolysis based on the structure of an NAD^+/Mn^{2+}-dependent phospho-alpha-glucosidase from Bacillus subtilis, *Structure 12*, 1619-1629.

24. Yip, V. L., Varrot, A., Davies, G. J., Rajan, S. S., Yang, X., Thompson, J., Anderson, W. F., and Withers, S. G. (2004) An unusual mechanism of glycoside hydrolysis involving redox and elimination steps by a family 4 beta-glycosidase from *Thermotoga maritima*, *J Am Chem Soc 126*, 8354-8355.

25. Chakladar, S., Cheng, L., Choi, M., Liu, J., and Bennet, A. J. Mechanistic evaluation of MelA alpha-galactosidase from *Citrobacter freundii*: a family 4 glycosyl hydrolase in which oxidation is rate-limiting, *Biochemistry 50*, 4298-4308.

26. Yip, V. L., and Withers, S. G. (2006) Mechanistic analysis of the unusual redox-elimination sequence employed by *Thermotoga maritima* BglT: a 6-phospho-beta-glucosidase from glycoside hydrolase family 4, *Biochemistry 45*, 571-580.

27. Yip, V. L., Thompson, J., and Withers, S. G. (2007) Mechanism of GlvA from *Bacillus subtilis*: a detailed kinetic analysis of a 6-phospho-alpha-glucosidase from glycoside hydrolase family 4, *Biochemistry 46*, 9840-9852.

28. Nishiyama, T., Ichikawa, Y., and Isobe, M. (2004) Glycocinnasperimicin D Synthetic Studies: Synthesis of Cinnamoyl Glycosides via Iodination-Heck Reaction Sequence Starting from Phenyl Glycosides, *SYNLETT 1*, 89-92.

29. Jacquemard, U., Be´ne´teau, V., Lefoix, M., Routier, S., Me´rour, J. Y., and Coudert, G. (2004) Mild and selective deprotection of carbamates with Bu_4NF, *Tetrahedron 60*, 10039-10047.

30. Lemieux, R. U., Nagabhushan, T. L., and Gunner, S. W. (1968) Synthesis of 3,4,6-tri-*O*-acetyl-2-oximino-a-D-hexopyranosides, *Can J Chem 46*, 405-411.

31. Smith, D. R., Maienthal, M., and Tipton, J. (1952) Reduction of oximes with lithium aluminium hydride, *Journal of Organic Chemistry 17*, 294-297.

32. Serjeant, E. P., and Dempsey, B. (1979) *Ionisation constants of organic acids in aqueous solution*, Pergamon Press, Oxford, New York.

33. Lemieux, R. U., Nagabhushan, T. L., and O'Neill, I. K. (1964) A new approach to the synthesis of 2-amino-2-deoxyglycosides, *Tetrahedron Lett*, 1909-1916.

34. Cleland, W. W. (1982) Use of isotope effects to elucidate enzyme mechanisms, *CRC Crit Rev Biochem 13*, 385-428.

35. Koshland, D. E., Jr. (1953) Stereochemistry and the mechanism of enzymatic reactions, *Biological Reviews 28*, 416-436.

36. Knowles, J. R. (1976) The intrinsic pK_a-values of functional groups in enzymes: improper deductions from the pH-dependence of steady-state parameters, *CRC Crit Rev Biochem 4*, 165-173.

37. Cornish-Bowden, A. (1976) Estimation of the Dissociation Constant of Enzyme-Substrate Complexes from Steady-State measurements., *Biochemical Journal 153*, 455-461.

38. Alberty, R. A., and Massey, V. (1954) On the interpretation of the pH variation of the maximum initial velocity of an enzyme-catalyzed reaction, *Biochim Biophys Acta 13*, 347-353.

39. Feuer, H., and Braunstein, D. M. (1969) Reduction of oximes, oxime ethers, and oxime esters with diborane. Novel synthesis of amines, *Journal of Organic Chemistry 34*, 1817-1821.

40. Hoffman, C., Tanke, R. S., and Miller, M. J. (1989) Synthesis of a-Amino Acids by Reduction of α-Oximino Esters with Titanium(III) Chloride and Sodium Borohydride, *Journal of Organic Chemistry 54*, 3750-3751.

41. Cleland, W. W. (1995) Isotope effects: determination of enzyme transition state structure, *Methods Enzymol 249*, 341-373.

4.6. Supporting Information.

^{1}H NMR datas for the all the synthesized substrates, both labelled and unlabelled, are provided in this section.

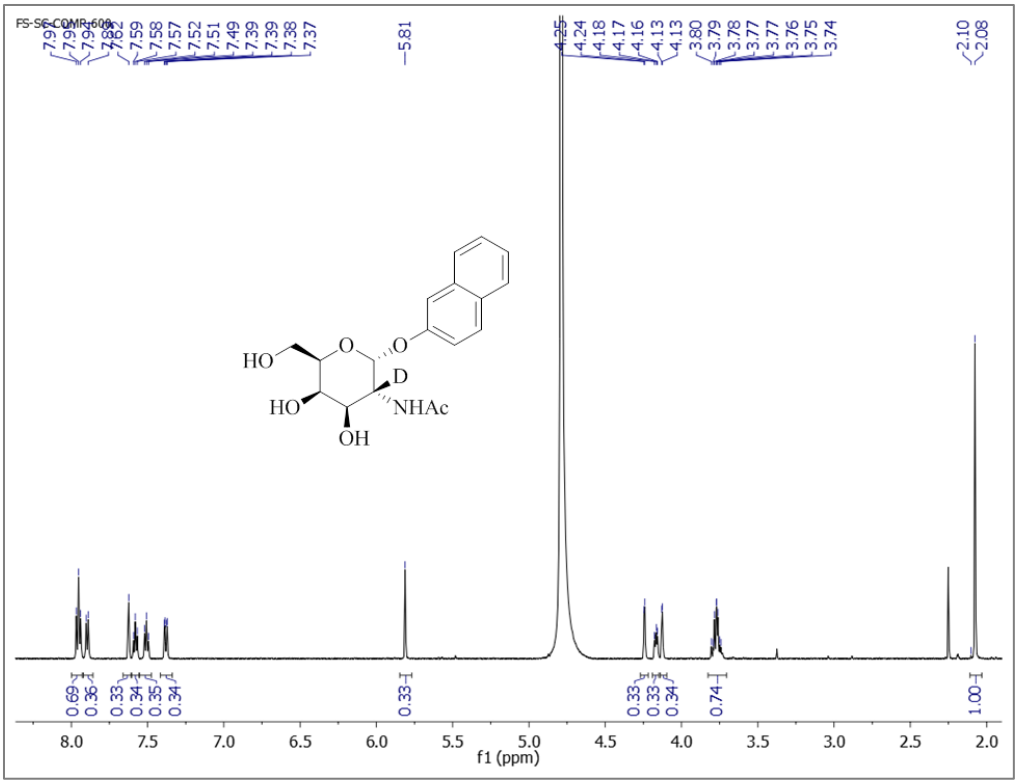


Figure 4S 1. 2-Naphthyl 2-acetamido-2-deoxy-α-D-(2-^{2}H)galactopyranoside (600 MHz, D_2O).

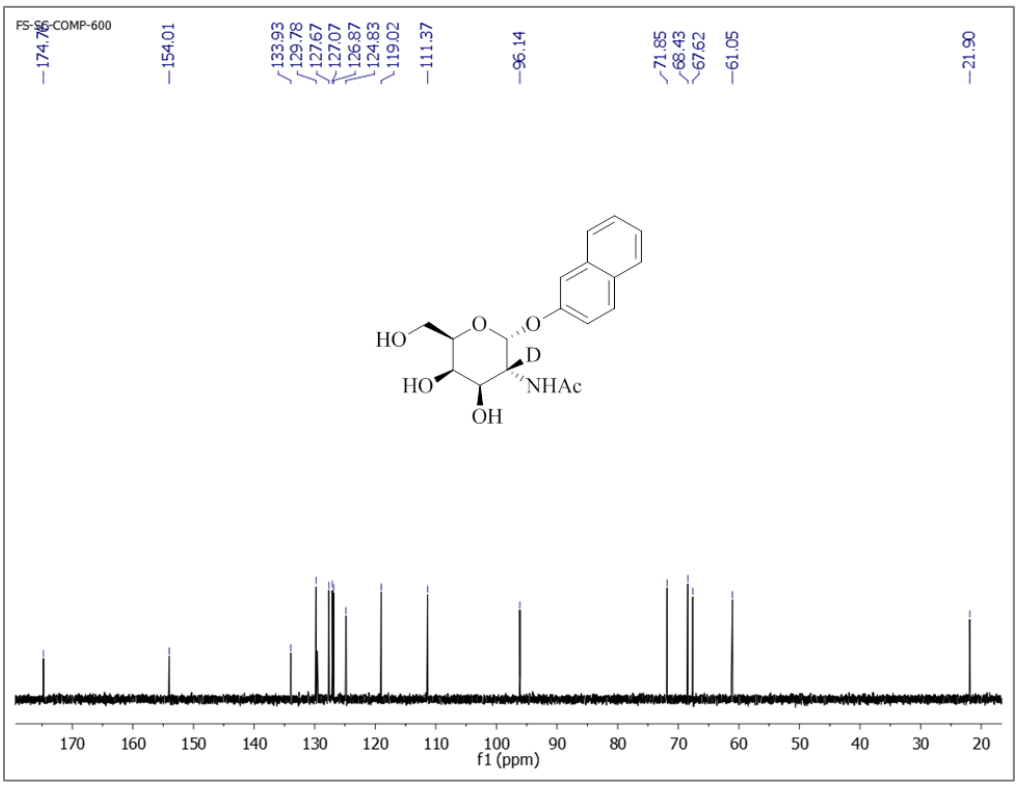


Figure 4S 2. 2-Naphthyl 2-acetamido-2-deoxy-α-D-(2-^{2}H)galactopyranoside (600 MHz, D_2O).

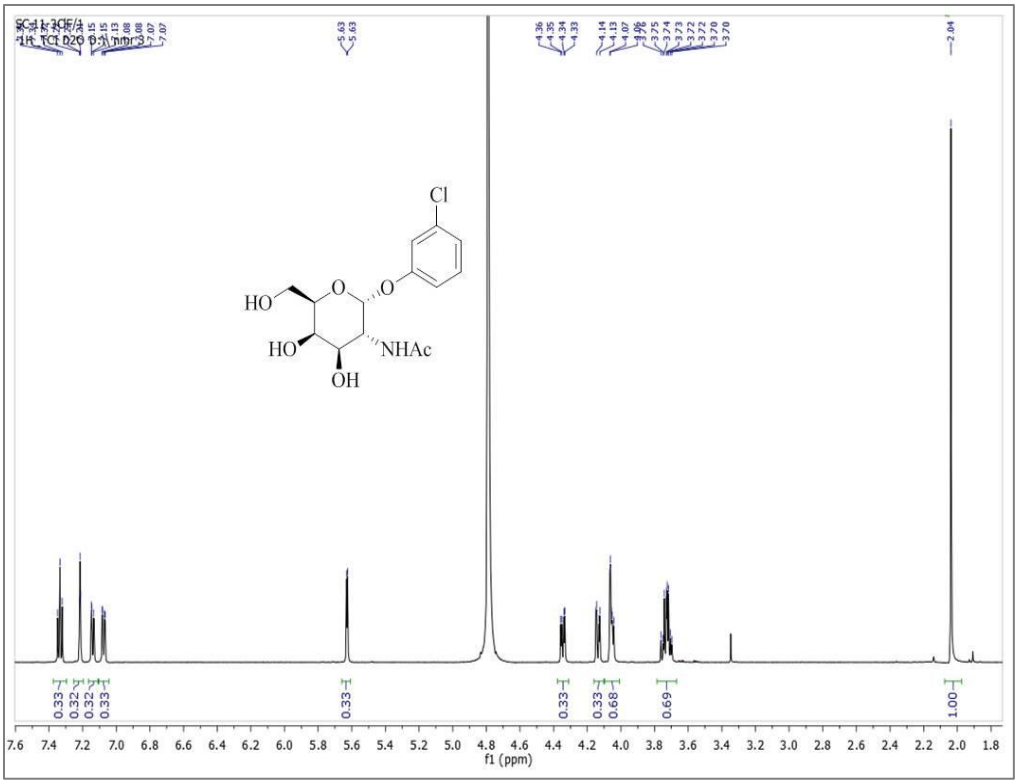


Figure 4S 3. 3-Chlorophenyl 2-acetamido-2-deoxy-α-D-galactopyranoside (600 MHz, D_2O).

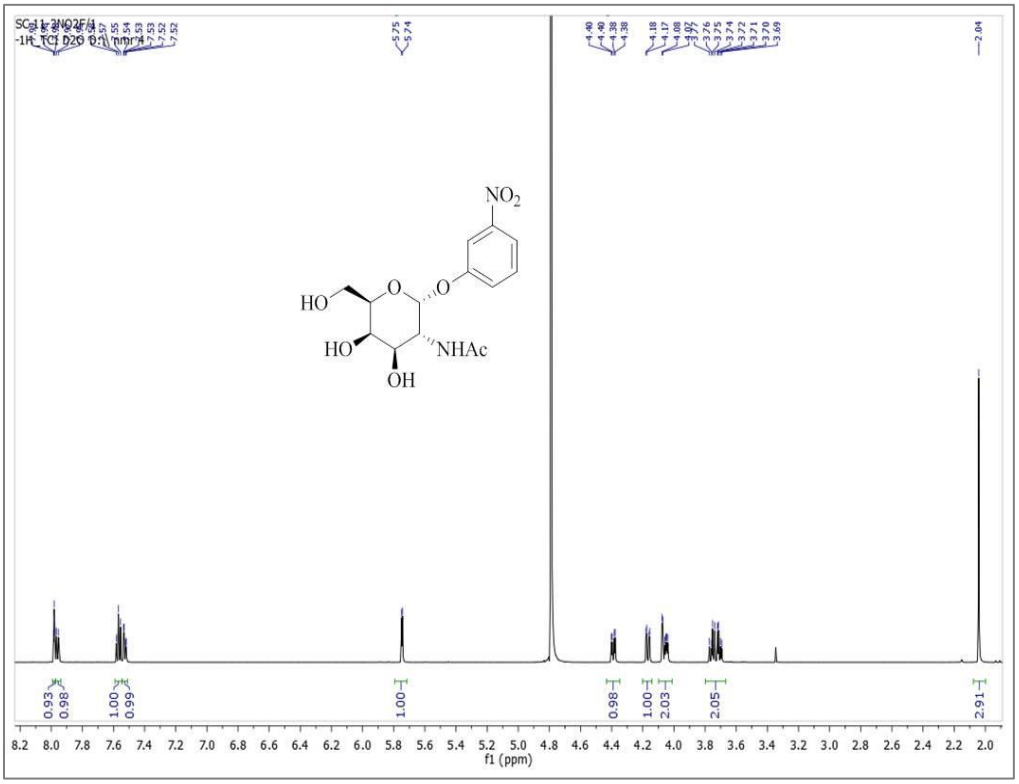


Figure 4S 4. 3-Nitrophenyl 2-acetamido-2-deoxy-α-D-galactopyranoside (600 MHz, D_2O).

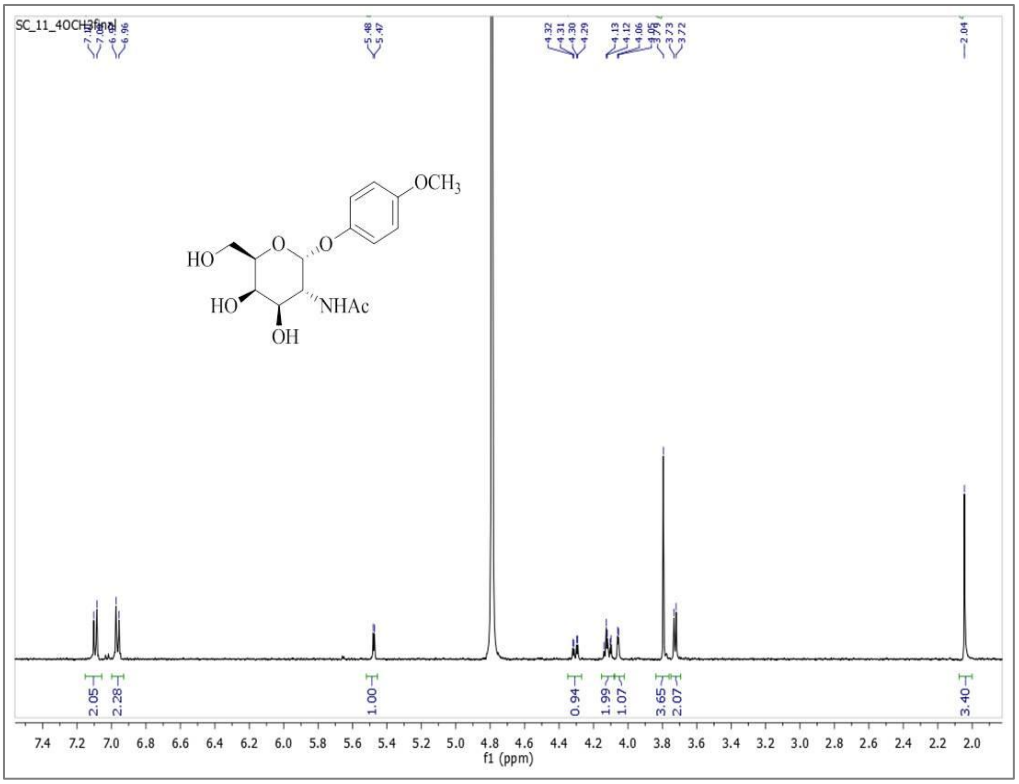


Figure 4S 5. 4-Methoxyphenyl 2-acetamido-2-deoxy-α-D-galactopyranoside (600 MHz, D_2O).

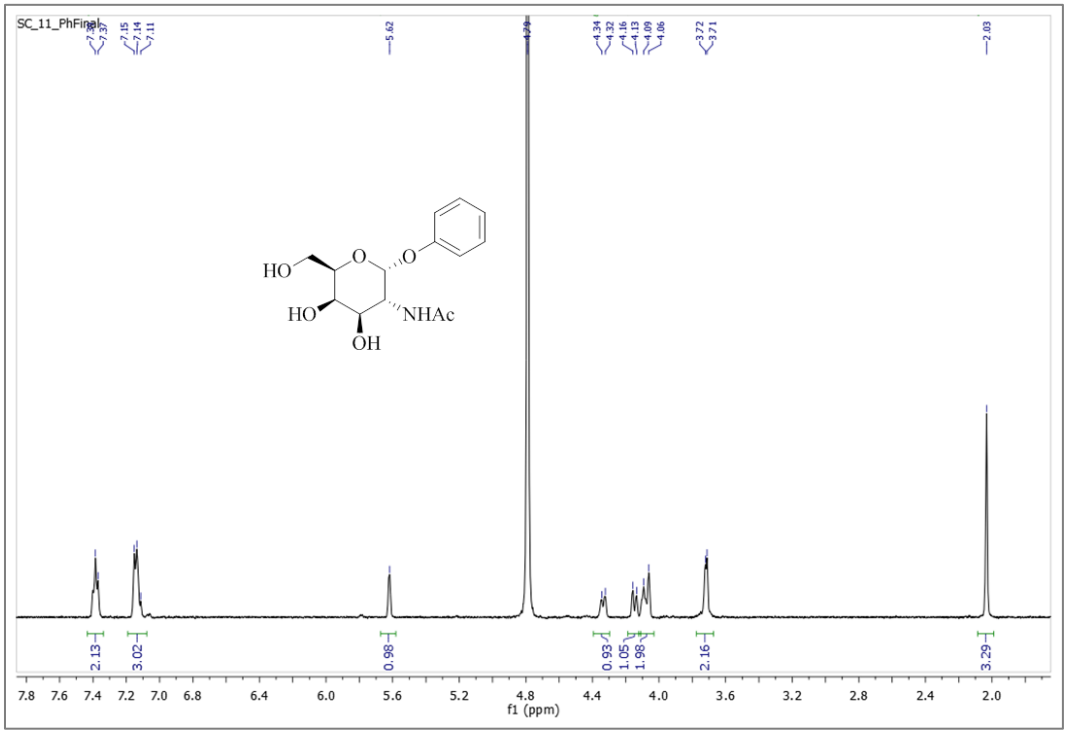


Figure 4S 6. Phenyl 2-acetamido-2-deoxy-α-D-galactopyranoside (600 MHz, D_2O).

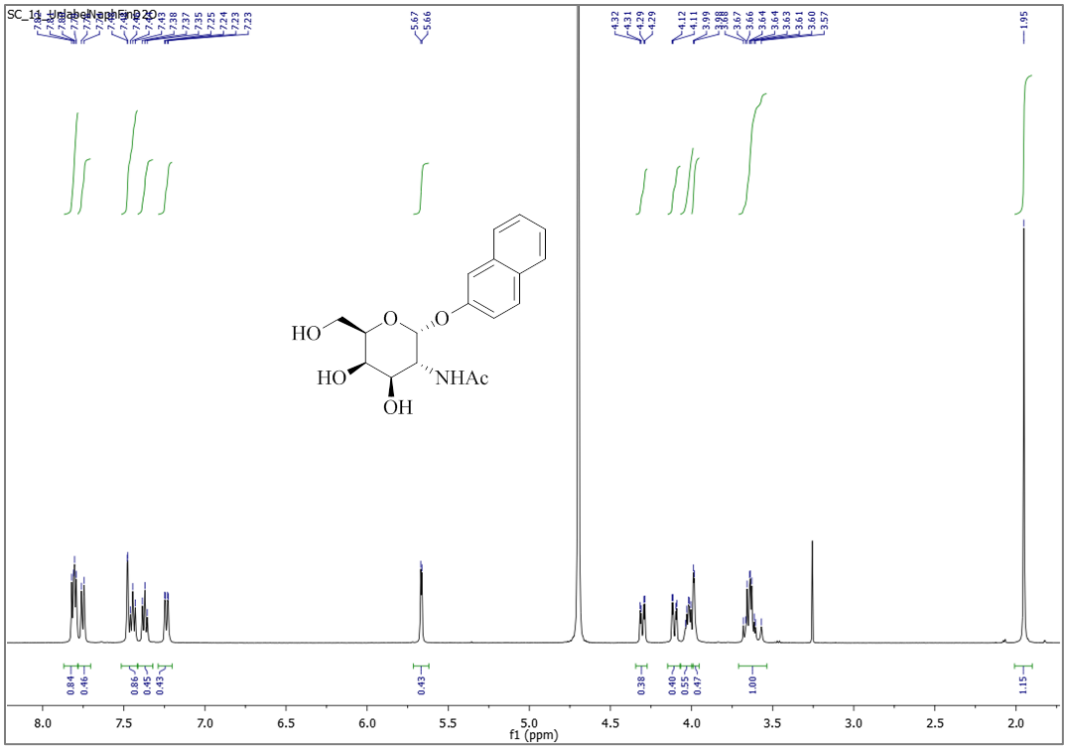


Figure 4S 7. 2-Naphthyl 2-acetamido-2-deoxy-α-D-galactopyranoside (600 MHz, D_2O).

5. Summary and Future Work

Glycosyl hydrolases belonging to GH4 and GH109, do not follow the "classical" mechanisms followed by glycosyl hydrolases, but instead utilize NAD^+ in a transient reduction/oxidation reaction with leaving group elimination. The similarity in the active site architectures of both of these NAD^+ dependent GH families, on the basis of their single crystal X-ray crystallographic structures, motivated me to undertake a detailed mechanistic elucidation of one enzyme from each family, in order to demonstrate a possible parallelism in their catalytic mechanisms. GH36 α-galactosidase, on the other hand, follows the "classical" retaining GH mechanism. With regard to this family, we desired to identify the key catalytic residues, through an elaborate mechanistic investigation using an irreversible inhibitor. In summary, the goal of my doctoral research work can be stated as follows:

To probe the catalytic mechanisms adopted by GH4 α-galactosidase from C. freundii and GH109 α-N-acetylgalactosaminidase from E. meningosepticum, on the basis of deuterium KIE and Brønsted plots that would enable us to establish a possible link between the two unusual families catalyzing through redox mechanism. Further, we developed an irreversible inhibitor of GH36 α-galactosidase, in order to gain insight into the catalytic mechanism followed by this enzyme, through identification of the key catalytic residues.

5.1. Mechanistic Evaluation of MelA α-Galactosidase from Citrobacter Freundii: A Family 4 Glycosyl Hydrolase in which Oxidation is Rate-Limiting

Prior to my thesis, an elaborate mechanistic elucidation of GH4 α-galactosidase was not known. To begin with, I chemically synthesized a series of α-Gal substrates in order to obtain the Brønsted plots. On the basis of these Brønsted plots, we were able to prove that for this α-galactosidase departure of the leaving group was not kinetically significant. Furthermore, with the goal of performing deuterium KIE studies for this enzyme, deuterium substrate isotopologues of phenyl α-D-galactoside at positions, 1-^{2}H, 2-^{2}H and 3-^{2}H were

synthesized. The magnitude of $^{D}V/K$ on both C-2 and C-3, implicated that both oxidation at C-3 and the subsequent proton removal at C-2 are kinetically significant steps in the catalytic cycle of GH4 α-galactosidase from *C. freundii*. Of note, I also synthesized the di-deuterated isotopologue (2-^{2}H, 3-^{2}H) in order to gain insight into the inter-dependence of the two rate-limiting steps. The measured KIEs ($^{2D,3D}V/K$) enabled us to propose that the two partially rate-determining steps on C-3 and C-2 were, in reality, concerted for this enzyme, as shown in **Figure 5-1**. To our knowledge, this is for the first time such in-depth mechanistic findings have been reported for this family of hydrolases.

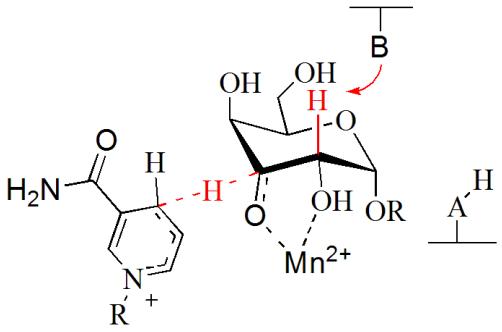


Figure 5-1. Proposed concerted mechanism for GH4 α-galactosidase.

5.1.1. Future Work

At present, the deuterium KIEs were determined using the non-competitive techniques. In order to improve upon the precision of the reported KIE results, the measurement of the KIEs based on competitive techniques instead would be desirable. In that way, both the labelled and unlabelled isotopologues could be present in the same reaction vessel (for example NMR tube). This could help in eliminating various systematic errors, like that of temperature variation, differences in concentration and presence of inhibitors or activators, which would otherwise be present in non-competitive techniques. With regards to an improvement in the deuterium isotopologues employed for the current study, we could change the leaving group phenol which possesses small absorbance changes, to that of 2-napththol. This would enable us to use fluorescence rather than absorbance spectroscopy, a situation that should result in more sensitive assay for the enzymatic hydrolysis of labelled substrates.

5.2. A Mechanism-Based Glycoside Inhibitor with a High Inhibitory Proficiency: Generation of a Cyclopropylmethyl Carbenium Ion in an Enzymatic Active Site

The reported crystal structure of GH36 α-galactosidase from *T. maritima*, motivated us to conduct the inhibition studies on this enzyme with our already synthesized carbocyclic irreversible inhibitors- *galacto*- analogue (**1**) (1*R*,2*S*,3*S*,4*R*,5*S*,6*S*)-5-(3,5-difluorophenoxy)-1-(hydroxymethyl)bicyclo[4.1.0]heptan-2,3,4-triol, *altro*- analogue ((1*S*,2*S*,3*S*,4*R*,5*S*,6*R*)-5-(3,5-difluorophenoxy)-1-(hydroxymethyl)-bicyclo[4.1.0]heptan-2,3,4-triol) (**2**) and the allyl analogue (1*R*,2*S*,3*S*,6*S*)-4-(hydroxymethyl)-6-(3,5-difluorophenoxy)-cyclohex-4-ene-1,2,3-triol) (**3**) (**Figure 5-2**).

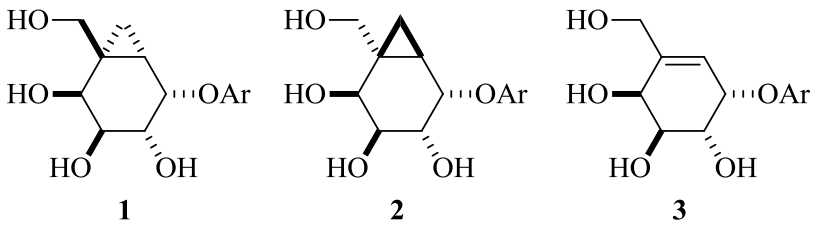


Figure 5-2. Compounds designed and synthesized as irreversible inhibitors of GH36 α-galactosidase.

Based on the results from the inhibitory assays, the *galacto* analogue proved to be an active-site directed irreversible inhibitor of GH36 α-galactosidase, whereas the *altro* analogue did not show any noticeable inhibition of enzymatic activity. Of note, the allyl analogue was kinetically determined to be a tight-binding competitive inhibitor for the same enzyme. We have further shown that the inactivation of GH36 α-galactosidase by the *galacto* inhibitor (**1**) and the corresponding reactivation, both were pH independent. Upon conducting the mass spectrometric studies on the alkylated enzyme, it was concluded that the inhibition results from the formation of a oxacarbenium ion in the enzymatic site that is trapped rapidly by an active site enzymatic residue which in this case was shown to be the general acid catalyst D387. Additionally, in order to estimate the enzymatic free energy stabilizations for the catalytic and the inhibitory TSs for this enzyme, three model compounds (**Figure 3**) were synthesized, and the activation parameters for their spontaneous hydrolysis at elevated temperatures were determined.

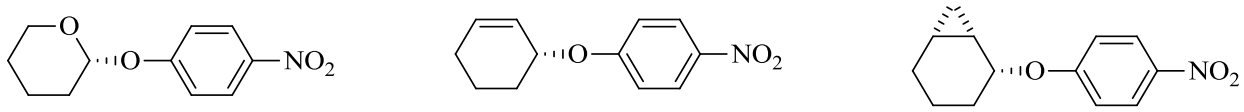


Figure 5-3. Model compounds synthesized to determine the catalytic proficiencies of inactivation by GH36 α-galactosidase.

5.2.1. Future Work

An improved synthetic strategy towards the synthesis of our key cyclohexenol intermediate would be ideal. A synthetic scheme comprising of fewer steps and convenient purification protocols, would enable us to design and synthesize comparatively more reactive inhibitors, for example, by changing the leaving group to that of 6,8-difluoro-4-methyl-methylumbelliferyl, instead of 3,5-difluorophenol.

5.3. Mechanistic Analysis of Family 109 α-N-acetylgalactosaminidase from Elizabethkingia Meningosepticum: A Glycosyl Hydrolase possessing Multiple Substrate specificity and Redox Mechanism

In order to prove the likelihood of an oxidation-reduction reaction mechanism followed by GH109 α-NAGAL, we decided to perform a detailed mechanistic study incorporating the following key experiments: 1) cloning of gene following by expression of the recombinant protein; 2) synthesis of substrates; 3) pH profile; 4) Brønsted plots; and lastly, 5) deuterium KIE studies. Despite the inherent synthetic difficulties associated with αGalNAc substrates, I was able to chemically synthesize a series of aryl αGalNAc substrates and carried out the associated Brønsted analysis. The resultant β_{lg} values ~ 0, for both k_{cat} and k_{cat}/K_m, demonstrated that the departure of the aglycone is indeed silent for GH109 α-NAGAL enzyme. The variation of k_{cat} with solution pH demonstrated a better fit to a modified bell-shaped curve incorporating Hill coefficients for the ionization events, with the pH optimum at 7.5. Importantly, the diverse substrate specificity of GH109 α-NAGAL enabled me to carry out the KIE studies on the previously synthesized phenylαGal deuterium substrate isotopologues. On the basis of the measured KIEs, I was able to demonstrate that the two kinetically significant steps are happening on positions C-2 and C-3, with associated KIE values of $^{2D}V = 2.36 \pm 0.18$ and $^{3D}V = 2.12 \pm 0.18$, respectively. These findings,

supposedly for the first time, present the experimental evidences for a possible mechanistic parity between the two NAD^+ dependent families of hydrolases-GH4 and GH109.

5.3.1. Future Work

The future endeavours in this project would include further experiments in order to establish the mechanism adopted by GH109 α-NAGAL persuasively. The primary requirement would be to synthesize deuterium labelled substrate isotopologues possessing better leaving groups other than phenol, for example-2,4-dinitrophenol, thereby improving the sensitivity and accuracy of the kinetic assays. We would also need to perform the KIE studies for the di-deuterated isotopologue of phenyl αGal ($^{2D,3D}V$) in comparison to the mono-deuterated isotopologues (^{2D}V and ^{3D}V), in order to gain insight into the interdependency of the two rate-limiting steps. Furthermore, in order to ascertain the reproducibility of the obtained deuterium KIEs on ^{D}V, similar isotope effect measurements at another saturating substrate concentration (> 200 μM), for example, 400 μM (10 × K_m), would be desirable.

6. Appendix

After this thesis had been written, the site of alkylation, for the inactivation of *Tm*GalA by the cyclopropyl *galacto* compound **3.4** (Chapter 3), was determined by mass spectrometry. Specifically, Mr. Thomas Clark reported to us that based on MS/MS data from tryptic digests of wild-type and inactivated enzyme. The alkylation site on GH36 α-galactosidase (*T. Maritima*) has been shown to be the general acid/base residue D 387 (IGP**D**TAPFWGEHIEDNG), as shown in the following mass spectroscopic data of the labelled peptide. We tentatively assign the mechanism of inactivation to be pathway a (**Scheme 3.3**, page 137)