MECHANISTIC STUDIES OF XANTHOMONAS MANIHOTIS β-GALACTOSIDASE AND THE DEVELOPMENT OF A RAPID TRANSGLYCOSYLATION SCREEN

by

Jan E. Blanchard B.Sc., M^cMaster University, 1996

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE

IN

THE FACULTY OF GRADUATE STUDIES DEPARTMENT OF CHEMISTRY

We accept this thesis as conforming to the required standard

March 2000 THE UNIVERSITY OF BRITISH COLUMBIA © Jan E. Blanchard 2000 In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of <u>Chemistry</u>

The University of British Columbia Vancouver, Canada

Date <u>April 05, 2000</u>

Abstract

The Xanthomonas manihotis β -galactosidase (BgaX) is a 66 kDa retaining glycosidase that hydrolyzes glycosidic bonds through a double displacement mechanism involving a covalent glycosyl-enzyme intermediate. Characterization of a recombinant form of this enzyme showed that it hydrolyzes pNPGal, DNPGal, pNPFuc, pNP- α -L-Ara, pNPGalNAc, pNPGlc and pNPXyl with kinetic parameters in the range of $k_{cat} = 0.0217$ -36 s⁻¹ and K_m = 0.050-4.3 mM. The mechanism based inactivator 2,4-dinitrophenyl 2-deoxy-2-fluorogalactopyranoside was shown to inhibit BgaX through the accumulation of a 2-deoxy-2-fluorogalactosyl-enzyme intermediate with the kinetic parameters $k_{inact} = 0.030 \pm 0.004 \text{ s}^{-1}$ and $K_i = 0.031 \pm 0.005 \text{ mM}$. The fluorogalactosyl-enzyme intermediate was long lived, with a half life of 40 hr. Peptic digestion of this labeled enzyme and analysis by HPLC/mass spectrometry allowed the elucidation of Glu²⁶⁰ as the catalytic nucleophile involved in the formation of the glycosyl-enzyme intermediate during catalysis.

Retaining glycosidases are capable of catalysing the formation of glycosidic bonds through transglycosylation to an acceptor bound in the aglycone site of the enzyme. The second part of this study involved the development of a strategy to rapidly screen compounds for their potential as acceptors in transglycosylation reactions. This methodology was based on the premise that the reactivation, or turnover, of a glycosidase trapped as a 2-deoxy-2-fluoroglycosyl-enzyme is accelerated in the presence of a compound which productively binds to the aglycone site. The approach involved incubation of samples of the fluoroglycosyl-enzyme in the presence of a number of potential acceptors, followed by monitoring of the amount of enzyme reactivated due

ABSTRACT

to transglycosylation at a fixed time. Using a 96-well plate format, seven glycosidases were screened in this manner using 46 different potential acceptors. Of the glycosides tested, 16-36% were positively identified as candidates to act as good acceptors for the given enzyme. Further evaluation of relative reactivation rates with these candidates was performed by monitoring the extent of reactivation of the fluoroglycosyl-enzyme species at a series of time points. The acceptors were ranked according to the observed initial velocity of reactivation. Generally, aryl glycosides and disaccharides were the preferred acceptors. Validation of the screening strategy as a method by which to identify good acceptors was carried out by the identification of products formed from some of the positively screened acceptors in four different cases.

Table of Contents

A	bstra	ct		ii
Ta	able (of Con	tents	iv
Li	st of	Table	S	vii
Li	st of	Figur	es	viii
Li	st of	Abbre	eviations	x
A	cknov	wledge	ements	xv
D	edica	tion		xvi
1	Intr	oducti	ion	1
	1.1	Glycos	sidases	 1
	1.2	Cataly	ysis by glycosidases	 3
		1.2.1	The mechanisms of hydrolysis and transglycosylation	 3
		1.2.2	2-Deoxy-2-fluoro-glycosides: mechanism-based inactivators .	 5
		1.2.3	The identification of the catalytic nucleophile of a	
			glycosidase	 6
	1.3	Glycos	sidases in oligosaccharide syntheses	 7
		1.3.1	Transglycosylation	 7
		1.3.2	Glycosynthases	 10
	1.4	The β	B-galactosidase from Xanthomonas manihotis	 13
	1.5	Aims	of study	 16

TABLE OF CONTENTS

2	The	e Char	acterization of X. manihotis β -Galactosidase	17
	2.1	Introd	$\mathbf{Luction}$	17
		2.1.1	Kinetic analysis of the hydrolysis reaction	18
		2.1.2	Kinetic analyses of inactivation	18
	2.2	Result	s and discussion	20
		2.2.1	Purification	20
		2.2.2	pH Stability	22
		2.2.3	pH Dependence	22
		2.2.4	Hydrolysis kinetics	24
		2.2.5	Glycone specificity	26
		2.2.6	Inactivation by 2FDNPGal	30
		2.2.7	The identification of the nucleophile	
			of X. manihotis β -galactosidase	34
		2.2.8	The nucleophile mutant E260A BgaX	38
	2.3	Concl	usions	39
3	Scr	een of	Potential Transglycosylation Acceptors	41
3	Scr 3.1	e en of Introd	Potential Transglycosylation Acceptors	41 41
3	Scr 3.1	een of Introd 3.1.1	Potential Transglycosylation Acceptors luction	41 41 41
3	Scr 3.1	een of Introd 3.1.1 3.1.2	Potential Transglycosylation Acceptors luction	41 41 41 43
3	Scr 3.1 3.2	een of Introc 3.1.1 3.1.2 Resul	Potential Transglycosylation Acceptors luction	41 41 41 43 47
3	Scr 3.1 3.2	een of Introd 3.1.1 3.1.2 Resul ² 3.2.1	Potential Transglycosylation Acceptors luction	41 41 43 47 47
3	Scr 3.1 3.2	een of Introc 3.1.1 3.1.2 Result 3.2.1 3.2.2	Potential Transglycosylation Acceptors luction	41 41 43 47 47
3	Scr 3.1 3.2	een of Introc 3.1.1 3.1.2 Result 3.2.1 3.2.2	Potential Transglycosylation Acceptors luction	41 41 43 47 47
3	Scr 3.1 3.2	een of Introd 3.1.1 3.1.2 Result 3.2.1 3.2.2 3.2.3	Potential Transglycosylation Acceptors luction	 41 41 43 47 47 53
3	Scr 3.1 3.2	een of Introc 3.1.1 3.1.2 Result 3.2.1 3.2.2 3.2.3	Potential Transglycosylation Acceptors luction . Screening of potential acceptors . Preparative-scale transglycosylation reactions . ts and discussion . Large scale screening of potential transglycosylation acceptors . The kinetics of enzyme reactivation via . transglycosylation . Positively screened glycosides as transglycosylation acceptors .	41 41 43 47 47 53 65
3	Scr 3.1 3.2 3.3	een of Introc 3.1.1 3.1.2 Result 3.2.1 3.2.2 3.2.3 Concl	Potential Transglycosylation Acceptors luction	41 41 43 47 47 53 65 68
3	Scr 3.1 3.2 3.3 Ma	een of Introc 3.1.1 3.1.2 Result 3.2.1 3.2.2 3.2.3 Concl terials	Potential Transglycosylation Acceptors luction	 41 41 41 43 47 47 53 65 68 70
3	Scr 3.1 3.2 3.3 Ma 4.1	een of Introc 3.1.1 3.1.2 Result 3.2.1 3.2.2 3.2.3 Concl terials The c	Potential Transglycosylation Acceptors luction	 41 41 41 43 47 47 53 65 68 70 70

TABLE OF CONTENTS

١

		4.1.2	The molar absorptivity of 2,4-dinitrophenol and	
			p-nitrophenol	71
		4.1.3	Source, expression and purification of BgaX	72
		4.1.4	pH Stability	74
		4.1.5	pH Dependence	75
		4.1.6	Kinetics	75
		4.1.7	Identification of the catalytic nucleophile of BgaX	78
		4.1.8	The mutant H6-E260A BgaX	79
		4.1.9	Measurement of activity of H6-E260A BgaX	79
	4.2	Large	-scale screening of potential acceptors	80
		4.2.1	General	80
		4.2.2	Screening A: large scale screening	81
		4.2.3	Screening B: analylsis of enzyme reactivation	82
		4.2.4	Preparative scale transglycosylation reactions	84
Bi	ibliog	graphy	7	88
۸	Kin	otic A	nalveos	02
А		Basic	anguma kinatics	9 0
	Δ 2	The I	nterpretation of k and k / K	95
	Δ3	Inacti	$\frac{1}{1} \frac{1}{1} \frac{1}$	90 00
	Δ.4	The T	Determination of K. for a	99
	11.1	comp	\mathbf{A}_{i}	100
	A 5	Prote	ction from inactivation	100
	Δ.6	React		102
	A.0	Iteact		103
в	Gra	phical	l Representation of Data	105
в	Gra B.1	a phica l Gener	l Representation of Data	105 105
В	Gra B.1 B.2	phica l Gener Enzyr	l Representation of Data ral	105 105 106
в	Gra B.1 B.2 B.3	ophica l Gener Enzyr Enzyr	I Representation of Data ral me kinetics me reactivation	105 105 106 109

vi

List of Tables

2.1	Kinetic parameters of hydrolysis by BgaX	25
3.1	Positive hits from screened glycosides	50
3.2	Glycosides that are negative hits for all enzymes.	51
3.3	Tranglycosylation reaction products from this study and the literature.	66
4.1	Values of ε for pNP and DNP	72
4.2	Conditions for the measurement of activity of the enzymes screened	80
4.3	Conditions for the screen of potential acceptors.	82
4.4	Conditions for the analysis of the effect of each postive hit on enzyme	
	reactivation	83
4.5	Conditions for preparative scale transglycosylation reactions.	84

[•] List of Figures

			-
	1.1	General oligosaccharide hydrolysis	2
	1.2	The general mechanism of catalysis by a retaining β -galactosidase	3
	1.3	The general mechanism of catalysis by an inverting β -galactosidase	4
	1.4	The oxocarbonium ion-like transition state	5
	1.5	Fluoro-glycosides: mechanism-based inactivators	6
	1.6	Transglycosylation.	8
	1.7	Azide rescue of the activity of a glycosynthase	11
	1.8	Glycosidic bond formation via glycosynthases	12
	1.9	Sequence similarities among family 35 glycosidases.	14
1	1.10	Sequence similarities among family 35 and family 10 glycosidases	15
	2.1	General scheme of enzyme catalysis	18
	2.2	General scheme of enzyme inactivation	19
	2.3	Scheme showing protection from inactivation	19
	2.4	The stability of BgaX with respect to pH.	22
	2.5	The dependence of k_{cat} and k_{cat}/K_m of BgaX on pH	23
	2.6	Structures of the p NP-glycoside substrates of Table 2.1	25
	2.7	The inactivation of BgaX by 2FDNPGal	30
	2.8	Galacto-configured pyridoimidazole (galacto-imidazole).	31
	2.9	Protection of BgaX from inactivation	32
	2.10	The spontaneous reactivation of 2-deoxy-2-fluorogalactosyl-BgaX	33
	2.11	TIC of control and labelled peptides	35
	2.12	MS/MS of the parent control and labelled peptides	37
	2.13	Azide rescue of the activity of E260A BgaX.	38
	3.1	Reactivation of a glycosidase and subsequent re-inactivation by its	
		transglycosylation product.	49
	3.2	The reactivation of BgaX inactivated by 2FDNPGal	54
	3.3	The reactivation profile of Man2A with p NP β -xyloside	56
		- 4 7 4	

3.4	Summary of the reactivation profiles of BgaX	57
3.5	Summary of the reactivation profiles of BgaC	58
3.6	Summary of the reactivation profiles of Man2A	59
3.7	Summary of the reactivation profiles of Abg	30
3.8	Summary of the reactivation profiles of CelB	31
3.9	Summary of the reactivation profiles of Cex	32
3.10	Summary of the reactivation profiles of HBG	33
4.1	Primers for BgaX PCR	73
A.1	The Michaelis-Menten curve	94
A.2	The Lineweaver-Burk (double reciprocal) plot.	95
A.3	Reaction coordinate diagram for enzymatic catalysis.	97
A.4	Lineweaver-Burk plots in the presence of a competitive inhibitor 10)1
A.5	General plot of enzyme reactivation)4
B.1	The determination of ε of pNP and DNP)5

List of Abbreviations

$\underline{\text{General}}$

§	section
A_{400}	absorbance at 400 nm
Ac	acetyl
BSA	bovine serum albumin
Da	daltons
DNP	2,4-dinitrophenyl
ε	molar absorptivity (cm ^{-1} mM ^{-1})
H6	6-histidine tag
HPLC	high-performance liquid chromatography
LC/MS	mass spectrometry in the single quadrupole scan mode
Me	methyl
MS/MS	mass spectrometry in the daughter ion scan mode
m/z	mass to charge ratio
NMR	nuclear magnetic resonance
OD_{600}	optical density at 600 nm
Ph	phenyl
pNP	para-nitrophenyl
SDS PAGE	sodium dodecyl sulfate polyacrylamide
	gel electrophoresis

TIC total ion chromatogram

TLC thin layer chromatography

Kinetic Parameters

k_{cat}	first order catalytic constant
k_{cat}/K_m	second order catalytic constant
\mathbf{K}_m	dissociation constant between enzyme and substrate
$\cdot \mathbf{K}_i$	dissociation constant between enzyme and inhibitor
k_{inact}^{obs}	the observed rate of inactivation
k_{inact}	first order inactivation rate constant
k_{react}^{obs}	the observed rate of reactivation
V_o^{react}	the initial velocity of reactivation

Glycosidases

- Abg Agrobacterium sp. β -glucosidase
- Arthro Arthrobacter β -galactosidase
- Bcx Bacillus circulans xylanase
- BgaC Bacillus circulans β -galactosidase

BgaX Xanthomonas manihotis β -galactosidase

CelB Streptomyces lividans endoglucanase

LIST OF ABBREVIATIONS

.

Cex	Cellulomonas fimi xylanase/glucanase
Man2A	Cellulomonas fimi β -mannosidase
H6-E260A BgaX	the 6-histidine tagged nucleophile mutant (Glu \rightarrow Ala) of
	BgaX
HBG	human β -glucuronidase
Strep	Streptomyces coelicolor β -galactosidase

Carbohydrates

.

2-deoxy-Glc	2-deoxy-D-glucopyranose
2-deoxy-Gal	2-deoxy-D-galactopyranose
DNPCell	2,4-dinitrophenyl β -cellobioside
DNPGal	2,4-dinitrophenyl $\beta\text{-D-galactopyranoside}$
$2,5$ -DNPX $_2$	2,5-dinitrophenyl β -D-xylobioside
2FDNPCell	2,4-dinitrophenyl 2-deoxy-2-fluorocellobioside
2FDNPGal	2,4-dinitrophenyl 2-deoxy-2-fluorogalactopyranoside
2FDNPGlc	2,4-dinitrophenyl 2-deoxy-2-fluoroglucopyranoside
2FGlcAF	2-deoxy-2-fluoro- β -D-glucopyranuronyl fluoride
2FManF	2-deoxy-2-fluoro- β -D-mannopyranosyl fluoride
α -Gal-F	α -D-galactosyl fluoride
$Gal-\beta-(1-3)-Glc-\beta-SPh$	phenyl β -D-galactopyranosyl-(1 \rightarrow 3)-1-thio-
	β -D-glucopyranoside
$Gal-\beta-(1-4)-Glc-\beta-SPh$	phenyl β -D-galactopyranosyl-(1→4)-1-thio-
	eta-D-glucopyranoside

.

LIST OF ABBREVIATIONS

galacto-imidazole	galacto-configured pyridoimidazole
α -Man-F	α -D-mannosyl fluoride
Me- α -Cell	methyl α -D-cellobioside
Me- β -Cell	methyl β -D-cellobioside
MeXyl	methyl β -D-xylopyranoside
N-Ac-glucalamine	2-acetamido-1,5-anhydro-2-deoxy-D-arabino-hex-1-
	enitol
N -Ac-glucosamine- β -NH ₂	2-acetamido-2-deoxy- β -D-glucosyl amine
oNPXyl	ortho-nitrophenyl β -D-xylopyranoside
pAcPhGlc	para-acetylphenyl β -D-glucopyranoside
PhGlc	phenyl β -D-glucopyranoside
p NP- α -L-Ara	$para$ -nitrophenyl α -L-arabinopyranoside
pNPCell	para-nitrophenyl β -cellobioside
pNPFuc	para-nitrophenyl β -D-fucopyranoside
pNPGal	para-nitrophenyl β -D-galactopyranoside
pNPGalNAc	para-nitrophenyl 2-acetamido-2-deoxygalacto-
	pyranoside
pNPGlc	para-nitrophenyl β -D-glucopyranoside
pNPGlcA	para-nitrophenyl β -D-glucuronide
pNPMan	para-nitrophenyl $\beta\text{-}\mathrm{D}\text{-}\mathrm{mannopyranoside}$
p NP- α -Xyl	$para$ -nitrophenyl α -D-xylopyranoside
pNPXyl	$para$ -nitrophenyl β -D-xylopyranoside

/

Amino Acids

А	Ala	alanine	М	Met	methionine
С	Cys	cysteine	Ν	Asn	asparginine
D	Asp	aspartic acid	Р	Pro	proline
Е	Glu	glutamic acid	Q	Gln	glutamine
F	Phe	phenylalanine	R	Arg	arginine
G	Gly	glycine	S	Ser	serine
Η	His	histidine	Т	Thr	threonine
Ι	Ile	isoleucine	V	Val	valine
Κ	Lys	lysine	W	Trp	tryptophan
\mathbf{L}	Leu	leucine	Υ	Tyr	tyrosine

xiv

Acknowledgements

First and foremost, great thanks goes to my supervisor Dr. Stephen Withers for his guidance and immeasurable patience. Thank you to Janine Foisey for cloning BgaX, and Dr. Laurent Gal for his help with protein expression as well as numerous insightful discussions. I greatly acknowledge Karen Rupitz for technical assistance, Shouming He for his expertise with mass spectrometry, Marietta Austria for generating my NMR spectra, and Helen Merkins and Drs. Tony Warren and Doug Kilburn for the use of their 96-well plate readers. I am very appreciative to all members of the Withers' group for the provision of various enzymes and inhibitors as well as many helpful discussions and suggestions. I thank them also for their friendship and support.

Great thanks also go to my family who have encouraged me every step of the way, especially when that light at the end of the tunnel seemed so far away.

Last, but far from least, I would like to thank Dan Melconian, my beau and best friend who has always been my inspiration. I never would have finished this without him.

Thank you all.

For my parents. Je t'aime.

CHAPTER 1

Introduction

1.1 Glycosidases

Carbohydrates, the most prevalent of all biological molecules [1], assume integral roles *in vivo* that are essential to a large number of cellular processes. Carbohydrate monosaccharides are linked via glycosidic bonds to form a fantastic number of oligomers and polymers. These macromolecules serve a variety of functions both as polysaccharides (eg: structural enhancement, energy storage) and coupled with proteins and lipids in more complex conjugates (eg: recognition elements, extracellular matrix).

Glycosyl hydrolases, or glycosidases, are enzymes that catalyse the hydrolysis of glycosidic bonds, and as such, play crucial roles in the catabolism of carbohydrates (Figure 1.1). Glycosidases are essential for processes which include post-translational modification of proteins in the rough endoplasmic reticulum, and digestion. The importance of glycosidases to proper cellular maintenance is evidenced by human disorders that are associated with the critical mutation of several of these enzymes, such as Tay-Sachs disease (deficiency in lysosomal hexosaminidase A), and human lysosomal storage disease mucopolysaccharidosis type VII (mutation of β -glucuronidase).

1



FIGURE 1.1: General glycosidase catalysed oligosaccharide hydrolysis. Sugar hydroxyls have been omitted for clarity.

As well as being an essential class of enzymes in vivo, glycosidases are also valuable diagnostic tools in the elucidation of carbohydrate structure and function. Glycosidases are specific for the glycone glycosyl unit and its configuration at the anomeric center. The structure of an oligosaccharide can be determined by discovering which glycosidases are required to successively remove terminal glycosyl units from the oligomer. Once the overall structure of the oligosaccharide under study is determined, this same strategy can be applied to determine its function by the removal of glycosyl units and subsequent observation of any corresponding effects in vivo. Glycosidases have also proved useful in the synthesis of oligosaccharides through transglycosylation and the development of glycosynthases (vide infra). Large numbers of glycosidases have been isolated and characterized from a wide range of organisms thereby demanding the development of a well-ordered classification system. The primary distinction made within this class of enzymes is dictated by substrate specificity (core moiety and anomeric configuration) of the glycone portion of the active site [2]. More recently, a distinct classification of glycosidases has been developed that is based on amino acid sequence similarities [3, 4, 5]. Over two thousand glycosidases have been assigned to over seventy families which share a common consensus pattern. Similarities in primary sequence demand that these proteins will fold and catalyse hydrolysis in a similar manner, thus making it possible to speculate as to the relatedness of en-

 $\mathbf{2}$

zymes within and between families. It also allows intelligent speculation regarding the identities of key catalytic residues, since candidates for these roles should be highly conserved.

Interest in glycosidases is well founded and numerous studies, particularly over the past decade, have provided much insight into the characterization and mechanism of this class of enzymes.

1.2 Catalysis by glycosidases

1.2.1 The mechanisms of hydrolysis and transglycosylation

The mechanism of glycosidase catalysis was first postulated by Koshland in 1953 [6] and has since been supported by structural, kinetic and mechanistic studies [7, 8, 9]. Substrates are hydrolyzed with either net retention or inversion of configuration at the anomeric center of the scissile bond with the aid of two strategically placed carboxyl groups in the enzyme active site from either glutamic or aspartic acid residues. Retaining glycosidases function through a double displacement mechanism as shown



FIGURE 1.2: The general mechanism of catalysis by a retaining β -galactosidase.



FIGURE 1.3: The general mechanism of catalysis by an inverting β -galactosidase.

in Figure 1.2. In the first step of catalysis (glycosylation) one of the carboxyl groups attacks the anomeric centre of the glycone sugar while the other provides general acid catalysis. This step results in the formation of a covalent glycosyl enzyme intermediate that is then broken down by general base catalysed nucleophilic attack of either water (deglycosylation) or an 'acceptor' sugar (transglycosylation).

Inverting glycosidases operate via a direct displacement mechanism with the active-site carboxyl groups functioning as general acid/base catalysts (Figure 1.3).

The transition states during catalysis by both retaining and inverting glycosidases have substantial oxocarbonium ion-like character (Figure 1.4). In the case of the retaining *Agrobacterium* sp. β -glucosidase (Abg), secondary deuterium kinetic isotope effects have demonstrated that the transition state for deglycosylation is more oxocarbonium ion-like as compared to that of glycosylation [10].



FIGURE 1.4: The oxocarbonium ion-like transition states for hydrolysis by (a) retaining and (b) inverting glycosidases.

1.2.2 2-Deoxy-2-fluoro-glycosides: mechanism-based inactivators

A class of mechanism based inactivators for retaining glycosidases has recently been developed in this laboratory [11, 12, 13]. These are glycosides that are substituted with fluorine at the 2- or 5- position and have an activated leaving group, such as 2,4-dinitrophenyl (DNP) or fluoride, at the anomeric center. The 2- or 5-fluorine substituent destabilizes the oxocarbonium ion-like transition states, slowing both steps of catalysis. However, the inclusion of a good leaving group at the aglycone site accelerates glycosylation. The net result of the attempted hydrolysis of such substrates by a retaining glycosidase is the accumulation of the glycosyl enzyme intermediate, rendering the enzyme inactive (Figure 1.5). It has been demonstrated through mass spectrometry, ¹⁹F-NMR, and X-ray crystallography of glycosidases inactivated in this manner that the inhibited enzyme species is the covalent fluoro-glycosyl enzyme in-



FIGURE 1.5: The inactivation of a β -galactosidase with 2,4-dinitrophenyl 2deoxy-2-fluoro- β -galactopyranoside (2FDNPGal) and subsequent reactivation, where $k_{inact} \gg k_{react}$.

termediate [12, 14, 15]. The existence of this intermediate is also suggested by the catalytic competence of the inhibited species. That is, once inactivated, a glycosidase will regain its catalytic capability over time as the result of the eventual turnover of the fluoro-glycoside. This reactivation can occur by the nucleophilic attack of water, as in the regular hydrolysis mechanism, or through attack by a suitable acceptor, as in transglycosylation.

It is thus readily apparent that, since inverting glycosidases do not form a covalent glycosyl enzyme intermediate during hydrolysis, fluoro-glycosides are not effective inactivators of this class of enzymes.

1.2.3 The identification of the catalytic nucleophile of a glycosidase

Fluoro-glycosides with good leaving groups have been shown to be excellent reagents with which to label the nucleophilic carboxylate of a retaining glycosidase, and thereby allow the identification of the amino acid involved in nucleophilic catalysis. Once the enzyme is inactivated, the glycosyl enzyme species is partially digested by a protease and the resultant peptide fragments separated, and analyzed by mass spectrometry. The peptide covalently attached to the fluoro-glycoside is identified by comparison of the mass spectrum chromatogram to that of a control sample of unlabelled enzyme similarly treated, and sequenced to identify the catalytic amino acid residue. This procedure has been used successfully to assign the position of the catalytic nucleophile in many glycosidases [16, 17, 18, 19].

1.3 Glycosidases in oligosaccharide syntheses

The investigation of the functional roles of carbohydrates *in vivo* necessitates the synthesis of oligosaccharides for study. As simple as this statements sounds however, a synthetic scheme can quickly become quite cumbersome if the target in question is complex. Monosaccharides have several hydroxyls which may participate in glycosidic bonds, thereby requiring numerous protection and deprotection steps to attain the correct regio- and stereochemistry between units. Alternatively, an enzyme that catalysed the specific formation of glycosidic bonds could link glycosyl units in a single step reaction, thereby greatly facilitating oligosaccharide syntheses.

1.3.1 Transglycosylation

Although retaining glycosidases preferentially hydrolyse their substrates *in vitro*, they have been exploited for their natural ability to transglycosylate. If transglycosylation can be enhanced relative to hydrolysis, glycosidases can catalyse the formation of glycosidic bonds in reasonable yields. For this reason, these enzymes have been studied extensively to determine their utility in the synthesis of a number of target molecules (for an excellent review see [20]).

The desired outcome of a transglycosylation reaction is the formation of a glycosidic linkage between a carbohydrate and another moiety, typically, although not necessarily, another sugar. The 'donor' in such a reaction is the source of the glycosyl unit to be transfered, while the 'acceptor' is the compound that becomes glycosylated.

There are two basic approaches to inducing transglycosylation by glycosidases. In the 'thermodynamic' approach a large excess of donor and acceptor are used to favour the reverse of the hydrolysis mechanism to result in glycosidic bond formation. The yields associated with such reactions, however, are typically low (< 15%) owing to the inherent 55 M concentration of water that favours hydrolysis. Methods to favor transglycosylation involve the reduction of water content by the use of organic co-solvents, or the continuous removal of the synthesized product through adsorption onto activated charcoal or crystallization *in situ*. A second approach is that of synthesis under 'kinetic' control in which the glycosyl enzyme intermediate is intercepted



FIGURE 1.6: Transglycosylation occurs when a suitable acceptor intercepts the glycosyl enzyme intermediate. For this illustrative example, donor = galactose (Gal) and acceptor = N-acetyl-D-galactosamine (GalNAc.)

by the acceptor. Product yields using this strategy are often higher than those under thermodynamic control and this approach is therefore generally preferred. For the purposes of this report, 'transglycosylation' refers to those reactions that follow this latter approach.

Potential donors for the kinetic approach are readily apparent from Figure 1.6 as they are necessarily those glycosides that will act as substrates for hydrolysis for the retaining enzyme in question. Donors are most often monosaccharides that are activated at the anomeric center, or simple disaccharides.

As in hydrolysis, the newly formed linkage in the transglycosylation product has the same anomeric configuration as that of the donor sugar. The first criterion regarding which glycosidase should be selected as a catalyst for the synthesis of a specific linkage is thus readily apparent from the natural specificity of each glycosidase: for example, the coupling of a galactose residue to the non-reducing end of an oligosaccharide in a β -linkage can be accomplished using an exo- β -galactosidase. However, the further choice of an appropriate glycosidase based on aglycone specificity is not as simple.

In order for product formation to occur, the acceptor moiety must bind productively to the enzyme's aglycone site to stabilize the transition state of the transglycosylation step, thereby intercepting the glycosyl enzyme intermediate. Therefore, the greater the favorable interactions between the aglycone site and an acceptor, the greater will be the chance that transglycosylation will be favored over hydrolysis. This has been clearly demonstrated by the observation that the product yield of transglycosylation reactions varies widely with the identity of the acceptor employed [20, 21]. Further, the precise mode of interaction between the aglycone site and acceptor will determine which hydroxyl group attacks the glycosyl enzyme intermediate, thereby defining the regiochemistry of the product. It is therefore evident that the specificity of the aglycone site of a glycosidase is a strong determinant of the regiochemical outcome and yield of a transglycosylation reaction.

The aglycone specificity of glycosidases has traditionally been examined using an approach in which the rates of hydrolysis of different substrates that share a common glycone moiety are compared. While this provides insight into the natural substrates of a glycosidase, it is not an adequate portrayal of acceptor specificity with respect to transglycosylation. Although substrate specificity with respect to hydrolysis may suggest several potential acceptors which could be used in transglycosylation reactions, it is not necessary that the *best* acceptors for a glycosidase be dictated by the susceptibility of their respective products. Further, the use of this strategy to examine aglycone specificity is not generally feasible as it requires access to a wide range of oligosaccharide structures due to the large number of unique acceptors used in various synthetic schemes. An alternative method to examine aglycone binding properties would be to carry out a series of transglycosylation reactions with various acceptors and determine which led to product formation via HPLC analysis. This process, however, would be very time consuming.

A more reliable and speedy method by which to examine the specificity of the aglycone site of a glycosidase is therefore needed in order to ensure the choice of the most appropriate glycosidase for a desired transglycosylation reaction.

1.3.2 Glycosynthases

Although transglycosylation by glycosidases has been shown to be useful in the synthesis of a number of targets, there are still several drawbacks to this methodology. The products of transglycosylation are necessarily substrates, and as such, may be



FIGURE 1.7: Azide rescue of the activity of a glycosynthase.

rapidly hydrolyzed as they are formed, resulting in relatively low yields.

The recent development of glycosynthases in this laboratory has afforded a new angle by which to maximize the utility of glycosidases for oligosaccharide synthesis. Glycosynthases are glycosidases that have been specifically mutated so as to replace the nucleophilic active site carboxylate with another amino acid residue, typically alanine or serine. X-ray crystallographic analyses have demonstrated that such mutant enzymes fold into the same wild-type tertiary structure and retain their active site architecture to successfully bind substrates as in the parent wild type protein [22]. Once bound however, the substrate cannot cleaved due to the alteration of the catalytic carboxyl.

Activity with these mutants can be 'rescued' by supplying an appropriate nucleophile such as azide or formate (Figure 1.7) [23, 24, 25]. These small anions mimic the role of the missing carboxylate to aid in the cleavage of glycosides with activated aglycone moieties such as DNP.

The utility of glycosynthases arises from their inability to catalyse hydrolysis coupled with their capacity to transglycosylate. If a glycosyl fluoride donor with the anomeric configuration opposite to that of the natural substrate is supplied, along



FIGURE 1.8: Glycosidic bond formation via glycosynthases.

with a suitable acceptor, transglycosylation can occur as shown in Figure 1.8. The glycosyl fluoride donor mimics the glycosyl enzyme intermediate in the regular transglycosylation mechanism. Thus, if the appropriate donor and acceptor are supplied, a glycosynthase can catalyse transglycosylation, but not the hydrolysis of the product, ensuring a higher overall reaction yield.

To date, 'successful' glycosynthases have been engineered by the replacement of the glutamic acid residue that functions as the catalytic nucleophile in three wild type glycosidases. Agrobacterium sp. β -glucosidase (Abg) Glu358Ala [26] and Bacillus licheniformis 1,3-1,4- β -glucanase Glu134Ala [27] have been shown to catalyse transglycosylation in excellent yields (70%–90%), while Cellulomonas fimi β -mannosidase (Man2A) Glu519Ala [28] appears to be much less efficient, with yields below 10%. Current studies with Abg are focused on the optimization of efficiency by the replacement of the catalytic nucleophile with other amino acids. Results with the Abg Glu358Ser mutant show a 24-fold improvement in synthetic rates as compared to the alanine counterpart [29]. This increase in efficiency is presumably due to a stabilizing interaction between the hydroxyl of the side chain of serine and the departing fluorine of the donor sugar. This study also provides hope for the engineering of successful glycosynthases from other glycosidases whose alanine mutants failed to catalyse transglycosylation.

As seen with wild type glycosidases, the efficiency of product formation using glycosynthases varies with the chosen acceptor. Therefore, in order to maximize product formation, the most appropriate moiety with respect to aglycone specificity should act as the acceptor. Since glycosynthases retain the active site architecture of their parent wild type glycosidase, the aglycone specificity determined for the wild type enzyme should be applicable to its corresponding glycosynthase. Therefore, the evaluation of the specificity of the aglycone site of a wild type glycosidase would yield results that would also be applicable to syntheses using glycosynthases.

1.4 The β -galactosidase from Xanthomonas manihotis

The recently isolated β -galactosidase from the pathogenic bacterium Xanthomonas manihotis (BgaX) is a 66 kDa exoglycosidase that hydrolyses (1 \rightarrow 3) and, to a lesser degree (1 \rightarrow 4), linked terminal β -galactose residues from oligosaccharides [30, 31]. Based on primary amino acid sequence similarities, BgaX has been classified as a family 35 glycosyl hydrolase (Figure 1.9) [31, 32]. This family consists of over 30 retaining β -galactosidases, the vast majority of which are eukaryotic in origin. To date, there are only four bacterial members: BgaX, and β -galactosidases from Bacillus circulans (BgaC) [32] Arthrobacter (Arthro) [33] and Streptomyces coelicolor (Strep) [34].

The significant sequence similarity between these prokaryotic enzymes and other β -galactosidases of family 35 is of particular interest given the distant evolutionary relatedness of prokaryotic and eukaryotic species. Taron and coworkers have specu-

									\mathbf{T}						
BgaX	153	AYLDA	LAK	Q V Q P	LLN	H N	GGP	IIAV	VENEY	G -	SYAD	DHA	YMADN	198	
Bcir	126	AYFDV	LFEI	RLRP	LLS	s N	GGP	IIA L	DENEY	G –	SFGN	ΟΟΚ	YLQYL	170	42%, 57%/553
Arthro	128	SYMEA	LAP	I L V P	RQI	DR	GGP	IILV	2 I E N E Y	G -	AYGS	DMH	YLEQL	171	35%, 51%/398
Strep	130	RWFDA	LVP	RIAE	ΓQΑ	G R	GGP	/ V A V	IENEY	G -	SYGD	DRA	YVRHI	174	37%, 51%/559
Human	158	KWLAV	LLP	KMKP	LLY	QN	GGP	ÍľtV	VENEY	GS	YFAC	DFD	YLRFL	203	41%, 52%/574
Mouse	159	KWLGV	LLP	KMKP	LLY	QN	GGP	I I T V (VENEY	GS	YPAC	DYD	YLRFL	204	40%, 53%/574
Aspar	151	KF T E K	IVS	MMKA	EGL	ΥΕΤΟ	GGP	ILS	1 ENEY	GΡ	νεчч	DGA	AGKSY	199	34%, 47%/318
Cat	159	KWLGV	LLP	KMKP	LLY	QN	GGP	ΙΙτν	VENEY	GS	YFTC	DYD	YLRFL	204	40%, 52%/576

														L													
		- 1-1-1-1-			1									.	-	-					· · · · ·						_
BgaX	242	SAFD	ĸ	LIK	jr i	RP	DQ	PR	-			:M;V	G	Ъ.	W J	A G	WF	DH	We	- K	PI	I]A	АТ	D	AR		2 281
Bcir	215	SAFA	Q	LK Q	Y (Q [P])	NA	ΡL]-			MC	м	ΕE	W	H G	WF	DH	WG	- E	E	н	TR	s	A E	s v	254
Arthro	216	SQRN	A	LRP	сı	ASI	нн	РТ	G	A 7	C B	I V L	G	ΕĽ]G 1	LA	GF]E P	LG	ΩD	н	ЯR	TI	S	νQ	ES	3 260
Strep	218	RAAA	L	LRS	RI	RP	АE	PF	-		-	FC	A	EI	W	N G	WF	DH	WG	- D	ĸ	Т.H	V R	P	A P	S P	257
Human	250	DAFL	s	QRK]C 1	EPI	KG	P.L]-			IN	s	EF]¥[r G	WL	DH	WG	- Q	PI	ĪS	T I	ĸ	T E	A V	284
Mouse	251	QAPL	v	QRK]F 1	E P	KÇ	ΡL]-			IN	S	EE]¥[r G	WL	DH	WG	- K	PJ	IS	τv	ĸ	тĸ	TI	290
Aspar	236	CDY-	-	FSP	NI	KDI	N K	ΡK	-			MW	т	D A	W	r G	WF]т G	FG	GΑ	v 1	P Q	RP	A	ED	M	\$ 275
Cat	249	AA Q	Ι	QRK	s I	E P I	RG	ΡL	-		• -	V N	S	EF	Y.	r G I	МГ	DH	WG	- Q	PI	Is	RV	R	ТЕ	v v	290

FIGURE 1.9: Sequence similarities among family 35 glycosidases using standard one letter amino acid abbreviations; codes for the origin of the enzymes are in the text or are self explanatory, except for Aspar, which is for asparagus. Identical amino acids are outlined with a solid line, similar with a dotted line, and conserved in white with a black background; dashes indicate gaps in the sequence. The top and bottom arrows indicate the position of the putative catalytic acid/base and nucleophilic amino acids respectively. Numbers at the beginning and end of the sequences indicate the amino acid numbering relative to the entire protein. Numbers to the right of the upper block of alignments indicate the total % identity and % similarity to BgaX, with the last value being the number of amino acids in the homologous region. 14

BgaX 72 GLNTVETYVFWNLVEPQQGQFDFSGN 97 55 GFNTVETYVAWNLHEPEEGQFVFEGI 80 BgaC Strep 48 GLNAVDTYVPWNFHERTAGDIRFDGP 73 Arthro 46 GLNTIETYVAWNLHAPSEDVFDTSAG 71 65 DAITPENEMKWEVVEPTEGNFDFTGT 89 Cal 67 GALTPENSMKWDATEPSRGQFSFSGS 92 Aka 41 GQ VTPENSMKWDATEPSQGNFNFAGA 66 Tau 71 GQLSPENSMKWDATEPSQGQFSFAGS 96 Pch 79 NLVVAENAMKWDATEPSQNSFSFGAG 104 Cfi

FIGURE 1.10: Sequence similarities among family 35 and family 10 glycosidases. Identical amino acids are outlined with a solid line, similar with a dotted line, and conserved in white with a black background. Numbers at the beginning and end of the sequences indicate the amino acid numbering relative to the entire protein. $Cal = Cryptococcus \ albidus \ xylanase \ A; Aka = Aspergillus \ kawachii \ xylanase \ A; Tau = Thermociscus \ aurantiacus \ xylanase; Pch = Penicillium \ chrysogenum \ xylanase; Cfi = Cellulomonas \ fimi \ cellulase/xylanase.$

lated that since X. manihotis is a plant pathogen, BgaX may have arisen as the result of gene transfer from a host [31]. However, the other three bacterial β -galactosidases originate from non-pathogenic microorganisms and, as such, are less likely to undergo gene transfer with a eukaryotic species [33]. Another evolutionary possibility, also suggested by Taron and coworkers, is based on seven proposed regions of sequence homology of the enzymes of family 35. The first of these regions of the bacterial β -galactosidases are similar to a domain found in family 10 glycosyl hydrolases (Figure 1.10). These similarities therefore suggest that the prokaryotic enzymes may have evolved from a related xylanase. It is apparent that the inclusion of these bacterial enzymes within a family of eukaryotic proteins poses interesting evolutionary questions as to how such similarities arose in such distantly related organisms.

As noted in Figure 1.9, BgaX has a significant degree of sequence similarity with a human lysosomal β -galactosidase. Mutations at several sites in this human enzyme result in the neurological disorders GM1-gangliosidosis and Morquio B. syndrome [35]. Due to the degree of similarity between the two enzymes, mechanistic or structural studies of BgaX are of particular interest as they could contribute to an understanding of the human β -galactosidase and its associated disorders.

Hydrophobic cluster analysis, coupled with primary amino acid sequence similarities, has predicted the position of the nucleophilic and acid/base amino acid residues of the glycosidases of family 35 [36]. As shown in Figure 1.9, these analyses predict the catalytic nucleophile of BgaX as Glu²⁶⁰. The assignment of the postition of the corresponding carboxylate in the human β -galactosidase has been experimentally confirmed in this laboratory by the previously described labeling methodology utilizing a 2-deoxy-2-fluoro-glycoside (§1.2.3) [18]. However, given the diversity of this family, it seemed wise to confirm the identity of the nucleophile in an enzyme from a prokaryotic source.

1.5 Aims of study

The aims of this study were two-fold:

(i) Identify the catalytic nucleophile of BgaX through a labeling study utilizing a 2-deoxy-2-fluoro-galactoside.

(ii) Develop a simple, yet rapid, screening procedure that identifies the most suitable acceptors for a given glycosidase or glycosynthase for transglycosylation reactions.

CHAPTER 2

The Characterization of X. manihotis β -Galactosidase

2.1 Introduction

It was noted in §1.4 that the assignment of the catalytic nucleophiles of family 35 glycosidases has been experimentally demonstrated to be accurate in the case of the human lysosomal β -galactosidase. Although there is a significant degree of global primary sequence homology among enzymes of this family, the local region bracketing the putative active site nucleophilic carboxylates appears to be somewhat variable (Figure 1.9). Greater sequence conservation can be seen within the region encompassing the residue that functions as the catalytic acid/base.

Sequence similarity to the human β -galactosidase was used to assign the putative sites of the catalytic residues of BgaX as Glu²⁶⁰ (nucleophile) and Glu¹⁸⁰ (acid/base). Given the weak sequence conservation in the region of the catalytic nucleophile, it was of interest to experimentally verify the position of this active site residue.

As well as identifying a catalytic residue of BgaX, we desired to examine the aglycone specificity of this enzyme with respect to transglycosylation reactions. These

results are reported in the following chapter which discusses the development of the aglycone screening method. We also briefly examined the specificity of the glycone site of BgaX through the kinetic analysis summarized below (see Appendix A for details).

2.1.1 Kinetic analysis of the hydrolysis reaction

Enzyme catalysis can be described by the following scheme:

$$E + S \xrightarrow{k_{-1}} E \cdot S \xrightarrow{k_{cat}} E + P \qquad K_m = \frac{k_{-1} + k_{cat}}{k_1}$$

FIGURE 2.1: General scheme of enzyme catalysis, where E is the free enzyme, S the substrate, $E \bullet S$ the Michaelis-Menten complex, and P the final product.

Values for k_{cat} and K_m are determined by measuring the initial reaction velocity at several substrate concentrations, and subsequent fitting of the data to the classical Michaelis-Menten expression:

$$v = \frac{k_{cat}[\mathsf{E}]_T[\mathsf{S}]}{\mathsf{K}_m + [\mathsf{S}]}$$
(2.1)

The parameters k_{cat}/K_m and k_{cat} are rate constants that are related to the first irreversible step of catalysis and the rate determining step respectively.

Comparison of these kinetic parameters for different substrates therefore provides insight into substrate specificity and the energy barriers associated with catalysis.

2.1.2 Kinetic analyses of inactivation

The inactivation of an enzyme by a mechanism-based inactivator can be described by the following general scheme:

$$\mathsf{E} + \mathsf{I} \xrightarrow{k_{-1}} \mathsf{E} \cdot \mathsf{I} \xrightarrow{k_{inact}} \mathsf{E} \mathsf{I} \qquad \mathsf{K}_{i} = \frac{k_{-1} + k_{inact}}{k_{1}}$$

FIGURE 2.2: General scheme of enzyme inactivation where E and I are the free enzyme and inhibitor, E•I the Michaelis-Menten complex, and EI the inactivated enzyme species.

The parameters k_{inact} and K_i are calculated from the observed rates of inactivation (k_{inact}^{obs}) at several different inhibitor concentrations through fitting of this data to a relationship analogous to the Michaelis-Menten expression (Equation 2.1):

$$k_{inact}^{obs} = \frac{k_{inact}[\mathsf{I}]}{K_i + [\mathsf{I}]} \tag{2.2}$$

The illustration that an inactivator is active site-directed can be achieved by monitoring the rate of the inactivation of the enzyme in the presence of an additional competitive inhibitor. This scenario is summarized in Figure 2.3 where P is the competitive inhibitor.

$$E + I \xrightarrow{k_{-1}} E \cdot I \xrightarrow{k_{inact}} EI$$

$$+ P$$

$$K_{i} = \frac{k_{-1} + k_{inact}}{k_{1}}$$

$$K_{i} = \frac{k_{-1} + k_{inact}}{k_{1}}$$

$$K_{p} = \frac{k'_{-1}}{k'_{1}}$$

FIGURE 2.3: Scheme showing the protection of an enzyme from inactivation by a mechanism-based inactivator in the presence of a competitive inhibitor.

The validity of applying this model to the inhibition of a glycosidase by a 2-deoxy-2-fluoro-glycoside can be determined by the ratio of the observed rate of inactivation in the presence (k_{ip}^{obs}) and absence (k_i^{obs}) of a suitable competitive inhibitor:
$$\frac{k_{ip}^{obs}}{k_i^{obs}} = \frac{K_i + [I]}{K_i \left(1 + \frac{[P]}{K_p}\right) + [I]}$$
(2.3)

2.2 Results and discussion

2.2.1 Purification

A tag of six histidine residues was incorporated into the BgaX protein at the carboxy terminus through direct manipulation of the gene. This modification made it possible to employ a Ni²⁺-chelating column in the first step of the purification of overexpressed BgaX from *E. coli* thereby eliminating much extraneous cellular material in a single step. The remaining impurities were removed by pooling those fractions containing β -galactosidase activity and subsequent application to a cation-exchange column. The resultant fractions that contained BgaX were combined into a single preparation that had a specific activity of 28.3 mol/min/mg with *p*NPGal at pH 6.0, and was homogeneous by silver stained SDS-PAGE.

In the first step of purification, BgaX was eluted from the Ni^{2+} -chelating column by a linear gradient of 0–100 mM imidazole. Although imidazole was effective at eluting BgaX from the column, its presence appeared to induce the aggregation of a portion of the desired enzyme (as determined by silver stained SDS-PAGE of the observed precipitate). Even over the time course needed to remove the imidazole by dialysis, a precipitate was observed to form. Once dialysis was complete, no additional precipitation was observed in the subsequent purification step or in the final concentrated enzyme preparation. Exactly why imidazole would cause this aggregation is unclear, but it is possible that a high concentration of this compound could have raised the ionic strength of the preparation to such a degree as to salt the enzyme out of solution.

Although some BgaX was lost through precipitation, the amount of purified enzyme isolated from the second step of the protocol (typically 30 mg/2L culture of *E. coli*) was sufficient for our purposes. Therefore, although high concentrations of imidazole appear to result in the loss of some enzyme, the use of a Ni²⁺-chelating column was seen as an appropriate initial step by which to eliminate most of the extraneous proteins from the preparation. It is apparent however that if this procedure is to be followed to purify BgaX, the imidazole used for protein elution should be removed as quickly as possible to minimize loss of the enzyme through aggregation.

The mass of BgaX, as determined from mass spectrometry, was 64849 Da, a value that is slightly higher than the expected mass of 64836 Da as predicted from the primary amino acid sequence. However, N-terminal sequencing of this histidine-tagged protein was consistent with the predicted sequence (ATPESW...). The 13 Da difference between the experimental and theoretical masses is most likely an artifact from cloning.

The values of K_m and k_{cat} for the hydrolysis of pNPGal by the 6-histidine tagged BgaX at pH 7.0 ($K_m=53.5 \ \mu M$, $k_{cat}=87 \ s^{-1}$) were shown to be essentially the same as those for the non-tagged enzyme ($K_m=55.7 \ \mu M$, $k_{cat}=105 \ s^{-1}$)¹. It was therefore concluded that although the mass of the protein was slighter higher than expected, the incorporation of this tag did not affect the enzyme in any way that altered its activity. Hence, for the purposes of this report, 'BgaX' refers to the 6-histidine tagged wild type enzyme unless noted otherwise.

¹Karen Rupitz, unpublished data. Kinetics were done with 90-95% purified enzyme as determined by silver stained SDS-PAGE.

2.2.2 pH Stability

The stability of BgaX with respect to pH was determined by the measurement of residual enzyme activity after a 2 minute incubation at pH values ranging from 4–8 (Figure 2.4). A two minute incubation period was chosen since this was the typical assay time. From the results depicted in Figure 2.4, it can be seen that the enzyme is most stable at a pH of 6.0 or greater and activity is rapidly lost at lower pH values.

2.2.3 pH Dependence

The dependence of k_{cat} and k_{cat}/K_m of BgaX on pH was determined using DNPGal as a substrate (Figure 2.5). The bell shape of these plots indicates that both kinetic parameters are dependent on two ionizations of the enzyme. The upper pK_a value of each plot is deemed as more reliable as compared to that of the lower pK_a value due to the limited data that were obtained below pH 5.0 as the result of the rapid



FIGURE 2.4: The stability of BgaX with respect to pH. The curve shown is for illustrative purposes only.





FIGURE 2.5: The dependence of (a) k_{cat}/K_m and (b) k_{cat} of BgaX on pH. Both data sets were fit to $y = \frac{L*10^{(pH-pK_a1)}}{10^{(2*pH-pK_a1-pK_a2)}+10^{(pH-pK_a1)}+1}$ (see Materials and Methods).

inactivation of BgaX at these pH values.

The first irreversible step of catalysis (glycosylation) is reflected by k_{cat}/K_m and is dependent on the protonation state of both active site carboxylates (Figure 2.5 (a)). The lower pK_a of 4.9 ± 0.2 is therefore attributed to the catalytic nucleophile, and the upper value of 6.4 ± 0.2 to the general acid catalyst.

The rate determining step for the hydrolysis of DNPGal by BgaX is deglycosylation (*vide infra*) and is therefore dependent on the protonation state of the general base catalyst. Therefore the lower ionization of the k_{cat} versus pH plot (Figure 2.5 (b)) with a pK_a of 3.7 ± 0.3 is presumably representative of this catalytic residue.

It is illustrated in Figure 1.2 that a single carboxyl group functions as both the general acid and the general base during catalysis. The values extrapolated from Figure 2.5 indicates that the pK_a of this residue changes *ca.* 2.7 units during catalysis, presumably due to local electrical changes within the enzyme active site as the result of the removal of charge from the catalytic nucleophile upon forming the covalent

glycosyl enzyme intermediate. Similar shifts of the pK_a of this residue have been observed for Agrobacterium sp. glucosidase (Abg) [10] and Bacillus circulans xylanase (Bcx) [37].

The origin of the dependence of k_{cat} on a moiety with a pK_a of 7.1 ± 0.1 is not clear. This may be the result of dependence on the side chain of another amino acid residue close to or in the enzyme active site that is involved in crucial hydrogen bonds or similar non-covalent interactions during catalysis. The activity of Abg with a deglycosylation rate limiting substrate has also shown to be dependent on an ionization with a pK_a of *ca.* 8.0 [10], suggesting the participation of a mechanistically important residue similar to that for BgaX. It is possible that the upper pK_a dependence of the k_{cat} versus pH plots of these enzymes is attributed to the ionization of a tyrosine residue which hydrogen bonds to the catalytic nucleophile, as was shown to occur in Bcx through crystallographic studies [38].

All subsequent analyses of BgaX were done at pH 6.0 in order to maximize enzymatic activity while maintaining protein integrity.

2.2.4 Hydrolysis kinetics

Kinetic parameters for the hydrolysis of several pNP-glycosides by BgaX are listed in Table 2.1. With the exception of pNPGal, the kinetic analysis of all pNP-glycosides was limited by substrate solubility, making it impossible to measure hydrolysis rates at concentrations significantly higher than K_m . However, the measured initial velocities corresponding to the highest possible substrate concentrations appeared to approach a maximum, enabling an adequate fit of equation 2.1 and extrapolation of realistic values of K_m and k_{cat} . Yet, since enzyme saturation was only definitively achieved with pNPGal, the kinetic parameters determined for all other pNP-glycosides are

Substrate	$\mathrm{K}_m(mM)$	$k_{cat}(s^{-1})$	$\frac{k_{cat}/\mathrm{K}_m}{(s^{-1}mM^{-1})}$	$\Delta\Delta { m G}^{\ddagger st}$ (kJ(kcal)/mol)	
DNPGal	0.070 ± 0.014	36 ± 2	600 ± 100	0.06(0.14)	
pNPGal	0.050 ± 0.002	35.8 ± 0.3	700 ± 30		
pNPFuc	3.5 ± 0.3	7.2 ± 0.3	2.1 ± 0.2	15 (3.6)	
p NP- α -L-Ara	1.3 ± 0.4	0.34 ± 0.05	0.27 ± 0.08	20 (4.8)	
pNPGalNAc	0.71 ± 0.14	0.0217 ± 0.0012	0.030 ± 0.006	26 (6.2)	
pNPGlc	2.9 ± 0.5	0.080 ± 0.004	0.027 ± 0.005	26(6.3)	
pNPXyl	4.3 ± 1.4	0.050 ± 0.006	0.012 ± 0.004	28 (6.8)	

TABLE 2.1: Kinetic parameters for the hydrolysis of a range of nitrophenyl-glycosides by BgaX.

 $^{*}\Delta\Delta G^{\ddagger} = \Delta G^{\ddagger}(pNP-glycoside) - \Delta G^{\ddagger}(pNPGal);$

values of ΔG^{\ddagger} were calculated with Equation 2.4



FIGURE 2.6: Structures of the pNP-substrates of Table 2.1. Shaded areas indicate how each glycoside differs from pNPGal.

viewed as approximate values.

The rate determining step of catalysis with DNP- and pNPGal

There are two steps associated with the catalytic mechanism of hydrolysis by BgaX: glycosylation and deglycosylation (Figure 1.2). The determination of which step is rate limiting for a particular substrate can be achieved by comparison of the k_{cat} values for different nitrophenyl derivatives. The pK_a values of the activated leaving groups of DNP- and pNPGal differ by 3.4 units $(pK_a^{DNP} = 3.8, pK_a^{pNP} = 7.2)^2$. Table 2.1 reveals that k_{cat} is essentially the same for DNP- and pNPGal, indicating that the leaving group ability of the aglycone has no effect on the rate of hydrolysis. This could only be true if glycosylation is not rate limiting; since this step involves the displacement of the aglycone, the rate should be dependent on its pK_a. The rate determining step for the hydrolysis of DNP- and pNPGal with BgaX is therefore most likely deglycosylation, but factors such as the initial association of the substrate with the enzyme, and conformational changes of the protein during catalysis may also play significant roles. A more detailed investigation involving kinetic studies with a series of substrates and measurement of kinetic isotope effects would, however, be needed to confirm this hypothesis.

2.2.5 Glycone specificity

Although the rate determining step for the hydrolysis of DNP- and pNPGal by BgaX was examined, a similar study was not done for the remaining glycosides used as substrates (Figure 2.6) as their DNP- derivatives were not available. Since the rate determining step was not known for these remaining substrates, the comparison of k_{cat}

²See Materials and Methods.

values associated with the hydrolysis of each is of little use since this parameter may reflect different steps of catalysis in each case. Similarly, since k_{cat} is a component of K_m (Figure 2.1), the pseudo dissociation constant should also not be considered as an adequate value on which to base analyses.

A better parameter by which to examine the glycone specificity of BgaX is k_{cat}/K_m since it reflects the overall efficiency of catalysis; that is, it is a measure of the energy barrier that exists between the free enzyme plus substrate and the transition state of the first irreversible step (ΔG^{\ddagger}). Values of $\Delta \Delta G^{\ddagger}$ (relative to ΔG^{\ddagger} of pNPGal) were calculated with the use of k_{cat}/K_m and Equation 2.4 in order to examine the effect of changes in substrate structure on enzyme catalysis (Table 2.1).

$$\frac{k_{cat}}{K_m} = \frac{\mathbf{k}T}{h} e^{-\Delta G^{\ddagger}/RT} \tag{2.4}$$

where	\mathbf{k}	=	Boltzmann's constant
	R	=	the molar gas constant
	h	=	Planck's constant
and	T	=	temperature

There are two predominant factors with respect to the association between substrate and enzyme that contribute to the efficiency of catalysis: (i) electronic factors affect the formation and breakdown of the oxocarbonium ion-like transition state, and (ii) various non-covalent interactions between the enzyme and substrate directly affect binding. For the substrates used in the study of BgaX, it is difficult to determine the contribution of each of these effects arising from a change in substrate structure. Therefore, the calculated values of $\Delta\Delta G^{\ddagger}$ are representative of the combination of both electronic and binding factors that affect catalysis. The greatest values of $\Delta\Delta G^{\ddagger}$ were seen for pNPGlc and pNPXyl ($\Delta\Delta G^{\ddagger} = 26-28 \text{ kJ/mol}$, 6.3–6.8 kcal/mol). These results illustrate a determinant for the specificity of a glycosidase. That is, the epimerization at a single position has a very significant effect on the efficiency of catalysis. These effects are likely the result of the loss of stabilizing interactions between the enzyme active site and the epimerized alcohol, and the subsequent introduction of steric interference.

The ability of BgaX to use pNPGlc and pNPXyl as substrates, albeit in a diminished capacity, may reflect a distant ancestral relationship to a cellulase or xylanase as was previously postulated by Taron and co-workers based on amino acid sequence similarities (§1.4) [31]. It should be noted however, that it has been reported that BgaX is not capable of hydrolyzing terminal β -(1 \rightarrow 4) linked glucose or xylose or β -(1 \rightarrow 3) linked glucose from oligosaccharides [31], indicating that the activity of BgaX with pNPGlc and pNPXyl is likely greatly facilitated by the inclusion of a good leaving group.

The comparison of $\Delta\Delta G^{\ddagger}$ associated with pNPGlc and pNPXyl with that of pNP- α -L-Ara illustrates the care with which kinetic data should be analysed. These substrates differ structurally only by the presence or absence of the hydroxymethyl at the 6-position, and epimerization at the 4-position (Figure 2.6). A preliminary analysis of the $\Delta\Delta G^{\ddagger}$ values calculated for pNPGlc and pNPXyl suggests that the epimerization at the 4-position contributes 26 kJ/mol (6.3 kcal/mol) to the reaction barrier, while elimination of the 6-position hydroxymethyl adds an additional 2 kJ/mol (0.5 kcal/mol). However, pNP- α -L-Ara, which only lacks the hydroxymethyl at the 6-position as compared to galactose, has a $\Delta\Delta G^{\ddagger}$ value of 20 kJ/mol (4.8 kcal/mol). Therefore, in this particular study, the contributions of each discrete structural change within the substrate to the magnitude of the reaction barrier are not

CHAPTER 2 The Characterization of X. manihotis β -Galactosidase 29

additive; indeed, they are somewhat compensatory. Values of $\Delta\Delta G^{\ddagger}$ must therefore be analysed in terms of alterations in substrate structure as a whole.

The $\Delta\Delta G^{\ddagger}$ for pNPGalNAc was comparable to that for pNPGlc and pNPXyl (26 kJ/mol, 6.2 kcal/mol). It has been suggested that substitutions at the 2-position of a substrate are particularly detrimental to enzymatic hydrolysis due to conformational changes incurred during catalysis [8, 39]. Upon binding, the ground state chair conformation of the substrate is distorted into a half-chair to achieve the oxocarbonium ion-like transition state. This conformational change requires the direct repositioning of the substituent at the 2-position. Therefore, any alterations made at this site would have an intimate effect on the formation of the required transition state. The substantial value of $\Delta\Delta G^{\ddagger}$ calculated for pNPGalNAc therefore most likely reflects this increased sensitivity of BgaX with respect to alterations at the 2-position of the substrate. The larger size of this substituent as compared to that of an hydroxyl is presumably responsible for disfavoring the conformational alteration of the substrate within the enzyme's active site to impede the formation of the transition state.

The lowest values for $\Delta\Delta G^{\ddagger}$ were seen for *pNPFuc* (15 kJ/mol, 3.6 kcal/mol) and *pNP-* α -L-Ara (20 kJ/mol, 4.8 kcal/mol). The reduced catalytic efficiency observed with these substrates is most likely dominated by the loss of hydrogen bonding interactions between the 6-hydroxyl and the active site of BgaX.

2.2.6 Inactivation by 2FDNPGal

The activity of BgaX in the presence of 2FDNPGal decreased in a time-dependent fashion according to pseudo first order kinetics (Figure 2.7 (a)):

$$\frac{v}{V_o} = A\left(e^{-k_{inact}^{obs}t}\right) + C \tag{2.5}$$

where v/V_o = the measured fraction of enzymatic activity at time t k_{inact}^{obs} = the observed rate of inactivation and A, C = constants

Saturation of BgaX with respect to inactivator was not achieved since, at high concentrations of 2FDNPGal, inactivation was too rapid to monitor reliably. However, values of $K_i = 0.031 \pm 0.005$ mM and $k_{inact} = 0.030 \pm 0.004 \text{s}^{-1}$ were extrapolated



FIGURE 2.7: (a) The time-dependent first order (Equation 2.5) decrease in activity of BgaX upon incubation with • 1.15 μ M, □ 2.30 μ M, ■ 11.5 μ M, \triangle 23.0 μ M, and ▲ 46.0 μ M 2FDNPGal. (b) The double reciprocal plot $\left(\frac{1}{k_{inact}^{obs}} = \frac{1}{[2FDNPGal]} \left(\frac{K_i}{k_{inact}}\right) + \frac{1}{k_{inact}}\right)$ using the k_{inact}^{obs} values from (a).

from the double reciprocal plot of Figure 2.7 (b). While these individual values are not too reliable, the ratio (k_{inact}/K_i) , which can be determined from the slope, is. The consequences of replacement of the alcohol at the 2-position by fluorine can be estimated by the comparison of the rate of glycosylation, reflected by k_{inact}/K_i $(1.0\pm0.2 \text{ mM}^{-1}\text{s}^{-1})$ and k_{cat}/K_m for DNPGal (500±100 mM^{-1}\text{s}^{-1}). This comparison illustrates that the fluorine substitution effectively reduces the rate of the first step of catalysis by 500-fold.}

Protection from inactivation

The galacto-configured pyridoimidazole (galacto-imidazole) (Figure 2.8), an analog of nagstatin, a naturally occurring inhibitor isolated from *Streptomyces amakusaensis* [40], has been shown to be a good competitive inhibitor of β -galactosidases from almond and *E. coli* (IC₅₀ = 0.0016–0.10 µg/mL) [41]. This inhibitor was demonstrated to also be a competitive inhibitor of BgaX. The Lineweaver-Burk plot of the initial reaction velocity at five different concentrations of *p*NPGal in the presence and absence of galacto-imidazole is shown in Figure 2.9 (a). To within error (intercepts = 15 ± 4 with galacto-imidazole and 22 ± 7 for the control), the two lines intersect on the y-axis, a pattern characteristic of competitive inhibitors [42].

An approximate K_i value of 15 nM for galacto-imidazole with BgaX was calculated



FIGURE 2.8: Galacto-configured pyridoimidazole (galacto-imidazole).





FIGURE 2.9: (a) Lineweaver-Burk plot of initial reaction velocity versus concentration of PNPGal, in the presence \Box and absence \bullet of 7.2 nM galacto-imidazole, where slope_i and slope_o are the respective slopes of each linear fit. (b) Fractional loss of activity of BgaX over time caused by 46.0 μ M 2FDNPGal in the presence \Box and absence \bullet of 40 nM galacto-imidazole; data sets were fit to a first order expression (Equation 2.5).

from the following relationship³ using the resultant slopes of the data of Figure 2.9 (a):

$$K_i = \frac{[I]}{slope_i/slope_o - 1} \tag{2.6}$$

Galacto-imidazole was shown to effectively protect BgaX from inactivation by 2FDNPGal. The fractional loss in activity of BgaX by 2FDNPGal in the presence and absence of galacto-imidazole (Figure 2.9 (b)) yielded a rate constant ratio (k_{ip}^{obs}/k_i^{obs}) of 0.43, which is in close agreement with the value of 0.48 predicted by Equation 2.3. Since the rate of inactivation by 2FDNPGal decreased by the expected amount in the presence of a competitive inhibitor, the 2-fluoro-glycoside is presumably active-site directed.

³See Appendix A for derivation.

Catalytic competance of the inactivated enzyme

Once BgaX was inactivated by 2FDNPGal and the excess inhibitor removed, the enzyme spontaneously regained hydrolytic activity over time according to pseudo first order kinetics (Equation 2.7) with a measured rate constant (k_{react}) of 0.018 hr^{-1} (Figure 2.10).

$$\frac{v}{V_o} = A(1 - e^{-k_{react}t}) + C$$
(2.7)

The observed spontaneous enzyme reactivation supports the belief that inactivated BgaX exists as a catalytically relevant species; that is, reactivation is attributed to the eventual turnover of the inhibitor-enzyme complex. Therefore, k_{react} can be described as the rate constant for the deglycosylation step of the 2-fluoro-glycosyl enzyme intermediate, and as such, can be compared to k_{cat} for DNPGal since it too is related to



FIGURE 2.10: The spontaneous reactivation of 2-deoxy-2-fluorogalactosyl-BgaX. Data was fit to Equation 2.7.

this step of catalysis. The value for k_{react} is 8 x 10⁶ times smaller than k_{cat} , thereby illustrating the effect of the fluorine substitution at the 2-position on deglycosylation. The decrease in the rate of this step upon this substitution is significantly greater than that observed for glycosylation (500-fold). This difference is comparable to that documented for other glycosidases [17, 43, 44], and is presumably due to the greater leaving group ability of DNP as compared to that of the active site carboxylate of the glycosyl enzyme intermediate. The greater oxocarbonium ion-like character of the transition state of deglycosylation versus glycosylation [10] may also contribute to the increased sensitivity of deglycosylation with respect to the fluorine substitution.

The catalytic competence of the inactivated enzyme species was also demonstrated by the increase in the relative rate of enzyme reactivation in the presence of an acceptor (k_{trans}^{obs}) as compared to k_{react} . The greater value for k_{trans}^{obs} indicates that reactivation was accelerated due to transglycosylation. These results are outlined in detail in Chapter 3.

The observed protection by galacto-imidazole of BgaX from inactivation, and comparison of the values of k_{trans}^{obs} and k_{react} , supports the belief that once inactivated by 2FDNPGal, BgaX exists as the 'trapped', yet catalytically competent fluoro-glycosyl enzyme intermediate.

2.2.7 The identification of the nucleophile of X. manihotis β galactosidase

The expected increase in mass of BgaX if 2-deoxy-2-fluoro-galactose is covalently attached at the active site is 165 Da. The mass spectrum of inactivated wild type BgaX yielded an apparent mass of 65558 ± 5 Da, which was 171 Da greater than that of a control of non-inactivated enzyme. This mass difference suggests that the 2-



FIGURE 2.11: TIC of the control (a) and labelled (c) BgaX after peptic digestion. The mass spectrum of the peak at ca. 28 min. of the TIC for the control (b) and labelled (d) enzymes showing the m/z of the major species.

fluoro-glycosyl enzyme intermediate is formed in a 1:1 stoichiometric reaction between BgaX and 2FDNPGal to form a relatively stable species, as anticipated.

Digestion of inactivated histidine-tagged BgaX with pepsin resulted in numerous peptide fragments as shown in the total ion chromatogram (TIC) HPLC trace of Figure 2.11. The TIC of the digested inactivated enzyme was similar to that of the non-inactivated control except for the peaks with an elution time of *ca.* 28 min. The mass to charge ratio (m/z) profiles of the peptide ions in these fractions are similar except for two major signals: m/z = 869.0 in the control, and m/z = 951.0 in the inactivated samples (Figure 2.11 (b), (d)).

The difference between the m/z values for these two peptides is 82.0 which is,

within error, one half of the expected mass increase if 2-deoxy-2-fluoro-galactose is covalently linked to the enzyme (165 Da), suggesting that z = +2 for these peptides. Hence, $951.0 \times 2 - 2H^+ = 1900.0$ Da and $869.0 \times 2 - 2H^+ = 1736.0$ Da were the masses of the 'labelled' and control peptides originating from the inactivated and control enzymes respectively. Therefore, the peptide sequence that bracketed the catalytic nucleophile had a mass of 1736 Da. A search of the primary amino acid sequence of BgaX for a peptide with this mass showed that there are two candidate sequences that contain the putative nucleophile, Glu^{260} (underlined): ²⁴⁸IKFRPDQRPMVG<u>E</u>Y²⁶¹ and ²⁵⁸VG<u>E</u>YWAGWFDHWGK²⁷¹. Both of these sequences contain a single Asp and Glu residue, either of which could potentially act as the nucleophile, based on the conservation of such catalytic residues in other glycosyl hydrolases.

The precise identity of the position of the nucleophile in BgaX was achieved through further analysis of the m/z = 951.0 and 869.0 ions of the labelled and control peptides respectively in the daughter ion scan mode (MS/MS). These parent peptides were each separately fragmented further into two separate segments at three distinct peptide bonds (Figure 2.12). Two of the three daughter fragments of these parent peptides that were cleaved at the carboxyl end (B_n ions) were identified in both the labelled and control samples as ²⁴⁸IKFR²⁵¹ (B₄, m/z = 545.0) and ²⁴⁸IKFRPD²⁵³ (B₆, m/z = 757.0). These two daughter ions originating from the labelled and control species had the same m/z value, indicating that Asp²⁵³ was not attached to the 2F-galactose moiety, and hence, does not function as the catalytic nucleophile. Three daughter ions that were produced by cleavage of the labelled parent peptide at the amino end (Y^{labelled}_n) were larger than fragments cleaved at the same site in the control peptide (Y^{control}_n) by 164.0 ± 0.5: ²⁵²PDQPRMVGEY²⁶¹ (Y^{labelled} m/z = 1355.5; Y^{control}₄ m/z = 1191.5), ²⁵⁴QPRMVGEY²⁶¹ (Y^{labelled} m/z =





FIGURE 2.12: MS/MS of the parent (a) m/z = 869.0 peptide of the control, and (b) m/z = 951.0 peptide of the labelled species. (c) Fragmentation patterns of peptides in general and specifically for the parent peptides from BgaX.

1144.5; $Y_6^{control}$ m/z = 980.0), and ²⁵⁵PRMVG $\underline{E}Y^{261}$ ($Y_7^{labelled}$ m/z = 1015.0; $Y_7^{control}$ m/z = 851.4). Therefore, the amino acid labelled with the 2-fluoro-galactose moiety was contained within residues 255–261. As Glu²⁶⁰ is the only residue with a carboxylic acid side chain in this sequence, it was concluded that this amino acid functioned as the catalytic nucleophile of BgaX.

2.2.8 The nucleophile mutant E260A BgaX

The mutant E260A BgaX was constructed through direct gene manipulation and was overexpressed in *E. coli*. The enzyme was purified to *ca.* 90% purity as estimated by silver stained SDS-PAGE. The BgaX mutant lacked hydrolytic activity, as was expected since the anticipated nucleophilic carboxyl of glutamic acid was replaced with the non-nucleophilic methyl side chain of alanine. Although E260A BgaX did not hydrolyze nitrophenyl glycoside substrates on its own, activity with DNPGal was recovered by the provision of azide. This finding indicates that the active site architecture remains intact in the mutant enzyme to productively bind the substrate. The addition of azide presumably rescues activity by taking the role of the mutated nucleophile of BgaX as shown in Figure 2.13.



FIGURE 2.13: Azide rescue of the activity of E260A BgaX.

Although a thorough analysis of E260A BgaX was not undertaken, the observation that activity was recovered in the presence of azide supports the identification of Glu^{260} as the catalytic nucleophile of BgaX.

2.3 Conclusions

The brief study of the glycone specificity of BgaX illustrates the effect that small changes in substrate structure have on enzymatic activity. Each alcohol group in the substrate influences the productive association with the enzyme to contribute to substrate specificity and the overall efficiency of catalysis. Further, the ability of BgaX to hydrolyze pNPGlc and pNPXyl, coupled with the homology of a region of this enzyme with that of family 10 cellulases and xylanases (Figure 1.10), suggests that this β -galactosidase may have evolved from the shuffling of discrete domains from ancestral glycanases.

The glycoside 2FDNPGal was demonstrated to be an active site-directed, mechanism-based inactivator of BgaX. The substitution of fluorine at the 2-position of DNP-Gal decreased the rates of both steps of catalysis, with the rate of deglycosylation decreasing several orders of magnitude more compared to that of glycosylation. The net effect of the activity of BgaX with 2FDNPGal was a trapped, yet catalytically competent 2-fluoro-glycosyl enzyme intermediate with a half life of *ca*. 40 hr. The longevity of this intermediate allowed its analysis by mass spectrometry both in the intact and digested enzyme to permit the elucidation of Glu^{260} as the active site nucleophile of BgaX. The assignment of this catalytic residue was further confirmed by the construction of the E260A BgaX mutant and the rescuing of its activity in the presence of azide. Future studies should fully investigate the capability of this and other nucleophile mutants of BgaX (eg. serine) to function as a glycosynthase.

This finding confirms the previous prediction of the location of this active site residue that was based on primary sequence similarities with other members of family 35, and agrees with the assignment of the nucleophile of the human β -galactosidase. Therefore, adequate comparisons can be made between such diversely related organisms as human and X. manihotis in order to identify catalytically important residues, provided of course that significant sequence similarities exist elsewhere in the enzymes. The aforementioned weak sequence homology among members of glycosyl hydrolase family 35 in the vicinity of the catalytic nucleophile is interesting given the importance of this active site residue. Studies that would further investigate how enzymes from such a diverse collection of organisms could be grouped in the same family would be of interest.

CHAPTER 3

Screen of Potential Transglycosylation Acceptors

3.1 Introduction

3.1.1 Screening of potential acceptors

It was noted previously (§1.3.1) that in order to maximize yields of transglycosylation reactions, it is desirable to be able to probe the specificity of the aglycone site of a glycosidase to identify the most suitable acceptors for the particular wild type enzyme or glycosynthase. This could be accomplished by the screening of a number of compounds to determine their degree of favorable association with the aglycone site. We have developed such a procedure with the use of 2-deoxy-2-fluoro-glycosides that is simple and has the capacity to rapidly screen a large number of potential acceptors.

A glycosidase inhibited by a 2-deoxy-2-fluoro-glycoside will eventually be reactivated through the nucleophilic attack on the trapped glycosyl enzyme intermediate by water or an acceptor. It has been previously demonstrated that the rate of such reactivation varies in the presence of different acceptors. The 2-fluoro-glucosyl enzyme species of Agrobacterium sp. β -glucosidase (Abg) was previously demonstrated in this laboratory to turn over 2700-fold faster in the presence of $pNP \beta$ -glucoside as compared to buffer alone, and only 600-fold faster in the presence of cellobiose [12]. Since the trapped fluoro-glycosyl-enzyme species very closely resembles the same intermediate formed in a regular transglycosylation reaction, the influence an acceptor has on the rate of enzyme reactivation should be indicative of the specificity of the aglycone site with respect to that particular acceptor. That is, the greater the degree of productive binding of the acceptor with the aglycone site, the greater the developed screening methodology: the specificity of the aglycone site of a glycosidase with respect to transglycosylation can be examined by comparison of the rates of reactivation of the inhibited enzyme in the presence of various potential acceptors. Once the most suitable acceptors for an enzyme are determined, they can be used with non-fluorinated donors for transglycosylation reactions. By identifying the best acceptors for a glycosidase, the most appropriate enzyme can be chosen as a catalyst for a desired reaction.

The basis of this screening methodology was briefly examined by D. Stoll in collaboration with this laboratory [28]. He compared the rates of reactivation of *Cellulomonas fimi* β -mannosidase Man2A that had previously been inactivated with 2deoxy-2-fluoro- β -mannosyl fluoride in the presence of nine glycosides. It was found that of these potential acceptors, only four induced a significant reactivation rate relative to a control in buffer alone. It is evident from this acceptor 'hit' rate of *ca*. 50% that in order to fully examine aglycone specificity, it is desirable to screen a large number of compounds. Furthermore, there are vast numbers of compounds that one may wish to use as acceptors, even if only oligosaccharides are taken into consideration. However, virtually any compound that requires glycosylation can be viewed as a potential acceptor for a transglycosylation reaction. For both of these reasons, it was desirable to be able to evaluate the aglycone specificity of a glycosidase in a high throughput screen. The development of 96-well plate readers for kinetic analyses has made this a relatively simple task. A glycosidase inhibited by a 2-deoxy-2-fluoroglycoside is added to each well of a 96-well plate that contains a different potential acceptor to be screened. After incubation for a period of time, each well is assayed for enzyme activity arising from reactivation by the addition of a nitrophenyl-glycoside substrate. Those wells that show greater enzyme activity relative to a control incubated in buffer alone indicate a greater reactivation rate, and thus contain those compounds that would be the best candidates to act as acceptors in a transglycosylation reaction.

Once these 'positive hits' were identified, we desired a more detailed look at the kinetics of enzyme reactivation in the presence of each positively screened compound. This was accomplished by the same strategy described above for the 96-well plate screening process. In this instance however, aliquots are removed from each reactivation mixture over time and individually assayed for enzyme activity. In this way, the effect of each positive hit on enzyme reactivation could be more accurately determined.

3.1.2 Preparative-scale transglycosylation reactions

Once the described acceptor screening methodology was applied to a glycosidase, it was necessary to demonstrate the utility of the identified positive hits with respect to functioning as acceptors in transglycosylation reactions. Four of the seven enzymes that were used in the screening process were used as catalysts in transglycosylation reactions with one of the identified positive hits acting as an acceptor in each case. The reactions were performed on a preparative scale and the resultant transglycosylation products isolated and characterized and compared to other such products from the literature.

The characterization of the transglycosylation products involves the determination of two parameters: (i) the number and identity of monosaccharide units; and (ii) the nature of the glycosidic linkage between these units. The first parameter can be easily determined via ¹H NMR, as the types of possible monosaccharides in the product are known, and there are numerous spectral data in the literature available for comparison and subsequent identification.

The determination of the linkage between saccharide units is not as easily determined. It was mentioned briefly in §1.3.1 that the precise mode of interaction between the aglycone site of a glycosidase and the acceptor in a transglycosylation reaction will determine which hydroxyl of the acceptor is in a position to attack the glycosylenzyme intermediate. Depending on the positioning of the acceptor in the aglycone site, a glycosidic bond may be formed in the transglycosylation product that does not correspond to the preferred linkage of the enzyme with respect to hydrolysis. There is also the possibility of a glycosidase catalyzing the formation of *several different linkages* in a single transglycosylation reaction. This phenomenon has been demonstrated for a number of enzymes, and again, is linked to the particular acceptor used in each reaction [12, 20, 28, 45].

The developed acceptor screening strategy identifies those acceptors that are best suited for a particular glycosidase for transglycosylation reactions. It does not however predict the linkages that will be formed when such acceptors are used in actual reactions. It was therefore necessary to identify the types of glycosidic bonds formed in the preparative-scale transglycosylation products by conventional means.

Two of the most common ways in which to identify the linkage between saccha-

ride units involve the comparison of HPLC retention times with those of standards, or methylation analysis. If a transglycosylation reaction is monitored by HPLC, the formation of the products is readily apparent. However, in order to identify a product in this manner, it is necessary to have a standard sample of the product in question in order to make comparisons between column retention times. While it may be possible to have access to a variety of such compounds, the question of regiochemistry with respect to transglycosylation means that any one reaction has the potential to result in the formation of a number of isomers. It would be very time consuming to chemically synthesize all potential products for their use as standard compounds. In methylation analysis, the linkage type is identified by methylating all free hydroxyls in the oligosaccharide in question and subsequent acid hydrolysis. The non-anomeric methyl ethers are stable to this hydrolysis, whereas glycosidic bonds are readily cleaved, leading to methylated monosaccharides that specifically reflect the linkage and sugar units of the parent oligosaccharide. Analysis of these resultant fragments is done by HPLC or GC. However, as in the case with straight HPLC analysis of transglycosylation reactions, it is necessary to have access to a large number of standards with which to make comparisons for identifications. If the necessary standards were available for this study, either of these strategies to identify the glycosidic linkages in the synthesized transglycosylation products may have been suitable. Yet the lack of such standards necessitated an alternative method.

Over the past decade, developments in carbohydrate NMR have made it possible to determine the linkage types in an oligosaccharide through non-destructive means without the explicit need for standards with which to make comparisons. Spectra are commonly taken of the peracetylated oligosaccharide in a deuterated organic solvent, which results in narrower, more resolved signals as compared to the non-derivatized sample in aqueous solution [46]. The following discussion refers to oligosaccharides that have been derivatized in this manner.

A common method to determine the site of a glycosidic bond is to compare the signals of each ring proton of a peracetylated monosaccharide contained within an oligosaccharide to that of the isolated unit. Glycosylation of a sugar residue typically shifts the chemical signal of the proton at the linkage site by -0.2 to -0.26 ppm, while the rest of the protons are relatively unaffected [47]. These 'glycosylation shifts' can thus be used to determine the site of a glycosidic bond, provided that *several* substantial proton shifts are not observed, as is often the case. In the case that glycosylation shifts do not provide adequate insight into linkage types in oligosaccharides, 2D NMR experiments are often of great help. Heteronuclear multiple bond correlation (HMBC) NMR experiments identify 3- and 4- bond correlations within compounds, a sensitivity that is sufficient to trace the connectivity of two sugars via coupling between ring protons at the glycosidic bond.

Recently a simple, yet powerful, methodology to elucidate linkage sites in carbohydrates has been developed that is based on ${}^{1}H{-}{}^{13}C$ couplings [48, 49, 50, 51]. The natural abundance of ${}^{13}C$ is too low to normally detect any such couplings between ring protons and the carbonyl carbon of the adjacent acetyl group. However, the use of an isotopically enriched reagent to derivatize the oligosaccharide, such as $1{-}^{13}C{-}$ acetyl chloride, enables the detection of a J-coupling between the ring protons and their adjacent carbonyl carbons. These additional splittings are in the range of 2.5-4.7 Hz and are thus distinguishable in 1D- ${}^{1}H$ spectra [51]. Protons that were not initially adjacent to an hydroxyl in the non-derivatived sugar, such as the anomeric proton of the non-reducing end monosaccharide and those adjacent to glycosidic linkages, do not experience this additional splitting. Therefore, the connectivity between carbohydrate units can be elucidated by the determination of which proton signals did not experience an additional splitting from the ¹³C-enriched acetyl groups.

The methods of glycosylation shift, and derivatization with ¹³C-enriched acetyl groups were used to determine the glycosidic linkages in the transglycosylation products synthesized in this study.

3.2 Results and discussion

3.2.1 Large scale screening of potential transglycosylation acceptors

The described transglycosylation acceptor screening methodology was applied to seven glycosidases: BgaX, *Bacillus circulans* β -galactosidase (BgaC), *Agrobacterium* sp. β -glucosidase (Abg), *Cellulomonas fimi* β -mannosidase (Man2A), *Streptomyces lividans* endoglucanase (CelB), *Cellulomonas fimi* xylanase/glucanase (Cex), and human β -glucuronidase (HBG). The complete list of carbohydrates that were screened and the positive hits associated with each are provided in Tables 3.1 and 3.2. The mono- and disaccharides that were chosen were selected for several reasons: they are representative of carbohydrates that comprise many common synthetic targets; they are interrelated by simple substitutions or epimerizations, thereby illustrating the effect of subtle structural changes in the acceptor on aglycone binding; and they were readily available in the laboratory.

Compounds were listed as a positive hit if they induced a higher degree of enzyme reactivation at the time of assaying as compared to that of a control in buffer alone. However, some of those glycosides screened induced a lower amount of enzyme reactivation with respect to a control. Two of the possible explanations as to why this may occur follow:

(i) Since the K_m of each potential acceptor with the aglycone site of each enzyme was not known, the inactivated glycosidases were incubated with a high concentration of the glycoside to be tested (~30 mM or 60% saturation)¹. Upon measurement of enzymatic activity to determine the amount of 2-fluoro-glycosyl enzyme intermediate that had turned over, it is possible that reactivated enzyme was inhibited to some degree by the excess of acceptor present. This inhibition could make it appear as though less enzyme had been reactivated as compared to a control. A simple way by which to test this theory would be to perform a control using non-inactivated enzyme. After incubation, a decrease in enzyme activity would identify those potential acceptors that inhibit the enzyme in question.

(ii) The turnover of the 2-fluoro-glycosyl enzyme species may appear to be slower in the presence of a potential acceptor compared to that of a control due to the re-inactivation of the enzyme by the synthesized transglycosylation product. If the glycosidase in question has a high affinity for the transglycosylation product, it is possible that this product will not diffuse out of the enzyme's active site, but will remain bound. The reactivated glycosidase may then remain inhibited in a competitive manner by the transglycosylation product. Alternatively, the enzyme may cleave the newly formed glycosidic bond to again inactivate the enzyme as shown in Figure 3.1.

Whether points (i) and/or (ii) hold true, the observation that the incubation of a potential acceptor with a 2-fluoro-glycosyl enzyme species decreases the rate of en-

¹Stock solutions of ≈ 50 mM in water were made for all acceptors; some of these solutions were saturated. Dilution of these stocks for the screening procedure resulted in final acceptor concentrations of ≈ 30 mM or 60% saturation.



FIGURE 3.1: Reactivation of a glycosidase and subsequent re-inactivation by its transglycosylation product.

zyme reactivation is a clear indication that this compound would not be among the best candidates to act as acceptors in a regular transglycosylation reaction. If the presence of the acceptor inhibits the glycosidase, the transglycosylation pathway will be blocked. Also, if the enzyme in question has a high affinity for the transglycosylation product it may be inhibited by active site competition, or the newly formed glycosidic bond may be hydrolyzed, lowering the overall product yield. It should be stressed that for the purposes of developing the screening procedure only those compounds that would act as the *best* acceptors for a particular glycosidase were considered. Therefore, those potential acceptors that induced a degree of reactivation that was less than that seen for the control were treated as negative results and were not analyzed further.

The total number of positive hits for each glycosidase (Table 3.1) were relatively low (7-16 out of a possible 45). Several general trends were seen with these positively screened carbohydrates:

(i) Of the potential acceptors screened, any glycosides were generally better reactivators than their parent carbohydrates. This was evident through observations ٢

POTENTIAL ACCEPTORS ^a	Glycosidase						
	BgaX	BgaC	Abg	Man2A	Cex	CelB	HBG
arabinose fructose galactose pNP eta -galactoside		+	+				+++++++++++++++++++++++++++++++++++++++
2-deoxygalactose N-Ac-galactosamine glucose p NP β -glucoside	+	+ +	+++			+	+
Ph β -glucoside $pAcPh \beta$ -glucoside D-glucal 1-cyanoglucal	+++++++++++++++++++++++++++++++++++++++	+ + + +	+	+ +	+	+	
$N ext{-}\operatorname{sc-glucosamine} N ext{-}\operatorname{Ac-glucalamine} N ext{-}\operatorname{Ac-glucosamine-} eta ext{-}\operatorname{NH}_2 p ext{NPGlcA}$	+	+ +					+
2-deoxyglucose p NP β -mannoside p NP- α -xyloside p NP β -xyloside	++	. +	+	+ + +	+	+	
$o{ m NP}\ eta{ m s}$ -xyloside Me eta -xyloside xylose $p{ m NP}\ eta$ -cellobioside			. · +	+ + +	+		+
DNP β -cellobioside Me α -cellobioside Me β -cellobioside lactose		+ + +	, .+	+ +	+ +	+ +	
$p{ m NP}\ eta-{ m lactoside}\ { m maltotriose}\ { m gentiobiose}\ { m Gal-eta-(1-3)-Glc-eta-SPh}$	+	+ +	+ +			+	+
$\begin{array}{c} \operatorname{Gal-}\beta\text{-}(1\text{-}4)\text{-}\operatorname{Glc-}\beta\text{-}\operatorname{SPh}\\ \mathrm{DNP}\ \beta\text{-}\operatorname{fucoside}\\ 2\text{,}5\text{-}\operatorname{DNP}\ \beta\text{-}\operatorname{xylobioside} \end{array}$	· +			+		+	+

TABLE 3.1: Positive hits from screened glycosides.

^aAll glycosides have the D- configuration unless noted otherwise; see the List of Abbreviations for full carbohydrate names.

and .

Glycoside
p NP α -L-arabinoside
α -L-fucose
$o \mathrm{NP} \ \beta$ -galactoside
D-galactal
D-galacturonic acid
acetyl β -glucoside
mannose
$p \text{NP} \alpha$ -mannoside
N-Ac-mannosamine

TABLE 3.2 :	Glycosides	that	are	negative	hits	for	all	enzymes.
---------------	------------	------	-----	----------	------	-----	-----	----------

such as $pNP \beta$ -galactoside being a positive hit for BgaC and Abg, yet galactose was not. Similar trends were seen for aryl glucosides and aryl xylosides; these carbohydrates were positively screened with BgaX, BgaC, and Abg, whereas negative results were obtained for all three enzymes with the parent compounds glucose and xylose. This general trend was seen previously in this laboratory for the reactivation of Abg and Man2A. The turnover of the 2-fluoro-glucosyl enzyme species of Abg increased 2700-fold in the presence of $pNP \beta$ -glucoside, yet an increase in reactivation rate was not seen with glucose [12]. Similarly, the reactivation of Man2A was significantly increased relative to a control in the presence of $pNP \beta$ -mannoside, but not with the parent compound mannose [28]. This finding is most likely the result of favorable interactions between the aryl substituent and the enzymes' aglycone site. One exception to this trend was seen with HBG, which showed positive results for galactose and xylose, but not for any of their aryl derivatives, suggesting that similar interactions are disfavoured in the aglycone site of this enzyme.

The finding that in general, aryl glycosides are better enzyme reactivators, and hence most likely the best candidates to act as acceptors, is especially useful from a synthetic viewpoint. If a nitrophenyl glycoside is used as an acceptor in a transglycosylation reaction, the product will necessarily have a nitrophenyl substituent at its reducing end. This provides an activated center for further chemistry if the desired target requires additional synthetic steps.

(ii) Although aryl glycosides were generally preferred as acceptors by comparison with their parent carbohydrates, the reactivation of some enzymes was sensitive to the substitution of the phenyl substituent. It appears that *para*-nitro substituted phenyl glycosides were generally better reactivators relative to their *ortho*- derivatives: $pNP \beta$ -galactoside and $pNP \beta$ -xyloside were positive hits for BgaC and Abg, yet $oNP \beta$ -galactoside and $oNP \beta$ -xyloside were not. Similarly, BgaX accepted the *para*substituted nitrophenyl β -xyloside, but not the *ortho*-substituted derivative. These examples illustrate the significant effect that small changes in acceptor structure have on productive binding to the aglycone site of the glycosidases studied. If an activated glycoside is to be used as a transglycosylation acceptor, it would be beneficial to determine which activated group would be best accommodated in the enzyme's aglycone site to ensure maximal product yields.

Again, however, an exception to this general trend was seen with Man2A, which had positive hits with both p- and o-NP β -xyloside, as well as with p-NP α -xyloside. Thus, Man2A either has a high tolerance with respect to substitution on the phenyl moiety, or associates with xylose in such a productive manner as to overcome any unfavorable interactions that may be incurred by an *ortho*- substituted aryl group or a change in the anomeric configuration of the acceptor.

3.2.2 The kinetics of enzyme reactivation via transglycosylation

Those compounds that were identified as positive hits in the large scale screen were studied further to determine the effect of their presence on the kinetics of enzyme reactivation. The fractional gain in activity of the 2-fluoroglycosyl-enzyme species with each potential acceptor over time (Figure 3.2) followed the same first order rate profile (Eqn. 3.1) seen for spontaneous reactivation (Figure 2.10).

$$V/V_o = A(1 - e^{-k_{react}^{obs}t}) + C$$
 (3.1)

where
$$V/V_o$$
 = the fractional gain in
enzyme activity at time t
 k_{react}^{obs} = the observed rate of reactivation
due to transglycosylation
and C and A = constants

It is important to note that when comparing first order reactivation profiles related to different acceptors for a given glycosidase, there are two dominant parameters: both k_{react}^{obs} and the upper limit of reactivation (A). The rationale behind the inadequacy of considering only k_{react}^{obs} when ranking acceptors is illustrated by the reactivation profiles of BgaX depicted in Figure 3.2.

In the presence of pNPGal the reactivation rate of BgaX is slightly greater than that for pAcPhGlc ($k_{react}^{obs} = 0.072 \text{ hr}^{-1}$ and 0.068 hr^{-1} respectively). However, from the examination of enzyme reactivation in the presence of these glycosides, it is readily apparent that the upper limit of reactivation for pNPGal is substantially less than that for pAcPhGlc. That is, in the presence of pNPGal, much less BgaX overall appears to be reactivated. There are several possible explanations as to why this may



FIGURE 3.2: The reactivation of inactivated BgaX in the presence of \bullet pAcPhGlc and \circ pNPGal.

occur:

(i) Over the course of the experiment the excess acceptor may be hydrolyzed by enzyme that has been reactivated. It was discussed in the previous section that aryl glycosides were significantly better acceptors than their parent glycosides. Therefore, if such an aryl glycoside is hydrolyzed, the produced sugar would not be as good an acceptor. Therefore, the concentration of 'good' acceptor is constantly decreasing during the experiment. While this result is obvious for such acceptors as pNPGalwith BgaX, this scenario may also be applicable in other cases where the hydrolytic activity of the enzyme under study was not tested with the potential acceptors. Therefore, there is one drawback to this screening strategy: it may not reliably identify a glycoside as an acceptor if it is a substrate for the enzyme. It would therefore be beneficial to know if the potential acceptors to be screened are substrates to avoid false negative results. Interestingly, this shortcoming does not apply to *all* substrates for a particular enzyme, as is demonstrated by pNPGal, pNPXyl and pNPMan being identified as positive hits for BgaC, Abg and Man2A respectively. This apparent discrepancy is most likely a consequence of a higher rate of reactivation versus hydrolysis at high concentrations of substrate. For BgaC and Abg, at concentrations of substrate that are several times K_m (*ca.* x 8) the initial reaction velocity versus substrate curve deviates from the classical Michaelis-Menten expression, an indication that transglycosylation is occurring [10, 52]. The reactivation profiles were generated in the presence of 30 mM of acceptor, a concentration that would have been sufficient for substantial transglycosylation (reactivation) to occur versus hydrolysis.

(ii) As was noted in the previous section, the newly formed transglycosylation product may be a good inhibitor of the enzyme under study. The transglycosylation product may be re-bound by the newly reactivated enzyme to again re-inactivate the glycosidase. This would make it appear as though less enzyme overall was reactivated.

Therefore, there are several reasons why the upper limit of reactivation may differ between reactivation profiles involving different acceptors. It is therefore inadequate to directly contrast values of k_{react}^{obs} to make comparisons between the compounds screened since this paramter may reflect additional steps aside from reactivation. In order to circumvent this problem, the initial velocity of each reactivation at t = 0 (V_o^{react}) was calculated from each reactivation profile as the product of the upper limit of reactivation and k_{react}^{obs} ².

Several reactivation profiles deviated from a first order expression in that after increasing over time, instead of reaching a limit, it decreases, in some cases quite dramatically (Figure 3.3). As the observed deviation occurs after several hours, it could be the result of the newly synthesized transglycosylation product inhibiting the reactivated enzyme. Further experimentation is required however to confirm this

²See Appendix A for the full derivation of V_o^{react} .


FIGURE 3.3: The reactivation profile of Man2A with pNP β -xyloside.

hypothesis.

Although the transglycosylation products associated with reactivation profiles such as that depicted in Figure 3.3 may inhibit the enzyme after a significant amount is synthesized, the glycosides associated with these profiles were nonetheless considered to be good acceptors. In these cases then, a first order expression was fit to the lower portion of the reactivation profile, and the extrapolated value of V_o^{react} ranked among the other 'well behaved' acceptors. For this reason, there is a substantial amount of error associated with some of the values of V_o^{react} .



FIGURE 3.4: Summary of the reactivation profiles of BgaX. Values of V_o^{react} have been scaled relative to the maximum.



FIGURE 3.5: Summary of the reactivation profiles of BgaC. Values of V_o^{react} have been scaled relative to the maximum.



FIGURE 3.6: Summary of the reactivation profiles of Man2A. Values of V_o^{react} have been scaled relative to the maximum.



FIGURE 3.7: Summary of the reactivation profiles of Abg. Values of V_o^{react} have been scaled relative to the maximum. It was necessary to use a logarithmic scale in order to adequately compare V_o^{react} values of the acceptors.



FIGURE 3.8: Summary of the reactivation profiles of CelB. Values of V_o^{react} have been scaled relative to the maximum.



FIGURE 3.9: Summary of the reactivation profiles of Cex. Values of V_o^{react} have been scaled relative to the maximum.



FIGURE 3.10: Summary of the reactivation profiles of HBG. Values of V_o^{react} have been scaled relative to the maximum.

For each glycosidase, excluding HBG, the V_o^{react} values associated with the disaccharide acceptors (and especially those derivatized with an aryl group) generally ranked highest among each group of positive hits. This was anticipated, as the additional sugar moiety is expected to provide further favorable contacts with the aglycone site of each enzyme. The glycosidic linkage between sugar units in the disaccharide acceptors did not always significantly influence the degree of productive binding between acceptor and aglycone site. For the glycosidases CelB and BgaX, values of V_o^{react} associated with enzyme reactivation are similar for the acceptors Gal- β -(1-4)-Glc- β -SPh and Gal- β -(1-3)-Glc- β -SPh. A dependence was seen however on the internal acceptor linkage with the reactivation of BgaC and HBG. These enzymes showed an initial positive hit for the β -(1-3) linked Gal-Glc thiophenyl glycoside, yet reactivation was less than that of the control for the β -(1-4) isomer in the large scale screen (Table 3.1).

A more substantial effect on the reactivation of BgaC was seen in the presence of maltotriose whose value for V_o^{react} was *ca.* 34-fold greater than that of the control, and *ca.* 4-fold greater than any other acceptor. This particular result is interesting in that it illustrates the need to screen a large number of compounds when determining the most appropriate transglycosylation acceptors for a given glycosidase. It was not anticipated that there would be substantial binding between the α -(1-4) linked glucose units of maltotriose and the aglycone site of an enzyme optimized for the hydrolysis of β -(1-3) linked terminal galactose residues.

The best acceptors for the remaining glycosidases studied were much less surprising: all are derivatives of cellobiose, or for the case of BgaX, the β -(1-3) and β -(1-4) linked Gal-Glc thiophenyl glycoside isomers.

The values of V_o^{react} associated with D-glucal and pNPCell for the cellulase Cex

are slightly lower than that for the control. These findings are believed to be the result of the aforementioned problem of the acceptors also acting as substrates for the glycosidase. Over the course of reactivation, pNPCell was likely hydrolyzed, and hence converted to less effective acceptor. Upon assaying for enzyme reactivation, the presence of cellobiose likely competed for the active site of Cex, resulting in an apparent lower activity. The acceptor D-glucal likely behaved in a similar manner. Cex has been shown to be capable of hydrolyzing aryl glucosides [53], and is therefore likely to be able to use D-glucal as a substrate, based on the finding that Abg, a glucosidase, is able to hydrate this sugar to form 2-deoxyglucose [54]. However, as was mentioned earlier for those acceptors whose reactivation profiles deviated from a first order expression after a period of time, pNPCell and D-glucal were still identified as good acceptors.

3.2.3 Positively screened glycosides as transglycosylation acceptors

The true test of the utility of the screen for transglycosylation acceptors comes from being able to perform transglycosylation reactions with the postitively screened glycosides. Four of the enzymes studied (BgaX, BgaC, Abg, and Man2A) were used in preparative scale reactions with a pNP-glycoside specific for their glycone site as the donor, and one of the positively screened glycosides acting as the acceptor. The isolated products from each reaction were acetylated, isolated, characterized by NMR, and compared to transglycosylation products from the literature (Table 3.3). Results of all but one of the experiments listed in Table 3.3 agree with the results of the screening procedure. That is, most of these sugars were identified as being capable of acting as acceptors in transglycosylation reactions. One exception does exist however,

ENZYME [REF]	DONOR ^a	ACCEPTOR ^a	PRODUCTS (% YIELI	
			β -(1-3)	β -(1-4)
			linked	linked
Abg	pNPGal	gentiobiose	21.4%	
Abg glycosynthase [26]	α -Gal-F	pNPGlu		84%
Abg glycosynthase [26]	$lpha ext{-Gal-F}$	pNPXyl	81%	
Abg glycosynthase [26]	α -Gal-F	pNPMan		66%
Abg glycosynthase [26]	$lpha ext{-Gal-F}$	$p{ m NPCell}$		92%
Man2A	pNPMan	PhGlc		11.8%
Man2A [28]	pNPMan	$p{ m NPMan}$	3.5%	6.7%
Man2A glycosynthase [28]	α -Man-F	pNPMan	1.4%	6.5%
Man2A glycosynthase [28]	α -Man-F	pNPGentiobiose	$\approx 80\%^{b}$	
Man2A glycosynthase [28]	α -Man-F	$p{ m NPCell}$	$\approx 30\%^{b}$	
BgaC	pNPGal	maltotriose	30.0%	
BgaC [55]	pNPGal	GlcNAc	12.2%	
$\operatorname{BgaC}[55]$	pNPGal	GalNAc	10.1%	
BgaX	pNPGal	pNPXyl	6.1%	
BgaX [45]	pNPGal	GlcNAc	22.4%	
BgaX [45]	pNPGal	GalNAc	17.3%	7.41%

TABLE 3.3: Tranglycosylation reaction products from this study and the literature.

^{*a*}All glycosides, excluding maltotriose (α -D-Glu-(1-4)- α -D-Glu-(1-4)-D-Glu) are β -D- configured unless noted otherwise.

^bYield was estimated from TLC analysis and represents all products formed.

which is discussed below.

It is difficult to draw direct correlations between the initial rate of enzyme reactivation in the presence of a particular compound and the product yield of a transglycosylation reaction in which that same compound is used as an acceptor. The yield of such a reaction is dependent on other factors including the concentration of donor and acceptor used, the time period over which the reaction is performed, and, if a glycosidase is used, the susceptibility of the products to hydrolysis. However, the results of Table 3.3 support the basic notion on which the screening procedure is founded: that is, adequate acceptors for a given glycosidase or glycosynthase can be identified by comparison of the rate of reactivation of an inhibited glycosidase in the presence of the compound in question. Note that analysis of enzyme reactivation in the presence of an acceptor does not give any indication as to the number of discrete products or the glycosidic linkages contained therein that are formed in a corresponding transglycosylation reaction. This must be deduced by actually performing the desired reaction.

It was discussed in the previous section that one drawback to the screening strategy is the inability to screen some glycosides that are substrates for the enzyme under study. This shortcoming is illustrated in Table 3.3 by the ability of pNPMan to act as an acceptor for Abg glycosynthase when this glycoside was tested as a negative hit. Wild type Abg is capable of using pNPMan as a substrate ($k_{cat}/K_m=6 \ s^{-1}mM^{-1}$ [10]), which suggests that this sugar was most likely hydrolyzed over the course of the reactivation study to result in a negative hit.

The data of Table 3.3 indicates that pNPXyl, GlcNAc, and GalNac are all capable of acting as acceptors for BgaX. The first two of these sugars were identified as positive hits in the screening procedure; however, GalNAc was observed to be a negative hit. It is unclear as to why GalNAc would register as negative; since it is not derivatized with a pNP group, it should not be degraded over the course of the experiment through enzyme-catalyzed hydrolysis. It is possible that the transglycosylation product formed from enzyme reactivation by GalNAc inhibited active enzyme during the experiment, causing a false negative result. Clearly, the combined yield of 24.7 % of the transglycosylation reaction in the presence of GalNAc indicates that this glycoside is capable of acting as an adequate acceptor. This false negative result was the only such observed for the screening process, excluding the aforementioned possibility of enzyme substrates.

3.3 Conclusions

Although the detection of a false negative result is detrimental to the integrity of the developed methodology, it does not compromise the original intention, to screen a large number of compounds to identify those that will act as acceptors in transglycosylation reactions. Such identification has been readily demonstrated by the ability of the several chosen positively screened glycosides to act as acceptors in the transglycosylation reactions performed, and by the reactions documented from the literature.

In general, aryl glycosides were shown to act as better acceptors as compared to their parent carbohydrates. Some specificity was shown regarding the substitution of the aryl moiety of such glycosides and appears to be dependent on the particular enzyme under study. Similarly, disaccharides were shown to generally act more favorably as acceptors compared with monosaccharides, a consequence believed to be the result of additional favourable interactions with the enzyme's aglycone site. Note however, that the sensitivity towards the linkage type of disaccharides with respect to aglycone specificity varies between enzymes.

The described screening procedure has one drawback in that it is not a reliable approach to evaluate the potential of a substrate for a given glycosidase to also act as a transglycosylation acceptor. Considering that there are a vast number of compounds (not necessarily glycosides) that one may wish to use as acceptors, this setback, while detrimental, does not detract from the overall utility of the methodology.

The generally low number of positive hits resulting from all potential acceptors examined (16-36%) illustrates the necessity of screening a large number of compounds to identify the maximal number of acceptors for a given glycosidase. The developed screening strategy is optimal for such a task; it is easily performed in a reasonable amount of time, and requires the use of few specialized reagents and equipment. The synthesis of fluorinated glycosides in this laboratory and others has provided a foundation for the development of mechanism based inactivators for numerous retaining glycosidases which could be used in the screening process.

Although the methodology described in this study was used to screen for potential transglycosylation acceptors, it also provides a strategy by which to generally probe the specificity of the aglycone site of a retaining glycosidase. This information would provide insight into architechture of an enzyme's aglycone site, which could be useful in future studies such as those that correlate structure and function.

CHAPTER 4

Materials and Methods

4.1 The characterization of BgaX

4.1.1 General

All buffers, substrates, and reagents were purchased from either Sigma-Aldrich Co. or Aldrich Chemical Company unless noted otherwise. Galacto-imidazole was a gift from A. Vasella (Institut für Organische Chemie, ETH-Zürich). *p*NPGalNAc was a gift from S.J. Williams (Department of Chemistry, UBC).

Enzyme activity assays were done with a Unicam UV4 ultraviolet/visible spectrometer; the spectrometer and cuvettes were pre-equilibrated to 37°C by the use of a circulating water bath. Polystyrol and acryl 1 mL disposable cuvettes were used for all spectroscopic measurements, excluding the measurement of activity of BgaX with $pNP-\alpha$ -L-Ara, pNPFuc, pNPGalNAc, pNPGlc, and pNPXyl; for these experiments, 200 μ L quartz cuvettes were used.

Protein sequence alignments were done with the NCBI BLAST server site [56, 57]. N-terminal protein sequencing was performed by the Nucleic Acid-Protein Service Unit, UBC. Curve fitting was done with Grafit [58].

4.1.2 The molar absorptivity of 2,4-dinitrophenol and *p*-nitrophenol

A sample of pNP was dried *in vacuo* and used to make a stock solution in water (3.31 mM), aliquots of which were added via syringe to 800 μ L of the appropriate 50 mM buffer to final pNP concentrations ranging from 41.0-80.7 μ M (pH 4.0-6.5 = citrate buffer; pH 6.5-8.0 = sodium phosphate buffer; pH 8.0-9.0 = glycine buffer). The absorbance at 400 nm of each sample was measured and used to calculate the molar absorptivity (ε) from Beer's Law:

$$A = \varepsilon dc \tag{4.1}$$

where
$$A =$$
 absorbance at 400 nm
 $\varepsilon =$ molar absorptivity (cm⁻¹ mM⁻¹)
 $d =$ the cell pathlength (1 cm)
 $c =$ the concentration of pNP (mM)

The resultant values of ε were fit to the following equation [58]:

$$\varepsilon = \frac{L1 + L2 \left(10^{(pH-pKa)}\right)}{10^{(pH-pKa)} + 1}$$
(4.2)

where L1 and $L2 = \varepsilon$ for the fully protonated and deprotonated forms of pNP respectively

The pK_a of pNP and ε at pH 6.0 was extrapolated from the curve fit to the above equation.

The determination of ε of DNP at pH values of 3.5–6.5 was performed in the same manner as that described for pNP. Values of A₄₀₀ at DNP concentrations of 20.0–90.0 μ M in 150 mM NaCl, 50 mM buffer (as above) were used to construct the ε

versus pH plot. The curve fit to Equation 4.2 was used to determine the pK_a of DNP and ε values at each pH.

The calculated values of ε for pNP and DNP (Table 4.1) were used in successive studies to convert V_{max} from units of $\Delta A_{400}/min$ to mM substrate hydrolyzed/min.

TABLE 4.1. Values of a for piver and Diver								
	pH							
	4.5	5.0	5.5	6.0	6.5	7.0	7.5	8.0
ϵ of pNP (cm ⁻¹ mM ⁻¹)	2.30	2.37	2.611	3.32	5.21	9.05	13.84	17.17
ϵ of DNP (cm ⁻¹ mM ⁻¹)	13.67	15.58	16.29	16.52	16.60	16.63	16.64	16.64

TABLE 4.1: Values of ε for pNP and DNP

4.1.3 Source, expression and purification of BgaX

Modification of gene

All gene manipulations were done by Janine Foisey (Department of Microbiology, UBC). The gene encoding BgaX without its amino-terminal 21 amino acid signal sequence coding region (pCT β -gal-100) was obtained as a construct in the New England Biolabs in-house vector pAGR3 from Dr. Christopher Taron (Department of Biochemistry, University of Illinois at Urbana-Champaign). A new pCT β -gal-100 construct that incorporated a 6-histidine tag at the carboxy-terminal of the protein was engineered in the vector pTug10N18 through the polymerase chain reaction (PCR) using the primers of Figure 4.1. The BgaX gene construct was transformed into *E. coli* JM101 via electroporation.

5' GTC GAA AAC AGC GGC CGC ATC AAT 3'

3' GTG GGG TCG CAC GCG CCG GTA GTG GTA GTG GTA GTG ATT CGA TCG TTC GAA TCT GAT 5'

FIGURE 4.1: The primers used in PCR to fuse a 6-His tag to BgaX.

Expression

Expression of BgaX was done by L. Gal (Department of Microbiology, UBC). E. coli cells containing the BgaX construct were plated onto LB/Amp media (10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl, 100 μ g/mL ampicillin), and a single resultant colony used to inoculate a 30 mL Typ/Amp starter culture (16 g/L tryptone, 16 g/L yeast extract, 5 g/L NaCl, 100 μ g ampicillin, pH 7.0). After incubation for 5.5 hr at 37°C, the starter culture was used to inoculate a 2 L Typ/Amp culture. After incubation at 30°C for 3 hr, isopropyl- β -D-thiogalactoside was added to a final concentration of 0.1 mM, and the culture was incubated at 30°C until $OD_{600} = 0.5$. The cells were harvested and spun at 4°C for 15 min at 44,000 x g to produce the cell paste.

Purification

The cell paste was suspended in 50 mM sodium phosphate buffer, 0.5 M NaCl, pH 6.0 (start buffer) (3.8 mL/g). Phenylmethylsulfonyl fluoride (final concentration = 0.25 mM) and β -mercaptoethanol (final concentration = 6.7 mM) were added and the suspension was French pressed 3x using an SLM Aminco French Pressure Cell Press in a 20-K cell. The lysed cells were centrifuged at 44,000 x g for 1 hr, and the supernatant decanted and filtered through a 45 μ m membrane (final volume *ca.* 45 mL). The clarified supernatant was loaded onto a Pharmacia Biotech X16 column packed with Pharmacia Biotech Chelating Sepharose Fast Flow gel (18 mL) charged with Ni²⁺ according to the manufacturer's instructions, and pre-equilibrated with start buffer. BgaX was eluted over 10 column volumes with a linear imidazole gradient of 0–100 mM. Fractions that showed β -galactosidase activity (determined as described below) were pooled ($V_T = 130$ mL) and dialyzed against 2 x 2.5 L of 20 mM sodium phosphate buffer, 50 mM NaCl, pH 6.0 (SP buffer) then concentrated in 30,000 molecular weight cut-off Amicon Centriprep concentrators and loaded onto two pre-packed 5 mL Pharmacia Biotech Hi-Trap SP columns connected in tandem, and pre-equilibrated with SP buffer. BgaX was eluted with a linear NaCl gradient of 50–350 mM over 15 column volumes. The purified BgaX appeared as a single band in silver-stained SDS-PAGE. N-terminal sequencing of the wild type BgaX (a gift from J. Foisey, Department of Microbiology, UBC) and the purified 6-His tagged BgaX proteins (ATPESW) were consistent with the SWISS-PROT entry (P48982) [59] for this enzyme .

4.1.4 pH Stability

Very preliminary preparations of BgaX were kept in 50 mM sodium phosphate, pH 7.0. Enzyme activity remained constant over at least 5 hr at room temperature at this pH. The activity of BgaX at pH 7.0 was therefore used as a reference by which to determine the stability of this enzyme over a range of pH values. Buffers used for this study were 150 mM NaCl in 50 mM citrate (pH 4–6) and sodium phosphate (pH 6–8). BgaX (0.12 mg) was added to each pH solution (V_T = 200 µL), incubated at room temperature for 2 minutes, and then assayed for activity as described below. The residual activity as compared to activity at pH 7.0 was plotted against pH.

4.1.5 pH Dependence

The parameters k_{cat} and K_m for the hydrolysis of DNPGal by BgaX were determined as described below (Hydrolysis kinetics) at pH values ranging from 4.5–8.0. The initial velocity of hydrolysis was determined after the addition of BgaX (0.94 mg) to aliquots of DNPGal at five different concentrations (0.016–1.39 mM) in 0.1% BSA (w/v) (total volume = 810 μ L). Buffers used at each pH value (pH 4.5–6.0 = citrate buffer, pH 6.0–8.0 = sodium phosphate buffer) were 50 mM and contained 150 mM NaCl. The initial velocity of each reaction for a given pH value was plotted against substrate concentration and fit to the classical Michaelis-Menten equation to derive V_{max} and K_m , and ultimately, k_{cat} . The kinetic parameters k_{cat} and k_{cat}/K_m were then plotted against pH and fit to $y = \frac{L*10^{(pH-pK_a1)}}{10^{(2+pH-pK_a1-pK_a2)}+10^{(pH-pK_a1)}+1}$ where $y = k_{cat}$ or k_{cat}/K_m and L is the upper limit.

4.1.6 Kinetics

Measurement of enzyme activity

The rate of hydrolysis of nitrophenyl-glycoside substrates were determined by the continuous spectrometric monitoring of the release of pNP or DNP at 400 nm $(\Delta A_{400}/\text{min})$ over 1–2 minutes. Spectrometer cuvettes contained 0.1% BSA (w/v) in 50 mM sodium phosphate buffer, pH 6.0 (buffer A) unless noted otherwise. Assays were initiated by the addition of a 5 μ L enzyme aliquot. Unless stated otherwise, the activity of BgaX was determined with 1.2 mM pNPGal at 37°C.

Hydrolysis kinetics

The activity of BgaX was determined as described above at at least five different concentrations of 0.007–3.00 mM pNPGal, 0.01–0.52 mM DNPGal, 0.3–2.2 mM pNP- α -L-Ara, 0.5–12.6 mM pNPFuc, 0.4–3.4 mM pNPGalNAc, 1.5–49.7 mM pNPGlc, and 1.8–15.0 mM pNPXyl. The resultant measured velocities ($\Delta A_{400}/min$) were plotted against substrate concentration and fit to standard Michaelis-Menten kinetics $\left(v = \frac{V_{max}[S]}{K_m + [S]}\right)$ to determine K_m and V_{max} , and ultimately, k_{cat} .

Inactivation by 2FDNPGal

2FDNPGal (1.15–46.0 μM) in BSA-free buffer A (final volume = 100 μ L) was incubated in a 20°C water bath for 5 minutes. BgaX (9.7 μ g) was added to initiate inactivation. Aliquots of 5 μ L of this inactivator mix were removed over time and residual enzyme activity was assayed as described. The fractional loss in activity was plotted against time and fit to a first order exponential decay $\left(\frac{v}{V_o} = Ae^{-k_{inact}^{obs}} + C\right)$. The observed rate constants (k_{inact}^{obs}) so derived were replotted in double reciprocal form with inhibitor concentration and fit to $k_{inact}^{obs} = \frac{k_{inact}[I]}{K_i+[I]}$ to determine the inactivation rate constant (k_{inact}) and inhibitor dissociation constant (K_i) of 2FDNPGal with BgaX.

Protection from inactivation

The K_i of galacto-imidazole with BgaX was determined by the measurement of the initial velocity of hydrolysis, as described above, by the addition of BgaX (9.7 μ g) to five concentrations of pNPGal (7.0-91.0 μ M) in the presence and absence of 7.2 nM galacto-imidazole (P) (volume = 805 μ L). The measured initial velocities and substrate concentrations were plotted in Lineweaver-Burk form and and the slopes fit to $K_i \approx \frac{[P]}{slope_i/slope_o-1}$ to derive K_i (where slope_o = slope of the Lineweaver-Burk plot in the absence of P, and slope_i = slope of the Lineweaver-Burk plot in the presence of P).

Protection of BgaX from inactivation by 2FDNPGal was shown by the determination of k_{inact}^{obs} (determined as described above) of BgaX (9.7 µg) with 46 µM of 2FDNPGal (in 100 µL BSA-free buffer A) in the presence (k_{ip}^{obs}) and absence (k_i^{obs}) of 40 nM galacto-imidazole and comparison of k_{ip}^{obs}/k_i^{obs} to the value predicted by the following equation:

$$\frac{k_{ip}^{obs}}{k_i^{obs}} = \frac{\mathbf{K}_i + [\mathbf{I}]}{\mathbf{K}_i \left(1 + \frac{[\mathbf{P}]}{\mathbf{K}_p}\right) + [\mathbf{I}]}$$
(4.3)

Catalytic competence of inactivated species:

Generation of the inactivated enzyme or 'I-mix': BgaX (2.42 mg) was incubated with 2FDNPGal (207 mmol) in BSA-free buffer A (volume = 1mL) until < 5% of the original activity remained. The inactivated enzyme was concentrated using a 30,000 molecular weight cut-off Millipore Ultrafree 0.5 centrifugal filter and tube assembly to a volume of 10-20 μ L, and fresh BSA-free buffer A was added to a final volume of 200 μ L. This process of concentration and dilution was repeated until all excess inactivator was removed, as determined by the maintenance of full activity of an aliquot (5 μ L, 0.01 mg) of non-inactivated enzyme sample diluted one half with the inactivated species or 'I-mix'. The concentration of inactivated enzyme in the I-mix was determined by absorption at 280 nm (ε = 1.97 mL mg⁻¹ cm⁻¹)¹, and used to calculate the maximal activity possible (V_o) if the enzyme were completely active.

 $\varepsilon = ((\# \text{ Trp})(5,500) + (\# \text{ Tyr})(1,490) + (\# \text{ Cys})(125))/\text{MW} [60]$

¹As determined by amino acid sequence:

Determination of catalytic competence: A dilution of the BgaX I-mix (24 μ g in 50 μ L) was incubated at 20°C in buffer A and aliquots were removed over time and assayed as described. The resultant gain in fractional activity was fit to a first order rate equation $\left(\frac{v}{V_o} = A\left(1 - e^{-k_{react}t}\right) + C\right)$ and the rate constant for turnover of the glycosyl-enzyme intermediate (k_{react}) determined.

4.1.7 Identification of the catalytic nucleophile of BgaX

Digestion of inactivated enzyme:

BgaX (14.5 ng) was incubated with 2FDNPGal (25.8 mmol) in BSA-free buffer A (volume = 65 μ L) until < 5% of the original activity remained. An aliquot (10 μ L) of the inactivated enzyme was frozen and kept at -20°C until analyzed by mass spectrometry. The inactivated species (50 μ L) was added to pepsin in a ratio of *ca.* 12:1 (enzyme:pepsin) in 200 mM sodium phosphate, pH 2.0 (volume = 78 μ L). The digestion was done at room temperature for 2 hrs; the sample was then frozen and kept at -20°C until analysed by mass spectrometry.

Electrospray mass spectrometry

All mass spectrometry was done by S. He (Department of Chemistry, UBC). Mass spectra were recorded on a Perkin-Elmer Sciex API 300 triple quadrupole mass spectrometer with an ion spray ion source. The spectrometer was interfaced with a Michrom BioResources Inc. Ultrafast Microprotein Analyzer which separated peptides via reverse phase HPLC. Digested proteins were loaded onto a Reliasil C18 column (1 x 150 mm) equilibrated with solvent A (0.05% trifluoroacetic acid (TFA), 2% acetonitrile in water). Peptides were eluted over a 0-60% gradient of solvent B over 60 minutes, followed by 100% solvent B over 2 minutes (solvent B: 0.045% TFA, 80% acetonitrile in water). The pump rate was 50 μ L/min. Spectra were obtained in the single quadrupole scan mode (LC/MS) or the tandem MS daughter ion scan mode (MS/MS).

In LC/MS, the quadrupole mass analyser was scanned over a mass to charge ratio (m/z) of 300-2400 Da with a step size of 0.5 Da and a dwell time of 1 ms per step. The ion source voltage (ISV) was 5 kV and the orifice energy (OR) was 50 V. In MS/MS, the spectra were obtained by selectively introducing one of the parent peptides of m/z=869.0 and 951.0 from the first quadrupole (Q1) into the collision cell (Q2) and observing the daughter ions in the third quadrupole (Q3). The scan range of Q3 was 50-1920 Da; step size, dwell time, and OR as above; ISV was 4.8 kV.

4.1.8 The mutant H6-E260A BgaX

E. coli cell paste containing H6-E260A BgaX was a gift from J. Foisey (Department of Microbiology, UBC). The mutant protein was purified by a Ni²⁺-chelating column as described for H6-BgaX. The protein was $\approx 90\%$ pure as estimated by silver stained SDS-PAGE.

4.1.9 Measurement of activity of H6-E260A BgaX

The activity of H6-E260A BgaX was determined with and without azide present. Spectrometer cuvettes contained 790 μ L 50 mM 2-(*N*-morpholino)ethane sulfonic acid (MES), pH 5.0, 0.1% BSA (w/v), 1.0 mM DNPGal, and either 0 or 2 M azide. The background hydrolysis of DNPGal was determined over 1–2 minutes, then a 5 μ L aliquot of the mutant enzyme was added and the initial velocity of the reaction determined.

4.2Large-scale screening of potential acceptors

4.2.1General

All enzymes, excluding BgaX, were expressed and purified by other laboratory members or collaborators (Table 4.2). The activity of each enzyme was determined as described for BgaX (§4.1, Measurement of activity) in the buffer system and with the substrates and concentrations of Table 4.2 (in 775–900 μ L). Enzyme inactivators were gifts from members of this laboratory: 2FDNPGal, 2FDNPGlc, and 2FManF (L. Mackenzie), 2FDNPCell (D. Zechel), 2FDNPGlcA (A. Wong). Potential acceptors were purchased from Sigma-Aldrich Co. or Aldrich Chemical Company or were gifts

TABLE 4.2: Conditions for the measurement of activity of the enzymes screened.					
ENZYME (SOURCE)	BUFFER	SUBSTRATE			
[PURIFICATION PROTOCOL]					
BgaX	50 mM NaPi ^a , pH 6.0	1.2 mM p NPGal			
BgaC (L. Gal ^{$\dagger b$})	50 mM NaPi, pH 7.0	0.94 mM p NPGal			
Abg (K. Rupitz [‡] [21])	50 mM NaPi, pH 7.0	0.50 mM p NPGlc			
Man2A (D. Stoll ^{\dagger} [28])	50 mM NaPi, pH 7.0	0.35 mM p NPMan			
Cex (K. Rupitz ^{\ddagger} [61])	50 mM NaPi, pH 7.0	1.2 mM pNPCell			
CelB (C. Dupont [62])	50 mM NaPi, pH 6.5	0.50 mM DNPGal			
HBG (J.H. Grubb [¶]	100 mM NaAc ^{c} , pH 4.5	1.83 mM p NPGlcA			
and W.S. Slv^{\P} [63, 64])					

^aSodium phosphate.

 b L. Gal incorporated a 6-His tag to the carboxy terminal of BgaC and purified the enzyme in an unpublished protocol that was very similar to that reported here for BgaX.

^cSodium acetate.

[†]Department of Microbiology, UBC.

[‡]Department of Chemistry, UBC.

[§]INRS Institut Armand-Frappier, Centre de recherche en microbiologie et biotechnologie [¶]E.A. Doisy Department of Biochemistry and Molecular Biology, St. Louis University School of Medicine

from members of this laboratory: DNPFuc (J. McCarter), $pAcPh \beta$ -Glc (Q. Wang) CO₂Me β -Glc (Q. Wang), D-glucal (L. Mackenzie), 1-cyano-D-glucal (E. Lai), 1,2dicyano-D-glucal (T. Harvey), DNPCell (H. Prade), Me α -Cell (V. Ferro), Me β -Cell (V. Ferro), Gal- β -(1-3)-Glc- β -SPh (H. Prade), Gal- β -(1-4)-Glc- β -SPh (H. Prade), 2,5-DNPX₂ (D. Zechel). Stock solutions of potential acceptors were made in water to a final concentration of *ca*. 50 mM; some of these solutions were saturated.

Thin layer chromatography (TLC) was done with aluminum backed silica 60 F_{254} from E.M. Science. All TLC's were examined under UV light and then stained with 10% H_2SO_4 in MeOH and charred. Chromatography was done with 230-400 mesh silica in *ca.* 15 cm columns.

NMR spectra were recorded in CDCl_3 on a Bruker AC 200 or Bruker WH 400 and are referenced to the solvent peak. Where required, interpretations were supported by COSY experiments.

4.2.2 Screening A: large scale screening

I-mixes of each enzyme were prepared as described for BgaX (§4.1.6, Catalytic competence of inactivated species) with the enzyme and inactivator amounts listed in Table 4.3 (volume = 0.5–1.0 mL; buffer system as in Table 4.2). An aliquot of each I-mix (Table 4.3) plus buffer and each compound to be screened (Tables 3.1 and 3.2) were added to the wells of a 96-well plate so the final concentration of the potential acceptors was 30 mM or 60% saturation (final volume = 50–100 μ L). Controls of (i) non-inactivated enzyme and (ii) inactivated enzyme, each in buffer alone, were included. The plates were incubated at room temperature for the times indicated (Table 4.3). The amount of enzyme reactivated after incubation was determined by the measurement of enzyme activity: 100 μ L of the corresponding assay substrate

ENZYME FOR	INHIBITOR	ENZYME IN	INCUBATION	Assay
INACTIVATION		EACH WELL	TIME	$\mathbf{SUBSTRATE}^{a}$
$340~\mu{\rm g~BgaX}$	258 mmol 2FDNPGal	$1.60 \ \mu { m g}$	19 hr	2.5 mM p NPGal
$440~\mu{\rm g~BgaC}$	155 mmol 2FDNPGal	$5.67~\mu{ m g}$	17 hr	2.7 mM p NPGal
470 μg Abg	122 mmol 2FDNPGlc	$1.73~\mu{ m g}$	18 hr	2.0 mM p NPGlc
$80~\mu{\rm g~Man2A}$	187 mmol 2FManF	$0.26 \ \mu { m g}$	$2.5 \ hr$	$0.35~\mathrm{mM}~p\mathrm{NPMan}$
$310~\mu{ m g}~{ m Cex}$	189 mmol 2FDNPCell	$4.55 \ \mu \mathrm{g}$	$16.5 \ hr$	2.3 mM p NPCell
$60~\mu{\rm g~CelB}$	60 mmol 2FDNPCell	$0.75~\mu{ m g}$	$3.5~\mathrm{hr}$	0.75 mM DNPCell
121 $\mu { m g}~{ m HBG}$	100 mmol 2FGlcAF	$1.39 \ \mu { m g}$	180 hr	2.05 mM p NPGlcA

TABLE 4.3: Conditions for the screen of potential acceptors.

^aFinal concentration in each well upon assaying for activity.

(Table 4.3) was added to each well to the final concentrations indicated, and the release of the pNP or DNP phenolate ion was measured with a 96-well plate reader either by: (i) taking single absorbance measurements of each well at 400 nm over 0.5-6 hr (using a Bio-Rad Model 2550 EIA Reader), or (ii) measuring the continuous change in absorbance at 400 nm of each well at 400 nm over 10-20 minutes (using a Molecular Devices Spectra Max 190 Reader), the latter procedure being preferred. Tested compounds were classified as a positive hit if they induced greater enzyme reactivation as compared to the control of reactivating enzyme in buffer alone.

4.2.3 Screening B: analylsis of enzyme reactivation

BgaX, BgaC, Man2A, Abg, Cex, CelB, and HBG I-mixes were prepared as described in §4.1.6 (Catalytic competence of inactivated species) with the conditions described in Table 4.4. I-mixes were added to buffer and those compounds identified as positive hits (from the large scale screening described above) in eppendorf tubes so the final concentration of each tested compound was 30 mM or 60% saturation (volume= 50-100 μ L). All such mixtures were incubated in a water bath at 37°C, excluding BgaX (incubation at 20°C). Controls of non-inactivated enzyme were included to account for enzyme death over time. Aliquots of 5 μ L of each sample were removed over time and assayed for activity as described above for BgaX (§4.1.6 Measurement of activity) with the buffers and substrates of Table 4.2. The fractional gain in activity was fit to a first order rate equation $\left(\frac{v}{V_o} = A\left(1 - e^{-k_{trans}^{obs}t}\right) + C\right)$ and values of k_{trans}^{obs} and A were extrapolated and used to calculate the initial rate of reactivation ($V_o^{react} = Ak_{trans}^{obs}$). For BgaC and Cex, reactivation was so slow that over the course of the experiment, the data collected was contained within the linear portion of the first order reactivation expression; therefore, values of V_o^{react} for these enzymes were obtained from a direct linear fit of the data.

ENZYME FOR	INHIBITOR	ENZYME IN
INACTIVATION		EPPENDORF
2.4 mg BgaX	207 mmol 2FDNPGal	$50 \ \mu { m g}$
$5.5 \mathrm{~mg~BgaC}$	103 mmol 2FDNPGal	$175~\mu{ m g}$
$1.8 \mathrm{~mg~Abg}$	612 mmol 2FDNPGlc	$40 \ \mu g$
0.1 mg Man2A	312 mmol 2FManF	$4 \ \mu { m g}$
0.6 mg Cex	94 mmol 2FDNPCell	$140 \ \mu { m g}$
$0.3 \mathrm{~mg~CelB}$	40 mmol 2FDNPCell	$2~\mu{ m g}$
2.8 mg HBG	98 mmol 2FGalAF	$140 \mu { m g}$

 TABLE 4.4: Conditions for the analysis of the effect of each postive hit on enzyme

 reactivation.

4.2.4 Preparative scale transglycosylation reactions

Enzyme in buffer (Table 4.5) was incubated with acceptor (A) and a portion of donor (D) (D:A ≈ 1.5) at room temperature so the final concentration of A was 40-50 mM (volume = 5-8 mL). Reactions were monitored by TLC (solvent: 7:2:1=EtOAc:MeOH:H₂O). The rest of D was successively added in *ca.* 1/4 amounts after approximately 90% of the D in the reaction mixture was depleted (estimated by TLC). After total depletion of D (4-14 hr), reactions were lyophilized and acetylated.

Acetylation of transglycosylation products

Transglycosylation products were acetylated with either (i) acetic anhydride (Ac_2O) or (ii) 1-¹³C-acetyl chloride (1-¹³C-AcCl) according to the following methods:

(i) Ac_2O (3 mL) was added under N_2 to the mixture of transglycosylation products in dry pyridine (6 mL) at room temperature. Once the reaction was complete by TLC (solvent: 1:1=EtOAc:hexanes), the solvent was removed by the addition of toluene and evaporation *in vacuo*.

(ii) 1-¹³C-AcCl (150-200 μ L) was added under N₂ to the mixture of transglycosylation products in dry pyridine (2 mL) at room temperature. Once the reaction

Enzyme	Donor ^a	ACCEPTOR	BUFFER SYSTEM
BgaX (0.194 mg)	73.2 mg	54.0 mg	50 mM NaPi, pH 6.0
BgaC (0.405 mg)	88.1 mg	220.6 mg	$50~\mathrm{mM}$ NaPi, pH 7.0
Abg (0.540 mg)	77.3 mg	71.5 mg	50 mM NaPi, pH 7.0
Man2A (0.160 mg)	48.7 mg	66.7 mg	50 mM NaPi, pH 7.0

TABLE 4.5: Conditions for preparative scale transglycosylation reactions.

^aTotal amount of donor added over the course of the reaction

was complete by TLC (solvent: 3:1=EtOAc:hexanes), the solvent was removed by the addition of toluene and evaporation *in vacuo*. Water and EtOAc were added to the residue (*ca.* 15 mL of each) and successively washed with 1 M HCl, 1M NaHCO₃ and saturated NaCl. The organic layer was recovered, and the solvent removed *in vacuo*.

Separation of transglycosylation products

Acetylated transglycosylation products were loaded onto a silica gel column for chromatography in the minimum amount of CH_2Cl_2 and eluted with a solvent system of 1:1 or 2:1 petroleum ether: EtOAc, or a combination thereof. Fractions of each product were pooled and the solvent removed *in vacuo*. The isolated products were dried *in vacuo* for *ca.* 48 hr, and submitted for NMR. Product yields are based on isolated products and are referenced to the total amount of donor used in each reaction.

NMR data of transglycosylation products

2,3,4,6-Tetra-O-acetyl- β -D-galactopyranosyl-(1 \rightarrow 3)-2,4,6-tri-O-acetyl- α -D-glucopyranosyl-(1 \rightarrow 4)-2,3,6-tri-O-acetyl- α -D-glucopyranosyl-(1 \rightarrow 4)-2,3,6-tri-O-acetyl- α -Dglucopyranose; ¹H NMR (400 MHz, CDCl₃): δ 5.71 (d, 1 H, J_{1,2} 8.1 Hz, H-1), 5.48 (dd, 1 H, J_{2,3} 10.1 Hz, J_{3,4} 9.0 Hz, H-3), 5.38 (dd, 1 H, J_{3',4'} 9.7 Hz, H-3'), 5.33 (d, 1 H, J_{3'',4'''} 3.3 Hz, H-4'''), 5.28 (d, 1 H, J_{1'',2''} 4.1 Hz, H-1''), 5.22 (d, 1 H, J_{1',2'} 4.0 Hz, H-1'), 5.04 (dd, 1 H, J_{2''',3'''} 10.4 Hz, H-2'''), 4.97–4.89 (m, 3 H, H-2, H-4'', H-3'''), 4.78 (dd, 1 H, J_{2'',3''} 10.2 Hz, H-2''), 4.69 (dd, 1 H, J_{2',3'} 10.3 Hz, H-2'), 4.61 (d, 1 H, J_{1''',2'''} 8.0 Hz, H-1'''), 4.47–4.38 (m, 3 H, H-6a, H-4', H-6a'), 4.24–4.12 (m, 5 H, H-6b, H-6b', H-6b'', H-6a''', H-6b'''), 4.09–3.99 (m, 3 H, H-4, H-5, H-3''), 3.95–3.81 (m, 4 H, H-4', H-5', H-6b'', H-5'''), 2.22, 2.14, 2.13, 2.11, 2.07, 2.03, 2.02, 1.99, 1.98, 1.97, 1.96, 1.93 (12 s, 42 H, Ac). Para-nitrophenyl 2,3,4,6-tetra-O-acetyl-β-D-galactopyranosyl-(1 → 3)-2,4-di-Oacetyl-β-D-xylopyranoside; ¹H NMR (400 MHz, CDCl₃): δ 8.20 (d, 2 H, $J_{a,b}$ 9.3 Hz, pNP), 7.06 (d, 2 H, $J_{c,d}$ 9.0 Hz, pNP), 5.37 (dd, 1 H, $J_{4',3'}$ 3.2 Hz, H-4'), 5.17 (dd, 1 H, $J_{2,3}$ 8.1 Hz, H-2), 5.16 (dd, 1 H, $J_{2',3'}$ 10.5 Hz, H-2'), 4.99 (dd, 1 H, $J_{3',4'}$ 3.3 Hz, H-3'), 4.96 (ddd, 1 H, $J_{4,5eq}$ 4.3 Hz, H-4), 4.56 (d, 1 H, $J_{1',2'}$ 7.8 Hz, H-1'), 4.47 (d, 1 H, $J_{1,2}$ 7.8 Hz, H-1), 4.19 (dd, 1 H, $J_{5ax,5eq}$ 12.1 Hz, H-5eq), 4.00–4.14 (m, 2 H, H-6a', H-6b') 3.94 (dd, 1 H, $J_{3,4}$ 5.6 Hz, H-3), 3.89 (dd, 1 H, $J_{5',6a'}$ 6.2 Hz, $J_{5',6b'}$ 7.3 Hz, H-5'), 3.59 (dd, 1 H, $J_{5ax,4}$ 7.0 Hz, H-5ax), 2.14, 2.10, 2.09, 2.05, 2.03, 1.97 (6 s, 18 H, Ac).

Phenyl 2,3,4,6-tetra-O-(1-¹³C-acetyl)- β -D-mannopyranosyl-(1 \rightarrow 4)-2,3,6-tri-O-(1-¹³C-acetyl)- β -D-glucopyranoside; ¹H NMR (400 MHz, CDCl₃): δ 5.47 (ddd, 1 H, J_{2',Ac} 3.2 Hz, J_{2',3'} 3.2 Hz, H-2'), 5.32 (m, 1 H, J_{4',Ac} 1.9 Hz, H-4'), 5.20 (ddd, 1 H, J_{3,Ac} 3.18 Hz, J_{3,4} 9.1 Hz, H-3), 5.13 (ddd, 1 H, J_{2,Ac} 3.41 Hz, J_{2,3} 10.0 Hz, H-2), 5.02 (ddd, 1 H, J_{3',Ac} 3.2 Hz, J_{3',4'} 6.4 Hz, H-3'), 4.66 (d, 1 H, J_{1',2'} 0.5 Hz, H-1'), 4.30 (ddd, 1 H, J_{6b,Ac} 3.1 Hz, J_{6b,6a} 12.4 Hz, H-6b), 4.26 (m, 1 H, H-6a), 4.10 (d, 1 H, J_{1,2} 7.1 Hz, H-1), 4.05–4.17 (m, 2 H, H-6a', H-6b'), 3.88 (dd, 1 H, J_{4,5} 9.6 Hz, H-4), 3.79 (ddd, 1 H, J_{5,6a} 2.44 Hz, J_{5,6b} 5.2 Hz, H-5), 3.62 (ddd, 1 H, J_{5',4'} 8.5 Hz, J_{5',6a'} 2.2 Hz, J_{5',6b'} 5.0 Hz, H-5'), 2.17, 2.15, 2.14, 2.09, 2.05, 2.04, 1.98 (7 s, 21 H, Ac).

2,3,4,6-Tetra-O-(1-¹³C-acetyl)-β-D-glucopyranosyl-(1 → 3)-2,4,6-tri-O-(1-¹³C-acetyl)-β-D-glucopyranosyl-(1→6)-1,2,3,4-tetra-O-(1-¹³C-acetyl)-β-D-glucopyranoside; ¹H NMR (400 MHz, CDCl₃): δ 5.68 (dd, 1 H, $J_{1,Ac}$ 3.3 Hz, $J_{1,2}$ 8.2 Hz, H-1), 5.32 (ddd, 1 H, $J_{3,Ac}$ 3.4 Hz, $J_{3,4}$ 9.1 Hz, H-3), 5.16 (ddd, 1 H, $J_{2'',Ac}$ 3.4 Hz, $J_{2'',3''}$ 11.2 Hz, H-2"), 5.14 (ddd, 1 H, $J_{3'',Ac}$ 3.3 Hz, H-3"), 5.14–5.03 (m, 2 H, H-4', H-4"), 5.05 (ddd, 1 H, $J_{2,Ac}$ 1.5 Hz, 1 H, $J_{2,3}$ 10.8 Hz, H-2), 5.04–4.89 (m, 1 H, H-4), 4.98 (ddd, 1 H, $J_{2',Ac}$ 0.1 Hz, $J_{2',3'}$ 10.8 Hz, H-2'), 4.51 (ddd, 1 H, $J_{6a',5'}$ 3.7 Hz, $J_{6a',Ac}$ 3.0 Hz, $J_{6a',6b'}$ 12.4 Hz, H-6a'), 4.47 (d, 1 H, $J_{1',2'}$ 7.8 Hz, H-1'), 4.33–4.20 (m, 3 H, H-6b', H-6a", H-6b"), 4.02 (d, 1 H, $J_{1'',2''}$ 8.3 Hz, H-1"), 3.93–3.80 (m, 2 H, H-6a, H-5'), 3.81 (ddd, 1 H, $J_{5,4}$ 8.8 Hz, $J_{5,6a}$ 2.5 Hz, $J_{5,6b}$ 3.7 Hz, H-5), 3.61 (m, 1H, H-6b), 3.57 (dd, 1 H, $J_{3',4'}$ 9.0 Hz, H-3').

Bibliography

- Voet, D. and Voet, J.G. Biochemistry. John Wiley & Sons, Inc., Toronto, 2nd edition, (1995).
- [2] Recommendations of the Nomenclature Committee of the International Union of Biochemistry. Academic Press, Inc., Toronto, (1984).
- [3] Henrissat, B. (1991) Biochem. J. 280, 309–316.
- [4] Henrissat, B. and Bairoch, A. (1993) *Biochem. J.* 293, 781–788.
- [5] Henrissat, B. and Bairoch, A. (1996) *Biochem. J.* **316**, 695–696.
- [6] Koshland, D.E (1953) *Biol. Rev.* 28, 416–436.
- [7] McCarter, J.D. and Withers, S.G. (1994) Curr. Opin. Struct. Biol. 4, 885–892.
- [8] White, A. and Rose, D. (1997) Curr. Opin. Struct. Biol. 7, 645–651.
- [9] Sinnott, M.L. (1990) Chem. Rev. 90, 1171–1202.
- [10] Kempton, J.B. and Withers, S.G. (1992) Biochemistry 31, 9961–9969.
- [11] Withers, S.G. and Street, I.P. (1988) J. Am. Chem. Soc. 110, 8551-8553.
- [12] Street, I.P., Kempton, J.B., and Withers, S.G. 1992 *Biochemistry* **31**, 9970–9978.
- [13] M^cCarter, J.D. and Withers, S.G. (1996) J. Am. Chem. Soc. 118, 241–242.
- [14] White, A., Tull, D., Johns, K., Withers, S.G. and Rose, D.R. (1996) Nat. Struct. Biol. 3, 149–154.
- [15] Tull, D., Withers, S.G., Gilkes, N.R., Kilburn, D.G., Warren, R.A.J. and Aebersold, R. (1991) J. Biol. Chem. 266, 15621–15625.

- [16] Mackenzie, L.F., Brooke, G.S., Cutfield, J.F., Sullivan, P.A. and Withers, S.G. (1997) J. Biol. Chem. 272, 3161–3167.
- [17] Zechel, D.L., He, S., Dupont, C. and Withers, S.G. (1998) Biochem. J. 336, 139-145.
- [18] McCarter, J.D., Burgoyne, D.L., Miao, S., Zhang, S., Callahan, J.W. and Withers, S.G. (1997) J. Biol. Chem. 272, 396–400.
- [19] Tull, D., Miao, S., Withers, S.G. and Aebersold, R. (1994) Anal. Biochem. 224, 509.
- [20] Vocadlo, D.J. and Withers, S.G.; ed. B. Ernst. Glycoside-catalyzed oligosaccharide synthesis; Book chapter for series 'Oligosaccharides in Chemistry and Biology: A comprehensive handbook'; In Press. Wiley/VCH.
- [21] Prade, H., Mackenzie, L.F. and Withers, S.G. (1998) Carb. Res. 305, 371-381.
- [22] Brayer, G., Rose, D. and Withers, S.G. Unpublished results.
- [23] MacLeod, A.M., Tull, D., Rupitz, K., Warren, R.A.J. and Withers, S.G. (1996) Biochemistry 35, 13165–13172.
- [24] Viladot, J.L., de Ramon, E., Durany, O. and Planas, A. (1998) Biochemistry 37, 11332–11342.
- [25] Wang, Q., Graham, R.W., Trimbur, D., Warren, R.A.J. and Withers, S.G. (1994)
 116, 11594.
- [26] Mackenzie, L.F., Wang, Q., Warren, R.A.J. and Withers, S.G. (1998) J. Am. Chem. Soc. 120, 5583-5584.
- [27] Malet, C. and Planas, A. (1998) FEBS Lett. 440, 208–212.
- [28] D. Stoll. Characterization of the mannan-degrading system of Cellulomonas fimi. PhD thesis, The University of British Columbia, Aug. 1998.

- [29] Mayer, C., Zechel, D.L., Reid, S.P., Warren, R.A.J. and Withers, S.G. (2000) FEBS Lett. 466, 40-44.
- [30] Wong-Madden, S.T., and Landry, D. (1995) *Glycobiology* 5, 19–28.
- [31] Taron, C.H., Benner, J.S., Hornstra, L.J., and Guthrie, E.P. (1995) *Glycobiology* 5, 603–610.
- [32] Ito, Y., and Sasaki, T. (1997) Biosci. Biotech. Biochem 61, 1270–1276.
- [33] Gutshall, K., Wang, K., and Brenchley, J.E. (1997) J. Bacteriol. 179, 3064–3067.
- [34] Redenbach, M., Kieser, H.M., Denapaite, D., Eichner, A., Cullum, J., Kinashi,
 H. and Hopwood, D.A. (1996) Mol. Microbiol. 21, 77-96.
- [35] Nishimoto, J., Nanba, E., Inui, K., Okada, S. and Suzuki, K. (1991) Am. J. Hum. Genet. 49, 556-574.
- [36] Henrissat, B., Callebaut, I., Fabrega, S., Lehr, P., Mornon, J.P. and Davies, G. (1995) Proc. Natl. Acad. Sci. USA 92, 7090–7094.
- [37] McIntosh, L.P., Hand, G., Johnson, P.E., Joshi, M., Korner, M., Plesniak, L.A., Ziser, L., Wakarchuk, W.W., and Withers, S.G. (1996) *Biochemistry* 35, 9958– 9966.
- [38] Wakarchuk, W.W., Campbell, R.L., Sung, W.L., Davoodi, J. and Yaguchi, M. (1994) Pro. Sci. 3, 467–475.
- [39] McCarter, J.D., Adam, M.J. and Withers, S.G. (1992) Biochem. J. 286, 721–727.
- [40] Aoyama, T., Naganawa, H., Suda, H., Uotani, K., Aoyagi, T., Takeuchi, T. and Takeuchi, T. (1992) J. Antibiotics 45, 1557–1558.
- [41] Tatsuta, K., Miura, S., Ohta, S. and Gunji, H. (1995) J. Antibiotics 48, 286–288.
- [42] Copeland, R.A. Enzymes; a Practical Introduction to Structure, Mechanism, and Data Analysis. VCH Publishers, Inc., New York, (1996).

- [43] Namchuk, M.N. and Withers, S.G. (1995) Biochemistry 34, 16194–16202.
- [44] Gebler, J.C., Aebersold, R., and Withers, S.G. (1992) J. Biol. Chem. 267, 11126–11130.
- [45] Fujimoto, H. (1997) J. Carb. Chem. 16, 967–970.
- [46] Allerhand, A., Doddrell, D. and Komoroski, R. (1971) J. Chem. Phys. 55, 189.
- [47] Agrawal, P.K. (1992) Phytochem. **31**, 3307–3330.
- [48] Goux, W.J. and Unkefer, C.J. (1987) Carb. Res. 159, 191–210.
- [49] Goux, W.J. (1988) Carb. Res. 184, 47–65.
- [50] W.J. Goux. NMR Applications in Biopolymers. Plenum Press:New York, (1990).
- [51] Bendiak, B. (1999) Carb. Res. **315**, 206–221.
- [52] Gal, L., Department of Microbiology and Immunology/ Biotechnology, UBC Unpublished results.
- [53] Tull, D. and Withers, S.G. (1994) *Biochemistry* **33**, 6363–6370.
- [54] Lai, E.C.K., Morris, S.A., Street, I.P. and Withers, S.G (1996) Bioorg. & Med. Chem. 4, 1929–1937.
- [55] Fujimoto, H., Miyasato, M., Ito, Y., Sasaki, T. and Ajisaka, K. (1998) Glycoconjugate J. 15, 155–160.
- [56] Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Nucleic Acids Res. 25, 3389–3402.
- [57] http://www.ncbi.nlm.nih.gov/cgi-bin/BLAST.
- [58] Leatherbarrow, R.J. (1994) Grafit. version 3.03, Erithacus Software Ltd.
- [59] http://expasy.hcuge.ch/cgi-bin/get-sprot-entry?BGLA XANMN.
- [60] Pace, N.C., Vajdos, F., Fee, L., Grimsely, G. and Gray, T. (1995) Pro. Sci. 4, 2411–2423.
- [61] Gilkes, N.R., Langford, M., Kilburn, D.G., Miller, R.C. and Warren, R.A.J. (1984) J. Biol. Chem. 259, 10455.
- [62] Sulzenbacher, G., Shareck, F., Morosoli, R., Dupont, C. and Davies, G.J. (1997) Biochemistry 36, 16032–16039.
- [63] Ho, K.J. (1991) Protein Express 2, 63-65.
- [64] Grubb, J.H., Kyle, J.W., Cody, L.B. and Sly, W.S. (1993) FASEB J. 7, 1255a.

1

APPENDIX A

Kinetic Analyses

A.1 Basic enzyme kinetics

v

The velocity of a typical enzymatic reaction is described by the Michaelis-Menten equation:

$$v = \frac{[\mathsf{E}_{\mathsf{T}}][\mathsf{S}]k_{cat}}{\mathsf{K}_m + [\mathsf{S}]} \tag{A.1}$$

where

the initial velocity of the reaction;
 a measure of the rate of depletion of substrate or formation of product

- $[E_T]$ = the total amount of enzyme present (free and associated with substrate)
- [S] = the substrate concentration

 k_{cat} = the catalytic constant

 K_m = the Michaelis constant: that substrate concentration where $v = V_{max}/2$ (half the maximal velocity)

This expression is applicable when two conditions are met: (i) $[E_T] \ll [S]$, which is generally true since enzymes are usually very efficient catalysts; (ii) v is a true measure of the *initial* reaction velocity; that is, v is measured over a time period in which a small (<10%) amount of substrate has been depleted, or product, formed.



FIGURE A.1: The Michaelis-Menten curve for a typical enzymatic reaction.

The reverse of the reaction can be then be ignored and [S] will change approximately linearly with time.

There are three distinct regions to the curve described by Equation A.1: (i) *First order in substrate* ($[S] \ll K_m$): most of the enzyme is not bound to substrate; therefore the amount of free enzyme ([E]) can be approximated as the total amount of enzyme present ($[E_T]$). Equation A.1 then becomes:

$$v = \frac{[\mathsf{E}][\mathsf{S}]k_{cat}}{\mathsf{K}_m} \tag{A.2}$$

The parameter k_{cat}/K_m is the second order rate constant which reflects the first irreversible step in reference to the free enzyme (E) and substrate (S).

(ii) Zero order in substrate ([S] $\gg K_m$): the initial reaction rate is independent of [S], and Equation A.1 becomes:

$$v = V_{max} = k_{cat}[\mathsf{E}_{\mathsf{T}}] \tag{A.3}$$

The catalytic constant, k_{cat} , reflects the rate limiting step of catalysis.

(iii) In the region of the Michaelis-Menten curve where $[S] \approx K_m$ the reaction is between first and zero order with respect to [S].

The Michaelis-Menten expression is often graphed in double reciprocal form as a Lineweaver-Burk plot. This transformation results in a linear function:

$$\frac{1}{v} = \frac{1}{[\mathsf{E}_{\mathsf{T}}]k_{cat}} + \left(\frac{\mathsf{K}_m}{[\mathsf{E}_{\mathsf{T}}]k_{cat}}\right)\frac{1}{[\mathsf{S}]} \tag{A.4}$$

or substituting in Equation A.3:

$$\frac{1}{v} = \frac{1}{V_{max}} + \left(\frac{\mathbf{K}_m}{V_{max}}\right) \frac{1}{[\mathsf{S}]} \tag{A.5}$$



1/Substrate Concentration

FIGURE A.2: The Lineweaver-Burk (double reciprocal) plot.

A.2 The Interpretation of k_{cat} and k_{cat}/K_m

The rate constants k_{cat} and k_{cat}/K_m reflect the free energy change associated with the rate determining step, and the first irreversible step of catalysis, respectively. This is demonstrated by the following argument.

Consider the following enzymatic reaction where the formation of $E \cdot S$ is the *association step* and the interconversion of $E \cdot S$ to $E \cdot P$ is the *chemical step*, followed by *product release*:

$$E + S \xrightarrow{k_1} E \cdot S \xrightarrow{k_2} E \cdot P \xrightarrow{k_3} E + P$$

The reaction coordinate diagram for this scheme is depicted in Figure A.3. It can be shown that the parameters k_{cat} , K_m and k_{cat}/K_m are defined as:

$$k_{cat} = \frac{k_2 k_3}{k_{-2} + k_2 + k_3} \tag{A.6}$$

$$K_m = \frac{k_3(k_{-1}+k_2)+k_{-1}k_{-2}}{k_1(k_{-2}+k_2+k_3)}$$
(A.7)

$$\frac{k_{cat}}{K_m} = \frac{k_1 k_2 k_3}{k_3 (k_{-1} + k_2) + k_{-1} k_{-2}}$$
(A.8)

Consider a reaction with rapid, reversible association followed by a rate determining chemical step, and product release $(k_2 \ll k_{-1}, k_{-2} \ll k_3, k_2 \ll k_3)$. Under these conditions, the kinetic parameters k_{cat} and k_{cat}/K_m are reduced to:

$$k_{cat} = k_2 \tag{A.9}$$

$$\frac{k_{cat}}{\mathbf{K}_m} = \frac{k_1 k_2}{k_{-1}} \tag{A.10}$$

The principles of transition state theory can be applied to Equations A.9 and A.10 to relate these rates constants to changes in free energy associated with the discrete



Reaction Coordinate

FIGURE A.3: Reaction coordinate diagram for a typical enzymatic reaction. Energy levels are arbitrary.

steps of catalysis:

$$k_{cat} = \frac{\mathbf{k}T}{h} \left(e^{[-(G_{\mathsf{ES}}^{\dagger} - G_{\mathsf{ES}})]/RT} \right)$$

$$= \frac{\mathbf{k}T}{h} \left(e^{-\Delta G_M^{\dagger}/RT} \right)$$
(A.11)
$$\frac{k_{cat}}{K_m} = \frac{\mathbf{k}T}{h} \left(e^{[-(G_{\mathsf{ES}}^{\dagger} - G_{\mathsf{E+S}})]/RT} \right)$$

$$= \frac{\mathbf{k}T}{h} \left(e^{-\Delta G_F^{\dagger}/RT} \right)$$
(A.12)

where

- \mathbf{k} = Boltzmann's constant T = temperature h = Planck's constant
- R = the gas constant

Both k_{cat} and k_{cat}/K_m reflect the free energy change associated with the rate determining chemical step (the formation of E·P), but k_{cat} uses the Michaelis-Menten complex (E·S) as a reference, and k_{cat}/K_m refers back to free enzyme (E) and substrate (S).

Now consider a reaction in which there is rapid, reversible association, followed by an *irreversible chemical step* and *rate determining product release* $(k_2 \ll k_{-1}, k_3 \ll k_2, k_{-2} \ll k_3)$. Under these conditions, k_{cat} and k_{cat}/K_m reduce to:

$$k_{cat} = k_3 \tag{A.13}$$

$$\frac{\kappa_{cat}}{\mathbf{K}_m} = \frac{\kappa_1 \kappa_2}{k_{-1}} \tag{A.14}$$

Again, from transition state theory, these expressions become:

$$k_{cat} = \frac{\mathbf{k}T}{h} \left(e^{\left[-(G_{\mathsf{EP}}^{\dagger} - G_{\mathsf{EP}})\right]/RT} \right)$$
$$= \frac{\mathbf{k}T}{h} \left(e^{-\Delta G_{\mathsf{EP}}^{\dagger}/RT} \right)$$
(A.15)

$$\frac{k_{cat}}{\mathbf{K}_m} = \frac{\mathbf{k}T}{h} \left(e^{-\Delta G_F^{\ddagger}/RT} \right) \tag{A.16}$$

In this case, k_{cat} reflects the energy barrier for product release, yet k_{cat}/K_m again reflects that for the irreversible chemical step from free E and S.

This argument demonstrates that k_{cat} is the rate constant for the rate determining step, while k_{cat}/K_m is that for the first irreversible step of catalysis with respect to E and S.

In the hydrolysis of glycosides by BgaX, the formation of E·P and subsequent product release is analogous to glycosylation and deglycosylation respectively. Since k_{cat}/K_m reflects the free energy change associated with the first irreversible chemical step, it is the rate constant for glycosylation. However, the first order rate constant, k_{cat} , may be associated with glycosylation or deglycosylation, depending on which step is rate limiting.

A.3 Inactivation

The inactivation of an enzyme (E) by a mechanism-based inactivator (IX) and subsequent reactivation by water is summarized by the following, where $k_3 \ll k_2$:

E + IX
$$\xrightarrow{k_1}$$
 E \cdot IX $\xrightarrow{k_2}$ EI $\xrightarrow{k_3}$ E + IOH
HX H₂O

The kinetic expression for this scheme is analogous tó the Michaelis-Menten equation (A.1):

$$v = \frac{[\mathsf{E}_{\mathsf{T}}][\mathsf{IX}]k_i}{\mathsf{K}_i + [\mathsf{IX}]} \tag{A.17}$$

where k_i = the inactivation rate constant K_i = the apparent dissociation constant between E and IX

If $[IX] \gg [E_T]$, then [IX] will change very little over the course of inactivation. This process will then be pseudo zero order with respect to [IX], and the reaction rate will be:

$$v = k_i^{obs}[\mathsf{E}_{\mathsf{T}}] \tag{A.18}$$

and Equation A.17 becomes

$$k_i^{obs} = \frac{[\mathsf{IX}]k_i}{\mathsf{K}_i + [\mathsf{IX}]} \tag{A.19}$$

The above equation is often presented in double reciprocal form as a linear function:

$$\frac{1}{k_i^{obs}} = \left(\frac{\mathbf{K}_i}{k_i}\right) \frac{1}{[\mathsf{IX}]} + \frac{1}{k_i} \tag{A.20}$$

Values of k_i^{obs} (the observed rate of inactivation) are obtained by fitting residual enzyme activity over time to a first order expression:

$$\frac{v}{V_o} = A\left(e^{-k_i^{obs}t}\right) + C \tag{A.21}$$

where

 $\frac{v}{V_o}$ = the fractional enzymatic activity at time t

A, C = constants

The values thus obtained for k_i^{obs} are then used with Equation A.20 to derive k_i and K_i .

A.4 The Determination of K_i for a

competitive inhibitor

The association of a competitive inhibitor (P) with an enzyme and substrate is summarized by the following:

$$E + S \xrightarrow{k_1} E \cdot S \xrightarrow{k_2} E + P$$

$$+ P$$

$$k'_1 \bigvee k'_{-1}$$

$$E \cdot P$$

It can be shown that in the presence of P, the equation describing enzyme catalysis (A.1) becomes:

$$v = \frac{V_{max}[\mathsf{S}]}{[\mathsf{S}] + \mathsf{K}_m(1 + [\mathsf{P}]/\mathsf{K}_i)}$$
(A.22)

where
$$K_i^{+} =$$
 the apparent dissociation constant
between E and P



1/Substrate Concentration

FIGURE A.4: The Lineweaver-Burk plots of enzyme catalysis in the presence $(slope_i)$ and absence $(slope_o)$ of P. The plots intercept on the y-axis, a pattern characteristic of competitive inhibition.

or taking the reciprocal of the equation:

$$\frac{1}{v} = \frac{1}{V_{max}} + \frac{1}{[\mathsf{S}]} \left(\frac{\mathsf{K}_m}{V_{max}}\right) \left(1 + \frac{[\mathsf{P}]}{\mathsf{K}_i}\right) \tag{A.23}$$

Comparison of Equations A.5 and A.23 indicates that in the presence of P, the apparent K_m of the substrate with the enzyme decreases by a factor of $(1 + [P]/K_i)$, but V_{max} is unchanged. Therefore, the slope of the Lineweaver-Burk plot of enzyme catalysis in the presence of P (slope_i) should differ from that in the absence of P (slope_o) by a factor of $(1 + [P]/K_i)$, but both plots will still have the same y-intercept. That is:

$$\frac{\text{slope}_i}{\text{slope}_o} = 1 + \frac{[\mathsf{P}]}{\mathrm{K}_i} \tag{A.24}$$

or with rearranging

$$\mathbf{K}_{i} = \frac{[\mathsf{P}]}{\frac{\mathrm{slope}_{i}}{\mathrm{slope}_{o}} - 1} \tag{A.25}$$

Therefore, an approximate value for the K_i of a competitive inhibitor P can be obtained by the determination of the slopes of the double reciprocal plots defined by Equation A.5 in the presence and absence of P, and application of Equation A.25.

A.5 Protection from inactivation

The illustration that an inhibitor (I) is active site-directed can be demonstrated by the protection of an enzyme from inactivation with a competitive inhibitor (P) as shown:

$$E + I \xrightarrow{k_1} E \cdot I \xrightarrow{k_2} EI$$

$$+ P$$

$$k'_1 | k'_{-1}$$

$$E \cdot P$$

In the presence of P, it can be shown that the expression describing inactivation (Equation A.19) becomes:

$$k_{ip}^{obs} = \frac{[\mathbf{I}]k_{ip}}{\mathbf{K}_i \left(1 + \frac{[\mathbf{P}]}{\mathbf{K}_p}\right) + [\mathbf{I}]}$$
(A.26)

where $K_p =$ the apparent dissociation constant between E and P

Combining Equations A.19 and A.26 and rearranging gives:

$$\frac{k_{ip}^{obs}}{k_i^{obs}} = \frac{\mathbf{K}_i + [\mathbf{I}]}{\mathbf{K}_i \left(1 + \frac{[\mathbf{P}]}{\mathbf{K}_p}\right) + [\mathbf{I}]}$$
(A.27)

Therefore, if the ratio of the observed rates of enzyme inactivation in the presence (k_{ip}^{obs}) and absence (k_i^{obs}) of P agrees with Equation A.27, then it can be concluded that

inactivation is affected by the presence of a competitive inhibitor, and the inactivator under study is active site-directed.

A.6 Reactivation

The reactivation of a glycosidase inhibited by a 2-deoxy-2-fluoro-glycoside (EI) by transglycosylation to a suitable acceptor (A) is summarized in the following scheme:

$$EI + A \xrightarrow{k_1} EI \cdot A \xrightarrow{k_2} E + IA$$

The observed increase in fractional enzyme activity (v/V_o) follows a first order expression:

$$\frac{v}{V_o} = A \left(1 - e^{-k_{react}^{obs} t} \right) + C \tag{A.28}$$

where k_{react}^{obs} = the observed rate of reactivation

In order for adequate comparisons to be made between the reactivation profiles associated with the various acceptors studied in Chapter 3, the initial rate of reactivation (V_o^{react}) was extrapolated from the above equation:

$$V_o^{react} = \frac{d}{dt} \left(\frac{v}{V_o} \right) \Big|_{t=0} = \frac{d}{dt} \left(A \left(1 - e^{-k_{react}^{obs}} \right) + C \right) \Big|_{t=0}$$
$$= -Ae^{-k_{react}^{obs}} (-k_{react}^{obs}) \Big|_{t=0}$$
$$= Ak_{react}^{obs} (e^{-k_{react}^{obs}}) \Big|_{t=0}$$
(A.29)

Values of V_o^{react} were thus calculated and used to rank those glycosides that were identified as positively screened acceptors.



FIGURE A.5: General plot of enzyme reactivation depicting the first order reactivation curve and the extrapolated slope of V_o^{react} .

APPENDIX B

Graphical Representation of Data

B.1 General



FIGURE B.1: The determination of the molar absorptivity (ε) of \circ pNP and \bullet DNP with respect to pH. The pK_a values for pNP and DNP were calculated to be 7.18 and 3.81 respectively from the equation $\varepsilon = \frac{L1+L2(10^{(pH-pK_a)})}{10^{(pH-pK_a)}+1}$.

B.2 Enzyme kinetics

The following seven plots depict the raw data for the initial reaction velocity versus substrate concentration for the hydrolysis of pNP- and DNP-glycosides by BgaX and fitting to $v = \frac{V_{max}[S]}{K_m + [S]}$.







108

B.3 Enzyme reactivation

The following plots depict the raw reactivation profiles (initial velocity versus time) in the presence of positively screened acceptors for the seven glycosidases studied after being corrected for enzyme death over time. The names of some of the acceptors have been abbreviated; consult the list of abbreviations for the full names of all glycosides. The profiles of each enzyme are separated into two or three plots for clarity. Note the scale of the y-axis ($\Delta A_{400}/min$) may differ between graphs for a single enzyme to better accommodate the data. Data was fit to $v = A \left(1 - e^{k_{trans}t}\right) + C$.

Reactivation profiles (3) for BgaX:

Note that the curve for pNPGlc was fit with the data up to t = 25 hr due to deviations from a first order expression at times greater than this.





Reactivation profiles (3) for BgaC:

Since the activity of BgaC failed to approach a maximum value over the course of the experiment, each data set was fit to a linear expression to directly extrapolate V_o^{react} . The value of V_o^{react} for maltotriose (which fits to a full first order equation) was estimated from the data up to *ca*. 40 hr. The V_o^{react} value for *p*NPGlc was estimated from the data up to 50 hr. Data collected after this time failed to correspond to a linear function.



•



Reactivation profiles (2) for Man2A:

The first order curve for pNPXyl and pNPMan were fit using the data up to ca. 2.5 hr.



0.15 DNPCell 0 • $v(\Delta A_{400}/min)$ 0 *p*NPGIc 0.10 × Control 0.05 0.00 2 Time (hr) 4 Ó 5 1 0.15 *p*NPCell • v(∆A₄₀₀∕min) PhGlc 0.10 Δ Gentiobiose 0 *p*NPXyI Glc ٠ *p*NPGal 0.05 Control Х 0.00 5 20 10 15 Time (hr) 25 Ó

Reactivation profiles (2) for Abg:

. · *

Reactivation profiles (2) for CelB:



Reactivation profiles (2) for Cex:

Each data set was fit linearly to directly determine V_o^{react} . The slope associated with reactivation in the presence of pNPXyl was fit with the data up to *ca.* 30 hr.





Reactivation profiles (3) for HBG:

