SURVEY OF THE GENOME OF *OPITUTUS TERRAE* AND OTHER ORGANISMS FOR NOVEL CARBOHYDRATE-ACTIVE ENZYME SPECIFICITIES

MARIACRISTINA BAWN

A thesis submitted in partial fulfilment of the requirements of the University of Northumbria at Newcastle for the degree of Doctor of Philosophy

Research undertaken in the
School of Life Sciences
University of Northumbria

May 2012

Dedication

This work is dedicated to my Mum and Dad, for without their support,

love and, most of all, their endless faith in me, this work would not have been possible.

Abstract

Lignocellulose is the major component of the plant cell wall and is a sustainable source of inexpensive abundant biomass. Efficient degradation of the lignocellulosic polysaccharides, cellulose and hemicellulose, is required if the plant cell wall is to be used as a resource for renewable biofuels. Microorganisms have the ability to catalyse the degradation of such bio-material through a cascade of enzyme activities into fermentable sugars and therefore are considered to be a major resource of biocatalysts for the emerging biofuel industry. The stability of the component polysaccharides and the complexity of the plant cell wall are reflected in the diverse range of functions and substrate specificities of lignocellulosic degrading enzymes.

Part of this work describes the identification of a novel GH8 endo-xylanase, OtXyn8A, from the soil bacterium, *Opitutus terrae*. GH8 is a family in which there is only a limited amount of data available on the xylanase substrate specificity in comparison to families GH10 and GH11 in which xylanases are well established. With this in mind, OtXyn8A is the only endo-xylanase characterised from GH8 that primarily releases xylobiose from its substrates.

Synergy between *O. terrae* enzymes was partially investigated within this study with the identification of a gene cluster within the bacterial genome. Genes organised within this cluster encoded products required for the degradation of xylan substrates and so the associated enzymes were cloned, expressed and subsequently determined for activity. Combined activities of gene products from the cluster exhibited synergy in the hydrolysis of 4-*O*-methyl glucuronoxylan.

While surveying the genome of *O. terrae*, the multiplication of genes encoding GH43 enzymes was also investigated. Genes encoding GH43 enzymes were cloned, expressed and investigated for catalytic activity. Three arabinofuranosidases from *O. terrae* and one from *Lactobacillus brevis* were characterised, including the characterisation of an exo-1,5-L-arabinofuranosidase. Furthermore, a β-xylosidase from *O. terrae* was characterised which exhibited dual functionality as it catalysed the release of arabinose in addition to xylose from arabino-xylooligosaccharides.

Contents

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Abbreviation	Term
Å	angstrom
A_x	absorbance at x nm
A1	arabinose
A2	arabinobiose
A3	arabinotriose
A4	arabinotetraose
A5	arabinopentaose
A6	arabinohexaose
A7	arabinoheptaose
Asp	aspartic acid
bp	base pair
BSA	bovine serum albumin
°C	degree Celsius
C	cytosine
CAZY	Carbohydrate-Active Enzyme Database
CBM	carbohydrate binding module
CE	carbohydrate esterase
CFE	cell-free extract
CjGlcA67A	*Cellvibrio japonicus* glucuronidase 67A
CjXyn10A	*Cellvibrio japonicus* xylanase 10A
CmXyn10B	*Cellvibrio mixtus* xylanase 10B
CMC	carboxymethylcellulose
C-terminal	carboxy terminal

3D	three-dimensional
Da	dalton
DMSO	dimethyl sulphoxide
DNA	deoxyribonucleic acid
DNSA	dinitrosalicyclic acid
dNTPs	deoxynucleotide triphosphates
EC	enzyme commission
EDTA	Ethylene diamine tetraacetic acid, disodium salt
G	guanidine
gDNA	genomic DNA
GH	glycoside hydrolase
Glu	glutamic acid
h	hour(s)
HPAEC-PAD	High performance anion exchange chromatography - pulsed amperometric detection
IMAC	immobilised metal affinity chromatography
IPTG	isopropyl-beta-D-thiogalactopyranoside
IUBMB	International Union of Biochemistry and Molecular Biology
k_{cat}	turnover number
k_{cat}/ K_M	catalytic efficiency
Kb	kilo bases
kDa	kilo Dalton
K_M	Michaelis-Menten constant
L	litre(s)
LB	Luria-Bertani medium
LbAraf43A	*Lactobacillus brevis* arabinofuranosidase 43A
LbXyl43A	*Lactobacillus brevis* xylosidase 43A

M	mole(s)
mA	milliamps
mg	milligram(s)
min	minute(s)
mL	millilitre(s)
MW	molecular weight
nC	nanocoulomb
NCBI	National Center for Biotechnology Information
ng	nanogram(s)
NpXyn11A	*Neocallimastix patriciarum* xylanase 11A
NREL	National Renewable Energy Laboratory
N-terminal	amino terminal
OD_x	optical density at x nm
OtGlcA67A	*Opitutus terrae* glucuronidase 67A
OSX	oat spelt xylan
OtAraf43A	*Opitutus terrae* arabinanase 43A
OtAraf43B	*Opitutus terrae* arabinanase 43B
OtAraf43A	*Opitutus terrae* arabinanase 43C
OtXyl52A	*Opitutus terrae* xylosidase 52A
OtXyl43A	*Opitutus terrae* xylosidase 43A
OtXyn8A	*Opitutus terrae* xylanase 8A
PCR	polymerase chain reaction
PCW	plant cell wall
PhXyl	*Pseudoalteromonas haloplanktis* xylanase
PL	polysaccharide lyase
psi	pounds per square inch
REX	*Bacillus halodurans* exo-oligoxylanases

REXA	*Bifidobacterium adolescentis* exo-oligoxylanases
rpm	revolutions per minute
s	second(s)
ScAra43A	*Streptomyces coelicolor* arabinanase 43A
SDS-PAGE	sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SGD	Saccharomyces Genome Database
TAE	tris-acetate-EDTA
TAXI	*Triticum aestivum* xylanase inhibitor
TEMED	N,N,N',N'-Tetramethylethylenediamine
T_m	melting temperature
Tris	tris(hydroxymethyl)aminomethane
UV	ultraviolet
V_0	initial velocity
V	volts
V_{max}	maximum velocity
v/v	volume per volume
w/v	weight per volume
X1	xylose
X2	xylobiose
X3	xylotriose
X4	xylotetraose
X5	xylopentaose
X6	xylohexaose
X7	xyloheptaose
x g	times gravity
XIP	xylanase inhibitor protein
XYLY	*Bacillus* sp. KK-1 xylanase 8

Xyn8	uncultured bacterium xylanase 8
α	alpha
β	beta
ε	Molar absorptivity
Σ	sum
µg	micro gram(s)
µL	micro litre(s)
µM	micro molar
18.2 MΩ H_20	18.2 mega ohm water

Acknowledgements

First and foremost, I would like to thank Prof. Gary Black and Dr. Simon Charnock. Not only have they provided me with the amazing opportunity to do this Ph.D, but also for their invaluable support and mentoring. I could not have asked for better, more approachable supervisors and I will always cherish their guidance.

I am deeply grateful to Dr. Meng Zhang for providing endless advice and knowledge throughout this project. Also a special thanks to Dr. Ruth Lloyd and everyone at Prozomix Ltd. for making me part of their team over the years.

I would like to extend my acknowledgement to all my colleagues, past and present, especially Andrew, Vatsala and Stephen for sharing the good (and not so good) days in lab A321.

Financial support for this project was provided by Prozomix Ltd. and the Engineering and Physical Sciences Research Council (EPSRC).

Great appreciation goes to my friends for their support and in particular Sophia, for always being there, especially through the more difficult times over the past three years. I am thankful to Dave, for while also studying for his Ph.D, has provided constant, unselfish support that gave me the energy to carry on- hopefully I can return the favour.

Finally, I owe a lot to my wonderful, crazy family (including Toby and Oscar). I am thankful to my three brothers, Gino, Carlo and Gianni for their encouragement, but ultimately everything I have achieved would not have been possible without my Mum and Dad. A word of acknowledgement does not suffice for their role in all of this but they have always encouraged me to keep on going and are forever believing in me- something I am truly grateful and I will never forget.

Declaration

I declare that the work contained within this thesis has not been submitted for any other qualification and the work is my own.

NAME: MARIACRISTINA BAWN

SIGNATURE:

DATE: 31st May 2012

1 Introduction

1.1 Background of biomass conversion

Renewable energy is becoming a realistic alternative fuel resource as economical and environmental concerns create uncertainty within global markets due to increased energy requirements, rising oil prices and reduced fossil fuel reserves. In terms of a sustainable bio-based economy, the biorefinery concept (established by the National Renewable Energy Laboratory, NREL) has emerged to seek several innovative synthetic biofuels such as bioethanol and biodiesel from biomass. With this in mind biomass has become a maintainable prosperity in the future of renewable energy. Sugar cane and corn biomass are currently the most recognised sources for the potential production of first generation biofuels from sucrose and starch respectively (Kumar *et al*., 2009). However, the use of these and other crops such as wheat is currently receiving significant global political attention as they pose the ethical issue of using food source for fuel as well as the requirement of fertile land. Thus the focus now lies with the development of second generation biofuels, such as bioethanol from lignocellulosic biomass. High quality yield, minimal nutrient and water requirement (Dodd *et al*., 2009), make the perennial grasses, Switchgrass (USA), Miscanthus Grass (Europe) and Sorghum (Asia) the most likely short-term biomass crop candidates for their cellulose and hemicellulose components.

Lignocellulose is the major plant cell wall (PCW) component of wood, grasses and plants and is therefore a sustainable source of abundant and inexpensive biomass. Conversion of biomass into its subsequent fermentable sugars is generally achieved by biochemical or enzymatic hydrolysis. Biochemical hydrolysis is carried out at severe and harsh conditions such as extremely low pH and high temperatures which have many disadvantages including the degradation of the desired products (Taherzadeh *et al*., 2007). Enzyme catalysed hydrolysis does not show any adverse effects on the sugar product and is a much ‘greener’

approach. Therefore the interest and prominence of enzyme technology has increased over recent years. Enzyme hydrolysis still begins with a pretreatment stage due to the recalcitrant nature of lignocellulosic biomass, followed by the actual enzymatic hydrolysis of the modified substrate, with ultimate fermentation of the released monosaccharides. The biorefinery concept combines technologies such as biocatalysis to obtain a variety of end products according to chemical component of the stating material which can further be converted to platform chemicals and biofuels (Turner *et al*., 2007). However, the enzymatic hydrolysis step is still far from economically viable, given the complex nature of the substrate and is an area of intense research.

1.2 Overview of PCW polysaccharides and lignin

Optimising lignocellulose biomass for sufficient bioenergy production requires an understanding of the dynamics and structure of the PCW. The lignocellulosic PCW is a composite of closely associated polymer layers which provide structural integrity and functional support for the main plant cell. Lignocellulose comprises the three major components lignin, cellulose and hemicellulose which are thought to be connected by a series of covalent cross linkages between lignin and the latter polysaccharide structures.

1.2.1 Cellulose

Cellulose is the most abundant organic molecule in the biosphere. It is produced mainly by plants and exists as the main polysaccharide constituent of PCWs even though several other organisms, including bacteria and algae, also synthesise cellulose (Brown, 2004). Cellulose is a high molecular weight linear polymer comprising β-1,4-linked glucose monomers with repeating units of cellobiose, as illustrated in Figure 1.1. Cellulose purely fulfils a structural

role within the PCW and is synthesised in long chains by cellulose synthase (Saxena *et al*., 2001). Since cellulose chains are entirely linear with no substituted groups, they reside as parallel molecules which can become associated through extensive Van der Waals forces and intra and inter molecular hydrogen bonds creating microfibrils providing integral strength to the PCW (Volynets, 2011). Furthermore, these microfibrils have a crystalline structure of which six different crystal forms have been identified. Cellulose I is the predominant form and consists of alternating less ordered amorphous regions with highly ordered crystalline regions of the cellulose microfibril (Brown, 1994; Park *et al*., 2010). Depending upon the plant species, the ratio of amorphous to crystalline cellulose varies greatly and it has been well established that the amorphous domains are most accessible to reactants and enzymatic attack leaving crystalline regions unaffected (Bertran *et al*., 1984).

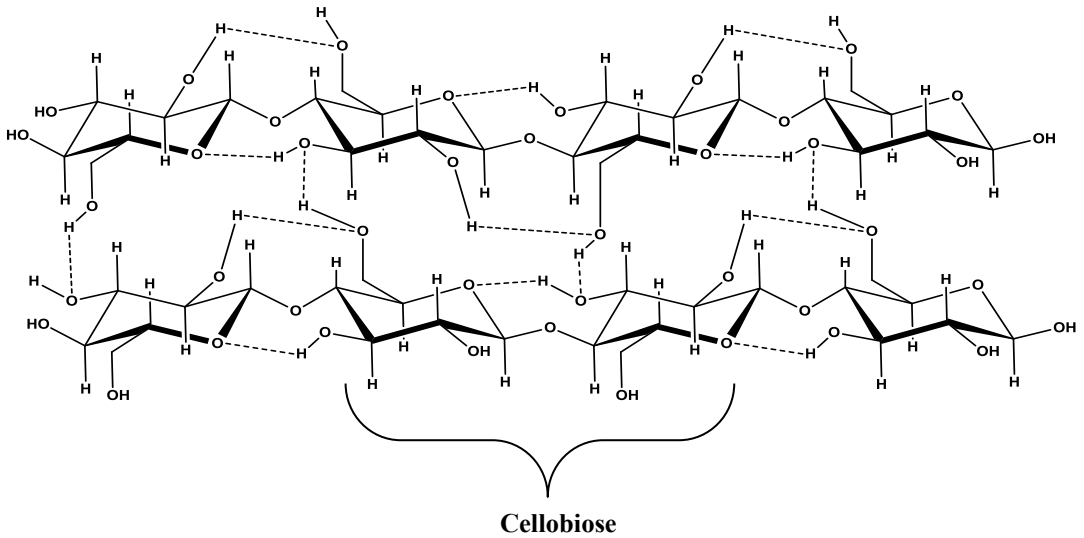


Figure 1.1 Structural representation of the structure of cellulose microfibrils. Schematic view of extensive intra and inter molecular hydrogen bonds between cellulose chains.

1.2.2 Hemicelluloses

In contrast to cellulose, the hemicelluloses are low molecular weight heterogeneous polysaccharides which are ultimately produced by a different biosynthetic route to that of cellulose (Sjöström, 1993). Originally hemicelluloses were thought by E. Schulze, 1892, to be precursors of cellulose, however it is now well recognised that the hemicelluloses comprise the pentose sugars, xylose and arabinose, the hexose sugars, glucose, mannose and galactose and a group of sugar acids. Hemicelluloses are quite variable in nature but can be classed as either hardwood or softwood polysaccharides. Hardwood hemicelluloses generally contain xylan which is the most abundant hemicellulose in nature (Saha, 2003). Alternatively softwood hemicelluloses exist mostly as glucomannans. The polysaccharide chains of such structures are linked to cellulose microfibrils by an assortment of weak hydrogen bonds and ether bonds that can be readily hydrolysed by acid or alkaline conditions. There are many potential uses for these hemicellulose derived carbohydrates and over recent years these products have been targeted for biomass energy sources.

1.2.2.1 Xylan

Xylan is a heteropolymeric polymer comprising a β-1,4-linked xylopyranose backbone which can usually be substituted with arabinofuranosyl, acetyl and 4-*O*-methyl glucuronyl groups (Figure 1.2). After cellulose, xylan is the second most abundant polysaccharide in nature accounting for 25-35% of total PCW material (Saha, 2003). The composition of xylan polymers differ depending upon the biomass resource with the differences varying between softwood and hardwood hemicelluloses.

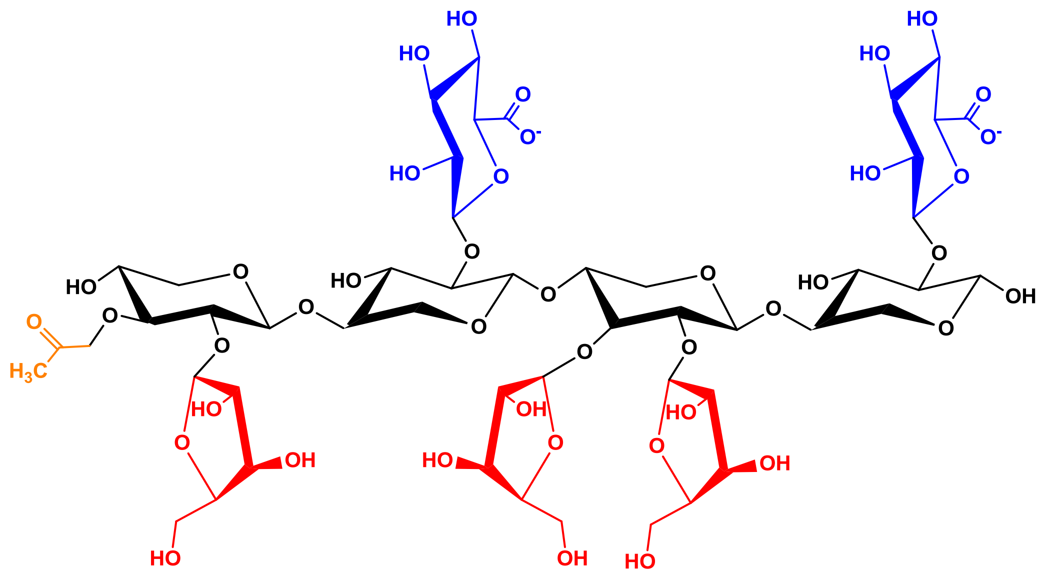


Figure 1.2 Structural representation of xylan polysaccharide. Schematic view of linkages present in arabinoxylan and glucuronoxylan, Black = β-1,4-xylose; Red = α-1,2- and α-1,3- arabinose ; Blue = α-1,2-glucuronic acid; Yellow = acetyl groups.

1.2.2.1.1 Arabinoxylan

Arabinoxylan is predominantly found in the PCWs of cereals and grasses (Heinze *et al*., 2005) including wheat and rye, generally consisting of a β-1,4-linked xylose backbone with arabinose residues linked to either the C-2 or C-3 position on the xylose (Figure 1.2). Depending on the tissue, arabinoxylan can be extracted by water or by alkaline solvent. This highly depends on the differences in the physical and chemical interactions with other cellular wall components and the degree of arabinose substituents (Frederix *et al*., 2004). Water extractable arabinoxylans are loosely connected at the cell wall surface whilst alkaline extracted arabinoxylans are strongly embedded in the cell wall matrix due to interactions with other components such as cellulose and lignin. These associations between arabinoxylan and cell wall components determine different functional and chemical properties of such sugar. The water extractable arabinoxylan has high viscosity forming potential while the alkaline solvent extracted arabinoxylans appear to have a high water

holding capacity (Frederix *et al*., 2004) and such properties dictates their uses in applied biotechnological processes.

1.2.2.1.2 Glucuronoxylan

Glucuronoxylan, also known as 4-*O*-methyl-D-glucuronoxylan, is the most prominent hemicellulose occurring in hardwoods and comprises a β-1,4-linked xylose backbone, substituted with α-1,2-linked 4-*O*-methyl-D-glucuronic acid approximately per tenth xylose unit (Gröndahl *et al*., 2003) as demonstrated in Figure 1.2. Furthermore, the xylopyranose units of the backbone can be acetylated at the C-2 and C-3 positions with an uronic acid group giving more structural diversity to glucuronoxylans.

1.2.2.1.3 Arabinoglucuronoxylan

Arabinoglucuronoxylan can usually be found within the supporting lignified tissues of softwood xylans. Even though no acetyl groups are present, there is a higher degree of polymerisation surrounding such softwood xylans in comparison to hardwood xylans. The structure of arabinoglucuronoxylan encompasses a β-1,4-D-xylopyranose backbone with 4-*O*-methyl-D-glucuronic acid and α-L-arabinofuranose residues at positions C-2 and C-3 (Achyuthan *et al*., 2010).

1.2.2.2 Arabinan

Arabinan is a hemicellulose polymer comprising α-1,5-linked L-arabinofuranose residues that can be further decorated with double or single α-1,2- and 1 α-1,3- linked arabinose side chains (Figure 1.3). Arabinan is a pectic polysaccharide in nature since it is generally linked to the pectic molecule, rhamnogalacturonan (Alhassid *et al*., 2009). Rhamnoglacturonan is a polymer consisting of L-rhamnopyranose and D-galactopyranouronic acid disaccharide repeating units making up the main chain to which arabinans are attached (Ochiai *et al*., 2007). Rhamnoglacturonan is extensively subsitituted with arabinan forming the desiganted 'hairy region' of rhamnoglacturonan withihn pectic substances.

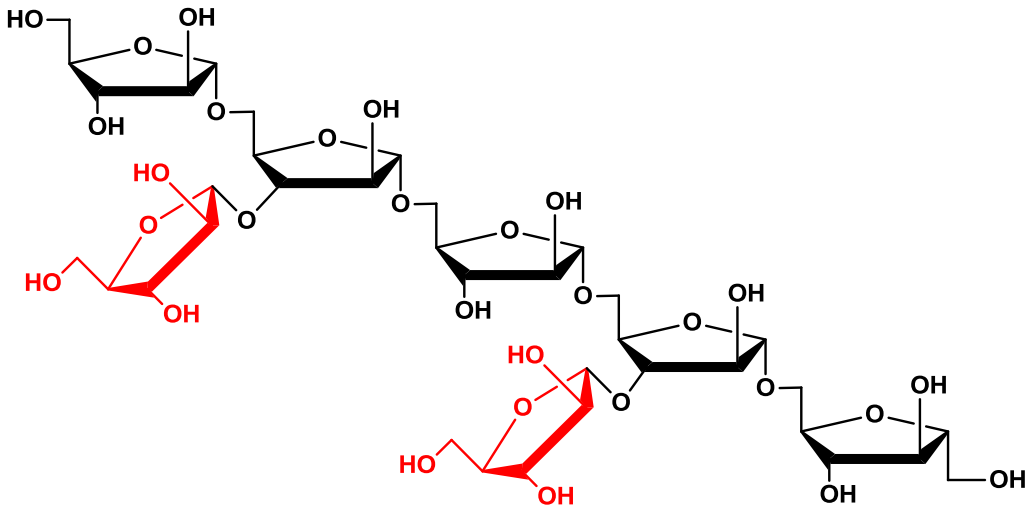


Figure 1.3 Structural representation of arabinan polysaccharide. Schematic view of linkages present in arabinan, Black = α-1,5-arabinofuranose; Red = α-1,2-arabinofuranose.

1.2.2.3 Mannan

Mannans are amoung the main hemicelluloses widespread in wood and plant seeds and are predominantly present as glucomannan, a major component of hardwood hemicelloses and galactoglucomannan, the main component of softwood hemicellulose (Agrawal *et al*., 2011). Glucomannans and galactoglucomannans are water soluble and are composed of a β-1,4-linked mannan chain interspersed wth glucose residues within the main chian (Figure 1.4). Furthermore, the mannose and glucose residues of the polysaccharide backbone of galactoglucomannan are partially substituted with galactosyl units via α-1,6-glycosidic bonds (Figure 1.4). Pure mannans comprising β-1,4-mannose units appear to be insoluble compared to glucomannans in which water solubility of the polymer increases (Do *et al*., 2009). Both the solubility and viscosity of mannans are influenced by the mannan to glucose ratio and the distribution of subsitutents along the mian chain. With this in mind, this structural variation of mannans presents a wide range of physio-chemical diversity which is currently being exploited in biotechnological processes within the food and textile industries (Ebringerová, 2005).

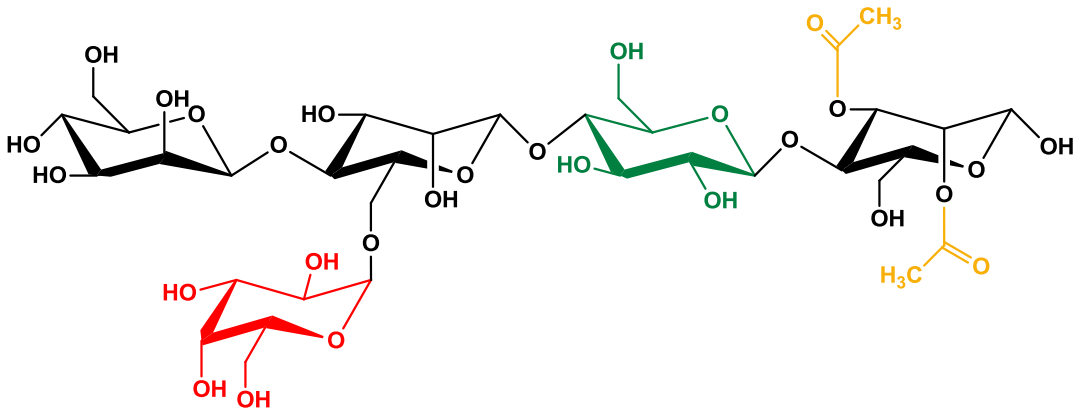


Figure 1.4 Structural representation of mannan polysaccharide. Schematic view of galactoglucomannan, Black = β-1,4-mannose; Green = β-1,4-glucose; Red = α-1,6-galactose; Yellow = acetyl groups.

1.2.3 Lignin

Lignin is a complex heteropolymer comprising aromatic hydroxycinnamyl alcohol monomers such as sinpyl alcohol, p-coumaryl alcohol and coniferyl alcohol (Boerjan *et al*, 2003). These alcohol subunits are usually designated as the monolignols and are synthesised by the phenylpropanoid specific pathway (Vanholme *et al*., 2010). The monolignol compounds have a basic structure encompassing a 3 carbon chain attached to an aromatic ring which originates from the structure of L-phenylalanine (Figure 1.5). Lignin is furthermore synthesised by the oxidative polymerisation of the monolignols (Hatfield *et al*., 2001) resulting in strong carbon-carbon and ether linkages.

Even given the heterogeneous complexity of PCW polysaccharides, lignin is predominantly responsible for the recalcitrant nature of lignocellulose due to its functional and chemical role within the PCW. The lignin polymer is usually deposited around the cell wall polysaccharides forming a matrix providing integral strength and rigidity to the PCW. Lignin helps protect the PCW polysaccharides from microbial degradation adding to the recalcitrance of the cell wall to enzymatic breakdown. From a chemical point of view, lignin is produced by an oxidative process in which a range of lignin peroxidases are required to break down such a structure. However, only a few microorganisms in nature are capable of degrading lignin with only a few being studied, one being white rot basidiomycete fungi (Geib *et al*., 2008). Therefore with lignin facilitating the protection of the PCW polysaccharides and with its own resistance to enzyme hydrolysis, this polymer remains one of the major limiting factors in enzymatic bioconversion meaning that delignification by pre-treatment of biomass is a critical step in biofuel production.

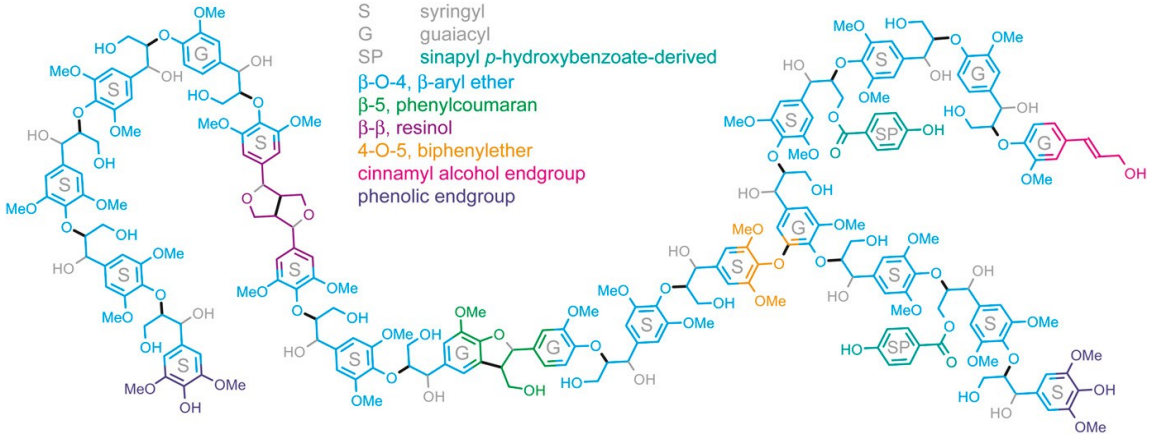


Figure 1.5 Schematic representation of the lignin polymer from poplar. NMR-based analysis of lignin reproduced from Vanholme *et al*., 2010.

1.3 General overview of PCW degrading enzymes

Together, microorganisms and plants play a major contribution to the carbon cycle. As plants fix carbon compounds into its cell walls, enzymes expressed by both saprophytic and pathogenic soil dwelling microbes breakdown such compounds into smaller units for transportation into the host organism as a major source of energy (St Leger *et al*., 1997). The extreme stability of component polysaccharides and the complexity of the PCW is emphasised by the diverse range of functions and substrate specificities of lignocellulosic degrading enzymes released by microorganisms. At present, the Carbohydrate Active Enzyme (CAZy) database is major resource tool for identifying such enzymes as well as categorising them into four main classes of enzyme activity, the classes being polysaccharide lyases, carbohydrate esterases, glycosyltransferases and glycoside hydrolases. This classification system is originally based on the grouping of catalytic domains of carbohydrate active enzymes (CAZymes) which furthermore has been subsequently divided into the sub-grouping of enzymes into amino acid sequence- based families (Henrissat, 1991) and structural similarities. Because of such organisation, these families allow the classification of uncharacterised proteins whose only known feature is the amino acid similarity to a characterised enzyme. Therefore, enzyme specificity may not necessarily correlate within a family. Traditionally enzymes have been classified in the International Union of Biochemistry and Molecular Biology (IUBMB) enzyme nomenclature as the type of reactions they catalyse. For example, glycoside hydrolases have been assigned the number EC 3.2.1.- which denotes their ability to hydrolyse glycosyl linkages, with the "-" representing the substrate specificity.

It should be noted that some carbohydrate active enzymes may require assistance for the interaction between enzyme and target substrate. This support usually occurs as a non-catalytic domain within the enzyme known as a carbohydrate binding module (CBM), thus

making the enzyme a modulated complex. CBMs are also classified into CAZy sequence based families in which there are currently 64 existing families.

The glycosyltransferases are not part of this research project and will not be considered further other than to state that they constitute a large group of enzymes that are involved in the biosynthesis of complex sugars generating oligosaccharides, polysaccharides and glycoconjugates (Breton *et al*., 2006).

1.4 Polysaccharide lyases

Polysaccharide lyases (PLs) (EC 4.2.2.-) recognise and cleave uronic acid residues-containing pyranoside substrates such as polygalacturonates, by β-elimination chemistry, producing unsaturated saccharides at their non-reducing end (Ochiai *et al*., 2007). Currently, PLs have been classified into 22 CAZy families based on relative sequence homologies however reviews in specific PL families appear limited in comparison to the extensive studies carried out on glycoside hydrolases and glycosyltransferases. The secretion of PLs from plant pathogens has been well documented and their action in bacterial pathogenesis within plants and fruit rotting are just one of the major manifestations of the pectinolytic enzymes. Polysaccharide depolymerisation by PLs has been characterised as β-eliminative cleavage of the glycosidic bond linking polymeric α-1,4-galacturonic acids and such a mechanism has been cited previously of PLs from *Erwinia chrysanthemi* (Scavetta *et al*., 1999), *Cellvibrio japonicus* (Charnock *et al*., 2002) and *Bacillus subtilis* (Seyedarabi *et al*., 2010).

1.5 Carbohydrate esterases

Ultimately, carbohydrate esterases (CEs) are involved in the removal of ester moieties present on substituted sugar residues, occurring as O- (ester) or N- (acetyl) groups on xylan, pectin and rhamnogalacturonan structures (Ospina-Giraldo *et al*., 2010). For instance, acetylxylan esterases are part of the microbial xylanolytic system in the breakdown of xylan in which they specifically hydrolyse the ester linkages of acetyl side groups found in positions C-2 and/ or C-3 of xylose moieties of the xylan backbone.

There are currently 16 existing families of CEs within the CAZy database, most of which contain enzymes included in the degradation of plant cell well polysaccharides. Three dimensional structures have been reported for at least 11 of the families to which the plethora of solved structure representatives display a typical α/β hydrolase fold with a Ser-His-Asp catalytic triad, except CEs from family 4 in which the catalytic domain exhibits a heavily distorted $(\beta/\alpha)_8$ barrel fold (Correia *et al*., 2008). A most recent review discusses a new acetylxylan esterase from *Geobacillus stearothermophilus* T-6, that does not share any sequence homology to any other carbohydrate esterase within the CAZy database and therefore possibly establishing a new family of carbohydrate esterases (Alalouf *et al*., 2011).

1.6 Glycoside hydrolases

Since most organisms utilise carbohydrates for carbon as their own energy source, they require an enzyme system which hydrolyses these very important glycosides. The enzyme class glycoside hydrolases (GHs) (EC 3.2.1.-) includes a very diverse range of PCW degrading enzymes targeting the complex carbohydrate structures. To date, over a hundred families of GHs have been classified within the CAZy database and are mainly responsible for the hydrolysis of the glycoside bond that link sugar residues within cellulose and hemicellulose polysaccharides. Additionally, GH families are grouped into structure-based

clans in which there are 14 existing clans. A clan consists of GH families that have similar tertiary structure, catalytic residues and utilise the same hydrolytic mechanism. There are currently seven tertiary folds displayed by GHs including $(\beta/\alpha)_8$, β-jelly roll, 6-fold β-propeller, 5-fold β-propeller, $(\alpha/\alpha)_6$, α+β and β-helix (Henrissat, 1991).

1.6.1 General structural and catalytic features of GHs

Depending on their mode of action, GHs can either be endo- or exo- acting and this is an essential balanced action required by all microbial enzyme systems. Exo-enzymes generally cleave the chain end of a polysaccharide or oligosaccharide whilst endo-acting enzymes are more internally specific cleaving randomly within the sugar residue. Such cleavage preference has been shown to be determined by the active site architecture of the enzyme. Endo-acting enzymes have been described to have an open cleft shaped active site (Proctor *et al*., 2005) for random interaction with the substrate (Figure 1.6). The active site of an exo-acting enzyme is a pocket-like structure enabling the enzyme to cleave the end of a polymeric chain (Figure 1.6) (Proctor *et al*., 2005). Furthermore, a tunnel shaped active site has been found in cellobiohydrolases in which cellobiose is cleaved and released from either the reducing end or non-reducing end of cellulose related oligosaccharides (Figure 1.6) (Varrot *et al*., 1999).

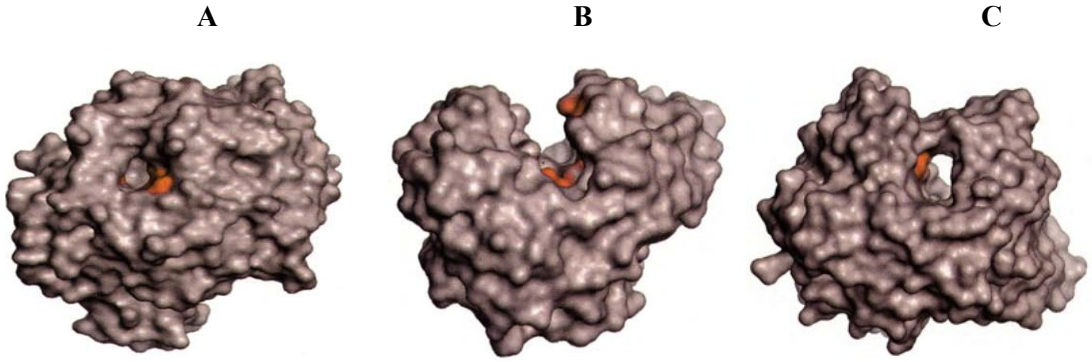


Figure 1.6 Surface topologies of glycoside hydrolases. 'A' shows the pocket-structured cleft found in exo-acting enzymes. 'B' is the open cleft topology exhibited by endo-acting enzymes. 'C' is the tunnel topology commonly found in processive cellobiohydrolases. Adapted from (Davies *et al*., 1995).

Since binding of numerous glycosyl residues occurs within the active site of a single glycoside hydrolase, a recommended nomenclature has been proposed which divides the binding surface into several subsites and represents the point of cleavage of the O-glycosidic bond (Davies *et al*., 1997). This nomenclature advises that subsites be labelled -3, -2, -1, +1, +2, +3 *etc*. in which non-reducing ends are represented in the negative direction while reducing ends are in the positive direction (Figure 1.7). The point of cleavage takes place between the -1 and +1 subsites indicating whether the non-reducing or reducing end is being enzymatically attacked.

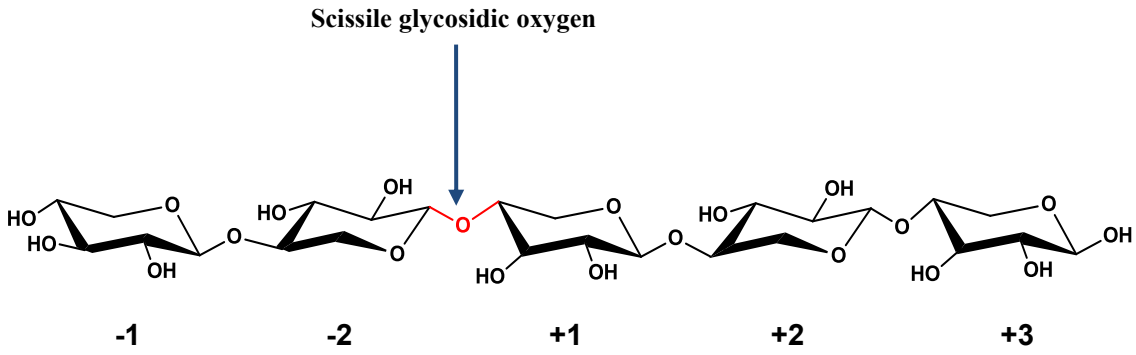


Figure 1.7 Schematic view of subsite nomenclature of GHs. Scissile glycosidic oxygen is shown in red while arrow points to site of bond cleavage.

Since the first enzyme X-ray structural determination in 1965 was the glycoside hydrolase lysozyme, the structure and function of GHs have been extensively studied and this has given insights into the enzymatic mechanism which glycoside hydrolases utilise. There is major evidence showing that there are two main mechanisms existing for the enzymatic hydrolysis of the glycosidic bond; the retention and inversion of the anomeric configuration. It should be noted that whether an enzyme is endo- or exo- acting, there is no correlation to this and the mechanism used by the enzyme.

1.7 Mechanisms

Knowledge of enzymatic stereochemistry variation has given rise to the establishment of two main mechanisms used by glycoside hydrolases in which the changes in anomeric configuration are the basis of inversion and retaining mechanisms. The processes of such mechanisms were initially proposed by Koshland (1953). Both mechanisms hydrolyse the glycosidic bond via an acid catalyst involving oxocarbenium-ion-like transition states and a pair of amino acid catalytic residues, typically either glutamate and aspartate, acting as a proton donor (acid) and a nucleophile (base) (Davies *et al*., 1995) but differ in the resulting stereochemical configuration of the anomeric carbon. However, contrary to hydrolytic cleavage of glycosidic bonds, non-hydrolytic mechanisms have been identified that are also utilised in nature (see section 1.8.3).

1.8.1 Retaining mechanism

Retaining glycoside hydrolases release sugar products with retention of the configuration at the newly exposed anomeric centre. The mechanism for a general retaining glycoside hydrolase proceeds as a double displacement reaction involving the side chain of a catalytic residue which acts as a nucleophile and another that acts as an acid/base. These catalytic residues have been cited to be situated ~5.5Å apart (Davies *et al*., 1995). The first step is glycosylation in which a nucleophilic attack on the glycoside anomeric carbon is carried out while an acid/base residue acts as an acid to protonate the glycosidic oxygen passing through oxocarbenium-ion-like transition states for the formation of a covalently bonded glycosyl-enzyme intermediate. In the second step, deglycosylation of such intermediates occurs as a general base catalysed attack of water at the anomeric C_1 carbon, resulting in the retention of anomeric configuration (Figure 1.8) (White *et al*., 1997).

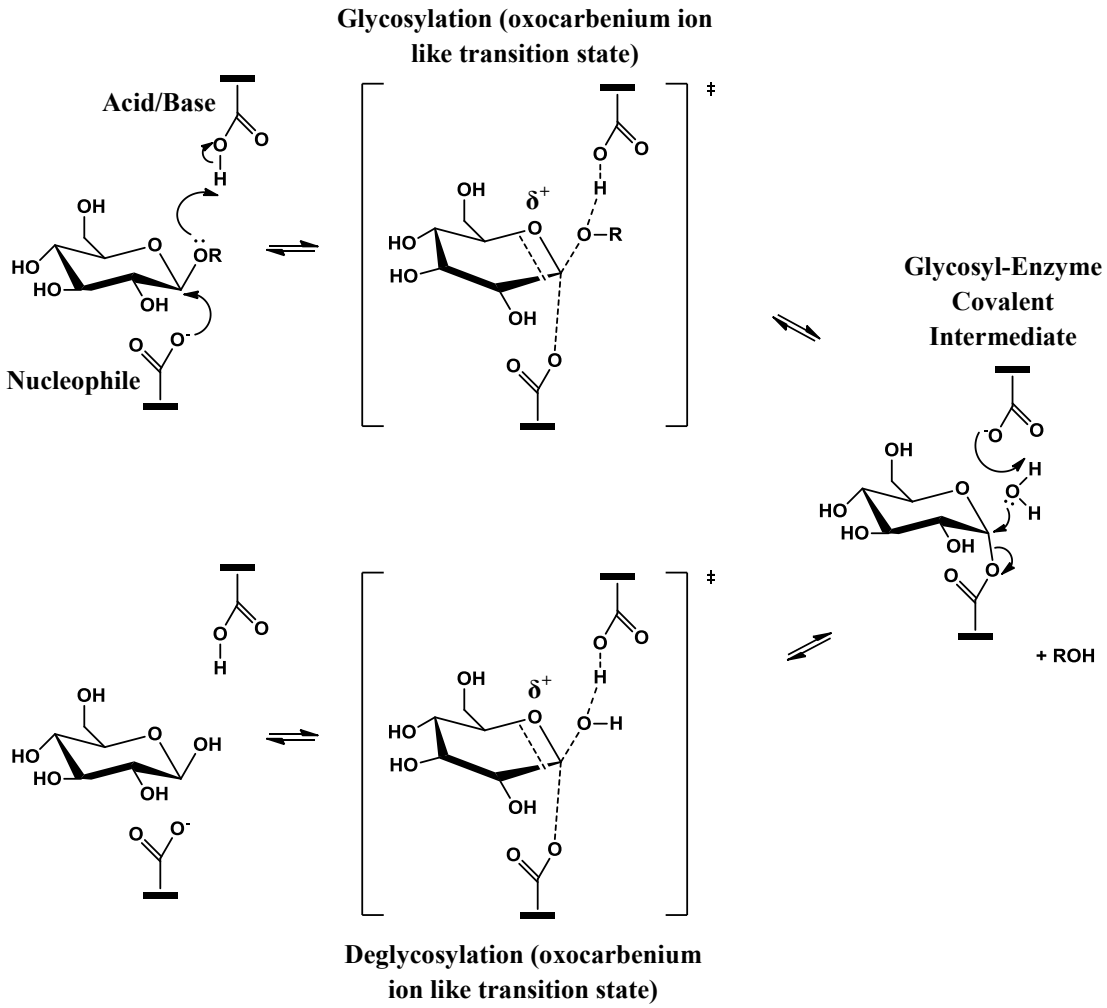


Figure 1.8 Schematic view of the β-Retaining Mechanism.

For the formation of a covalently bonded glycosyl-enzyme intermediate, a nucleophilic carboxylate within the glycoside hydrolase active site attacks the substrate at the anomeric C_1 carbon while simultaneously in a general acid reaction, the catalytic acid donates a proton to the glycosidic oxygen of the leaving group. Such intermediates have been recently demonstrated within crystal structures by trapping the enzyme-sugar complexes using

fluorosugars with good leaving groups to identify catalytic nucleophiles in retaining enzymes (Withers, 2001).

The structure of the cellobiohydrolase CBHI from *T. reesei*, a typical GH7 retaining glycoside hydrolase, naturally showed that the structure of the active site proved to contain acid residues of catalytic importance including Glu212, Glu217 and Asp214. The glutamate residues Glu212 and Glu217 were found to be suitably positioned to act as the acid/base catalyst in a double displacement nucleophilic substitution at the anomeric C1 carbon. The aspartate residue was shown to maintain the correct position and protonation state of Glu212 (Divne *et al*., 1998).

Early crystallography analysis of retaining enzymes showed that at the -1 subsite the saccharide ring is deformed from the usual $^4C_1$ 'chair' conformation to a 'twisted boat' conformation or a 'half chair' conformation during interactions with the transition state (Sinnott, 1990). The 'half chair' conformation is depicted in Figure 1.8. Such distortion of -1 subsite facilitates cleavage reactions by positioning the glycosidic oxygen closer to the glutamate residue involved in catalysis and the carbonyl oxygen near to the anomeric C_1 carbon, thus being the optimal positions for proton donation and nucleophilic attack.

Ring distortion has thought to play a crucial role in the retaining catalytic mechanism and distorted sugars have been reported for a few structures including that of the endoglucanse EGI from *Fusarium oxysporum* in which distortions resembled a *B*-boat conformation (Sulzenbacher *et al*., 1996). Such distortion within the -1 subsite has been described to be a result of interactions between the enzyme and sugar residue within the +1 subsite (Zechel *et al*., 2000). The boat and half chair conformation allow the atoms C5, O5, C1 and C2 of the sugar to achieve planarity which is required for the oxocarbenium-ion-like transition states. It should be noted that the two steps of retention occur via transition states with substantial oxocarbenium ion character. Data for the existence of oxocarbenium-ion-like transition states arises from secondary kinetic isotope effects and inhibition studies with transition

state analogues (Gloster *et al*., 2010). The partially planar arrangement of the distorted glycosyl ring mediates the oxocarbenium-ion-like positive charge due to the double bond character across the O5-Cl imposed by the sugar residue (Sabini *et al*., 1999). The $^4C_1$ chair conformation in which sugars usually reside is therefore unfavourable as it does not fit the stereochemical obligations for glycosyl transfer.

A typical GH11 β-1,4-xylanase from *Bacillus circulans* showed to heavily distort the xylose residue, bound to Glu78, within the -1 subsite from the conformational 'chair' to adopt the B boat distortion while the xylose residue found in the -2 subsite maintained the chair conformation (Sidhu *et al*., 1999). However, observations have been made that indicate some enzymes can tolerate or even favour undistorted sugars in the -1 subsite. This is demonstrated with a mutant of the cellobiohydrolase, CBHI from *T.reesei* in which there was no evidence of ring distortion (Divne *et al*., 1998). In the -1 subsite where Glu212 was expected to form the covalent enzyme-sugar intermediate, there was no indication of distortion of the glycosyl ring from the $^4C_1$ 'chair' conformation. This mode of binding has been described to 'by-pass' the catalytic machinery and bind in a non-productive manner (Fort *et al*., 2001). There is also evidence showing that furanose substrates such as arabinose polysaccharides, bind with relatively low distortion as energy barriers are low in comparison to pyranose rings such as xylose (Hövel *et al*., 2003). This may explain the difference in specificities for different substrates amongst the enzymes for example, the occasional low specificity in which α-L-arabinofuranosidases maintain for xylopyranosidic substrates.

Another form of retaining mechanism, termed substrate-assisted catalysis is used by various retaining glycoside hydrolases that hydrolyse substrates containing a N-acetyl group at the C-2 position (Jitonnom *et al*., 2011). These enzymes do not require a catalytic nucleophile but rather use the substrate itself as an intramolecular nucleophile.

1.8.2 Inverting mechanism

Inverting glycoside hydrolases cleave the glycosidic bond via a single displacement mechanism inducing a net inversion of the configuration at the anomeric C1 carbon involving an oxocarbenium-ion-like transition state (Figure 1.9). Again this process involves the action of two catalytic groups, mainly aspartate and glutamate carboxyls, that act as a general acid (proton donor) and general base (nucleophile) and these are typically situated ~10Å apart. However, within this single-step mechanism, a water molecule activated by the catalytic base attack as the anomeric centre hydrolyses the glycosidic residue through protonation with the aid of a general acid. This results in an inversion of the anomeric configuration (Honda *et al*., 2008).

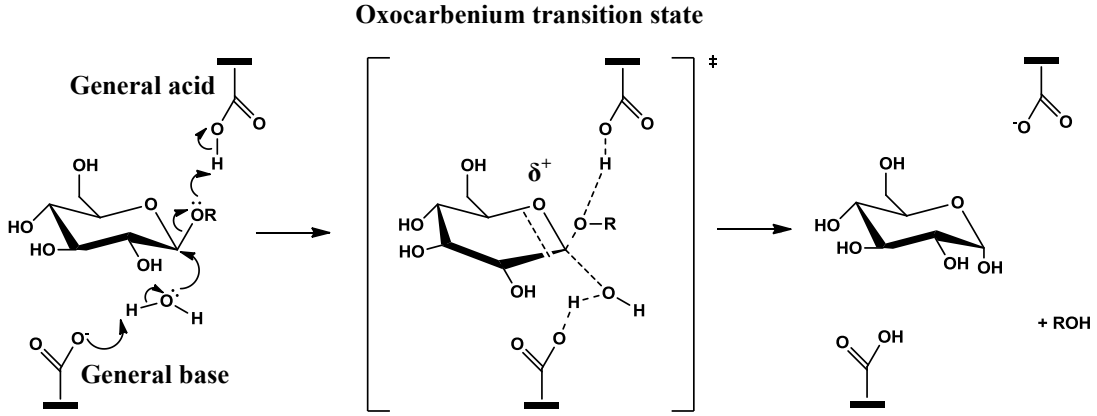


Figure 1.9 Schematic view of the β-Inverting Mechanism.

The structure of an inverting family GH43 arabinanase, AbnB, from *Geobacillus stearothermophilus* T-6 has been reported to use three catalytic residues unlike most inverting glycoside hydrolases which utilise only two acidic catalytic residues. This requirement of three conserved catalytic residues has been reported for several members of

GH43 as well as members of families GH32, GH62 and GH68 (Alhassid *et al*., 2009; Pons *et al*., 2004). The residues in AbnB have been resolved as Asp27, as the general base, Glu201, as the general acid and Asp147, as the pKa modulator adjusting acid dissociation. The position of the Asp27 residue allows it to activate the water molecule for a single displacement attack on the anomeric carbon resulting in the inversion of the anomeric configuration. The Glu201 residue is in the correct position in accordance to the glycosidic oxygen allowing it to protonate the leaving group while Asp147 enables it to raise its pK_a value and orientating the proton donor for the reaction to occur at physiological pH values (Alhassid *et al*., 2009).

The inverting mechanism has also been highlighted with CelS, a cellobiohydrolase from *Clostridium thermocellum*, which was the first member of the GH family 48 (Guimarães *et al.,* 2002). On the basis of crystal structure, Glu87 was assigned as the proton donor of CelS in which one of its carboxylate oxygen atoms was seen to form a hydrogen bond with glycosidic oxygen of the sugar ring. The Asp255 of CelS was proposed to be the candidate residue acting to stabilise the boat conformation of the sugar ring at position -1. A notable feature of this mechanism however was the absence of the base catalyst in which the role was originally assumed to be Asp255. The identity of the catalytic base is still not conclusively recognised in GH family 48 (Ficko-Blea *et al*., 2011).

1.8.3 Non-hydrolytic glycosidic bond cleavage

Even though hydrolysis is the most common enzymatic mechanism for the cleavage of glycosidic bonds, reactions involving non-hydrolytic cleavage do occur. Pectate lyases are such enzymes in which utilise a general acid/ base β-elimination mechanism (Figure 1.10). The β-elimination reaction is believed to involve three chemical processes. Firstly, the carboxyl group of the substrate, i.e. the group adjacent to the scissile glycosidic bond, is neutralised. This is followed by the abstraction of the C-5 proton in which an enolate anion

intermediate is created. Lastly, the proton is transferred to the glycosidic oxygen forming a double bond between C-4 and C-5 resulting in the cleavage of the C-4-O-1 glycosidic bond (Scavetta *et al*., 1999; Garron *et al*., 2010).

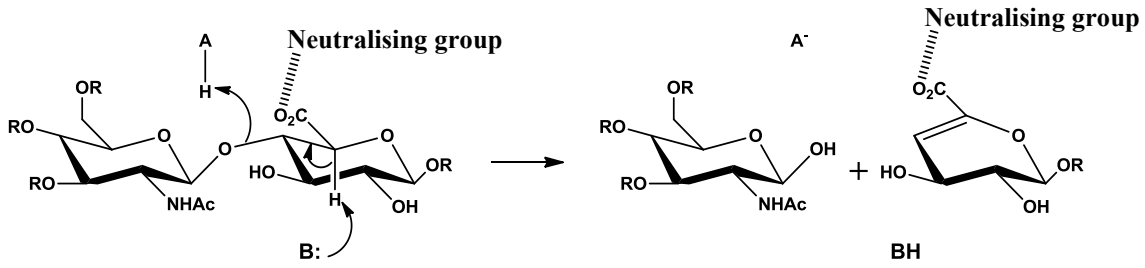


Figure 1.10 Schematic view of the β-Elimination Reaction.

1.8 Cellulases

Cellulases are a large group of glycoside hydrolases and are the enzymes responsible for cellulosic biomass degradation into glucose. Cellulases are categorised into three main groups; endoglucanases, cellobiohydrolases and β-glucosidases. Within various microbes such as *Bacillus subtilis* and *Clostridium thermocellum*, collectively this complex of cellulolytic enzymes is referred to as the cellulosome which is ultimately arranged around and bound together by non-catalytic polypeptides called scaffoldin (Shoham *et al*., 1999). These cellulases are involved in both the endo- and exo- cleavage of glycosidic bonds suggesting that there is a degree of synergism between enzymes in nature to fully hydrolyse the complex cellulose structure.

Endoglucanases are endo-acting cellulases hydrolysing cellulose at multiple sites through a series of random attacks. It has been suggested that endoglucanases mainly act on the amorphous regions of cellulose producing cellulo-oligosaccharides (Quay *et al*., 2011) along with glucans and some cellulose derivatives such as carboxymethyl cellulose (CMC).

Cellobiohydrolases are exo-cellulases that predominantly cleave cellobiose from reducing and non-reducing ends of cellulose chains and cellulo-oligosaccharides. Cellobiohydrolases have been shown to slowly degrade crystalline cellulose chains but activity on modified structures such as CMC is blocked due to methylated residues occurring on the substrate. The active site topology differs between the endoglucanases and cellobiohydrolases. While endoglucanases have the typical endo-acting enzyme open cleft site, cellobiohydrolases tend to have a tunnel shaped active site suggesting that the cellulose structures can be ‘threaded’ into the active site and thus resulting in the processive action of the cellobiohydrolases in the cleaving of cellobiose (Zou *et al*., 1999).

The cellobiohydrolase Cel6A was the first cellulase in which a 3D structure was determined that initially clarified the active site was located within a tunnel structure formed from

pairing loops (Rouvinen *et al*., 1990). The topology of the tunnel assumes that the cellulose chain is 'threaded' into the active site and such a reduced space prevents the substrate molecule from reorientating.

Comparing this with an endoglucanase, the open site cleft of this enzyme, allows for more accessible points along the cellulose chain upon which these enzymes will act. Recent studies have identified the absence of a surface tunnel loop within the active site of endoglucanases including that of the endoglucanase Cel6B from *Humicola isolens* (Davies *et al*., 1999). Furthermore, the cellobiohydrolase Cel7A from *T. reesei* underwent genetic engineering to produce mutants varying in the structure of the tunnel loop of the active site. Deletion of the loop caused the mutant enzyme to have a reduced processive activity on crystalline cellulose and increased activity on amorphous cellulose (von Ossowski *et al*., 2003) which is characteristic of an endoglucanase.

β-Glucosidases cleave the 1,4-glycosidic bond of cellobiose and cellooligosaccharides releasing glucose monomer units usually from the non-reducing terminal, thus completing hydrolysis of cellulose. β-Glucosidases are regarded to be part of the cellulase system within various microorganisms such as *Sporotrichum thermophile* in which β-glucosidases stimulate the rate and extent of cellulose degradation by relieving cellobiose-induced inhibition and increasing efficient hydrolysis by the endoglucanses (Bhat *et al*., 1993). β-Glucosidases from both fungal and bacterial origin have been classified into several CAZy families in which most belong to either GH1 or GH3. Members of both families catalyse hydrolysis via the retaining mechanism, however GH1 and GH3 β-glucosidases are structurally diverse with different catalytic folds and dissimilar active site residues. Family GH1 β-glucosidases have a $(\beta/\alpha)_8$ barrel structure while GH3 β-glucosidases accommodate a two domain structure consisting of a $(\beta/\alpha)_8$ barrel followed by an α/β sandwich in which the catalytic site is situate between the domains (Cairns *et al*., 2010).

The cellulosome was first described in the anaerobic thermophilic bacterium, *Clostridium thermocellum* as an extracellular multi-enzyme complex (Lamed *et al*., 1983) which contains the action of synergistic components that catalyse the efficient degradation of cellulose. Furthermore, the numerous kinds of cellulases and their related sub-units are assembled into non-enzymatic scaffoldin. The identification of the cellulosome by Lamed *et al*., (1983) established a particularly well characterised and abundant cellobiohydrolase known as CelS. CelS is a GH48 enzyme using the inversion mechanism to liberate cellobiose and showing typical traits of an exo-cellulase (Kruus *et al*., 1995). It is understandable that this cellobiohydrolase is the most abundant cellulase within this cellulosome as research has identified that *Clostrium thermocellum* thrives on cellobiose as its carbon source for energy. This also applies to other Clostridia genus producing similar cellulolytic systems (Dror *et al*., 2003) with the regulation of CelS being broadly studied. The crystal structure of CelS has since been elucidated (Guimarães *et al*., 2002) showing the tunnel shaped binding region which, as described previously, is representative of cellobiohydrolases. Such a structure is comparable to that of another characterised cellobiohydrolase from *Trichoderma reesei*, Cel6A, formerly known as CBH II.

The filamentous fungi *Trichoderma reesei* has been studied extensively due to its efficiency in breaking down cellulose substrates. It has been suggested that *T. reesei* utilises at least five endoglucanases, two cellobiohydrolases and several β-glucosidases to successfully hydrolyse cellulose into glucose (Ng *et al.,* 2011). Furthermore this is evidence of endo- and exo- cellulases working synergistically within a microorganism to utilise cellulose efficiently. While the endoglucanases (EG1- Eg5) non-specifically hydrolyse the internal glycosidic bonds within the cellulose polymer, the cellobiohydrolases CBHI and CBHII attack the reducing and non-reducing ends respectively releasing cellobiose (Kleman-leyer *et al*., 1996). This is then followed by the complete hydrolysis of cellulose by a β-glucosidase by cleaving glucose from the cellobiose.

Research into the cellulosome has demonstrated that such a complex has great potential in the application of the attributed enzymes within biotechnical processes. Microbial cellulases have thus been successfully utilised in processes within a range of industries such as the food, agricultural, textile and namely cellulosic biomass conversion (Kuhad *et al*., 2011).

1.9 Hemicellulases

Hemicellulases are the enzymes responsible for the breakdown of hemicellulose substrates and due to the structural complexity of hemicellulose, there are numerous challenges to the hemicellulolytic microorganism. As xylan is the most abundant hemicellulose, xylanases and xylosidases are critical in the breakdown of this polymeric substrate. However, additional enzymes such as α-L-arabinofuranosidases, endo-1,5-α-L-arabinanases, or β-mannanases are required depending upon the structural composition of the hemicellulose polymer. Furthermore, acetylxylan esterases have been commonly found in hemicellulolytic systems of microorganisms, hydrolysing esterase linkages (Dashtban *et al*., 2009). Hemicellulases belong to at least 20 GH families (1, 2, 3, 4, 5, 8, 10, 11, 26, 27, 36, 39, 43, 51, 52, 53, 54, 57, 62 and 67) and 8 out of 16 CE families (1, 2, 3, 4, 5, 6, 7 and 12).

1.9.1 Arabinan degrading enzymes

The two enzymes involved in arabinan hydrolysis are endo-1,5-α-L-arabinanase and α-L-arabinofuranosidase and these can work synergistically to fully hydrolyse an arabinose containing substrate. Arabinanases are specific for of α-1,5-L-arabinofuranosidic linkages of an arabinan backbone acting in a random endo-action pattern. α-L-Arabinofuranosidases are exo- acting enzymes that are known to catalyse the hydrolysis of α-1,5 linkages of the arabinan polymer or arabinosyl side chains of xylan residues. These enzymes can cleave α-1,2 and α-1,3-L-arabinofuranosidic linkages which occur when main chain arabinan and xylan residues (arabinoxylans) are substituted with α-L-arabinose units at C-2 and/ or C-3 positions . The enzymes that attack arabinofuranose linked glycosidic bonds are usually located in GH families 43, 51, 54 and 62. GH43 is receiving significant attention due to the expansion of the family as it also contains many xylanolytic enzymes. It has been shown that many genome sequences contain genes encoding a high number of GH43 enzymes including

Cellvibrio japonicus and *Opitutus terrae*, with the latter being the main target of this current study.

Cellvibrio japonicus encodes a well characterised endo-1,5-α-L-arabinanase, Arb43A, belonging to GH43 which is able to hydrolyse linear arabinan into mostly arabinotriose. This implies a predominant exo fashion as well as an endo-acting pattern by the enzyme and a possible processive action taking place after an initial internal cleavage (McKie *et al*., 1997). The 3D structure of Arb43A from *C. japonicus* has since been solved indicating a new topology of binding site revealing a β-propeller fold in which a blockage occurs by a large loop between the third and fourth propeller blades (Figure 1.11) (Nurizzo, *et al*., 2002) .This loop is absent in endo-acting arabinanases further adding to exo-acting properties of the Arb43A enzyme, however substrate binding data observed for this enzyme was similar to a characterised endo-acting arabinase (AbnA) from *Aspergillus niger*.

Figure 1.11 Structure of Arb43A from *Cellvibrio japonicus*. The representation of the native crystal structure of CjArb43A revealing a five-bladed β-propeller fold. Colours are ramped N-terminus blue through to C-terminus re. The central ion is shown as a sphere. Reproduced from RCSB Protein Data Bank.

The arabinan degrading system of *Aspergillus niger* has been extensively studied and to date there have been two distinct α-L-arabinofuranosidases characterised (AbfA and AbfB) alongside and an endo-1,5-α-L-arabinanse (AbnA) (Veen *et al*., 1991). Also belonging to GH43 family, AbnA was initially characterised on sugar beet arabinan with its preferred substrate being linear 1,5 linked α-L-arabinan. Synergism between these three arabinan degrading enzymes has been observed within this fungal organism, as their synthesis was induced upon growth on arabinose containing polysaccharide substrates (Flipphi *et al*., 1994).

The biochemically characterised α-L-arabinofuranosidases, AbfA and AbfB, from *A. niger* belong to two different GH families, families GH 51 and GH 54, respectively. Even though both families utilise the inversion mechanism for catalysis, the difference in families most probably reflects the difference in substrate specificity between the two enzymes. While AbfA hydrolyses linear 1,5-α-linked-L-arabinofuranosyl oligosaccharides, AbfB acts on α-1,2-, α-1,3- and α-1,5- linkages in both oligosaccharides and polysaccharides (Gielkens *et al*., 1999). Alternatively and most recently, a GH43 family α-L-arabinofuranosidase from *C. japonicus* (*Cj*Abf43A) has been discovered which solely acts upon 1,2-α- linked arabinose units of doubly substituted arabinan polysaccharides (Cartmell *et al*., 2011). Therefore it can be suggested that there is likely to be an enzyme within this microorganism that solely acts upon α-1,3- linkages for the total hydrolysis of the suspected sugar residue. Similarly GH51 and GH54 α-L-arabinofuranosidases have been shown to attack α-1,2- and α-1,3- linkages while GH62 -L-arabinofuranosidases show high specificity for arabinoxylans (Taylor *et al*., 2005). Therefore, it is apparent that the several GH families display widely varying specificities for the arabinan degrading hemicellulases thus making generalisations and predictions about such enzymes quite difficult. However such diverse specificity does pose a number of industrial applications related to hemicellulose degradation.

1.9.2 Xylan degrading enzymes

Similarly to arabinan, the total hydrolysis of xylan requires different hydrolytic enzymes to work synergistically and this is observed in xylanolytic systems of both fungal and bacterial origin. As the xylan polymer may be substituted with arabinose residues and acid groups, debranching enzymes make xylan more accessible to the main depolymerising xylan degrading enzymes. These accessory enzymes include α-L-arabinofuranosidases, α-glucuronidases and acetylxylan esterases. This section 1.9.2 will however focus on the enzymes degrading the main xylan chain, endo-β-1,4-xylanases and β-1,4-xylosidases. Endo-β-1,4-xylanases internally cleave the β-1,4- bonds of the xylan backbone in a random action producing various sized oligosaccharides while β-1,4-xylosidases catalyse the hydrolysis of the oligosaccharides into xylose.

1.9.2.1 Xylanases

Endo-β-1,4-xylanases are produced by a variety of microorganisms including filamentous fungi as well as a plethora of bacterial microbes and as a consequence of the complexity of xylan, an abundance of xylanase diversities with varying substrate specificities, catalytic folds and primary sequences have evolved. Endo-β-1,4-xylanases were initially only found in GH families 10 and 11 based on their hydrophobic cluster and amino acid sequence, until the recent inclusions of xylanases in families 5, 7, 8 and 43, albeit studied to a much lesser extent.

1.9.2.1.1 Xylanase multiplicity

The complexity of xylans has resulted in an abundant diversity of xylanases that differ in physiochemical properties, mode of action, active site architecture and substrate specificity. It has become apparent that many microorganisms produce multiple xylanases to tackle the

heterogeneity of xylans and increasing the extent of xylan hydrolysis. *Aspergillus niger* has been reported to secrete at least 15 extracellular xylanases (Collins *et al*., 2005) belonging to family GH11 and *Cellvibrio japonicus* has been reported to produce five different xylanase. Xylanase multiplicity has been suggested to arise from post-translational processing of the same gene as well as gene products of several xylanase-encoding genes. Post-translational modifications resulting in xylanase heterogeneity may occur in the process of either glycosylation or proteolysis as demonstrated by Fernandez-Abalos *et al*., 2003. At least five serine proteases isolated from *Streptomyces* spp. were reported to have the ability to process the xylanase Xys1L, originally produced by *Streptomyces halstedii* JM8 (Fernandez-Abalos *et al*., 2003). Alternatively, *Aspergillus nidulans* is genetically well characterised and it is reported that its major xylanases X_{22} and X_{24} are encoded by the genes *xlnA* and *xlnB*, respectively (Pérez-Gonzalez *et al*., 1996) thus providing evidence of xylanases encoded by as distinct gene products.

1.9.2.1.2 GH10 and GH11 xylanases and substrate specificity

GH10 xylanases generally have higher molecular mass and lower isoelectric point compared to GH11 xylanases. On the basis of their crystal structures of both groups, GH10 xylanases have similar catalytic domain structures consisting of $(\beta/\alpha)_8$ barrels while GH11 xylanases assume a tightly packed β-jelly-roll architecture in which two large β-sheets and one α-helix form a long, deep cleft (Figure 1.12) (Biely *et al*., 1997) .The shapes of these active sites reflect their difference in substrate specificity as well as different specificity within families. The shallow groove of the GH10 xylanases is responsible for their specificity towards substituted xylose units and enabling to cleave decorated regions of xylans while GH11 xylanases show higher affinity towards unsubstituted xylose units due their cleft shape active site (Biely *et al*., 1997).

Endo-xylanase of family 10 **Endo-xylanase of family 11**

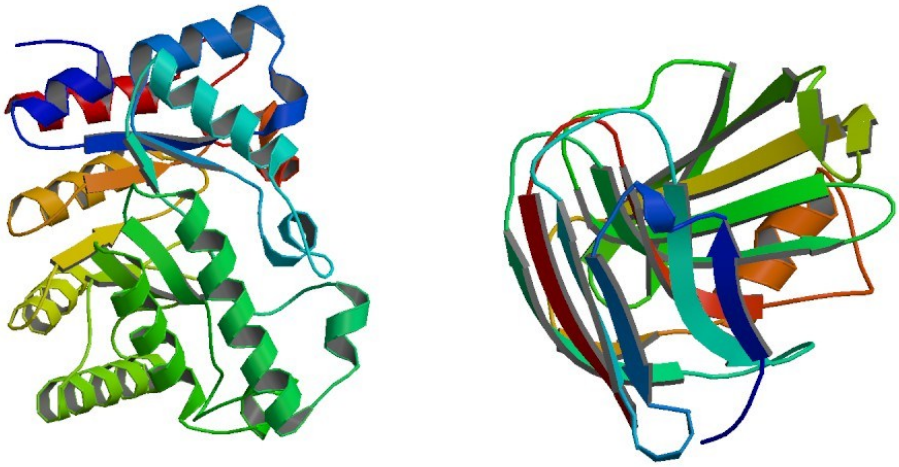


Figure 1.12 Structure of main structural fold of the catalytic domains of family 10 and 11 endo-xylanases. Ribbon representation of typical $(\beta/\alpha)_8$ and β-jelly-roll of GH10 and GH endo-xylanases respectively. Reproduced from RCSB Protein Data Bank.

Kinetic studies have revealed that the active sites of GH10 xylanases have at least four subsites, as they are almost identical from the -2 to +2 subsites, although the length of the substrate clefts of GH10 xylanases are variable with the possibility of more subsites in some cases (Charnock *et al*., 1998; Biely *et al*., 1997). However the majority of GH11 xylanases contain at least five subsites from -2 to +3 (Paës *et al*., 2011). As a result, shorter oligosaccharide products are released by GH10 xylanases and such active site structure may account for the low substrate specificity observed by GH10 xylanases in comparison to GH11 xylanases.

Previous studies have examined the mode of action between the GH10 and GH11 xylanases. Kolenová *et al*., (2006) observed the action of such enzymes on acidic xylooligosaccharides from beechwood 4-*O*-methyl-D-glucuronoxylan showing that GH10 xylanases produced shorter oligosaccharides and acquired a higher affinity to xylose units substituted with 4-

methyl-D-glucuronic acid compared to GH11 xylanases. This study also indicated that GH11 xylanases from numerous organisms such as *Streptomyces lividans* and *Sporotrichum thermophile* released aldopentaouronic acid from the glucuronoxylan polymer. This, being the smallest product, could not be hydrolysed further and thus reflecting the low rate of hydrolysis of short oligosaccharides by the GH11 xylanases, as such action was impeded by the decoration of xylose units with 4-*O*-methyl-D-glucuronic acid.

The *Cellvibrio japonicus* GH10 xylanase *Cj*Xyn10C, however is an example in which biochemical properties differ slightly within families, as the activity of this xylanase is 10-100 fold less than other GH10 on xylooligosaccharides (Pell *et al*., 2004). Also there are slight variations of substrate specificities demonstrated within families. For example, as *Cj*Xyn10C predominantly releases xylose and xylotriose from xylotetraose, a GH10 xylanase from *Cellulomonas fimi*, formerly known as *Cex*, generates xylobiose from the tetrasaccharide chain (Charnock *et al*., 1998). Therefore making the -3 subsite of *Cj*Xyn10C more significant than the +2 subsite due to the release of xylose and xylotriose hydrolysis products. Substrate specificity within families has been observed within this current research and is discussed further in Chapter 4. It is important to note that activity of *Cj*Xyn10C against variable decorated xylans is comparable to other xylanases within the GH10 family, as it acts as a typical endo-acting enzyme producing a range of xylooligosaccharides.

The existence of exo-xylanases is now well recognised and such enzymes have been discovered releasing single products such as xylobiose or xylotriose by an exo-type mechanism (Kubata *et al*., 1995). The GH10 xylanase, XynX secreted by *Aeromonas caviae* ME-1 was reported to exclusively produce xylobiose and xylotetraose from birchwood xylan with no other oligosaccharides detected (Usui *et al*., 1999). It was suggested that this xylanase only recognised the terminus of xylan substrates and was therefore acting as an exo-xylanase. Earlier reports claim that such exo-xylanases are difficult to distinguish from β-xylosidases, as both are capable of degrading xylooligosaccharides. However, exo-

xylanases are reported to attack xylan polymers unlike β-xylosidases. A proposed action of the exo-xylanases can be suggested through their potential synergy with endo-xylanases. For instance, endo-xylanases randomly attack the main xylan polymer generating xylooligosaccharides and free chain ends which exo-xylanases could act upon thus creating small oligosaccharides for the degradation into xylose by β-xylosidases. This is quite similar to cellulase system as described in section 1.8.

1.9.2.2 Xylosidases

For the degradation of xylooligosaccharides, β-xylosidases are exo-acting enzymes required for the cleavage of short chain xylooligosaccharides from the non-reducing end to release xylose. β-Xylosidases are produced by both bacteria and fungi and have been grouped into families GH3, GH39, GH43, GH52, GH54, GH116 and GH120 within the CAZy database. Noticeably, these families hydrolyse via the retaining mechanism, apart from family GH43 which hydrolyse by the inverting mechanism. The structure of a retaining family, GH 39 β-xylosidase, XynB1, from *Geobacillus stearothermophilus* was solved alongside the associated inactive mutant XynB1-E160A in which the catalytic acid-base residue, E160A, was inactivated by point mutation (Czjzek *et al*., 2005). Crystal structures were solved in complex with the chromogenic substrate molecule 2,5-dinitrophenyl-β-D-xyloside. Both structures revealed the xylosyl-enzyme covalent intermediate and the enzyme-uncleaved substrate interactions within the aglycon pocket (+1), therefore defining the xylosidases binding site as they cleave from the non-reducing end of the substrate.

Also from *Geobacillus stearothermophilus*, the structure of a GH43 β-xylosidase, XynB3, has been solved showing the active site to be mainly constructed from a five bladed β-propeller fold which is common to all GH43 members (Brüx *et al*., 2006).The β-propeller units form a cleft that is blocked at one side by a loop. This forms a pocket topology of the

active site which is very suitable for the exo-action observed by XynB3, which releases single xylose units.

Dual functional enzymes with both β-xylosidase / α-L-arabinofuranosidase activities have recently been cited in family GH43 (Wagschal *et al*., 2009; Zhou *et al*., 2011). L-Arabinofuranose and D-xylopyranose are spatially similar in that the hydroxyl groups and glycosidic bonds can be overlaid therefore justifying the existence of a bifunctional β-xylosidase / α-L-arabinofuranosidase. It has also been reported that microbes, predominantly bacteria, which secrete both β-xylosidase and α-L-arabinofuranosidase enzymes are often encoded by a single gene (Mai *et al*., 2000).

1.9.3 Other hemicellulases

Although xylan is the most prevalent hemicellulolytic polymer, other enzymes have been isolated and characterised that show specificity towards other major hemicellulases, including mannans and galactans. β-Mannanases randomly cleave β-1,4-mannosyl linkages depolymerising mannan in an endo acting fashion producing oligosaccharides that can be hydrolysed by β-mannosidases to release mannose units. In addition to these, α-galactosidases are needed for complete degradation, as they are required for the complete degradation of polymeric galactomannans. α-Galactosidases catalyse the hydrolysis of α-1,6-linked α-galactose residues from galactomannan oligosaccharides.

1.10 Non-catalytic polysaccharide domains

PCW degrading enzymes often exhibit modular architecture containing non-catalytic domains in addition to the catalytic modules. The majority of these non-catalytic modules are carbohydrate binding modules (CBM) which usually function by anchoring the enzyme to the substrate promoting efficient contact between the respective active site of the enzyme

and the target substrate (Hervé *et al*., 2010). CBMs were originally designated cellulose binding modules, as they were initially discovered to bind cellulose (Gilkes *et al*., 1988). However more domains in carbohydrate active enzymes such as xylanases, have been found to bind to other carbohydrate substrates such as xylans, mannans and pectins. CBMs have been located in either the C- or N- terminal of the parental protein and there are many accounts in the literature where truncation of modular enzymes containing such CBMs reduced catalytic activity, for example, on insoluble polysaccharides without affecting activity on soluble substrates (Charnock *et al*., 2000).

Similar to glycoside hydrolases, CBMs are divided into families, based upon amino acid sequence similarity, of which there are currently 64 CBM families. Members have been characterised to bind cellulose, xylan, chitin, glucans, mannans and starch. CBMs within families and from different families can have varying binding specificities (Boraston *et al*., 2001), therefore function predictions cannot be made.

Generally it is accepted that CBMs have three major roles; (I) a proximity affect (II) a targeting function (III) and a disruptive function upon the substrate structure (Boraston *et al.,* 2004). Furthermore a functional classification of CBMs has been determined based on ligand binding specificities, Type A, Type B and Type C. Type A CBMs are referred to as surface binding CBMs as they bind specifically to the surfaces of insoluble polysaccharides such as crystalline cellulose. In contrast Type B CBMs recognise single polysaccharide chains and Type C CBMs are responsible for small chain binding.

It is thought that Type A CBMs presents flat or planar surfaces that complement the planar surfaces of crystalline cellulose and this has been visualised by transmission electron microscopy on *Valonia* cellulose crystals (Sugiyama *et al*., 2002). These CBMs are found in a wide distribution of enzymes involved in PCW degradation and show little or no affinity for soluble polysaccharides.

Type B CBMS bind soluble polysaccharides and have evolved binding topologies that are able to bind to single glycan chains in contrast to the interaction with crystalline structures in the case of Type A CBMs. Type B CBMs have extended binding sites described as clefts or grooves comprising several subsites (Boraston *et al*., 2004) and they bind to the substrates for their cognate catalytic module of the enzyme. For example, those CBMs linked to β-glucosidases interact with soluble cellulo-oligosaccharides.

Type C CBMs that bind small oligosaccharides have lectin-like properties with an affinity for mono-, di-, or tri- saccharides (Boraston *et al*., 2004) and therefore do not acquire the extended binding site which Type B CBMs require.

Many enzymes including cellulases and xylanases have been shown to contain CBMs potentiating the effect of cellulases and hemicellulases to degrade PCW polysaccharides. A familiar example in which this applies, is the cellulosome with *Clostrium thermocellum* which contains the non catalytic polypeptide, scaffoldin and this has been established to mediate the attachment of nine catalytic subunits to cellulose by an internal family 3 carbohydrate binding module, CBM3 (Nataf *et al*., 2010). The xylanase, Xyn10B, from *C. thermocellum* also contains two CBMs belonging to CBM family 22 in which biochemical properties were investigated showing that these modules had thermostabilising and proteinase resistance functions (Xie *et al*., 2001; Charnock *et al*., 2000).

1.11 Tansglycosylation

It is generally acknowledged that GHs have a transglycosylating capacity resulting in the synthesis of an assortment of biologically important compounds including oligosaccharides and glucoconjugates (Eneyskaya *et al*., 2003). Transglycosylation has proved to be a useful strategy in preparing branched oligosaccharides while preserving the original sugar structure and therefore has obvious advantages over traditional organic synthesis. Even though, as

briefly discussed before, glycosyltransferases catalyse the formation of glycosidic linkages, they are not entirely sufficient for facilitating practical oligosaccharide synthesis due to the difficulty in obtaining a substantial amount of enzyme and their donor substrates (Murata *et al*., 2002). Additionally, commercial donor substrates are expensive limiting their large scale applications (Eneyskaya *et al*., 2003).

GHs that exhibit transglycosylation activity are those that utilise the retaining mechanism for glycosidic hydrolysis. Transglycosylation proceeds via the enzyme-glycosyl intermediate, within the catalytic site, through the action of a carbohydrate acceptor instead of a water molecule that would normally release the hydrolysed sugar residue (Kawai *et al*., 2004). Unlike the mechanism of retention, the inverting process is irreversible once the linkage is cleaved with the addition of the water molecule (Honda *et al*., 2008), hence the transglycosylating ability of retaining GHs.

Transglycosylating activity has been reported for many GH enzymes including several members of family GH3 (Kawai *et al*., 2004; Crombie *et al.,* 1998). A β-D-glucosidase from the cotyledons of nasturtium seedlings was shown to catalyse significant transglycosylation accompanying the early stages of cello-oligosaccharide hydrolysis (Crombie *et al*., 1998). There was evidence from this study to suggest that transglycosylation was sensitive to oligosaccharide size and structure as hydrolysis could also occur without transglycosylation of certain glucan-oligosaccharides. Similarly, newly synthesised products produced by the transglycosylation reaction catalysed by an endo-β-1,4-xylanase, Xyn1, from *Paenibacillus* sp. strain W-61 has been recorded in which also xylooligosaccharide size was an important factor in transglycosylation activity (Watanabe *et al*., 2008). Additionally, this process has been observed as a side reaction when a high concentration of xylooligosaccharide substrates accumulates and transglycosylation utilises donor xylooligosaccharides to produce larger xylooligosaccharide chains (Kolenová *et al*., 2006).

1.12 Proteinaceous inhibitors

In addition to structural support, the cell wall polysaccharides are the first line of defence, providing a physical barrier against microbial pathogens and such resistance has been reviewed (Vorwerk *et al*., 2004). Additionally, plants can resist hydrolytic attack by releasing inhibitor proteins to counteract the activity of PCW degrading enzymes. It is now recognised that plants deploy various inhibitor proteins including polygalacturonase-inhibiting proteins, pectin methylesterase inhibitors and most recently there is evidence of xylanase inhibitors. Such xylanase inhibitors are widely spread in cereals including wheat, rye and rice in which the main hemicellulose polymer occurring in such structures is xylan. However, inhibitors can also play an important role in the regulating of endogenous enzymes produced by these cereals, as endogenous xylanases have demonstrated physiological function in seed germination and fruit ripening (Simpson *et al*., 2003).

Two structurally distinct classes of inhibitors have been reported to inhibit xylanases, namely xylanase inhibitor protein (XIP) type and *Triticum aestivum* xylanase inhibitor (TAXI) type (Juge *et al*., 2004) and these have been isolated from various cereals. XIP-1 was first purified from wheat flour *T. aestivum* (also the source of TAXI) in which the 3D structure showed that the inhibitor possessed a $(\beta/\alpha)_8$ barrel fold and a notable sequence similarity to GH18 members. Family GH18 contains mainly chitinases in which some have been shown to be naturally inactive and it is these proteins that have been identified as xylanase inhibitors (Vasconcelos *et al*., 2011). True plant chitinases defend the plant against fungal infection by attacking chitin, a structural component of fungal cell walls (Punja *et al*., 1993). Structural differences were found in regions of XIP-1 corresponding to the active site of chitinases which is most likely to explain the lack of activity against chitin, for example, the Gly81 residue situated in the -1 subsite of hevamine, an active chitinase, is replaced by Try80 residue in XIP-1. Such a variation prevents the substrate accessing the catalytic glutamic acid residue which ultimately required to protonate the glycosidic oxygen in

catalysis. Upon interaction with xylanases, it was found that XIP-1 was effective against families GH10 and GH11 fungal xylanases and being able to distinguish between fungal and bacterial enzymes, XIP-1 was ineffective against xylanases of bacterial origin (Tokunaga *et al*., 2007; Durand *et al*., 2005; Juge *et al*., 2004).

In contrast to XIP-type proteinaceous inhibitors, TAXI-type inhibitors have demonstrated inhibition exclusively against both fungal and bacterial GH11 xylanases, but have no activity against family GH10 xylanases (Weng, *et al* 2010; Rasmussen *et al.,* 2010). This may not be surprising that such an inhibitor can distinguish between the two different families since GH10 and GH11 are structurally diverse with a $(\beta/\alpha)_8$ barrel and a β-jelly fold, respectively. However, this poses further questions about the binding interaction between XIP-1 and xylanases from both families, whether or XIP-1 can simultaneously inhibit both types of xylanase or if there is competitive inhibition. Also there is no evidence for either of these proteinaceous inhibitors inhibiting relatively newly confirmed xylanases belonging to family GH8.

1.13 Application

Progress in carbohydrate-active enzyme research has lead to recent advances in their effective application in several aspects of industry including the food, paper and pulp divisions. However, the potential of such enzymes in the bioconversion of lignocellulosic material has been an important research focus for more than four decades. As briefly mentioned before, the biorefinery has become a key modern day concept integrating multiple system platforms and technologies for the effective conversion of biomass into ethanol. Technologies required for the system platforms include the pretreatment of biomass, enzymatic hydrolysis, fermentation, distillation and subsequent solid-liquid separation much like petroleum refineries (NREL; Carvalheiro *et al*., 2008). Enzymatic hydrolysis is therefore just one aspect of the biorefinery concept and the continued research and

development into desired activities/ specificities for biomass degradation paves the way for more efficient enzymes.

1.14 Aims and objectives of this study

Bioconversion of lignocellulosic biomass into ethanol is significantly hindered by the structural and physiochemical complexity of biomass. However, the pretreatment of lignocellulose biomass is now a well recognised step into making lignocellulosic substrates more accessible to facilitate efficient hydrolysis by both acidic and enzymatic systems. The utilisation of enzymes for the degradation of lignocellulose into its subsequent fermentable sugars is considered prospectively to be the most viable strategy over chemical conversion routes. This alternative biological route employs mild operating systems with low energy requirements, higher yields with minimal by-products and ultimately environmentally friendly processing (Zheng *et al*., 2009)**.**

It is evident from literature that efficient enzymatic saccharification of lignocellulose biomass continues to be a challenge, especially if the prerequisite is the commercialisation of such a bioconversion process. Therefore, with respect to this research the primary aims were to identify and investigate lignocellulose degrading enzymes that have the potential for the utilisation in biomass conversion reactions in emerging industrial processes. Also to delve into the construction of rational recombinant enzyme cocktail libraries targeted towards efficient hydrolysis of various lignocellulosic materials. Considering the immense scale of variations of lignocellulosic structures, numerous enzyme activities were addressed from various microorganisms, including *Opitutus terrae* and were able to provide a source of catalysts to screen against challenging lignocellulosic substrates. However, the data generated provided valuable insights into substrate specificities utilised by the bacterium *Opitutus terrae* and has given rise to a new identified GH8 xylanase.

2 Materials and Methods

All solutions and buffers were made up using ultrapure 18.2 MΩ / cm H_2O from a Millipore Direct-Q™ Water Purification System and stored at room temperature. For all main molecular biology protocols, *Molecular Cloning: a laboratory manual* by Sambrook and Russel were used as a general reference. Table 2.1 lists manufacturers (and the abbreviated form) that were used to source enzymes, chemical, reagents and biological kits that were used within this study.

Manufacturer	Abbreviation
Bio-Rad	BR
Fisher Bio- Reagents	FIS
Invitrogen	INV
Megazyme	MEG
Melford Laboratories Ltd	MEL
New England Biolabs	NEB
Novagen	NOV
NZYTech	NZYT
Oxoid	OX
Qiagen	QIA
Sigma	SIG
Server	SER

Table 2.1 Names and abbreviations of manufacturers.

2.1 Materials

2.1.1 Bioinformatics

Gene sequences encoding putative carbohydrate active enzymes were selected from the Carbohydrate Active Enzymes Database, CAZy (http://www.cazy.org/). Bioinformatics were carried out on each gene sequence and the tools utilised are presented in the Table 2.2.

Bioinformatics Tool	**URL**	**Application**
National Center for Biotechnology Information (NCBI)	http://www.ncbi.nlm.nih.gov/	Resource of genome information.
SignalP 3.0 Server	http://www.cbs.dtu.dk/services/SignalP/	Predicts the occurrence and location of signal peptide cleavage sites within amino acid sequences.
Webcutter 2.0	http://rna.lundberg.gu.se/cutter2/	Provides map of restriction sites and non-cutting restriction enzymes within a DNA sequence.

Table 2.2 Table providing information on the bioinformatics tools used within this research.

UniProt	http://www.uniprot.org/	Resource of protein sequences and functionality information.
Saccharomyces Genome Database (SGD)	http://www.yeastgenome.org/	Program used to design primers.
ClustalW	http://www.ebi.ac.uk/Tools/msa/clustalw2/	Multiple sequence alignment for proteins.
ExPASy Proteomics Server	http://expasy.org/tools/	Providing access to proteomic software tools, i.e. ProtParam tool.

Table 2.2 Table providing information on the bioinformatics tools used within this research (continued).

2.1.2 Molecular biology kits

The molecular biology kits utilised in this research are presented in Table 2.3. All kits were stored at room temperature however the manufacturer recommended that a few reaction components were to be kept at 4°C.

Kit	Manufacturer	Application
NZY Midiprep	NZYT	Large scale purification of expression vector.
NZY Miniprep	NZYT	Purification of recombinant plasmid DNA.
NZYTech Gelpure	NZYT	Agarose gel purification of DNA.
DNeasy Blood and Tissue	QIA	For extraction and purification of genomic

Table 2.3 Molecular biology kits used within this research.

2.1.3 Enzymes

All enzymes and their corresponding buffers were stored at -20°C. Table 2.4 lists enzymes along with their supplier and application.

Enzyme	**Supplier**	**Application**
KOD Hot Start DNA Polymerase	NOV	PCR
T4 DNA Ligase	NEB	Ligation
Hen Egg White Lysozyme	FIS	Bacterial cell wall lysis.
Restriction enzymes: *Nde*1, *Xho*1, *Bam*H1, *Eco*R1, *Hin*d111, *Xba*1	NEB	Restriction enzyme digest.

Table 2.4 Table showing enzymes, supplier and application.

2.1.4 Bacterial strains

All bacterial strains were stored at -80°C in 50 % (v / v) glycerol.

2.1.4.1 E. coli Strains used for cloning and expression

Table 2.5 illustrates the *E. coli* strains required for cloning and expression.

***E. coli* strain**	**Relevant Features/ Genotypes**	**Application**
E .coli TOP10 (INV)	F– *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*ara leu*) 7697 *gal*U *gal*K *rps*L (StrR) *end*A1 *nup*G	Cloning host.
E. coli BL21(DE3) (NOV)	F– *omp*T *hsd*SB(rB–, mB–) *gal dcm* (DE3)	Hyper expression of soluble protein.

Table 2.5 *E. coli* strains with their relevance and application.

2.1.4.2 Bacterial strains required for extraction of genomic DNA

Different bacteria required for the extraction of genomic DNA included:- *Streptomyces avermitilis* MA 4680, *Streptomyces coelicolor* A3, *Lactobacillus brevis* ATCC 367, *Streptococcus pyogenes* MGAS315, *Rhizobium etli* CFN 42, *Bacteroides fragilis* NCTC 9343, *Bradyrhizobium japonicum* USDA 110, *Rhodopseudomonas palustris* CGA009, *Staphylococcus saprophyticus* ATCC 15305, *Bacillus subtilis* 168, *Leuconostoc mesenteroides*, *Bradyrhizobium sp.* BTAil, *Escherichia coli* K12 and *Opitutus terrae* PB90-1.

2.1.5 Antibiotics

Antibiotics were dissolved in sterile 18.2 MΩ / cm H_2O and filtered using a Minisart® single use filter. Antibiotics were diluted with media to achieve the required concentration. For the addition of antibiotics to molten agar, the media was left to cool to ~50°C. Table 2.6 shows the antibiotics used and their final concentrations.

Antibiotic	**Stock Concentration**	**Dilution of Stock using Media**	**Final Concentration in Media**
Kanamycin (SIG)	10 mg / mL	1:200	50 µg / mL
Ampilcillin (SIG)	10 mg / mL	1:100	100 µg / mL

Table 2.6 Stock and final concentrations of antibiotics.

2.1.6 Media

All liquid and solid media were prepared using distilled H_2O and sterilised by autoclaving at 121°C for 20 min. Liquid media was stored at room temperature and solid media at 4°C. The pH was adjusted using NaOH or HCl.

2.1.6.1 Liquid media

2.1.6.1.1 LB (Luria-Bertani) Broth (pH 7.0)

Per / L	
Tryptone (OX)	10.0 g
Yeast Extract (OX)	5.0 g
NaCl (SIG)	10.0 g

2.1.6.1.2 Auto-induction media

Auto- induction media was used to express proteins in *E. coli* BL21 cells without the requirement of IPTG to induce expression.

Per / L	
ZY	947.0 mL
1M $MgS0_4$	2.0 mL
1000 x Metals	1.0 mL
50 x 5052	20.0 mL

50 x M	20.0 mL
Kanamycin	10.0 mL

Stock solutions:

ZY Broth

Per / L

Tryptone (OX)	10.0 g
Yeast Extract (OX)	5.0 g

1000 x Metals (SIG)

Per 100 mL

H_2O	36.0 mL
0.1M $FeCl_3 \cdot 6H2O$ dissolved in 0.12M HCl	50.0 mL
1M $CaCl_2$ anhydrous	2.0 mL
1M $MnCl_2 \cdot 4H_2O$	1.0 mL
1M $ZnSO_4 \cdot 7H_2O$	1.0 mL
0.2M $COCl_2 \cdot 6H_2O$	1.0 mL
0.2M $CuCl_2 \cdot 2H_2O$	2.0 mL
0.2M $NiCl_2 \cdot 6H_2O$	1.0 mL
0.1M $Na_2MoO_4 \cdot 5H_2O$	2.0 mL

0.1M $Na_2SeO_3 \cdot 5H_2O$	2.0 mL
0.1M H_3BO_3 anhydrous	2.0 mL

50 x 5052

Per 100 mL

Glycerol (SIG)	25.0 g
Glucose (SIG)	2.5 g
α-Lactose (SIG)	10.0 g

50 x M

Per 100 mL

Na_2SO_4 anhydrous (SIG)	3.6 g
NH_4Cl anhydrous (SIG)	13.4 g
KH_2PO_4 anhydrous (SIG)	17.0 g
Na_2HPO_4 anhydrous (SIG)	17.7 g

2.1.6.1.3 YMG Media for growth of *Streptomyces* species (pH 7.2)

Per / L	
Yeast Extract (OX)	4 g
Malt Extract	10 g
Glucose	4 g

2.1.6.1.4 MRS (de Man, Rogosa, Sharpe) Broth (pH 6.2-6.5) for growth of *Lactobacillus brevis* ATCC 367

Per / L	
Casein Peptone	10 g
Meat Extract	10 g
Yeast Extract	10 g
Glucose	10 g
Tween 80	1 g
K_2HPO_4	2 g
Na- Acetate	5 g
Citrate	5 g
$MgSO_4 \times 7H_2O$	0.20 g
$MnSO_4 \times H_2O$	0.05 g

2.1.6.2 Solid media

2.1.6.2.1 BHI (Brain Heart Infusion) - glucose media for growth of *Streptococcus pyogenes* MGAS315

Per/ L	
Brain Heart Infusion	18.5 g
Agar (OX)	12 g

2.1.6.2.2 YEM (Yeast Extract, Mannitol) - *Rhizobium etli* CFN 42 Media (pH 7.2)

Per / L	
Yeast (OX)	1 g
Mannitol	10 g
Agar	15 g
Soil Extract	200 mL

2.1.6.3 Selective media

In addition to antibiotic, IPTG was added to *E. coli* BL21 strain carrying the pET plasmid to induce transcription of the recombinant gene.

2.1.7 Vectors

PCR products were ligated into vectors shown in Table 2.7 for high level expression and production of histidine tagged proteins. Restriction maps of the different vectors are located in Appendix A.

Vector	Supplier
pET-28a	NOV
pET-22b	NOV

Table 2.7 pET vectors utilised in this research.

2.1.8 Primer sequences (Eurofins MWG)

Open reading frames of chosen genes were checked for signal peptides and internal restriction sites using the bioinformatics tools Signal P and Webcutter 2.0 (Table 2.2). The nucleotides encoding any signal peptides were removed from the gene sequence when designing primers as signal peptides tend to interfere with protein expression and purification.

2.1.9 Isolation of plasmid DNA

2.1.9.1 STET Buffer

0.05 M Tris-HCl pH 8.0, 8 % w / v sucrose, 5 % v / v Triton X-100, 0.05 M EDTA.

Per / L

Sucrose (SIG)	80.0 g
Tris (SIG)	6.0 g
EDTA (SIG)	100.0 mL
Triton X-100 (SIG)	50.0 mL

2.1.9.2 TE/RNAse

TE Buffer (TRIS – HCl 10 mM, EDTA 1mM pH 8.0)

Per L

Tris (SIG) – HCl (0.5 M) pH 7.5	20.0 mL
EDTA (0.5 M) pH 8.0	2.0 mL

2.1.10 Protein purification

2.1.10.1 IMAC (Immobilised Metal Affinity Chromatography)

Affinity Resin

Chelating Sepharose Ni^{2+} FF (Fast Flow) stored in 20 % (v / v) ethanol.

Affinity Buffers

Start buffer (Na_2HPO_4 50 mM, NaCl 0.5 M, imidazole 10 mM) pH 7.4

Elution buffer (Na_2HPO_4 50 mM, NaCl 0.5 M, imidazole 500 mM) pH 7.4

2.1.11 Gel filtration chromatography of oligosaccharides

Gel filtration resin: Bio-Rad P2 gel stored in 20 % v / v ethanol

Gel filtration buffer: degassed 18.2 MΩ /cm H_2O

2.1.12 Electrophoresis

2.1.12.1 Agarose gel electrophoresis

Agarose gel

Agarose (SIG) 0.3 g

1 x TAE Buffer 30 mL

TAE Running buffer (50x Stock)

(2 M Tris-HCl, 1M Glacial acetic acid, 0.05 M EDTA)

Per/ L

Tris-HCl	242.0 g
EDTA (0.5 M) pH 8	100.0 mL
Glacial acetic acid (17.51 M)	57.1 mL

This buffer was diluted 1:50 with 18.2 MΩ / cm H_2O to a working concentration.

Bromophenol blue DNA loading buffer (6x) (stored at 4°C)

Per 10 mL

Bromophenol blue (SIG)	0.025 g
Glycerol (FIS)	3.0 g

This loading buffer used for DNA bands > 5Kb

Xylene DNA loading buffer (6 x) (stored at 4°C)

Per 10 mL

Xylene cyanol	0.025 g
Glycerol	3.0 g

DNA loading buffer was diluted 6 x with sample prior to loading on a 1 % (w / v) agarose gel.

DNA Size Standard

For agarose gels, HyperLadder ^{Tm} 1 (Bioline) was used. Size standard range M.W. 200, 400, 600, 800, 1000, 1500, 2000, 2500, 3000, 4000, 5000, 6000, 8000, 10000 bp. Size standard was initially stored at -20°C until first use when it would be then stored at 4°C.

Ethidium Bromide (FIS)

Ethidium bromide was diluted 1:1000 prior to use from an original concentration of 10 mg / mL to a working concentration of 10 µg / mL.

2.1.12.2 SDS polyacrylamide gel electrophoresis (PAGE)

Resolving Gel

% Gel	40 % (w / v) Acrylamide (SIG)	Buffer B*	H_2O	10 % (w / v) ammonium persulphate (APS)	TEMED (SIG)
10	2.50 mL	2.50 mL	5.00	50 µL	10 µL
12	3.0 mL	2.50 mL	4.50	50 µL	10 µL

Table 2.8 Components of resolving gel.

Stacking Gel

% Gel	40 % (w / v) Acrylamide (SIG)	Buffer C*	H_2O	10 % (w / v) ammonium persulphate (APS)	TEMED (SIG)
4	0.50 mL	1.00 mL	2.50	30 µL	10 µL

Table 2.9 Components of stacking gel.

Buffer B*

Per / 100 mL

2 M Tris pH 8.8	75.0 mL
10 % (w / v) SDS	4.0 mL

Buffer C*

Per / 100 mL

1 M Tris-HCl pH 6.8	50.0 mL
10 % (w / v) SDS	4.0 mL

SDS-PAGE running buffer 10x (Stock)

SDS-PAGE running buffer was made up to a 10 x stock concentration and then diluted 1:10 for a working concentration.

Per / L

Glycine (SIG)	144.0 g
Tris-HCl (MEL)	30.3 g
SDS (SIG)	10.0 g

The pH of glycine and Tris - HCl was adjusted to 8.8 in a volume of ~900 mL before addition of SDS.

SDS-PAGE loading buffer (stored at -20°C)

Per / 10mL

Tris-HCl (60 mM) pH 6.8	0.6 mL
Glycerol (50 % v / v) (SIG)	5.0 mL
SDS (10 % w / v)	2.0 mL
β- Mercaptoethanol (14.4 mM) (SIG)	0.5 mL
Bromophenol blue (0.1 % w / v) (SIG)	1.0 mL
H_2O	0.9 mL

Solubilisation buffer (stored at 4°C)

SDS-PAGE loading buffer	7.6 mL
Urea (SIG)	2.4 g

Protein Size Standards

Low molecular weight range (SIG) M.W. 20, 24, 29, 36, 45 and 66 kDa

High molecular weight range (SIG) M.W. 36, 45, 55, 66, 84, 97, 116 and 205 kDa

Lyophilised samples were resuspended in 200 µL 18.2 MΩ / cm H_2O and 25 µL

Low molecular weight range (NZYT) M.W. 96, 66, 48, 40, 32, 26 and 18.5 kDa

SDS loading buffer before storing the standards at 4°C.

Coomassie blue gel stain

Per / L

Coomassie Brilliant Blue R-250	1 g
Glacial acetic acid	100 mL
Methanol	450 mL

Coomassie gel destain solution

Per / L

Glacial acetic acid	100 mL
Methanol	100 mL

2.1.13 High performance anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD) chemicals

All chemicals required for HPAEC-PAD were degassed by a vacuum pump prior to use.

0.1 M NaOH

Per / L

NaOH (20 M)	5.0 mL

0.1 M NaOH containing 1 M Sodium Acetate

Per / L

NaOH (20 M)	5.0 mL
Sodium Acetate Trihydrate	136.8 g

2.1.14 Substrates and biochemical assays

Biochemical assays were carried out on recombinant enzymes using substrates presented in the following table, Table 2.10.

Substrate	**Supplier**	**Assay**
4-Nitrophenyl-β-D-xylopyranoside	Carbosynth	Xylosidase activity.
4-Nitrophenyl-α-L-arabinofuranoside	Carbosynth	Arabinofuranosidase activity.
Oat Spelt Xylan	SIG and SER	Xylanase and Xylosidase activity.
Birchwood Xylan	SIG	Xylanase and Xylosidase activity.
Wheat Arabinoxylan	MEG	Xylanase and arabinofuranosidase activity.
Rye Arabinoxylan	SIG	Xylanase and arabinofuranosidase activity.
Sugar Beet Arabinan	MEG	Arabinanase and arabinofuranosidase
Linear Arabinan	MEG	Arabinanase activity.
4-*O*-Methyl-D-	SIG	Xylanase and glucuronidase activity.
AZO-Oat Spelt Xylan	MEG	Xylanase activity.

Table 2.10 Table showing various substrates and their specific assay.

2.1.15 DNSA Reagent

Per / L

Dinitrosalicylic acid	10 g
Phenol	2.0 g
NaOH	10 g

Per / 10 mL

Sodium Sulphite	0.5 g
Glucose	2.0 g

2.2 Methods

2.2.1 Microbial methods

2.2.1.1 Growth of *Streptomyces avermitilis* MA480

YMG media (section 2.1.6.1.3) was used to grow *Streptomyces avermitilis*. The culture was incubated at 25°C with shaking at 100 rpm until sufficient growth was observed. The bacteria culture was then processed for the extraction of genomic DNA.

2.2.1.2 Growth of *Lactobacillus brevis* ATCC 367

To grow *Lactobacillus brevis*, MRS media (section 2.1.6.1.4) was used. The culture was incubated at 30°C until sufficient growth of bacterial cells was achieved and genomic DNA extraction could be completed.

2.2.1.3 Growth of bacteria for pET-28a/ pET-22b vector purification

The *E. coli* TOP10 strain was used to purify large quantities of pET DNA. A TOP10 glycerol stock containing the pET vector was used to inoculate a 5 mL starter culture supplemented with the appropriate antibiotic. This starter culture was then used to inoculate a 500 mL broth supplemented with antibiotic in a 2 L conical flask. The culture was incubated overnight at 37°C with shaking at 200 rpm.

2.2.1.4 Growth of bacteria for crude plasmid extraction

E. coli TOP10 cells containing recombinant plasmid were grown in 30 mL sterile glass universals containing 5 mL LB media supplemented with kanamycin or ampicillin, depending upon which vector was used. These cultures were inoculated from a single colony on an agar plate using a sterile tip applied to a Gilson pipette and were incubated overnight at 37°C with orbital shaking at 200 rpm.

2.2.1.5 Growth of bacteria for starter cultures

E. coli strains TOP10 and BL21 (DE3) were grown in 30 mL sterile glass universals containing 5 mL or 10 mL LB media and were supplemented with the appropriate antibiotic. The cultures were inoculated with 50 μL of glycerol stock or from an agar plate using a sterile loop and incubated overnight at 37°C with orbital shaking at 200 rpm.

2.2.1.6 Growth of bacteria for large scale expression and purification of protein

The *E. coli* BL21 (DE3) strain containing recombinant plasmid was grown on a fresh LB agar plate supplemented with the appropriate antibiotic. A colony was used to inoculate a 30 mL sterile universal containing 10 mL LB media also with the addition of antibiotic and incubated overnight at 37°C with orbital shaking at 200 rpm.

This 10 mL starter culture was used to inoculate a 2 L baffled conical flask containing 1 L of LB media supplemented with antibiotic. The culture was incubated at 37°C with shaking at 200 rpm for approximately 3 hours until the culture reached an OD_{600} of 0.6 - 1. The OD of the culture was verified by aseptically transferring 1 mL of culture to a cuvette and reading the OD using a spectrophotometer which had previously been set to 600 nm and zeroed with LB. Once the OD read between 0.6 – 1, the cells were induced by the addition of IPTG (0.1 M) at a stock concentration of 24 mg / mL, to give a final concentration of 0.24 mg / mL.

Post induction, the cells were incubated at either 23°C or 16°C with overnight orbital shaking at 100 rpm.

2.2.1.7 Growth of bacteria for small scale expression of protein to optimise growth conditions.

To optimise growth conditions for large scale protein expression, expression was initially observed on a small scale using the *E. coli* strain BL21. BL21 cells containing recombinant plasmid DNA were grown on a fresh agar plate supplemented with antibiotic. A small scrape of cells were picked using a sterile loop which were used to inoculate a 250 mL sterile conical flask containing 50 mL LB media supplemented with antibiotic. The culture was incubated at 37°C with orbital shaking at 200 rpm for approximately 2 hours. The cells were monitored by measuring the OD with a spectrophotometer which had previously been zeroed with LB media. When the culture had reached an OD of 0.6 - 1, 50 µL of IPTG (0.1 M) was added to the cells and then incubated at 23°C or 16°C with overnight shaking at 100 rpm. A replicate culture that was not induced by IPTG was also grown but followed all the same growth conditions.

2.2.1.8 Preparation of *E. coli* chemically competent cells

A sterile 30 mL glass universal containing 10 mL LB media was inoculated with a single fresh colony of TOP10 or BL21 *E. coli* cells. This culture was incubated for 16 h at 37°C with orbital shaking at 200 rpm. After this incubation period, 1.5 mL of culture was aliquoted into a sterile 2 L conical flask containing 500 mL LB media and incubated at 37°C, 200 rpm. When the OD_{600} reached 0.35 – 0.4, the cells were then poured aseptically into sterile ice cold centrifugal containers and centrifuged for 10 min at 2,000 x g, 4°C. The

pelleted cells were retained and resuspended in 100 mL sterile ice cold $MgCl_2$ by gentle vortexing. The cells were pelleted again by centrifugation at 2,000 x g, 4°C for 10 min and the supernatant later discarded. The cells were finally resuspended in 15 mL of sterilised ice cold $CaCl_2$ and incubated on ice for at least 1.5 h to become competent. This method produced a large supply of competent cells for many transformations therefore 4.5 mL of 50 % (v / v) glycerol was added to the cells before they were aliquoted into 150 µL samples within sterilised 1.5 mL microcentrifuge tubes and stored at -80°C for future use.

2.2.2 DNA Methods and molecular biology kits

The recipes for the buffers used within the kits are unknown unless stated otherwise.

2.2.2.1 Genomic DNA (gDNA) isolation from Gram-positive bacteria

The extraction of gDNA from Gram-positive bacteria was carried out using the DNeasy Blood and Tissue kit (QIA). Firstly, 0.5 mL of cultured cells (section 2.2.1) was aliquoted into a sterile 1.5 mL microcentrifuge tube and centrifuged for 10 min at 5000 x g. The supernatant was discarded and the pellet resuspended in 180 µL of enzymatic lysis buffer (20 mM Tris-HCl, pH 8, 2mM EDTA, 1.2% Triton™ X-100). The sample was incubated at 37°C with occasional vortexing so cells were fully resuspended. After 30 min of incubation, 25 µL of proteinase K and 200 µL of Buffer AL (for precipitation of gDNA) were added to the sample and this was mixed by vortexing. Subsequently, 200 µL of ethanol was added to the sample and again this was mixed thoroughly by vortexing before the sample was loaded into a DNeasy Mini spin column. The column was placed in a 2 mL collection tube and centrifuged at 6000 x g for 1 min. The flow through was discarded and the column was placed in a new 2 mL collection tube prior to the addition of 500 µL of Buffer AW1. The column was centrifuged again at 6000 x g for 1 min and the flow through discarded. The column was placed into a new 2 mL collection tube and 500 µL of Buffer AW2 was added. The column was centrifuged for 3 min at 20,000 x g to dry the column membrane and this step was repeated for 1 min to remove any residual ethanol. The column was then placed into a sterile 1.5 mL microcentrifuge tube and 200 µL of Buffer AE (5 mM Tris-HCl, pH 8) was directly pipetted onto the centre of the column membrane. This was then left to stand for 1 min at room temperature before being centrifuged at 6000 x g to elute the gDNA. For maximum gDNA yield, the elution step was repeated. To check whether gDNA extraction had been successful, 5 µL of the eluted sample was run on an agarose gel (section 2.2.9.1).

2.2.2.2 gDNA isolation from Gram-negative bacteria

The extraction of gDNA from Gram-negative bacteria was carried out using the DNeasy Blood and Tissue kit (QIA). Firstly, 0.5 mL of cultured cells (section 2.2.1) was aliquoted into a sterile 1.5 mL microcentrifuge tube and centrifuged for 10 min at 5000 x g. The supernatant was discarded and the pellet re-suspended in 180 µL Buffer ATL (cell re-suspension buffer) followed by the addition of 20 µL proteinase K (>600 mAU / mL). The sample was mixed and then incubated at 56°C with the occasional vortexing until the cells had completely lysed. After incubation, 200 µL of Buffer AL (for precipitation of gDNA) was added to the sample and this was mixed vortexing before the addition of 200 µL ethanol. Once the ethanol was added, the sample was vortexed once more and then was loaded onto the DNeasy Mini spin column. The column was placed in a 2 mL collection tube and centrifuged at 6000 x g for 1 min. The flow through was discarded and the column was placed in a new 2 mL collection tube prior to the addition of 500 µL of Buffer AW1. The column was centrifuged again at 6000 x g for 1 min and the flow through discarded. The column was placed into a new 2 mL collection tube and 500 µL of Buffer AW2 was added. The column was centrifuged for 3 min at 20,000 x g to dry the column membrane and this step was repeated for 1 min to remove any residual ethanol. The column was then placed into a sterile 1.5 mL microcentrifuge tube and 200 µL of Buffer AE (5 mM Tris-HCl, pH 8) was directly pipetted onto the centre of the column membrane. This was then left to stand for 1 min at room temperature before being centrifuged at 6000 x g to elute the gDNA. For maximum gDNA yield, the elution step was repeated. To check whether gDNA extraction had been successful, 5 µL of the eluted sample was run on an agarose gel (section 2.2.9.1).

2.2.2.3 Preparation of cloning vector

The preparation of vector was carried out using the NZYMidiprep Kit (NZY) and was the same protocol used for both pET28a and pET-22b vectors. The cell culture (section 2.2.1.5) was centrifuged at 6000 x g for 10 min at 4°C in which the remaining pellet was resuspended in 8 mL M1 buffer containing RNase (re-suspension buffer). A lysate was formed by the addition of 8 mL M2 Buffer (cell lysis buffer) and this was mixed by inverting the tube 4-6 times and incubated at room temperature. After 5 min incubation, 8 mL of M3 Buffer (precipitation of DNA) was added to the lysate and mixed immediately by inverting the tube 10-15 times. The suspension was then centrifuged at 10,000 x g for 15 min at ambient temperature after which the supernatant was removed instantly and re-centrifuged as residual precipitate had the potential to block the column. The supernatant containing the DNA was removed and applied to an NZYTech Plasmid Midi Column pre-equilibrated with 12 mL MEQ Buffer. The supernatant was then allowed to flow through the column by gravity. The Midi Column was then washed with 5 mL MEQ Buffer followed by a wash with 8 mL of Buffer MW. The plasmid DNA was eluted with 5 mL Buffer ME into a sterile plastic universal. The eluted DNA was precipitated by the addition of 3.5 mL isopropanol at room temperature which was subsequently aliquoted into 1.5 mL microcentrifuge tubes and centrifuged at 15,000 x g for 30 min at 4°C. The supernatant was carefully poured off and the remaining pellet was washed with 0.5 mL 70 % ethanol. This was centrifuged for 5 min at 15,000 x g at 4°C after which the supernatant was removed. The DNA was dried using a speed vac centrifuge until no residual ethanol remained. The pellets were resuspended in 500 µL of TE (5 mM Tris-HCl, pH 8) buffer and stored at -20°C.

2.2.2.4 Crude extraction of plasmids for screening recombinant DNA

A 3 mL volume of overnight grown bacterial cell culture was pipette into a 1.5 mL microcentrifuge by repeated centrifugation at 14,000 x g for 3 min. The pelleted cells were

resuspended in 150 µL STET buffer (section 2.1.9.1) and 10 µL of 10 mg / mL lysozyme was added to the cells and mixed by vortexing. Cells were incubated at room temperature for 10 min before being boiled for 40 sec. They were then immediately centrifuged at 14,000 x g for 15 min producing a viscous pellet that was removed using a yellow tip attached to a Gilson pipette and discarded. To the remaining supernatant, 150 µL of isopropanol was added and the sample was mixed by inverting the tube multiple times. The sample was then incubated at -20°C for 1 h to allow precipitation of DNA. Immediately after removal from -20°C, the sample was centrifuged at 14,000 x g for 5 min and the supernatant discarded. The pellet was then washed with 500 µL of ethanol before centrifugation at 14,000 x g for 3 min. After discarding the supernatant, the pellet was pulsed spun and the residual supernatant was removed with a yellow tip attached to a Gilson pipette. Finally, the sample was dried to remove any remaining ethanol, prior to resuspension of the pellet in 30 µL of TE buffer containing 10 mg / mL RNase.

2.2.2.5 Isolation of plasmid DNA

Isolation of plasmid DNA from transformed *E. coli* TOP10 (section 2.2.8) was achieved using NZYMiniprep kit. Firstly, 5 mL of *E. coli* culture was centrifuged at 11,000 x g for 1 min with the resulting supernatant discarded. The pellet was resuspended in 250 µL Buffer A1 (cell re-suspension buffer) by pipetting using a Gilson pipette. This was followed by the addition of 250 µL of Buffer A2 (cell lysis buffer) and mixed gently by inverting the tube 8 times. This was incubated at room temperature for 4 min before adding 300 µL of Buffer A3 (precipitation of DNA). Again the sample was mixed gently by inverting the tube 8 times followed by centrifugation for 10 min at 11,000 x g. The supernatant was then loaded onto an NZYTech spin column which had been placed into a 2 mL collecting tube. The column was centrifuged for 1 min at 11,000 x g. The flow through was discarded and 500 µL of Buffer AY was loaded onto the column which was centrifuged again at 11,000 x g for 1 min.

With the flow through being discarded, Buffer A4 proceeded to be loaded onto the column and centrifuged for a further 1 min at 11,000 x g. The flow through was discarded and the column was centrifuged once again to remove residual ethanol from the column. The spin column was then placed into a sterile 1.5 mL microcentrifuge tube and 30 µL of Buffer AE (5 mM Tris-HCl, pH 8) was added to the centre of the column. The column was left to incubate at room temperature for 1 min before centrifugation at 11,000 x g for 1 min. The eluted sample was passed through the column again by centrifugation to increase overall DNA yield. The DNA was essentially stored at -20°C.

2.2.2.6 Gel extraction of DNA

Prior to the extraction of a gel fragment, the gel was briefly visualised under a low intensity UV light and the location of the correct fragment was etched using a clean scalpel. The UV light was then switched off and using the scalpel, the fragment was cut out from the gel and placed into a pre-weighed sterile plastic universal tube. The tube was re-weighed and the weight of the gel fragment could be calculated. The purification of the DNA from the gel fragment was carried out using NZYGelpure kit. For every 100 mg of gel weight, 300 µL of Binding Buffer was added to the tube. The gel was then dissolved by incubating the tube at 50°C for 10 min, with occasional mixing using a vortex. Once the gel had completely dissolved, one gel volume of isopropanol was added to the sample and this was mixed thoroughly using a vortex. The sample was then loaded into the spin column which had been placed into a 2 mL collection tube and left to stand for 1 min. The spin column was centrifuged at 11,000 x g for 1 min and the collected flow through was discarded. Subsequently, 600 µL of Wash Buffer was loaded into the spin column and was left to stand for 2 min prior to centrifugation at 11,000 x g for 1 min. The flow through was also discarded. The centrifugation step was repeated to remove any residual ethanol and dry the column membrane. The spin column was then placed into a sterile 1.5 mL microcentrifuge

tube and 50 µL of Elution Buffer was pipette into the centre of the column membrane. This was incubated at room temperature for 2 min prior to centrifugation at 11,000 x g for 1 min to elute the DNA. The elution step was repeated to recover the maximum amount of DNA and the sample was stored at -20°C. To check purification was successful, 2 µL of sample plus 8 µL of sterile 18.2 MΩ / cm H_2O, was loaded and ran on an agarose gel (section 2.2.9.1).

2.2.2.7 PCR/ restriction digest clean up

PCR product and digested DNA were cleaned up using NZYGelpure kit. Firstly, the DNA was added to five volumes of Binding Buffer in a sterile 1.5 mL microcentrifuge tube. The reaction volume was mixed by inverting the tube a few times. After mixing, the DNA sample was loaded into a spin column which was placed in a 2 mL collection tube. This was left to stand at room temperature for 2 min before centrifugation at 11,000 x g for 1 min. The collected flow through was discarded. The DNA sample was then washed by the addition of Wash Buffer and centrifuged at 11,000 x g for 1 min. Again the flow through was discarded. Centrifugation was repeated to remove any residual ethanol from the membrane of the spin column. The spin column was then placed into a sterile 1.5 mL microcentrifuge tube and 50 µL of Elution Buffer was added to the centre of the spin membrane. The column was left to stand for 2 min at room temperature before being centrifuged at 11,000 x g for 1 min to elute the DNA. This step was repeated to recover the maximum amount of DNA. To determine if the clean up was successful, 2 µL of sample was added to sterile 18.2 MΩ / cm H_2O making up the volume to 10 µL and run on an agarose gel (section 2.2.9.1).

2.2.3 PCR

Genes were amplified using PCR in which components are listed in Table 2.11. The PCR protocol is presented in Table 2.12. However for the amplification of individual genes, different primary and secondary annealing temperatures were required due to varying primer sequences.

Reaction Component	Volume	Final Concentration
Sterile 18.2 MΩ / cm H_2O	32.5 µL	
Genomic DNA	1 µL	
dNTPs 2 mM (each)	5 µL	0.2 mM (each)
$MgSO_4$ 50 mM	3 µL	1.5 mM
Primer mixture (Primer mixture= Sense (5') primer (10 µM) + Anti-sense (3') primer (10 µM) + 8µL water)	1.5 µL	0.3 µM for each primer
KOD Hot Start DNA Polymerase (1 U/µL) (INV)	1 µL	0.02 U/ µL
10x buffer for KOD Hot Start DNA (INV)	1 µL	1 x

Table 2.11 Components required for one single PCR reaction.

Temperature (°C)	Duration (min:sec)	Cycles
95	2:00	Initial Denaturation
95	0:15	4
Primary Annealing Temperature	0:30	
72	Kb / min	
95	0:15	24
Secondary Annealing	00:30	
72	Kb / min	
72	10:00	Final Elongation
10	HOLD	

Table 2.12 General reactions conditions for PCR.

2.2.4 Restriction digest of DNA

All DNA restriction digests were performed at 37°C for 2 h, unless suggested otherwise. Each digest required the appropriate buffer (NEB) which was added with a final concentration of 1 x. BSA was also added to the reaction mix at a final concentration of 0.1 mg / mL. Restriction endonucleases were supplied in 50 % glycerol so care was taken to ensure that the glycerol concentration did not exceed 10 % in the final reaction volume as glycerol as a tendency to affect enzyme specificity.

Mini spin DNA restriction digests using 0.25 ng of plasmid DNA were required to confirm the presence of insert. From crude plasmid extraction, 7 µL of plasmid DNA was digested

with the appropriate restriction endonucleases to release the correct insert. From this, positive cultures were processed to isolate mini spin plasmid DNA.

PCR products were digested using the appropriate restriction enzymes and ligated into pET vector.

For ligation into pET vectors, 5 µg of pET vector was digested with the appropriate restriction enzymes and incubated for > 2 hours. Once fully digested the vector was loaded onto a 1 % agarose gel and left to run over night at 2 mA for sufficient separation of undigested and digested DNA, before gel purification was carried out.

2.2.5 Determination of DNA concentration

Plasmid DNA was quantitatively determined by transferring 50 µL of DNA to a 1 mL sterile cuvette (Eppendorf UVette ®) and scanned in a UV spectrophotometer previously zeroed with water. Optimal wavelength of DNA was A_{260} and with the extinction co-efficient of double stranded DNA being 50 (According to Sambrook and Russell), the DNA concentration could be calculated using the following equation:

[DNA] ng / µL = (A260) x 50 (extinction co-efficient)

2.2.6 DNA ligation

Following digestion with restriction enzymes, PCR products were ligated into pET-28a or pET-22b vectors. For maximum efficiency during ligation a 10:1 to 3:1 range of insert to vector was used. A standard ligation into pET vector was carried out using T4 DNA ligase and an optimised buffer system. The components of the ligation reaction are listed in Table 2.13.

Component	Volume/ Amount
PCR product	3:1 insert : vector ratio
Digested pET vector	100 ng
10 x T4 ligation buffer (NEB) 50 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM ATP, 10 mM DithiothreitoL	1 µL
T4 DNA ligase (4 U/µL) (NEB)	1 µL
18.2 MΩ/cm H_2O	Make up to a final volume of 10 µL

Table 2.13 Reaction components for ligation into pET vector.

After the addition of reaction components, the ligation was incubated overnight at 16°C prior to transformation into TOP10 *E. coli* cells (section 2.2.8).

2.2.7 Transformations in *E. coli* BL21/ TOP10 strains

Primarily, 10 µL of ligated product was added to 150 µL of competent cells that had previously been aliquoted and stored at -80°C and incubated on ice. After 30 min, the cells were heat shocked by incubation in a 42°C water bath for 90 s. The cells were placed back on ice to cool for 3 min before the aseptic addition of 200 µL LB media. The cells were then incubated for 45 min at 37°C prior to being plated out aseptically using a sterile glass spreader, onto a pre- dried 2 % (w / v) agar plate supplemented with the correct antibiotic.

Plates were then incubated inverted for 16 h at 37°C to allow the growth of bacterial cells. This procedure was identical regardless of the *E. coli* strain being used.

2.2.8 Isolation of cell free extract (CFE)

2.2.8.1 Isolation of CFE on a small scale

Two 1.5 mL aliquots of overnight culture (section 2.2.1.7) were centrifuged for 5 min at 14,000 x g. The pelleted cells from one of the aliquots was re-suspended in 150 µL re-suspension buffer (IMAC start buffer, section 2.1.10.1) and these cells were lysed by sonication on ice at an amplitude of 6 for 1.5 min in 10 s pulses with 10 s intervals. The lysate was centrifuged once more for 5 min at 14,000 x g and the supernatant collected in a clean microcentrifuge tube which was kept on ice as CFE. The remaining aliquot of pelleted cells was re-suspended in 150 µL solubilisation buffer (see section 2.1.12.2) to detect the occurrence of insoluble protein. A 20 µL sample of CFE and solubilised protein was analysed on a SDS-PAGE (section 2.2.9.2).

2.2.8.2 Isolation of CFE on a large scale

Overnight BL21 cultures were centrifuged in 200 mL centrifuge containers for 10 min at 4,000 x g. The supernatant was discarded and the cells were re-suspended in 5 mL re-suspension buffer (IMAC start buffer, section 2.1.10.1). The cells were then subjected to cell lysis by sonication at an amplitude of 14 for 1.5 – 2 min in 10 second pulses with 10 second intervals. The lysate was then decanted into a sterile 20 mL centrifuge tube and centrifuged at 4°C for 30 min at 24,000 x g. The soluble fraction was poured into a sterile plastic universal and kept on ice as CFE until protein purification (section 2.2.10.4) was carried out.

2.2.9 Electrophoresis

2.2.9.1 Agarose gel electrophoresis

Agarose gels (1 % w / v) were prepared with 1 x TAE buffer, heated over a Bunsen flame until the agarose had completely dissolved. After a few minutes of cooling, the gel was poured into a gel casting tray and the appropriate sized comb was inserted. The gel was then left to set for > 20 min.

After the gel had set, the tray was placed horizontally in the gel tank and submerged in 1 x TAE buffer. DNA samples to be analysed were loaded with the addition of either Bromophenol Blue Loading Buffer or Xylene DNA Loading Buffer loading buffer. Size standard (5 µL) was also loaded into one of the wells and the gel was electrophoresed for approximately 30 min at 120 milliamps to ensure good separation of DNA fragments.

After electrophoresis, the gel was stained by submersion in ethidium bromide (10 µg / mL) for 10 min. For analytical gels, ethidium bromide could be re-used until the stain became weak however for preparative gels, fresh ethidium bromide was made up each time.

After staining, the gel was washed thoroughly with distilled water and visualised under a Bio-Rad Universal Hood for a source of UV light.

2.2.9.2 SDS-PAGE

SDS-PAGE gels were cast between two clean glass plates with built in 0.75 mm spacers. The plates were assembled so that they were level at the bottom edges before being clamped together and secured in a casting stand on a rubber gasket. The components of the resolving gel (section 2.1.12.2) were pipetted within the space between the glass plates using a 5 mL Gilson, until the solution was 1.5 cm below the ridge of the smallest plate. A small volume

of water was pipetted to the top of the gel solution in order to stop oxygen from inhibiting polymerisation and for the gel edge to ultimately set as a straight line. The resolving gel was allowed to polymerise for > 20 min. The water from the top of the gel was poured off and any residues left behind were removed by absorption using filter paper.

The stacking gel (section 2.1.12.2) was pipetted into the remainder of the space between the two plates and a 10 well comb was inserted immediately. The stacking gel was allowed to polymerise for > 20 min prior to the removal of the comb and extensive washing of the wells with distilled H_2O. The gels were then placed vertically opposite one another into the gel tank and secured with a holder.

SDS–PAGE running buffer (section 2.1.12.2) was poured into the tank, ensuring no leaks, ready for the loading of samples.

SDS loading buffer was added to the sample and boiled for 2 min prior to being loaded into the wells using a Hamilton syringe. Low molecular weight and high molecular weight size standards (10 μL) were also loaded onto the gel along with 20 μL of sample. The gels were subjected to electrophoresis at 200 constant volts for 60 min.

After electrophoresis, SDS–PAGE gels were stained with Coomassie Blue stain (section 2.1.12.2) for 15 min on a flat bed orbital shaker at 40 rpm. After staining, gels were destained overnight using Coomassie destain solution (section 2.1.12.2), again shaking at 40 rpm on a flat bed shaker. The destaining solution was poured off and gels were rinsed with distilled H_2O and visualised using a Bio-Rad GS-800 Calibrated Densitometer scanner.

2.2.10 Protein purification

2.2.10.1 Immobilised metal affinity chromatography (IMAC)

IMAC was the method used to purify recombinant proteins with Chelating Sepharose^{TM} Fast Flow (FF) resin which was poured into open tubular mini columns. The Chelating Sepharose^{TM} was supplied and stored in 20 % (v / v) ethanol and therefore all traces had to be removed before use. This was carried out by 'slurrying' the resin with 5 x column bed volumes of 18.2 MΩ / cm H_2O within the mini column before stripping the resin with EDTA (0.5 M). The resin was washed again with 5 x column bed volumes of 18.2 MΩ / cm H_2O ensuring that no EDTA remained within the column. This was followed by a single wash with NaOH (1 M) to remove any traces of carbohydrates that could be present and then a wash with 18.2 MΩ / cm H_2O, in which 6 x column bed volumes were applied. Subsequently, 2 x column bed volumes of Nickel ions, in the form of $NiSO_4$ solution (0.1 M), were poured onto the resin before finally be washed with 18.2 MΩ / cm H_2O. The column was then equilibrated with 5 x column bed volumes of IMAC start buffer ready for the binding of the protein.

The CFE containing the desired recombinant protein was loaded onto the column and the column was washed with 50 mL of start buffer under the flow of gravity. The protein was eluted with elution buffer containing different concentrations of imidazole respectively, 50 mM - 500 mM. Proteins usually eluted after the addition of 500 mM imidazole elution buffer. All the different eluants from the column were collected in separate sterile 50 mL Falcon tubes and a sample of each were analysed on a SDS-PAGE gel (section 2.2.9.2).

Once the purified protein was eluted from the column, proteins were stored in 3.2 M ammonium sulphate allowing the stabilisation of the protein for storage at 4°C.

2.2.10.2 Gel filtration chromatography

Gel filtration was used to exchange the buffer from the IMAC purified protein to gel filtration buffer. The gel filtration column was prepared by equilibration at 1 mL / min with 120 mL of gel filtration buffer using the FPLC system (section 2.1.10.2). The sample was loaded onto the column and eluted over 120 min with gel filtration buffer running at 1 mL / min. Samples were collected in 5 mL volumes using a fraction collector. Protein was identified using an inline UV spectrophotometer set at A_{280}. Fractions were analysed by SDS–PAGE (section 2.2.9.2) to locate the desired protein. Fractions containing the required protein were pooled together and concentrated for analysis (section 2.2.10.3).

2.2.10.3 Concentration and buffer exchange of protein

Concentrating a protein was achieved by centrifugation at 4,000 x g, 4°C in a 30 kDa cut-off concentrator unit (manufacturer). Proteins purified by IMAC were concentrated to 0.5 mL and washed three times with 5 mL gel filtration buffer for loading onto the gel filtration column.

The resulting protein from gel filtration was concentrated to 0.5 mL then washed three times with 18.2 MΩ / cm H_2O to prepare the protein for crystallisation.

2.2.10.4 Protein quantification

Protein concentration was determined using a UV spectrophotometer set at A_{280} and zeroed with 18.2 MΩ / cm H_2O. Protein was usually diluted in a total volume of 100 µL. The dilution was transferred into a cuvette (Eppendorf UVette®) and scanned in the

spectrophotometer. The concentration of the protein was determined using the equation below:

$$[\text{Protein}]\ \text{mg / mL} = \frac{A_{280} \times \text{dilution factor} \times \text{molecular weight}}{\text{Extinction coefficient}}$$

For enzymatic assays the molar concentration of enzymes were required and were therefore calculated using the equation below:

$$[\text{enzyme}]\ \text{M} = [\text{Protein}] / \text{Molecular Weight}$$

or

$$[\text{enzyme}]\ \text{M / L} = Ab_{280} / \text{Extinction Coefficient}$$

2.2.11 Enzymatic assay methods

All assays were carried out in triplicate and performed at 25°C unless stated otherwise, with the addition of enzyme always performed last. BSA was added to all assays at a final concentration of 1 mg / mL. Standard curves were produced in triplicate and only assay points in the linear region were accepted. Initial rates of reaction were observed by linearity up to the first four points on a [product] versus time plot. For every substrate, a control assay was set up to determine any non-enzymatic release of product. This generally involved the incubation of substrate with BSA, buffer and pre-boiled enzyme. For calculating kinetic parameters and when producing linear points on a graph were imperative, the first time point was taken once the enzyme had been added to the reaction and a reaction sample was taken and boiled straight away. This first time point was designated the time zero sample.

2.2.11.1 *p*-Nitrophenol assay

For initial screening of enzyme activity, enzymes were added to the substrates 4-nitrophenyl-β-D-xylopyranoside an 4-nitrophenyl-α-L-arabinofuranoside diluted with sodium phosphate or sodium acetate buffer and 1 mg / mL BSA making up the volume to 500 µL. The reaction was incubated at room temperature and colour change was observed, with the liberation of the chromophore, 4-nitrophenol, from the glycosyl molecule.

2.2.11.2 Dinitrosalicylic acid (DNSA) reducing sugar assay

The rate of hydrolysis of polysaccharides was monitored by reducing sugar content which was determined by DNSA using Miller's method (Miller, 1959). Sugars with a free anomeric carbon can exist as a linear molecule with an aldehyde group that is capable of acting as reducing agent. Upon reduction DNSA undergoes a colour change which can be

detected at 575 nm. Firstly, enzyme was added to 8 mg / mL (final concentration) soluble substrate diluted in 0.05M (final concentration) sodium phosphate or sodium acetate buffer containing 1 mg / mL (final concentration) BSA. The reaction was made up to 1 mL within a microcentrifuge tube and incubated for an appropriate amount of time. The reaction was terminated by boiling for 5 min. A 1 in 5 dilution of the reaction was taken making the final volume 500 µL and this was added to 500 µL of DNSA reagent. The microcentrifuge tube lid was pierced with a needle and the reaction was boiled. After 20 min of boiling, tubes were placed on ice for 10 min before being allowed to equilibrate to room temperature. The absorbance of each assay was determined at 575 nm in a plastic 1cm light path cuvette using a spectrophotometer. Reducing sugar released was determined from a DNSA standard curve.

A DNSA standard curve (Figure 2.1) was generated by the addition of 0, 0.05, 0.1, .015, 0.2 and 0.25 mg of xylose to 500 µL DNSA reagent. The volume was made up to 1 mL with 0.05 M sodium phosphate buffer containing 1 mg / mL BSA.

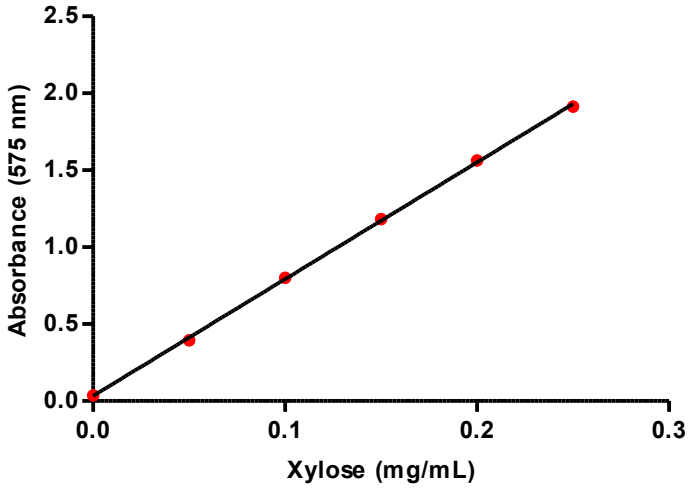


Figure 2.1 Standard DNSA curve for the determination of reducing sugar.

2.2.11.3 Preparation of soluble oat spelt xylan from xylan oat spelts (SIG)

Soluble oat spelt xylan was prepared using 10 g of xylan from oat spelts suspended in 100 mL 18.2 MΩ / cm H_2O. The pH was adjusted to 10 with 1 M NaOH and continually stirred for 1 h at room temperature using a magnetic stirrer. After an hour, 1 M acetic acid was added to the mixture until pH 7 was reached. The solution was then centrifuged at 10,000 x g at room temperature for 10 min. The supernatant was retained and stored at -20°C until further use. The insoluble xylan pellet was discarded.

2.2.11.4 AZO-Oat spelt xylan assay

This assay was used to detect xylanase activity which only allowed a qualitative comparison to relative enzyme specificities. Absolute values of activity could not be calculated through this method since this dyed polysaccharide was of unknown size and composition and therefore could not be quantified.

Enzymes were incubated with 8 mg / mL AZO-OSX diluted in sodium phosphate or sodium acetate buffer containing 1 mg / mL BSA. The reaction was made up to 1 mL within a microcentrifuge tube and incubated for an appropriate amount of time. Subsequently, 400 µL of the reaction was added to 1 mL of ethanol to which the reaction was stopped. The sample was then vortexed and centrifuged at 3,000 rpm for 10min. The absorbance of the supernatant was then determined at 590 nm.

2.2.11.5 Determination of kinetic parameters

For the determination of kinetic parameters of OtXyn8a against OSX, varying substrate concentrations between 1 mg/mL and 8 mg/mL (1 mg/mL, 1.2 mg/mL, 1.4 mg/mL, 1.6

mg/mL, 2mg/mL, 4 mg/mL and 8 mg/mL) were chosen and performed as described in section 2.2.11.2. A standard curve was generated for each substrate concentration. To all assays 0.05 M sodium acetate, pH 6 and BSA (1 mg/mL) were added. Aliquots of 100 µL were removed at 0, 0.5, 1, 2, 4, 8 and 10 min. For time 0, the enzyme had been added to the reaction and a sample was taken and boiled straight away to terminate the reaction. All assays were carried out at 45°C. In order to ascertain if these data reflected true Michaelis-Menten kinetics, a Lineweaver-Burk plot was generated and if a straight line was observed the data was accepted as true.

2.2.11.5.1 Calculation of K_M, k_{cat} and k_{cat} / K_M from raw data

All graphs were plotted in GraphPad Prism 5.2 and the kinetic constants were determined by non-linear regression of the plots of V_0 against [S]. K_M and k_{cat} were calculated from Lineweaver-Burk plot as follows:

K_M (mg ml-1) = 1 / (x intercept on Lineweaver-Burk plot)

V_{max} = 1 / (y intercept on Lineweaver-Burk plot)

k_{cat} (min-1) = V_{max} / [E] moles

$V_{0\ =}$ initial velocity, V_{max} = maximum reaction rate of the enzyme, K_M = concentration at which the reaction rate is half V_{max}, [S] = substrate concentration, [E] = enzyme concentration, k_{cat} = number of substrate molecules turned over per molecule of enzyme.

2.2.12 HPAEC-PAD analysis of hydrolysis products

To analyse sugar products from enzymatic hydrolysis, HPAEC-PAD was used. The HPAEC-PAD set up comprised a Dionex 500 system with a CarboPac PA-100 column. This

was suitable for detecting monosaccharides, disaccharides and oligosaccharides using Chromeleon software. Samples were loaded onto the system using an Automated Sampler AS40, which processed the sample through to the column by a 150 µL loop at 1 mL min^{-1} with a system pressure of ~2500 psi. Degassed solutions (section 2.1.13) were used to equilibrate the column and ultimately elute the oligosaccharides. An isocratic elution was carried out for 5 min with 100 mM NaOH to allow the separation of mono- and di-saccharides. This was followed by 5-15 min 100 mM NaOH with a 0-7.5 mM NaOAc linear gradient for the resolution of larger oligosaccharides. The column was washed for 10 min with 100 mM NaOH and 1 M NaOAc before it was equilibrated again with 100 mM NaOH for 20 min. It should be noted for the separation of arabinose and xylose, 50 mM NaOH solutions were used resulting in increased retention times of eluted saccharides. Eluted oligosaccharides were detected using pulsed amperometry from an electrochemical detector, Dionex model ED40. The pulsed amperometric detection settings were E_1 = +0.05, E_2 = +0.6 and E_3 = -0.6. The All data collection and analysis was performed through Chromeleon software version 6.8.

2.2.12.1 Quantification of xylooligosaccharides from HPAEC-PAD

For the quantification of xylooligosaccharides xylose was used as an absolute standard for higher sugars therefore this requires that all the other sugars are related to xylose in terms of peak area. Therefore correction factors had to be determined so the area of xylose equivalents could be calculated. Correction factors had previously been determined and kindly supplied by Dr Simon Charnock. The correction factors for peak are relative to xylose were 1.985, 2.144, 2.436, 2.755 and 3.091 for X2, X3, X4, X5 and X6, respectively. The peak area for a standard solution of xylose at 0.2 mg/mL (1300 µM) was determined for each HPAEC-PAD experiment.

The concentration of an xylooligosaccharide was calculated using the following equation:

[xylooligosaccharide] μM = (area x correction factor x 1300) / area X1 standard

2.2.13 Purification of xylooligosaccharides

Hydrolysis products of oat spelt xylan digestion with endo-xylanase, OtXyn8A, were separated and purified using a Bio-Rad Econo-Column (170 x 1.5 cm) packed with Bio-gel^{TM} P2 polyacrylamide gel resin. The column was equilibrated using 18.2 MΩ / cm H_2O at a flow rate of 0.1 mL min^{-1} for 24 h. A sample of 2 mL hydrolysed xylan was then loaded onto the column and eluted at a flow rate of 0.1 mL min^{-1} in 18.2 MΩ / cm H_2O. Fractions were collected using a Pharmacia Biotech RediFrac fraction collector and were subsequently analysed with HPAEC-PAD.

3 Results and Discussion

Cloning and Expression of Carbohydrate-Active Enzymes

3.1 Introduction

For the screening and ultimate isolation of novel specificity of carbohydrate-active enzymes, it was essential, within this research, to propose a strategy that rationally selected uncharacterised enzymes from the vast amount of sequence data available from accessible resources. Prokaryotes possess a great amount of sequence diversity and the exploitation of various bacteria allows the use of conventional methods to tap into such information. Therefore with regards to this concept, an in-house library of bacterial strains was applied to this current study, providing a compliant resource. Through the utilisation of the CAZy database, an initial task was undertaken to select uncharacterised carbohydrate-active enzymes that were to be investigated, from numerous CAZy families for each obtainable bacterial strain. It was important to elucidate uncharacterised enzymes from families that shared distinct diversity in substrate specificity amongst its members, such as family GH43.

The functional significance of family 43 glycoside hydrolases and the expansion of the family itself meant that the occurrence of novel specificities within GH43 seemed quite probable. Therefore, an opportunity to identify uncharacterised GH43 enzymes from the available bacterial strains became one of the fundamental objectives during this research. The well known soil bacterium *Cellvibrio japonicus* is categorised as one of the most efficient plant cell wall degraders (DeBoy *et al*., 2008), due to the repertoire of glycoside hydrolases the organism produces. In addition, *C. japonicus* encodes a large collection of GH43 enzymes. This intriguing feature has drawn much attention in characterising *C. japonicus* GH43 enzymes and understanding their requirement for various roles in nature. It therefore became apparent to look for alternative bacteria that shared this GH43 characteristic and this was a simple, although a relatively time consuming task of directly

searching the genomes presented within the CAZy database for GH43 enzymes. Numerous bacterial genomes have been found to constitute a large number of GH43 enzymes such as *Saccharophagus degradans* 2-40 which contains 13 GH43 members and *Bacteroides thetaiotaomicron* VPI-5482 in which 6 out of its 33 GH43 enzymes have been characterised. However, the bacterium *Opitutus terrae* PB90-1 harbours a large number of uncharacterised GH43 enzymes and was therefore a target of this research. *O. terrae* PB90-1 was first isolated from rice paddy soil and it was shown that growth was supported by mono-, di- and polysaccharides (Chin *et al*., 2001). The genome of *O. terrae* has been very recently sequenced (van Passel *et al*., 2011) identifying that this anaerobic bacterium harbours at least 156 genes predicted to encode glycoside hydrolases. *O. terrae* PB90-1 was an almost untouched genome, with only one characterised enzyme to date, a D-galactosyl-B1-4-L-rhamnose phosphorylase (Nakajima *et al*., 2010). Thus this genome provided a great resource of uncharacterised enzymes from numerous CAZy families including GH8, GH10, GH11, as well as GH43 in which 21 of the enzymes encoded by this bacterium belonged to this family (Figure 3.1). However, as the prerequisite was always to find and explore novel specificity, enzymes produced during this research were determined for activity no matter which family the enzyme originated from and this consequently lead to the finding of a novel GH8 xylanase from *Opitutus terrae* PB90-1 which is further discussed in Chapter 4. The characterisation of individual GH43 glycoside hydrolases from *O. terrae* PB90-1 was attempted to some degree and results are reported in Chapter 6 but once novel specificity was reported in the one single GH8 enzyme, it was important to focus on this enzyme.

This chapter will provide results and a discussion on the cloning and expression of numerous uncharacterised carbohydrate-active enzymes encoding genes from various bacteria, with a main focus on *Opitutus terrae* PB90-1. Putative activities of the enzymes, based on their primary sequence alignments and BLAST sequences, were provided by a range of bioinformatics resources. Due to the significant number of genes cloned in this research, only example gel pictures are shown in this chapter.

Glycoside Hydrolases Families of *O. terrae*

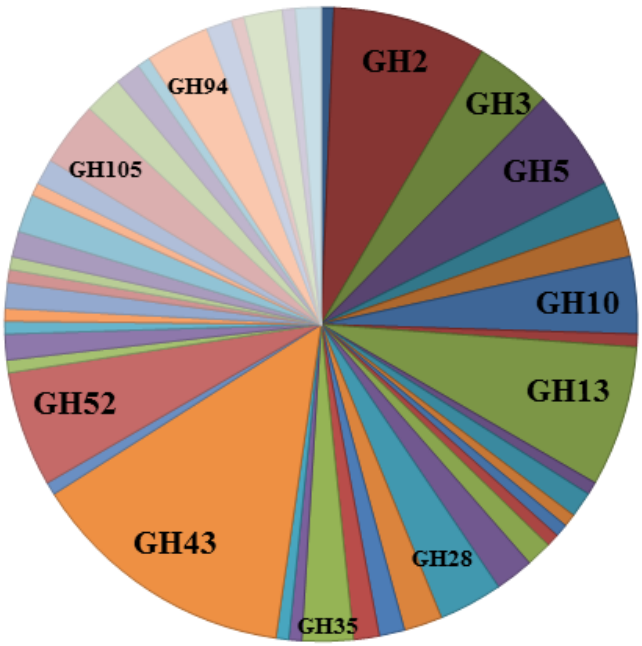


Figure 3.1 Distribution of GH families within *O. terrae*. The larger families are highlighted.

3.2 Bioinformatics analysis

In an attempt to identify and characterise novel lignocellulose degrading enzymes, putative hemicellulases and cellulases from a selected group of available bacteria were chosen from the CAZy database. Primary amino acid sequences and nucleotide sequences were obtained from the NCBI database and analysed using the bioinformatics programs detailed in section 2.1.1.

Analysis of the primary amino acid sequences using ExPASy and UniProt provided important information regarding protein parameters, such as molecular weight and molar absorptivity coefficient including the N-terminal hexahistidine tag and putative functionality information of selected proteins. ClustalW alignment program was used to determine primary sequence alignments of functionally related proteins and to illustrate whether they displayed similarity.

The primary amino acid sequences were checked for the presence of signal peptides using SignalP 3.0 server so that the nucleotide sequences encoding putative N-terminal signal peptides could be identified for removal prior to primer design. This was to prevent incorrect folding and/or export of the expressed proteins in *E. coli*. Many of the selected primary sequences showed the presence of signal peptides, which was entirely expected. The presence of a signal peptide is indicative of the protein being targeted for export from the cell (Antelmann *et al.,* 2001). Endo-acting hemicellulases and cellulases are typical extracellular proteins, catalysing the hydrolysis of the polysaccharide main chain into smaller molecules for efficient transport into the cell (Honda *et al*., 2004; Lagaert *et al*., 2010). Knowledge of whether a putative enzyme was intracellular or extracellular helped guide choice of polysaccharide and oligosaccharide substrates for enzyme assays.

The gene fragments to be amplified were subsequently probed for the presence of internal restriction sites which occurred within the cloning / expression vector pET-28a using Webcutter 2.0 sequence analysis server.

Two primers, a forward primer and a reverse primer, were designed for each gene under investigation for the generation of a PCR fragment. Primers were designed using the Web Primer tool of the Saccharomyces Genome Database in which primers were selected such that each primer was a minimum of 18 nucleotides in length.

For all gene fragments, the forward primer was identical to the first 18 - 22 nucleotide bases in the 5'→3' direction of the section of the 'sense' strand encoding the mature protein. Reverse primers were the reverse complementary to the last 18-22 bases of the 'sense' strand. To enable cloning of the PCR products into the pET-28a vector, pET-28a specific restriction sites were added to the 5' end of each primer, forward and reverse. Primers were also designed with an additional non-complementary sequence containing 8 nucleotides upstream of the restriction site in order for restriction enzymes to cleave their recognition sequence efficiently. Two primer pairs are tabulated in Table 3.1 as an example, with the calculated annealing temperatures in Table 3.2. Primers designed for target gene fragments and cloning regions of pET-28a are included in Appendix A.

Gene	Forward Primer	Reverse Primer
CAA99576.1 from *Bacillus subtilis* subsp.168	5'- GATCGATCCATATGTGAACATCAAGCAGTG -3'	5'- GATCGATCCTCGAGTTAAGAATCAGCACGCAGCG -3'
CAH06512.1 from *Bacteroides fragilis* NCTC 9343	5'- GATCGATCCATATGAACACATTCGGGAAGAAA -3'	5'- GATCGATCCTCGAGTTACATTTTAAAAGTATAT -3'

Table 3.1 Primer pairs for PCR amplification and cloning of two GH encoding genes into pET-28a vector. Restriction sites are underlined.

Gene	Primary Annealing Temperature (°C)		Secondary Annealing Temperature (°C)	
	C	O	C	O
CAA99576.1	46	49	58.6	60
CAH06512.1	49	42	56.4	64

Table 3.2 Calculated (C) and optimised (O) primary and secondary annealing temperatures for primer pairs. Temperatures were calculated by subtracting 5°C from the lower T_m value specific for each primer pair. T_m values were calculated from the equation: Tm = 69.3 + (0.41 x % GC) – (650 / primer length).

3.3 PCR Amplification and cloning of gene fragments

Primer pairs permitted PCR resulting in an amplified gene that was inserted into the pET-28a vector so that an N-terminally hexahistidine-tagged protein could be expressed thereby allowing protein purification by IMAC.

3.3.1 PCR amplification of target genes followed by ligation and transformation into *E. coli* TOP10 competent cells

Gene sequences hypothesised to encode putative hemicellulases and cellulases were amplified by PCR from genomic DNA (an example of PCR amplification can be observed in Figure 3.2). It was observed on many occasions that increasing initial primary and consequently secondary annealing temperatures generated purer PCR products due to more specific primer annealing.

All PCR products were gel purified using NZYTech Gelpure kit as described in section 2.2.2.6. PCR products were gel purified to eliminate non-specific amplification products. Once purified, the PCR products were subjected to a restriction digestion with their corresponding restriction endonucleases to cleave the DNA fragments. This left the amplified gene products with just the cohesive ends, compatible with pET-28a. PCR products were then further purified using the 'PCR Clean Up' provided by the NZYTech Gelpure kit to remove any salts from buffers and restriction enzymes that could inhibit ligation, leaving just the purified DNA sample resuspended in Elution Buffer.

Plasmid pET-28a was cleaved with restriction endonucleases corresponding to target PCR fragments in order to linearise the circular DNA template for the ligation of product and prevent the vector from being preferentially transformed into *E. coli* TOP10 competent cells.

The DNA concentration of purified PCR products was determined using the BioRad Quantity One software and quantifiable DNA HyperLadder ^{Tm} 1. PCR products were ligated into pET-28a vector with a maximum 10:1 and minimum 3:1 insert to vector ratio and transformed into *E. coli* TOP10 competent cells which were incubated overnight at 37°C on 2 % (w/v) LB agar plates supplemented with kanamycin antibiotic (section 2.2.7). To ensure *E. coli* TOP10 transformation efficiency, 2 µL of pET28-a vector DNA (0.01 ng / µL) was added to a sample of competent cells and was used as a positive control (>200 colonies indicated good level of competency). Additionally, to check that the competent cells and /or agar plates were not contaminated, a sample of competent cells was simply plated as a negative control.

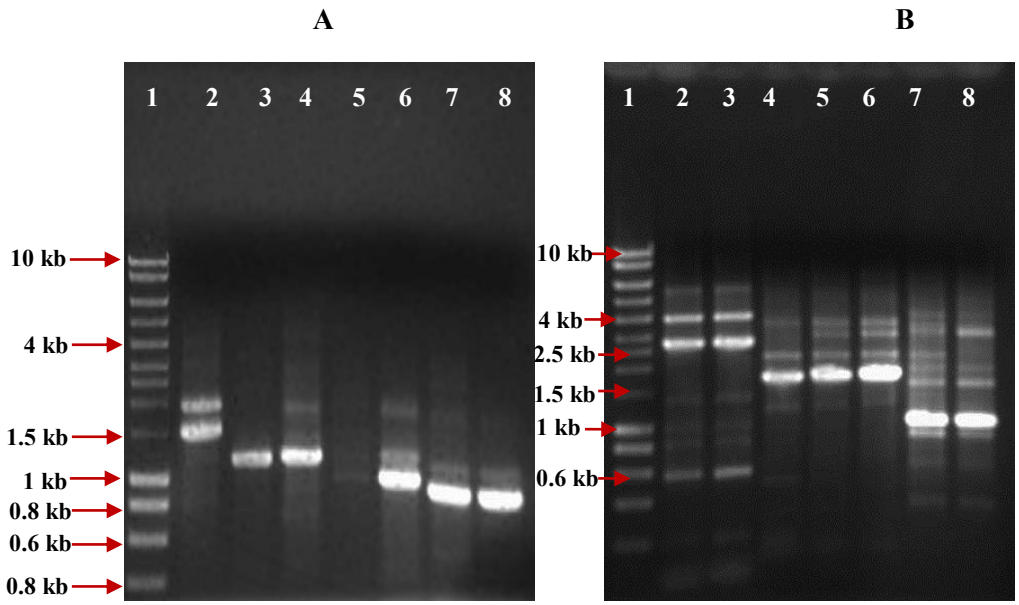


Figure 3.2 1% (w/v) agarose gels of amplified PCR products. A: 1 = Hyperladder 1 size standard; 2 = BAE17270.1; 3 and 4 = BAE17243.1; 5 and 6 = BAE17242.1; 7 and 8 = BAE19338.1. **B:** 1 = Hyperladder 1 size standard; 2 and 3 = CAH07509.1; 4 – 6 = CAH09839.1; 7 – 8 = CAH06920.1.

3.3.2 Confirmation of cloned target gene inserts

As previously mentioned gene fragments were cloned such that recognition sites flanked the upstream and downstream regions of the target sequence. Therefore the target sequence was confirmed by restriction endonuclease digestion of isolated recombinant plasmid DNA (see section 2.2.2.4 for crude extraction of plasmids) from a single clone with the restriction enzymes that cleave the recognition sites that flank the upstream and downstream regions of the target sequence. Figure 3.3 shows an agarose gel image of digested plasmids harbouring the target gene insert. Furthermore, to confirm the insert was in fact the target gene, a diagnostic digestion was carried out that cleaved and plasmid restriction sites and/or internal restriction sites present within the gene sequence as determined by Webcutter 2.0.

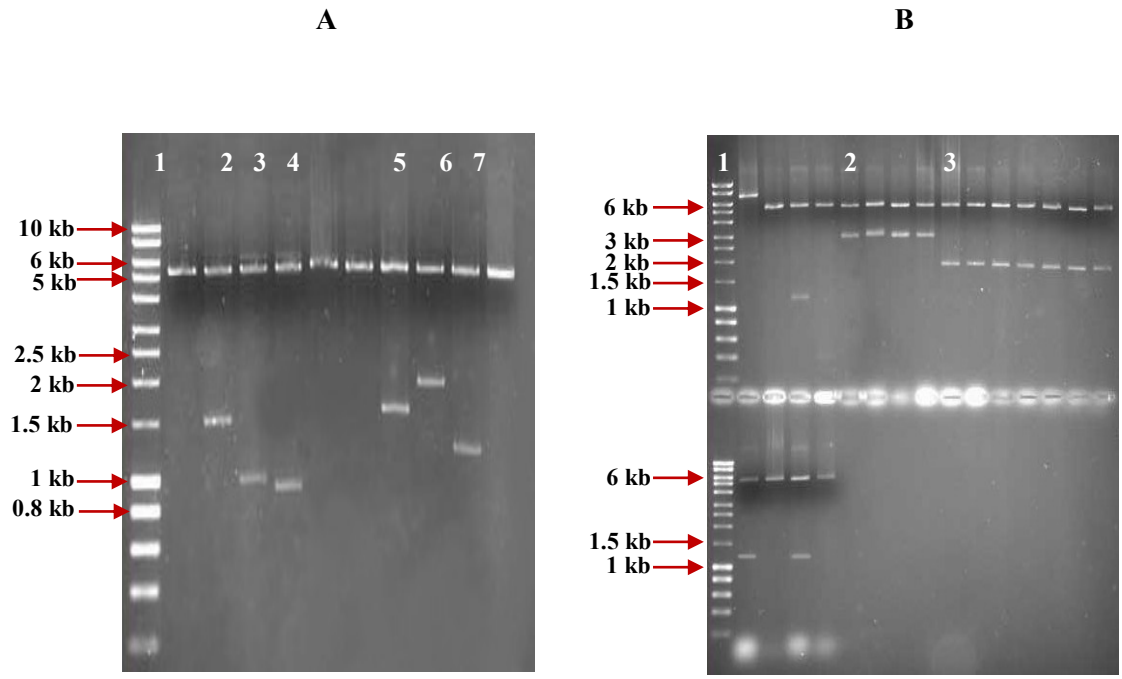


Figure 3.3 1% (w/v) agarose gels of digested plasmid DNA. A: 1 = Hyperladder 1 size standard; 2 = CAE27177.1 digested with *Nde*I and *Xho*I; 3 =BAC73310.1 digested with *Nde*I and *Xho*I; 4 = BAE17242.1 digested with *Nde*I and *Xho*I; 5 = CAD55266.1 digested with *Nde*I and *Xho*I; 6 = CAB56687.1 digested with *Nde*I and *Xho*I; 7 = ABJ63538.1 digested with *Nde*I and *Hin*dIII. **B:** 1 = Hyperladder 1 size standard; 2 = CAB53318.1 digested with *Nde*I and *Hin*dIII; 3 = ABJ65333.1 digested with *Nde*I and *Xho*I; 4 = ABJ64815.1 digested with *Nde*I and *Xho*I. All amplified genes were ligated with pET-28a.

3.4 Protein expression and purification

For the expression of target proteins, plasmid DNA was transformed into chemically competent *E. coli* BL21 (DE3) cells. In order to determine optimum growth conditions for maximum protein expression, *E. coli* cultures were grown on a small scale (see section 2.2.1.7) at temperatures of 16°C and 23°C post induction as a lower temperatures usually results in a change in the rate of protein synthesis which can increase soluble expression (Jones *et al*., 1987). For most cases, the optimum temperature for protein expression was found to be 23°C. Cultures were also grown in duplicate 'with' and 'without' the addition of IPTG for induced expression during the mid log phase.

Cultures were processed as described in section 2.2.8.1 and soluble / insoluble protein expression was investigated by SDS-PAGE (section 2.2.9.2). Once soluble recombinant protein was confirmed, protein expression was carried out at a much larger scale (see section 2.2.1.8) for subsequent purification of the protein.

Cell free extract (CFE) (section 2.2.8.2) was prepared from an overnight *E. coli* culture harbouring the recombinant plasmid. N-terminally tagged proteins were purified from the CFEs using gradient elution IMAC with Ni sepharose resin. A simple gradient elution technique, as detailed in section 2.2.10.1, was used in which a relatively low concentration (10 mM) was applied to the CFE loaded mini column to eliminate *E. coli* native proteins while the N-terminal hexahistidine tag engineered on the recombinant protein chelates to the Ni^{2+} ions at these conditions. The protein was eventually eluted by introducing 500 mM imidazole and this method proved highly efficient in protein purification for subsequent enzymatic assays. Purified proteins were visualised by SDS-PAGE as shown in Figures 3.4, 3.5 and 3.6. Once soluble protein was produced and processed, purified proteins were 'cut' with ammonium sulphate for storage at 4°C. It has been suggested that most precipitated proteins in ammonium sulphate solution retain their activity and native conformation. Furthermore, precipitated proteins could be re-dissolved easily at various concentrations for

use in enzymatic assays. Consequently before each assay, a sample of protein solution was aliquotted into a microcentrifuge tube and pelleted. The protein pellet was then resuspended in BSA solution (1 mg/mL) before the addition to individual assays.

Table 3.3 summarises the cloning and expression of target carbohydrate-active enzymes within this research. The following chapters (4 to 6) focus upon enzymes that ultimately exhibited detectable activity against natural substrates that could be subsequently monitored by HPAEC-PAD. It should be clarified that since there was a large number of genes targeted within this research, inevitably many were not successfully expressed in BL21 cells. Therefore, due to the time scale of the project, after two unsuccessful optimised attempts at expression, further work on the individual target was terminated. This gave the opportunity to analyse enzymes showing sufficient activity for a more in depth characterisation.

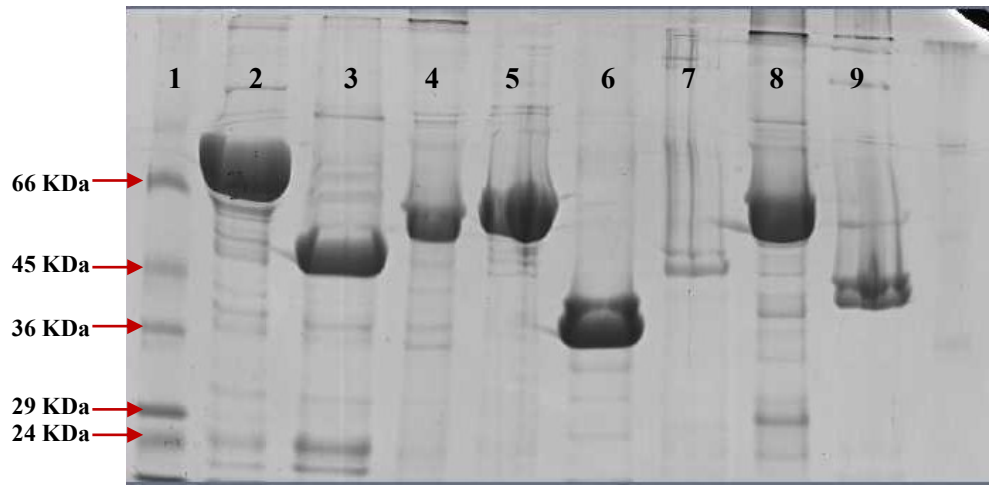


Figure 3.4 12% (w/v) SDS-PAGE of IMAC purified proteins. 1 = Low molecular weight marker (SIG); 2 = CAH06110.1 (83 kDa); 3 = ABC92395.1 (51 kDa); 4 = AAC73374.1 (60 kDa); 5 = LbXyl43A (63 kDa); 6 = LbAraf43A (37 kDa); 7= CAB13699.1 (54 kDa); 8 = AAB41091.1 (61 kDa); 9 = OtXyn8A (47 kDa).

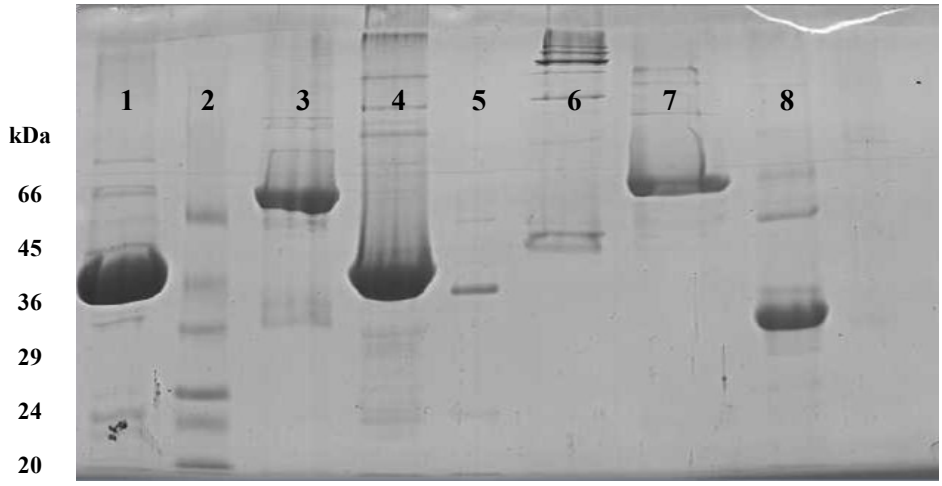


Figure 3.5 12% (w/v) SDS-PAGE of IMAC purified proteins. 1 = ACB76979.1 (47 kDa); 2 = Low molecular weight marker (SIG); 3 = ACB74536.1 (74 kDa); 4 = OtAraf43B (42 kDa); 5 = ACB76389.1 (38 kDa); 6 = OtXyl43A (61 kDa); 7 = OtXyl52A (81 kDa) = OtAraf43A (39 kDa).

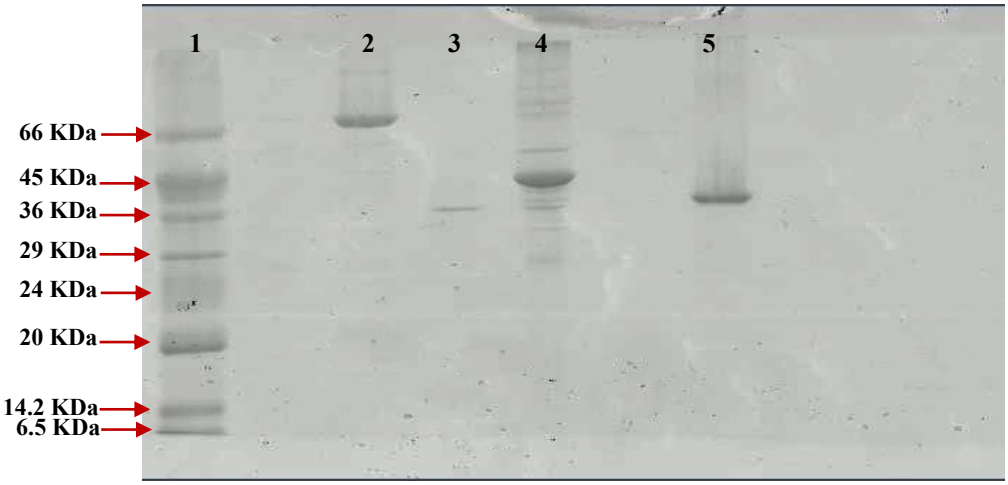


Figure 3.6 12% (w/v) SDS-PAGE of IMAC purified proteins. 1 = Low molecular weight marker (SIG); 2 = OtGlcA67A (82 kDa); 3 = ACB74675.1 (38 kDa); 4 = ACB73665.1 (36 kDa); 5 = ACB76487.1 (37 kDa).

CAZy Family	Number of Gene Targets amplified	Number of Genes Cloned	Number of Proteins Expressed (soluble)	Activity detected on Synthetic Substrates	Activity on Natural Substrates
GH1	20	12	8	8	8
GH3	10	4	1	1	1
GH5	1	1	1	-	-
GH8	2	2	1	0	1
GH10	7	4	4	4	3
GH11	1	1	0	-	-
GH26	1	1	0	-	-
GH39	3	2	2	0	2
GH43	29	24	17	12	12
GH51	8	3	2	0	2
GH52	1	1	1	1	1
GH53	2	2	2	-	-
GH54	1	0	-	-	-
GH62	1	0	-	-	-
GH67	1	1	1	-	-
GH78	2	2	1	-	-
Σ	90	60	36	25	28

Table 3.3 A summary of the genes cloned/expressed within this research. Genes were cloned/ expressed in the respective regions of pET-28a vector and subsequently expressed in *E. coli* BL21 cells. Once solubility was achieved, proteins were purified by IMAC and subsequently determined for activity.

4 Results and Discussion

The Characterisation of a Novel GH8 Xylanase from *Opitutus terrae* PB90-1

4.1 Introduction

The first xylanase to be isolated belonging to glycoside hydrolase CAZy family 8 (GH8) was that of XYLY from the thermophilic bacterium *Bacillus* sp. KK-1 reported by Yoon *et al*., (1998). A thorough hydrolytic analysis was not carried out but the data suggested activity upon various xylan substrates similar to that of typical endo-xylanases from families 10 and 11. Since this first publication, a further three GH8 endo-xylanases (EC 3.2.1.8) and two GH8 exo-oligoxylanases (EC 3.2.1.156) have been isolated to date. Three-dimensional structures have been solved for only two of the characterised xylanases from this small group of GH8 glycoside hydrolases which have shown subtle differences in their active sites resulting in diverse substrate specificities. Therefore, a greater understanding of the binding capacity of such enzymes and their substrate interactions remains to be established. This chapter provides an account of the identification and characterisation of OtXyn8A, a GH8 xylanase from *Opitutus terrae* PB90-1 that possesses unique substrate specificity to that of other GH8 xylanase members. The functional and biochemical properties of OtXyn8A were assessed using synthetic and natural substrates with hydrolysis products routinely analysed by HPAEC-PAD.

In keeping with the nomenclature proposed for plant cell wall degrading enzymes (Henrissat *et al*., 1998), the newly identified GH8 xylanase from *Opitutus terrae* has been denoted with the genus (capital O), the species (lower case t) and the activity/ GH family prefix (Xyn8). The letter ‘A’ has also been given indicating the first GH8 xylanase reported for *O. terrae*.

In an effort to establish if OtXyn8A possessed novel specificity, an approach was taken in which three of the previously characterised GH8 xylanases, Xyn8, REX and REXA, were

also produced within this research and substrate specificities were probed and compared side by side to that of OtXyn8A. The substrate specificities of the tested characterised GH8 xylanases are very different despite their identical protein folds. Xyn8 encoded by a gene from an environmental genomic DNA library is an endo-xylanase that selectively releases xylotriose from its substrates. REX and REXA from *Bacillus halodurans* and *Bifidobacterium adolescentis* respectively, are GH8 exo-oligoxylanases releasing xylose from the reducing end of xylooligosaccharides with a degree of polymerisation greater than 3. In conjunction with data published previously, the results accumulated from this study were used to evaluate diversity of substrate specificity within family 8 enzymes. Furthermore, a section of the work carried out was aimed at identifying related properties between GH8 xylanases and the comparatively greater studied xylanases belonging to GH10 and GH11.

4.1.1 Background to family 8 glycoside hydrolases

The GH8 family consists of various endo- acting glycoside hydrolases including endoglucanases (EC 3.2.1.4), chitosanases (EC 3.2.1.132) and licheninanases (EC 3.2.1.73). The three dimensional structure of six GH8 enzymes, have been determined revealing a $(\alpha/\alpha)_6$ double barrel architecture formed by six repeating helix-loop-helix motifs (Alzari *et al*., 1996; Petegem *et al*., 2003; Adachi *et al*., 2004; Fushinobu *et al*., 2005; Yasutake *et al*., 2006; Mazur *et al*., 2011). This catalytic fold topology is common to all GH8 and GH48 members, grouping both families into the CAZy clan GH-M. Although the $(\alpha/\alpha)_6$ fold is typical of GH8 glycoside hydrolases, differences between structures of individual enzymes occur resulting in various degrees of barrel distortion and numbers of structural helices and β-strands. The structure of the GH8 endo-xylanase from *Pseudoalteromonas haloplanktis*, PhXyl, consists of 13 α-helices forming the barrel and 13 β-strands forming irregular sheets that reside as the cleft (Petegem *et al.,* 2003). In comparison, the GH8 endoglucanase CelA

from *Clostridium thermocellum*, has only 12 α-helices with its barrel far less distorted than that of PhXyl forming quite a circular cross-section (Alzari *et al*., 1996). Figure 4.1 demonstrates an example of the typical $(\alpha/\alpha)_6$ double barrel architecture displayed by GH8 enzymes.

Figure 4.1 Overview of the structure of *C. thermocellum* endoglucanase. A top view of the $(\alpha/\alpha)_6$ barrel fold of *C. thermocellum* CelA (Alzari *et al*., 1996).

GH8 enzymes hydrolyse their corresponding substrates via a single displacement reaction aided by both a general base and a general acid, resulting in anomeric inversion of their yielded products (see section 1.8.2). It has been proposed that the GH8 family can be further divided into at least three sub-families (GH8a, GH8b and GH8c) depending upon the position of the general catalytic base residue (Adachi *et al*., 2004). The crystal structures of PhXyl and CelA, identified that the general catalytic base was an aspartate (Asp) residue

while a glutamate (Glu) residue acted as the general acid (Petegem *et al*., 2003; Alzari *et al*., 1996). The Asp residue of both was located at the N-terminus of helix α_8 of the $(\alpha/\alpha)_6$ and this conservation of the catalytic base has categorised the individual enzymes into sub-family GH8a. Alternatively, the GH8 chitosanase, ChoK, has been grouped into sub-family GH8b in which an Asn319 residue occupies the position of the Asp residue that would naturally be present in GH8a members such as CelA. Furthermore, it was revealed that a glutamic acid (Glu309) served as the general base of ChoK which could be located in the long loop inserted between α_7 and α_8 of the catalytic fold (Adachi *et al*., 2004). Lastly, the third sub-family, GH8c, has been suggested to be similar to GH8b in that the Asp residue is inactivated (Fushinobu *et al*., 2005) however no examples are available for this sub-family.

Compared to the GH10 and GH11 endo-xylanases, only a limited amount of data is available on the substrate specificity of the family 8 xylanase. The prospect of any newly identified GH8 xylanase would therefore pose valuable insights to this recently established group of enzymes and their hydrolytic capability. Additionally, there is great interest in discovering xylanases for various industrial processes in which GH10 and GH11 xylanases are routinely used (Lee *et al*., 2006)and the GH8 xylanases may offer an alternative for various biotechnological applications.

4.2 Cloning, expression and purification of OtXyn8A, REX, REXA and Xyn8

This section describes the cloning and expression of OtXyn8A, REX, REXA and Xyn8. It is important to clarify that identical cloning methods using the pET-28a vector system were utilised for all GH8 target enzymes. Proteins were also expressed and purified via the same conditions, therefore a rationally-based comparative analysis could be carried out.

4.2.1 Bioinformatics of GH8 xylanases

The four target GH8 sequences were analysed using bioinformatics described in section 2.1.1. On the basis of sequence similarity, an attempt to identify the putative function of OtXyn8A was analysed using a BLAST search. BLAST analysis revealed that OtXyn8A shared the highest sequence homology of 61% to a GH8 glycoside hydrolase from *Anaerolinea thermophilia*. No putative function was annotated for this enzyme therefore making it difficult to predict any related function for OtXyn8A. Sequence identity between OtXyn8A and other GH8 members ranged from 46-55 %, therefore showing quite low identities. This is discussed further in section 4.4.4.

The amino acid sequence obtained from NCBI for the GH8 xylanases under investigation were checked for the presence of an N-terminal putative signal peptide. The SignalP analysis of OtXyn8A revealed that there was a signal peptide present, therefore nucleotides encoding the signal peptide were not amplified. This result was not unexpected as OtXyn8A was annotated as a putative endo-xylanase and such enzymes usually reside as extracellular enzymes (Tsujibo *et al*., 2004). This however, contradicted the sequence information of REX, REXA and Xyn8 as no signal peptide was predicted for the individual sequences.

The deduced nucleotide sequence of OtXyn8A was then probed for restriction sites and Webcutter analysis confirmed that sites of the restriction endo-nucleases *Bam*HI and *Hin*dIII

could be added to the gene sequence for appropriate cloning with pET-28a vector. To keep cloning conditions similar, sequences of REX, REXA and Xyn8 were also checked for the sites corresponding to the restriction enzymes *Bam*HI and *Hin*dIII.

4.2.2 PCR Amplification and cloning of GH8 gene fragments

Gene fragments encoding the target GH8 xylanases were amplified by PCR, except for Xyn8, which was synthesised (GenScript Inc.) with the appropriate restriction sites ready for ligation into pET-28a, as the gene sequence was attained from an uncultured bacterium. For OtXyn8A, REX and REXA, two primers were designed for each sequence in which the forward primers were engineered with a 5' *Bam*HI restriction site and reverse primers were engineered with *Hin*dIII restriction site at the 5' ends, as shown in Table 4.1. The calculated and optimised primary and secondary annealing temperatures for PCR are summarised in Table 4.2.

PCR fragments were run on an agarose gel with DNA size standard. The target gene fragment was expected to be located at approximately 1175 bp for OtXyn8A, 1167 bp for REX and 1140 bp for REXA. Figure 4.2 illustrates successful PCR amplification of the target gene sequences. Amplified DNA was then digested with the *Bam*HI and *Hin*dIII restriction endo-nucleases and ligated into similarly restricted pET-28a vector generating plasmids encoding the protein of interest, including a hexahistidine tag at the N-terminus. Cloned fragments were confirmed by restriction endo-nuclease digestion of isolated plasmid DNA from a single TOP10 clone using *Bam*HI and *Hin*dIII (Fig 4.3).

Gene	Forward Primer	Reverse Primer
OtXyn8A	5’- TGCCATAGGGATCCATGCAAGGCGCCGCTGAAA -3’	5’- TGCCATAGAAGCTTTCATTTCGCAACGGGACC -3’
REX	5’- TGCCATAG GGATCC ATGAAGAAAACGACAGAAGGTGC -3’	5’- TGCCATAG AAGCTT CTAGTGTTCCTCTTCTTGGCCCT -3’
REXA	5’- TGCCATAG GGATCC ATGACAAATGCAACCGATACC -3’	5’- TGCCATAG AAGCTT TCATTCATACCGATACTTTCCGC -3’

Table 4.1 Primer pairs for PCR amplification. Engineered restriction sites are shown underlined.

Gene	Primary Annealing Temperature (°C)		Secondary Annealing Temperature (°C)	
	C	O	C	O
Xyn8A	51	50.9	63.3	68
REX	51	56.6	66	71.6
REXA	48	53.6	64.5	70.1

Table 4.2 Calculated (C) and optimised (O) primary and secondary annealing temperatures for primer pairs. Temperatures were calculated by subtracting 5°C from the lowest T_m value for each primer pair. T_m values were calculated with the equation: Tm = 69.3 + (0.41 x % GC) – (650 / primer length). The corresponding results are shown in Figure 4.2.

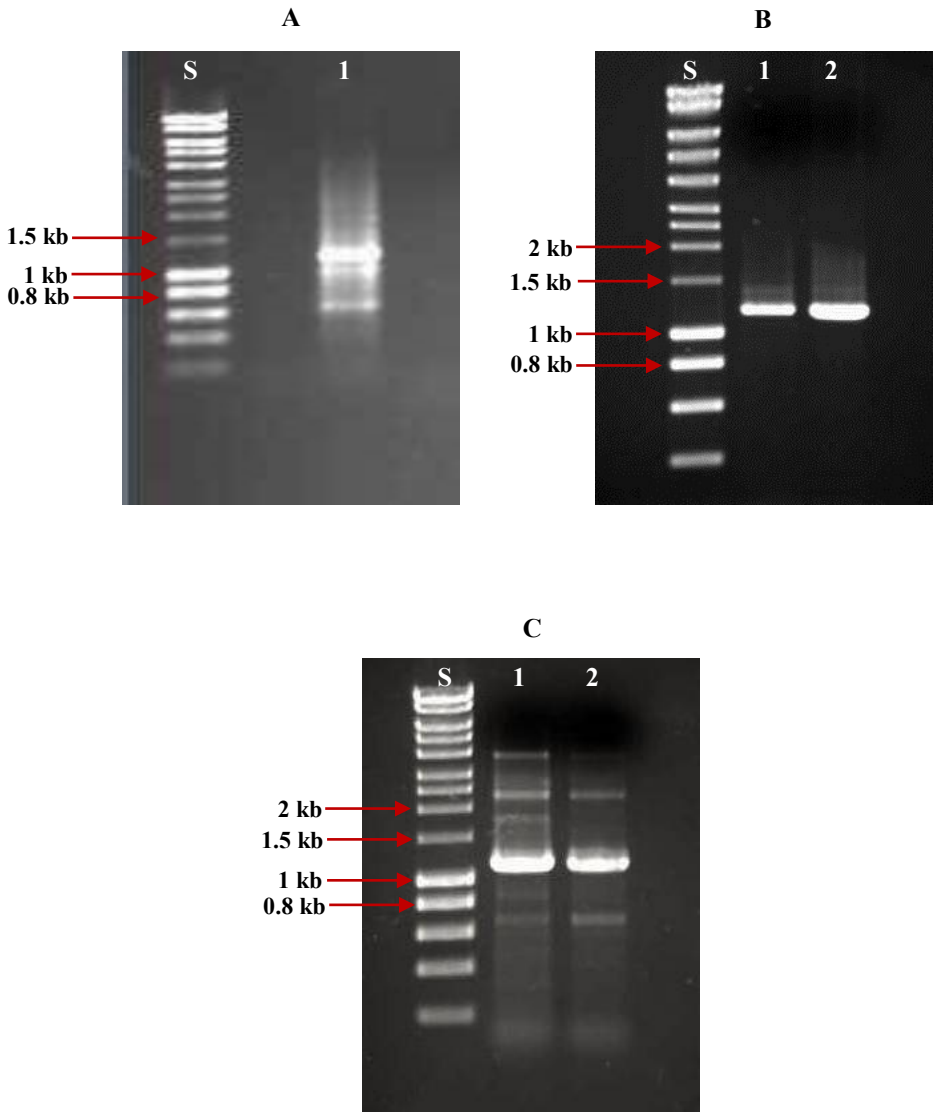


Figure 4.2 1 % (w / v) Agarose gel showing gene fragments following PCR on agarose gel. Annealing temperature sets used are listed in Table 4.2. **A:** Lane 1: PCR product for OtXyn8A (expected size 1175 bp) at optimised annealing temperature. **B:** Lanes 1 and 2: PCR product for REX (expected size 1167 bp) at annealing temperatures at calculated and optimised values, respectively. **C:** Lanes 1 and 2: PCR product for REXA (expected size 1140 bp) at annealing temperatures at calculated and optimised values, respectively. S = DNA size standard ranging between 200 bp and 10,000 bp.

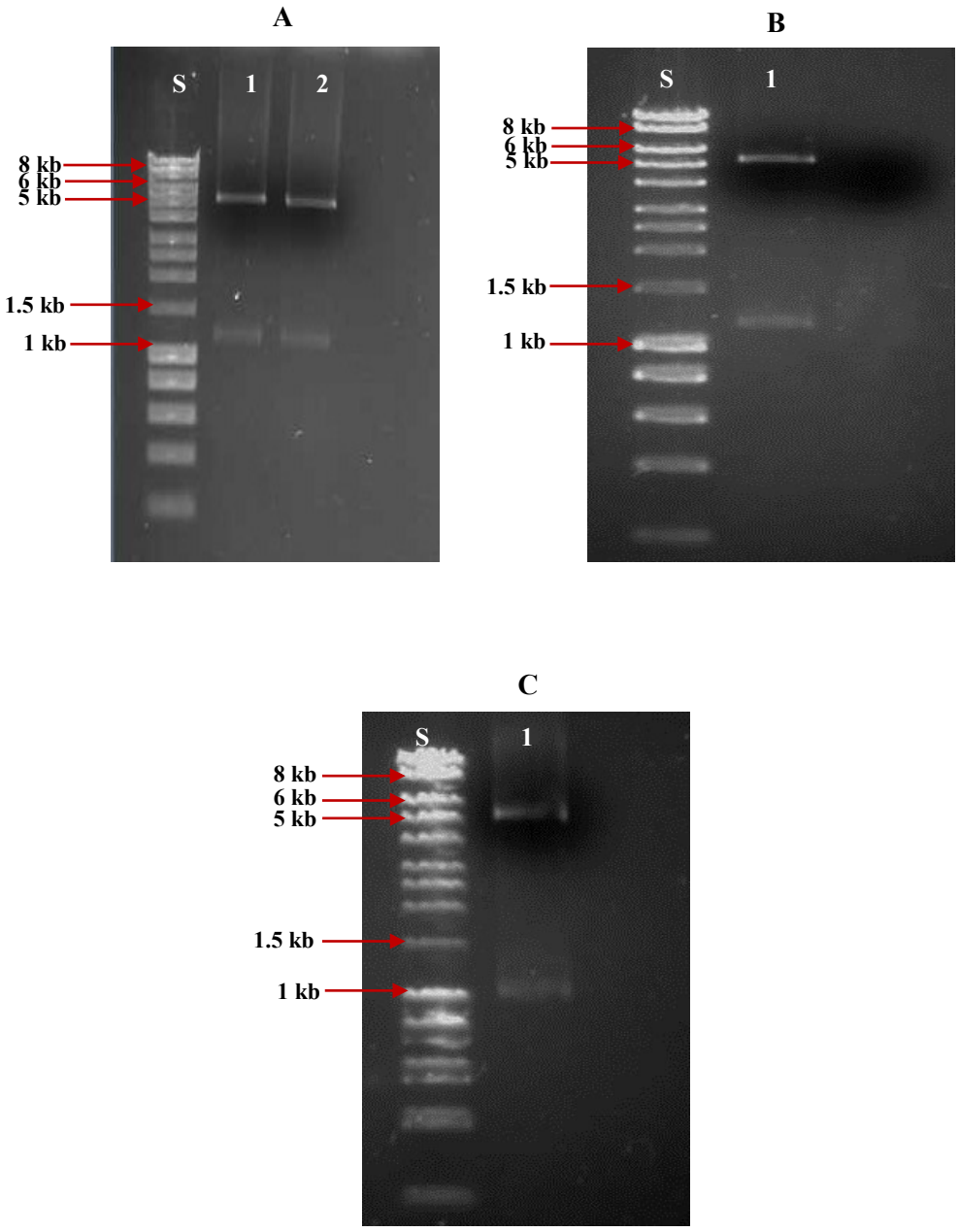


Figure 4.3 1 % (w / v) Agarose gel image showing gene inserts following restriction digest of plasmid DNA. A: Lanes 1 and 2: OtXyn8A fragment (expected size 1175 bp). **B:** Lane 1: REX fragment (expected size 1167 bp). **C:** Lane 1: REXA fragment (expected sixe 1140 bp). S = DNA size standard ranging from 200 bp to 10,000 bp.

4.2.3 Expression and purification

The recombinant plasmids of OtXyn8A, REX, REXA and Xyn8 were chemically transformed into *E. coli* BL21 (DE3) competent cells for protein expression and the cultures were grown at 37°C to mid-exponential (log) phase. Expression was induced by the addition of IPTG and the cultures were incubated overnight at 23°C. Trial expressions showed that the highest level of soluble expression of the GH8 proteins was achieved by IPTG induction and at a downshift of temperature to 23°C. Cells were harvested and cell free extracts (CFE) were prepared as described in section 2.2.8.2. The N-terminally histidine tagged proteins were purified from the CFE by IMAC using a chelating sepharose mini column and a gradient technique (see section 2.2.10.1). All proteins were eluted from the column in 500 mM imidazole and monitored by Bradford's reagent. The proteins were then checked for purity by SDS-PAGE (Figure 4.4) and the concentration was determined (see section 2.2.10.4). All proteins migrated as a single band in SDS-PAGE and molecular masses were in agreement with the theoretical masses of 47.8 kDa, 51.7 kDa, 48.6 kDa and 47.3 kDa for OtXyn8A, Xyn8, REX and REXA respectively. Purified proteins were precipitated with ammonium sulphate and stored at 4°C for further use.

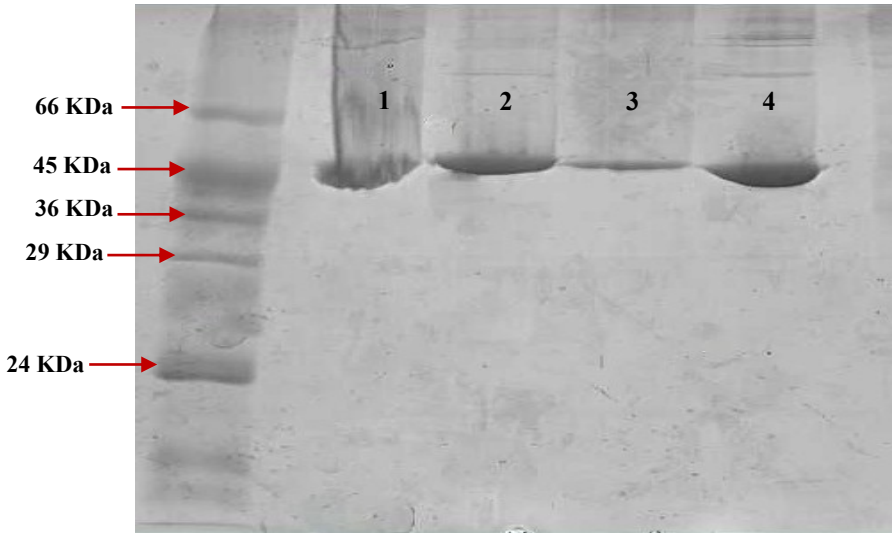


Figure 4.4 12% (w / v) SDS-PAGE analysis of IMAC purified GH8 proteins. Lane 1: Purified protein band for OtXyn8A (predicted MW 47756.6 Da). Lane 2: Purified protein band for Xyn8 (predicted MW 51731.7 Da). Lane 3: Purified protein band for REX (predicted MW 48553.8 Da). Lane 4: Purified protein band for REXA (predicted MW 47329.3 Da). LMW = Low molecular weight marker (SIG) (m.w. 20, 24, 29, 36, 45 and 66 kDa).

4.3 Characterisation of OtXyn8A

4.3.1 Functional characterisation

To investigate the activity of the N-terminally tagged OtXyn8A, the enzyme was tested against the synthetic substrates 4-nitrophenyl-β-D-xylopyranoside and 4-nitrophenyl-α-L-arabinofuranoside to indicate any xylosidase and/ or arabinofuranosidase activities, respectively. No colour change was observed upon incubation of the enzyme with either substrate indicating the chromophore 4-nitrophenol was not liberated by the enzyme. Therefore OtXyn8A was shown to be inactive against these synthetic substrates. The enzyme was then assayed for reducing sugar products with the xylan polysaccharide substrates wheat arabinoxylan and oat spelt xylan (OSX). While OSX is quite a linear polysaccharide, the β-1,4-linked xylopyranose backbone of wheat arabinoxylan is highly decorated with α-1,2-linked and α-1,3-linked arabinofuranose units. DNSA assay revealed that OtXyn8A was active against both polysaccharide substrates without the addition of any endo-acting enzyme generating more free ends. Therefore it was assumed at this initial stage that OtXyn8A was an endo-xylanase. In order to ascertain if this assumption was accurate, the sugar products of such reactions were analysed by HPAEC-PAD. Additionally, no activity was recorded against cellulose substrate, carboxymethyl cellulose (CMC), suggesting that OtXyn8A has no cellulase activity.

4.3.2 Biochemical characterisation

Once primary enzyme activity was confirmed through multiple reducing sugar assays and HPAEC-PAD analysis, a more biochemical characterisation of OtXyn8A could be carried out. Since no activity was detected against the synthetic substrate 4-nitrophenyl-β-D-xylopyranoside, all biochemical characterisation of OtXyn8A had to be resolved by data

obtained from enzyme and OSX reactions via DNSA assay. The effect of pH and temperature on enzyme activity and stability was confirmed with the release of reducing sugar measured over time. Reducing sugar content was determined using a standard DNSA curve based on xylose equivalents (section 2.2.11.2, Figure 2.1).

OtXyn8A has an optimum pH of 6 (Figure 4.5) with optimal activity detected at 45-55°C (Figure 4.6). The enzymatic reaction for investigating the pH optimum was carried out in a range of 0.05 mM sodium phosphate (pH 5.5 – 8.5) and sodium acetate (pH 4.5 - 6) buffers. The optimum temperature of activity was measured for 10 min under optimal pH and standard conditions (2.2.11.2) with temperature ranging from 25°C – 55°C.

Thermostability assays revealed that OtXyn8A retained the highest activity at 25°C with a rapid decrease in activity above 45°C (Figure 4.7). No activity was detected between 55°C and 75°C suggesting that OtXyn8A was completely inactive at such temperatures. The thermostability was determined by incubating OtXyn8A at various temperatures (25°C, 35°C, 45°C, 55°C, 65°C, 75°C) for 20 min in 0.05 mM sodium phosphate buffer (pH 6) prior to the addition to OSX, followed by measuring the activity under standard conditions at 25°C (section 2.2.11.2).

Most xylanases characterised to date are optimally active at temperatures below 50°C and are generally active in acidic or neutral pH (Khasin *et al*., 1993), depending upon the origin of xylanase; as bacterial xylanases have been found to have relatively higher pH optima than that of fungal xylanases (Subramaniyan *et al*., 2002). However, there are a few reports in which xylanases have showed optimal temperature above 90°C (Winterhalter *et al*., 1995) and quite alkalophilic conditions of around pH 9.0 (Gupta *et al*., 2000). These are usually described as thermostable alkaline-tolerant xylanases which can be of great potential in certain applications in industry.

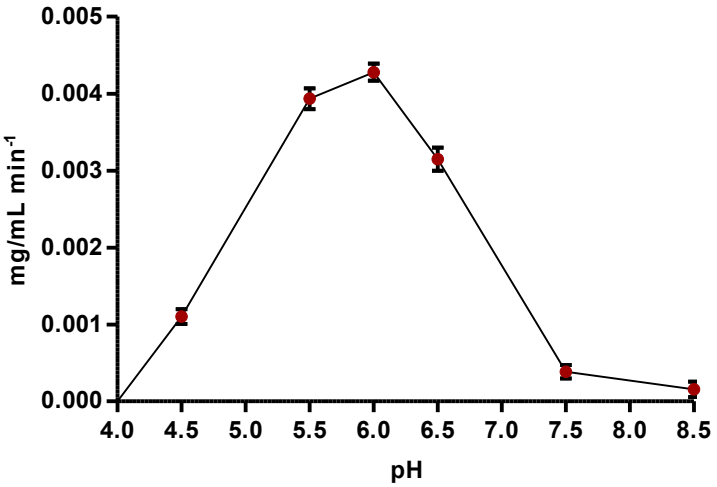


Figure 4.5 The effect of pH on the rate of OtXyn8A (0.2 µM) activity against OSX. The reaction was performed in duplicate, with error bars representing the standard deviation from the mean. Reactions were carried out at a final substrate concentration of 8 mg/mL and BSA (1 mg/mL final concentration). Sodium acetate and phosphate buffers (0.05 M final concentration) were used to cover maximum pH range (pH ranging from 4.5 to 8.5). Reaction mix was prepared and rate of reaction was analysed by DNSA using standard conditions described in section 2.2.11.2.

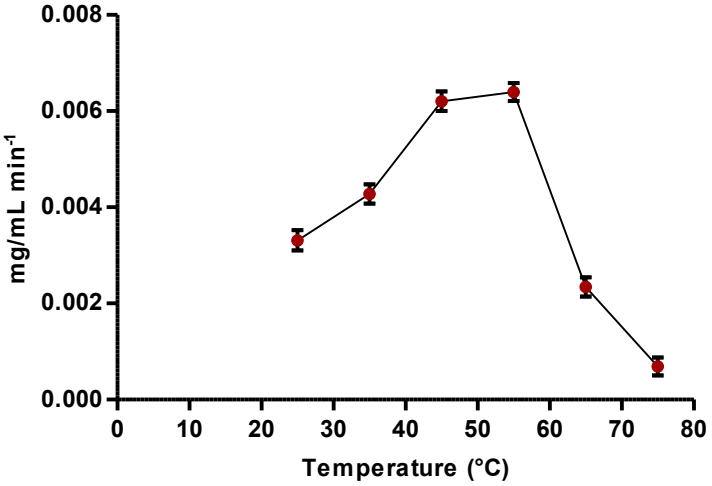


Figure 4.6 The effect of temperature on the rate of OtXyn8A (0.2 µM) activity against OSX. The reaction was performed in duplicate, with error bars representing the standard deviation from the mean. The reactions were performed at 25, 35, 45, 55, 65 and 75°C and pH 6 at a final substrate concentration of 8 mg/mL and BSA (1 mg/mL final concentration). Reaction mix was prepared and rate of reaction was analysed by DNSA using standard conditions described in section 2.2.11.2.

.

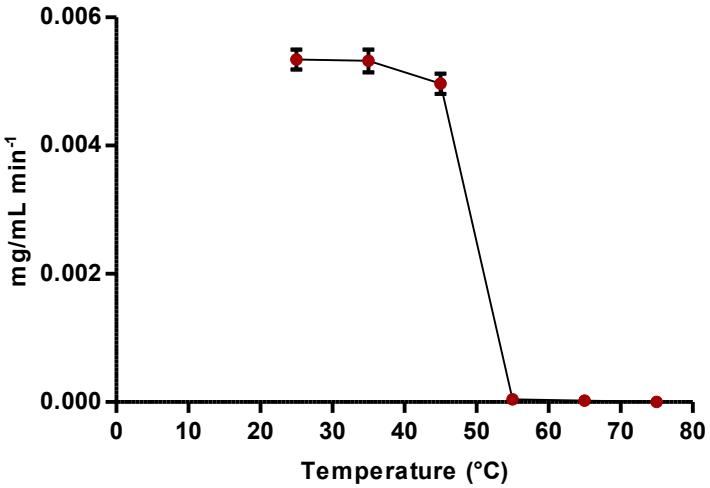


Figure 4.7 The effect of temperature on OtXyn8A (0.2 µM) stability as measured by rate of activity against OSX. The reaction was performed in duplicate, with error bars representing the standard deviation from the mean. Reactions were carried out at 25°C, pH 6 and at a final substrate concentration of 8 mg / mL, after pre-incubation of OtXyn8A in buffer (0.05 M final concentration) and BSA (1 mg/mL final concentration) for 25 min at 25, 35, 45, 55, 65 and 75°C. Rate of reaction was analysed by DNSA using standard conditions described in section 2.2.11.2.

4.3.3 Kinetic analysis

To determine apparent kinetic parameters, OSX was subjected to hydrolysis by OtXyn8A in 0.05 mM sodium acetate buffer (pH 6) at 45°C. The initial rates were measured as the increase in reducing sugar content over time (10 min). For all kinetic analysis, substrate concentrations that were determined empirically within the limiting range spanning the K_M, were chosen and reactions were performed in triplicate (section 2.2.11.5). In order to clarify if data reflected true Michaelis-Menten kinetics, Lineweaver-Burk plots from the three independent data sets were constructed as illustrated in Figure 4.8 and used to determine average values of V_{max}, K_M, k_{cat} and k_{cat} / K_M (Table 4.3) (section 2.2.11.5.1). Kinetic analysis of OtXyn8A was performed using GraphPad Prism version 5.2 and results are tabulated in Table 4.3.

	k_{cat} (min^{-1})	K_M (mg/mL)	k_{cat} / K_M (min/mg/mL)
OtXyn8A	7661 ± 23	6.84 ± 0.66	1120 ± 9

Table 4.3 Michaelis-Menten kinetic parameters for the hydrolysis of OSX by OtXyn8A. The reaction was performed in triplicate (see Figure 4.8) to ascertain average values. Reactions were carried out at 45°C in 0.05M sodium acetate buffer, pH 6 with 1 mg/mL BSA.

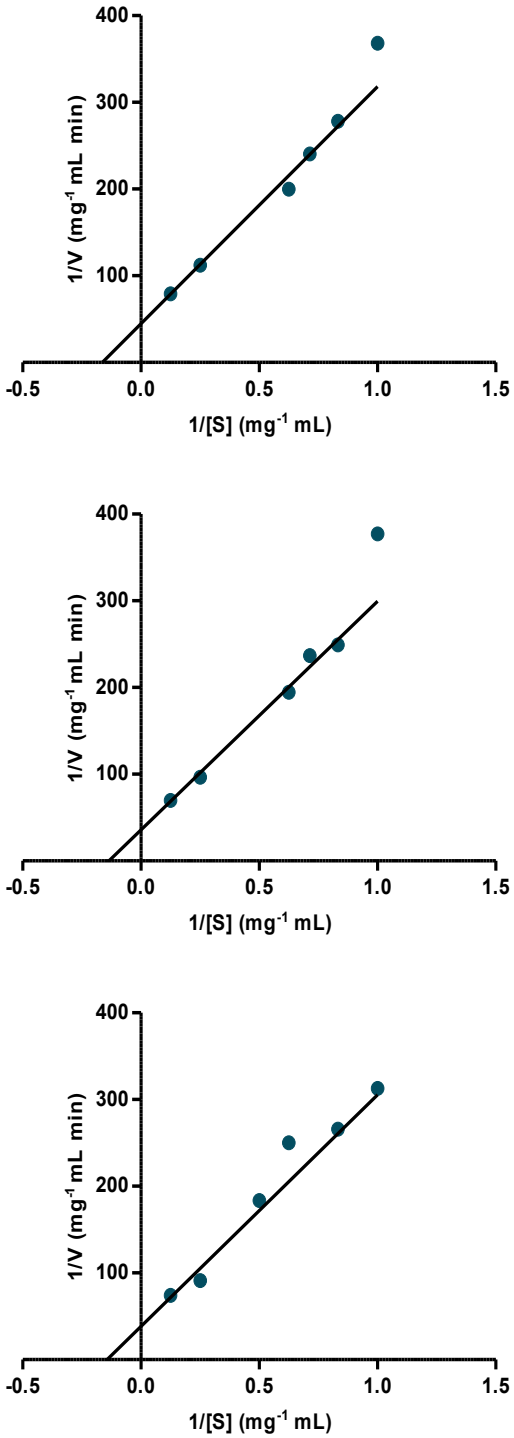


Figure 4.8 Lineweaver-Burk plots for OtXyn8A against OSX. The reaction was performed at 45 °C in 0.05M sodium acetate buffer, pH 6 including BSA at a final concentration of 1 mg/mL. Initial reaction rates were subjected to DNSA (section 2.2.11.2).

4.3.4 Mode of action of OtXyn8A

HPAEC-PAD analysis of the products generated by OSX digestion with OtXyn8A revealed a range of xylooligosaccharides from a degree of polymerisation of xylobiose (X2) and higher. More importantly, the observed amount of X2 released was the main product of initial hydrolysis with more X2 accumulating as the reaction progressed. After prolonged digestion, xylotriose (X3), xylotetraose (X4), xylopentose (X5) and xylohexose (X6) fully degraded giving rise to a large quantity of X2 with a low level of xylose (X1) in comparison (Figure 4.9). Preliminary analysis of the OSX hydrolysis products by HPAEC-PAD was indicative of endo-xylanase activity, as OtXyn8A randomly cleaved the xylan substrate releasing xylooligosaccharides of varying degrees of polymerisation. However, the significant release of X2 throughout the reaction suggested that this was a product of exo-xylanase activity, as X2 residues were continually being removed from the xylooligosaccharide chain ends after initial endo-action on the polymeric substrate.

OtXyn8A liberated X1 and X2 as the major end products from wheat arabinoxylan and 4-*O*-methyl-D-glucuronoxylan, although the enzyme only displayed trace activity in comparison to the activity against OSX. The data displayed in Figure 4.9 revealed that the reduced endo-activity of OtXyn8A on wheat arabinoxylan, reflecting the decrease in β-1,4-glycosidic bonds of the xylan backbone that are accessible for cleavage. This indicates that OtXyn8A is sterically hindered by arabinose substituents on the main xylan chain and the decorated xylooligosaccharide generated from arabinoxylan requires the action of an accessory enzyme, presumably an arabinofuranosidase for the action of OtXyn8A to be more effective. Similarly, a slower rate of depolymerisation was also observed with the single action of OtXyn8A on 4-*O*-methyl-D-glucuronoxylan (Figure 4.9). However the products liberated were X1 and primarily X2, hence similar hydrolytic patterns were observed for xylan degradation.

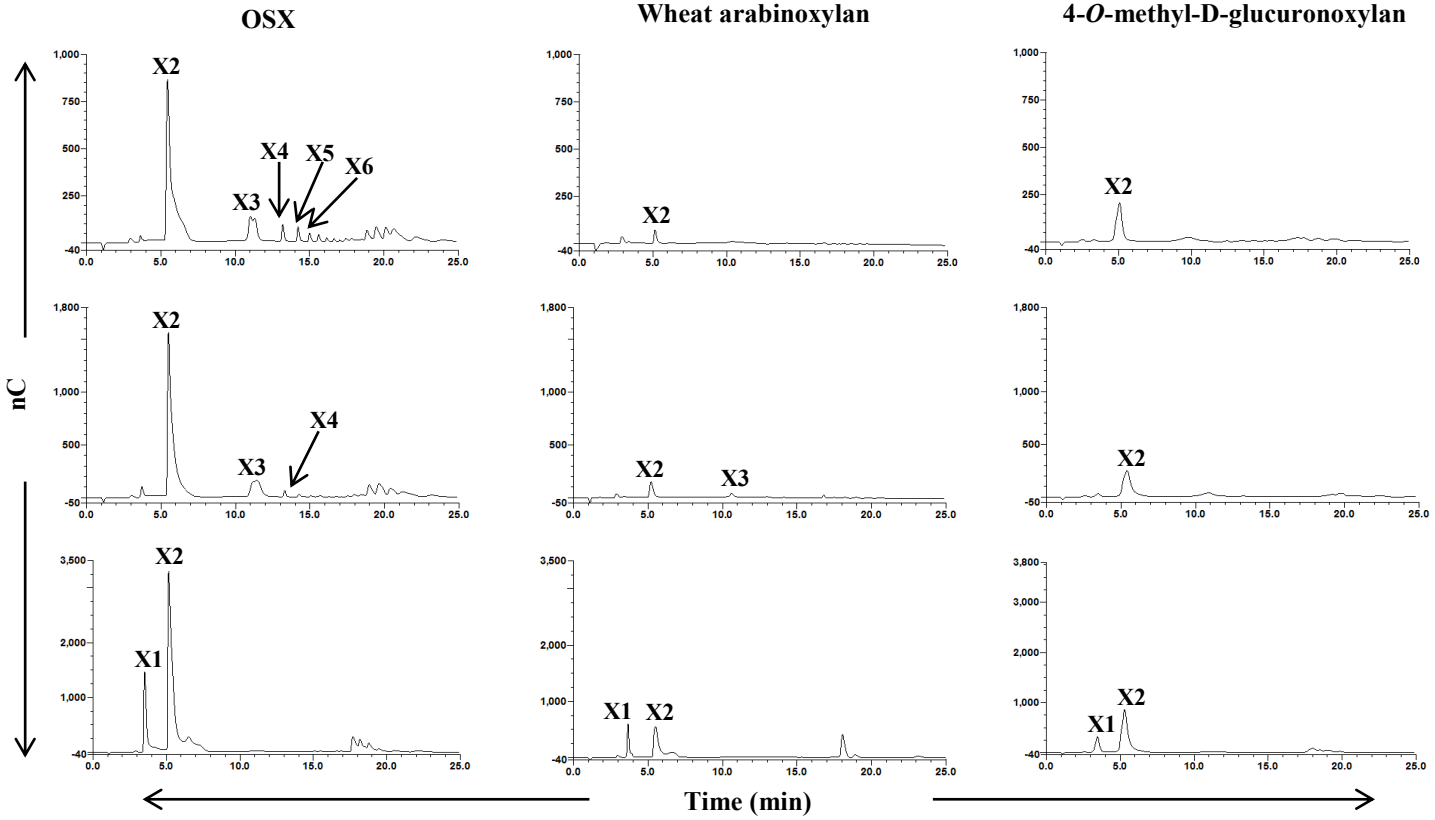


Figure 4.9 HPAEC-PAD analysis of the hydrolysis of OSX, wheat arabinoxylan and 4-*O*-methyl-D-glucuronoxylan. OtXyn8A (0.2 µM) was incubated with OSX, wheat arabinoxylan and 4-*O*-methyl-D-glucuronoxylan at a final substrate concentration of 8 mg/mL. Reactions were carried out in sodium acetate (0.05 M final concentration) pH 6 and BSA (1 mg/mL) at 25°C. At intervals 30 min, 60 min and 1440 min samples were removed and subjected to HPAEC-PAD analysis (see section 2.2.12).

The activity of OtXyn8A on various xylooligosaccharides was also investigated under standard conditions and analysis of the products released showed that X2 was preferentially formed. It was important to observe the inability of the enzyme to cleave X2, with X3 being the smallest xylooligosaccharide it was active upon. However, activity on X3 was negligible in comparison to larger xylooligosaccharides as only small quantities of X1 and X2 accumulated in the very early stages of X3 hydrolysis (Figure 4.10). The activity recorded upon X4 was significant with the production of X2 within the first 3 min of hydrolysis. After an extended hydrolysis of an hour, complete hydrolysis of X4 into X2 was detected (Figure 4.10). The activity on variable xylooligosaccharides gave an insight into the substrate binding cleft of OtXyn8A and it can be assumed at this stage that this enzyme comprises at least four subsites, -2, -1, +1, +2, as the xylanase was readily more active against X4 than X3 (see Figure 4.10). Although, it should be noted that specificity constants could possibly be determined more accurately with xylooligosaccharide standards with a higher degree of polymerisation than 4. At no point were xylooligosaccharides larger than the original substrates detected, possibly demonstrating the absence of a transglycosylation activity by OtXyn8A. These results are consistent with the mechanism utilised by OtXyn8A as it is well established that enzymes possessing a double displacement retaining mechanism have the potential to perform transglycosylation reactions (Sinnott, 1990), unlike inverting glycoside hydrolases such as OtXyn8A.

For further confirmation that OtXyn8A is an endo-acting enzyme, an assay with AZO-OSX was undertaken (see section 2.2.11.4), which is generally used to determine endo-xylanase activity. More of the action of OtXyn8A on AZO-OSX is detailed in section 4.4.1.3 but results revealed that OtXyn8A exhibited an endo-mode of action.

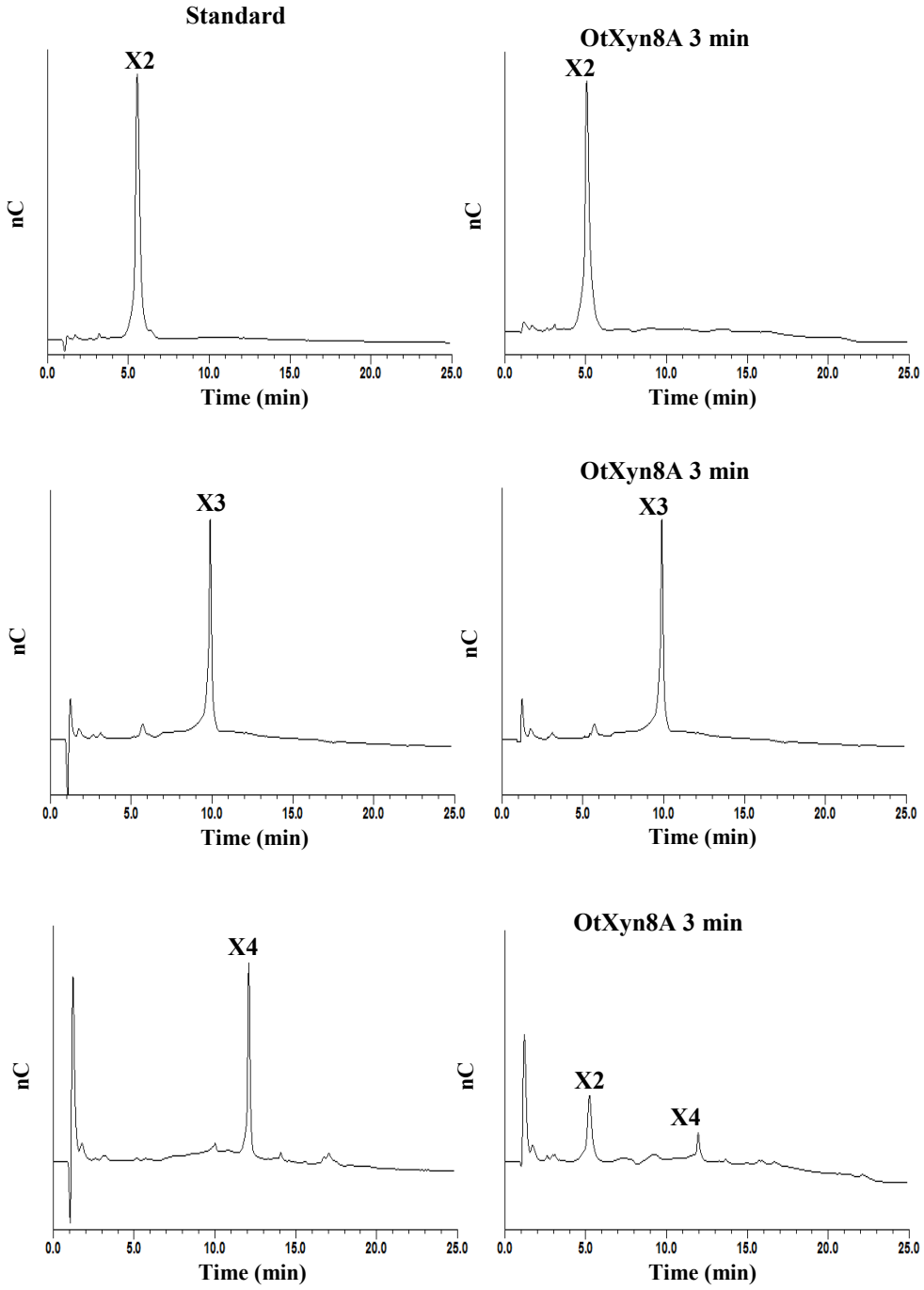


Figure 4.10 Activity of OtXyn8A against xylooligosaccharides. OtXyn8A (0.2 µM) was incubated with X2, X3 and X4. Reactions were carried out in sodium acetate (0.05 M final concentration) pH 6 and BSA (1 mg/mL) at 25°C. At 3 min, a sample was removed and subjected to HPAEC-PAD analysis (see section 2.2.12).

In attempt to resolve the question of whether OtXyn8A also possessed an exo-mode of action on polymeric xylan substrate, the enzyme was tested on pre-treated OSX in which the xylan termini were inaccessible to the action of an exo-acting enzyme. The xylan backbone contains side chain substituents in the form of arabinosyl groups and *O*-acetyl esters which must be removed to create potential xylan access points for endo-xylanases and β-xylosidases. As unsubstituted xylose residues are therefore considered preferential to xylan degrading enzymes, it was hypothesised that the removal of such unsubstituted residues from the reducing and non-reducing free ends of the polymeric substrate would inhibit the exo-action of OtXyn8A and just allow for the cleavage of internal bonds. For this experiment, REX was utilised to cleave the reducing end of the xylan polysaccharide and a β-xylosidase, from *O. terrae* produced as part of this project, was utilised to attack the unsubstituted xylose residues on the non-reducing end, thereby eliminating any free xylose ends and inhibiting an exo- attack on the polysaccharide. Figure 4.11 shows the results of initial hydrolysis by OtXyn8A on OSX and pre-treated OSX. The results indicate an endo-mode of action by OtXyn8A on the pre-treated OSX however there is still the familiar predominant release of X2 which is of a similar amount accumulated with the reaction on non-pre-treated OSX. Therefore it could be speculated that this release of X2 is an exo-product of xylooligosaccharide hydrolysis, in both cases and as OtXyn8A acts in an endo-fashion on the polysaccharide, it creates more free ends available for an exo-mode of attack.

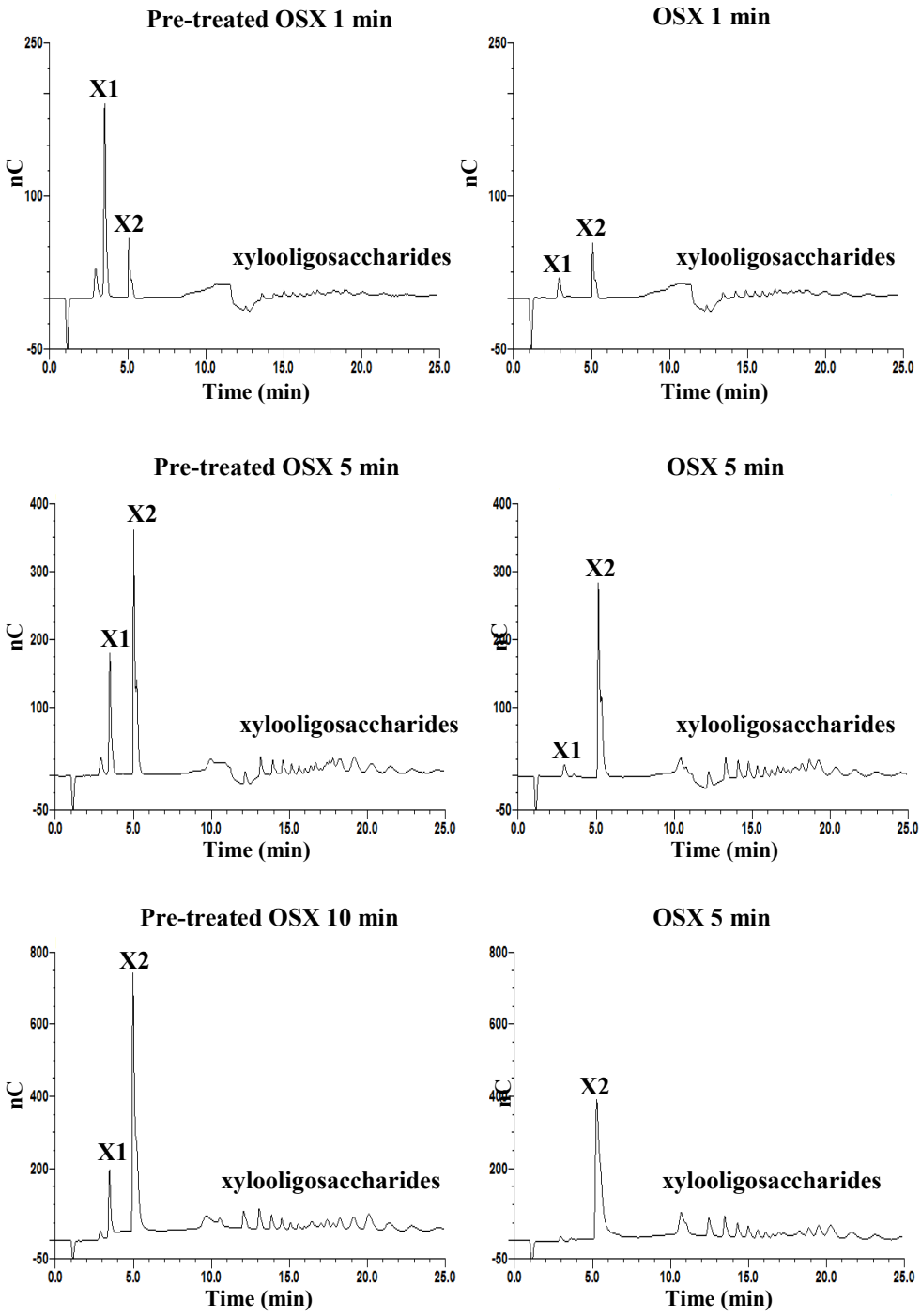


Figure 4.11 OtXyn8A activity against pre-treated OSX for the determination of exo-mode of action. OtXyn8A (0.2 μM) was incubated with pre-treated OSX and OSX. Reactions were carried out in sodium acetate (0.05 M final concentration) pH 6 and BSA (1 mg/mL final concentration) at 25°C. Pre-treated xylan underwent ethanol precipitation for the removal of xylose released by the action of REX and β-xylosidase. At 1 min, 5 min and 10 min samples were removed and subjected to HPAEC-PAD analysis (see section 2.2.12).

The data obtained in this study collectively suggests that OtXyn8A is an endo-xylanase with a certain amount of exo-activity, although, it is quite difficult to make this endo-/exo-activity assumption solely based on the evidence provided. It could be suggested that OtXyn8A is in fact a typical endo-xylanase due to its activity against polymeric xylan in which a mixture of xylooligosaccharides are generated during the initial stages of hydrolysis. A typical endo-xylanase would then progressively degrade the xylooligosaccharides yielding primarily xylose, xylobiose and xylotriose. However, OtXyn8A predominantly and continually produces X2 from the beginning (>1 min) of hydrolysis to the end of the reaction similar to an exo-acting enzyme which is therefore a significant feature. If OtXyn8A did possess exo-activity, it would require an understanding of the active site alongside conclusive analysis of the products released from individual larger xylooligosaccharides. This in turn would allow the determination of what level of polymerisation is required for OtXyn8A to act in an endo-fashion in addition to exo-activity releasing X2.

To identify if OtXyn8A possessed exo-activity as well as an endo-mode of action, it would be essential to determine topology of the enzyme. There are limited reports of glycoside hydrolases exhibiting both endo- and exo- activities but rather their topographies are consistent with either one or the other modes of action. It was reported that CjArb43A from *Cellvibrio japonicus* possessed both an endo- and exo- mode of action (McKie *et al*., 1997) as the arabinanase released arabinotriose as its major product from linear arabinan as well as mixture of arabinooligosaccharides, albeit quantitatively negligible in comparison to arabinotriose. Analysis of the reaction products of CjArb43A was compared to those released by a typical endo-arabinanase which initially generated a mixture of arabinooligosaccharides. However, viscosity measurements on linear arabinan showed that CjArb43A was predominantly an exo-enzyme.

It can be argued that for an enzyme to exhibit both an endo- and exo- modes of action, the active site must undergo substantial conformational changes since the binding domains are

structurally different for each action (see section 1.6.1). For example, cellobiohydrolases are categorically recognised as exo-enzymes releasing cellobiose in a processive manner that is related to their tunnel shaped binding cavity in which the catalytic site is concealed inside (Kurasin *et al*., 2011) (see section 1.8). However, it has been reported that the cellobiohydrolase Cel6A, from *Humicola insolens*, exhibited endo-activity on cellulose ribbons as it cleaved the bacterial cellulose into shorter fragments. This was due to a change in its conformation as structural surface loops covering the tunnel shaped active site, could open up to allow initial endo-activity on the substrate (Boisset *et al*., 2000). Furthermore, numerous studies have attempted to alter the mode of action preference of exo-enzymes to an endo-enzyme including that of CjArb43A, as discussed earlier. Interestingly, CjArb43A was described not to have a tunnel or pocket topology that is observed in exo-enzymes but rather an open groove which is generally displayed by endo-enzymes (Proctor *et al*., 2005). Procter et al., (2005) were able to introduce endo-activity to CjArb43A by compromising the -3 subsite which was ultimately responsible for the enzyme's exo-mode of action. The insertion of a surface loop into the active site altered the -3 subsite leaving the subsite unable to recognise the reducing end of the arabinan chain. This resulted in a switch from an exo- to an endo-mode of action.

In respect of xylanases, several exo-xylanases have been isolated and characterised which are active upon the xylan substrate, solely releasing either xylose or xylobiose. Knowledge concerning exo-acting xylanases is quite limited but such enzymes are becoming increasingly recognised and it has been suggested that this mode of action may account for the intense synergy in xylan degradation along with endo-xylanases (Gasparic *et al*., 1995). For example, Xylanase V from *Aeromonas caviae* ME-1 exclusively produced xylobiose from xylan (Kubata *et al*., 1994). Xylanase V was reported to demonstrate an exo-mechanism as it attacked both oat spelt and birchwood xylan and was also active upon X3. Similarly to OtXyn8A, Xylanase V was inactive on X2, its natural product. However, as hydrolysis progressed (>1 h) the concentration of X2 released by Xylanase V did not

increase but remained constant unlike the X2 generated by OtXyn8A in which X2 concentration increased over time due to the breakdown of xylooligosaccharides. Therefore, as Xylanase V exclusively releases X2, it is strictly classified as an exo-acting enzyme with no endo-activity. Furthermore, XynX also produced by *A.caviae* ME-1 was described to be a novel xylanase producing only X2 and X4 (Usui *et al*., 1999).

A fairly recent study claimed to report the activity of a xylanase from *Streptomyces bangladeshiensis* sp. exhibiting both an endo- and exo- pattern of hydrolysis (Al-bari *et al*., 2007). However the exo-mode of reaction was a result of the debranching potential in which this xylanase had upon arabinoxylan as it released arabinose substituents while the endo-action was the obvious release of xylooligosaccharides from the xylan backbone. Therefore it would be erroneous to call this xylanase an endo-/exo- xylanase, but rather a debranching xylanase.

In conclusion, at this stage of characterisation, OtXyn8A can be characterised as an endo-xylanase significantly releasing xylobiose as its main reaction product. Such an enzyme must play an important role with the organism *Opitutus terrae* PB90-1 while complementing the actions of other enzymes. Even though the direction in which OtXyn8A recognises X2 has not yet been confirmed, it would be important to observe if there are any structural determinants of the -2 and +2 subsites that may facilitate exo-activity. Additionally, knowledge of the substrate-binding conformation will ultimately solve the exo- and endo-hydrolytic capability of the enzyme.

4.4 Comparison of the activity OtXyn8A with other characterised GH8 xylanases

This next section will describe the substrate specificities and functional properties of the structurally related xylanases REX, REXA, Xyn8 and OtXyn8a belonging to GH8. In addition, the endo-activity of OtXyn8A will also be compared to literature accounts of the GH8 endo-xylanase, PhXyl from *P. haloplanktis*.

4.4.1 Comparing the mode of action of OtXyn8A, Xyn8, REX and REXA

Xyn8 is reported as an endo-xylanase but strictly releases xylotriose from its substrates (Lee *et al*., 2006). Alternatively, REX and REXA are both cited as reducing end xylose-releasing exo-oligoxylanase, showing no endo-mode of action but preferentially convey hydrolysis affinity for xylooligosaccharides whose degree of polymerisation is greater than or equal to 3 (Honda *et al*., 2004; Lagaert *et al*., 2007). As expected, REX and REXA primary sequences were predicted to not possess a signal peptide identifying that such enzymes were existing intracellularly. Many hemicellulases carrying out the breakdown of oligosaccharides are generally found to reside as intracellular enzymes. It is well recognised that the primary targets for intracellular and extracellular glycoside hydrolases are oligosaccharides and polysaccharides respectively (Shallom *et al*., 2003; Beylot *et al*., 2001). It can be hypothesised that xylooligosaccharides are imported into the cell by a transporter system for their efficient hydrolysis by REX/ REXA accordingly. However, the lack of a signal peptide of the endo-xylanase, Xyn8, was quite surprising as endo-acting glycoside hydrolases are usually present as extracellular proteins for the breakdown of large polysaccharide substrates. This contrasts with the extracellular endo-xylanase OtXyn8A showing a certain degree of diversity within the GH8 family.

4.4.1.1 GH8 activities on natural xylan substrates

Comparative analysis of the xylanase activities of OtXyn8A, Xyn8, REX and REXA were determined by DNSA reducing sugar assays with various natural polysaccharide substrates including OSX, birchwood xylan, rye arabinoxylan and wheat arabinoxylan using standard conditions described in section 2.2.11. Furthermore, subsequent hydrolysis was monitored by HPAEC-PAD. Incubation experiments were performed with equal moles of enzyme (0.2 μM) at a constant temperature of 25°C and at the optimum pH of each enzyme.

4.4.1.1.1 GH8 activities on OSX and birchwood xylan

The hydrolysis products formed upon incubation with OSX and birchwood xylan were monitored over a 24 h period with rates of hydrolysis by OtXyn8A and Xyn8 shown in Figures 4.12 and 4.13. OtXyn8A systematically produced X2 throughout the course of the reaction with larger xylooligosaccharides up to a degree of polymerisation of 16 detected mainly between 5-10 min of digestion. Xyn8 yielded high levels of X3 and significant amounts of X4 within the in initial stages of hydrolysis which corresponds to results produced by Lee *et al.,* 2006. The data presented in Figure 4.16 shows that Xyn8 displayed an endo-mode of action against xylan during the initial stages of hydrolysis. However, in comparison to OtXyn8A, the range of oligosaccharides generated by Xyn8 appeared to be less, yielding primarily X3 and X4. At 24 h, it is interesting to note the difference in progression of degradation as the products of OtXyn8A are simply X1 and X2, while the final products of Xyn8 are X1, X2 and X3 with residual substituted compounds residing from both reactions (Figures 4.17 and 4.16). Also by quantifying these xylooligosaccharides, there is evidence that OtXyn8A degrades more of the polymeric xylan substrate in comparison to Xyn8. Therefore, Xyn8 has an apparent lower affinity for the polymeric

substrate in comparison to OtXy8A possibly indicating that the polymeric substrate is not a main substrate of Xyn8 as similar results were observed with the activity on arabinoxylan.

Products of both OtXyn8A and Xyn8 activity differ to that reported for PhXyl as the hydrolysis of OSX and birchwood xylan by PhyXyl initially generated a mixture of xylooligosaccharides including X3, X4, X5, X6 and larger xylooligosaccharides (Collins *et al*., 2002). As the reaction continued X5 and X6 progressively degraded while X3 and X4 accumulated.

REX and REXA exhibited no significant activity on OSX when analysed via DNSA assay, thus identifying both enzymes do not have endo-acting properties as described previously by Honda *et al*., (2004) and Lagaert *et al*., (2007). However the HPAEC-PAD results illustrated a slight release of X1 from both enzymes after 24 h with no X1 detected during the first hour of hydrolysis as shown in Figures 4.14 and 4.15. This was most likely a result of the action of REX and REXA releasing xylose from the reducing termini of the xylan main chain which may have been prominent until the cleavage was blocked by a substituted group on the main backbone. The release of xylose after 24 h was not reported by Honda *et al*., (2004), as no apparent activity of REX was observed on OSX after progressed hydrolysis.

4.4.1.1.2 GH8 activities on wheat arabinoxylan and rye arabinoxylan

Upon incubation with wheat and rye arabinoxylan, a significant difference in hydrolytic progression was observed by HPAEC-PAD for both OtXyn8A and Xyn8. Comparing the product profiles of the two xylanases on arabinoxylan (chromatograms within Figures 4.16 and 4.17) revealed that OtXyn8A and Xyn8 were releasing similar end products to those generated via OSX degradation. However, after prolonged hydrolysis (24 h) the enzymes displayed only trace activity in that products were quantitatively less compared to those accumulated from OSX hydrolysis. After 24 h, X2 was still predominantly released by OtXyn8A however there were also substantial detection of X1 and larger xylooligosaccharides, possibly X7. Similarly, Xyn8 released X1, X2 and X3 with considerably amounts of larger xylooligosaccharides (>X7) (Figure 4.16). This reveals information on the substrate specificity of each enzyme when acting on highly decorated xylans. Hydrolysis of the larger xylooligosaccharides (>7) in both cases was most likely inhibited by an arabinosyl side group present on the main xylan chain, blocking the action of OtXyn8A and Xyn8. Also, rates of hydrolysis by reducing sugar analysis correlated with HPAEC-PAD analysis, proving that less sugar was released during arabinoxylan hydrolysis in comparison to xylan degradation suggesting that both the xylanases prefer OSX (Figures 4.12 and 4.13). A question now arises whether this applies to all GH8 endo-xylanases and if this group of enzymes are more active upon lower degree of substitution xylan as they are unable to govern substantial binding to highly substituted xylans, which is most definitely the situation portrayed by OtXyn8A and Xyn8. This specificity was also identified with the action of the GH8 PhXyl endo-xylanase by Collins *et al*., (2002), in which the action of PhXyl was hindered by arabinose substituents. However, interestingly within 60 min of hydrolysis, Xyn8 attacked more of the arabinoxylan structure than OtXyn8A. Pollet et al., (2010) created a theoretical structure of Xyn8 based on REX in which the structure was compared to PhXyl. This study found that a structural loop in PhXyl folded over the glycon

region decreasing the accessibility of the active to substituted xylans. This loop region was then found to be considerably shorter in the Xyn8 model and structure of REX. Therefore this may suggest that Xyn8 has the ability to accommodate substituted xylans or a higher affinity for such substrates in comparison to OtXyn8A and PhXyl which can be seen in reducing analysis (Figure 4.13). The structure of OtXyn8A may therefore reveal an extended structural loop similarly to that of PhXyl.

REX and REXA exhibited no activity on wheat arabinoxylan or rye arabinoxylan, which was expected as arabinose substituents would have inhibited the action of these enzymes on any of the free reducing ends present on the polysaccharide.

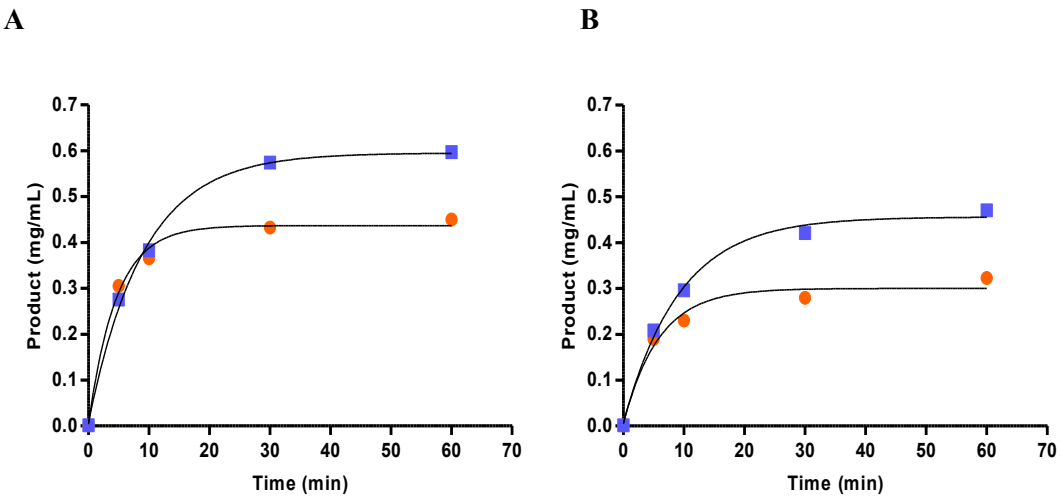


Figure 4.12 Rates of OSX (A) and Birchwood xylan (B) hydrolysis by OtXyn8A and Xyn8. OtXyn8A (0.2 µM) (■) and Xyn8 (0.2 µM) (●) were incubated with OSX **(A)** and birchwood xylan **(B)** at a final substrate concentration of 8 mg/mL. Reactions were carried out in sodium acetate (0.05 M final concentration) pH 6 and BSA (1 mg/mL final concentration) at 25°C. At intervals 5 min, 10min, 30 min and 60 min samples were removed and rates were measured by reducing sugar assay using DNSA reagent (see section 2.2.12).

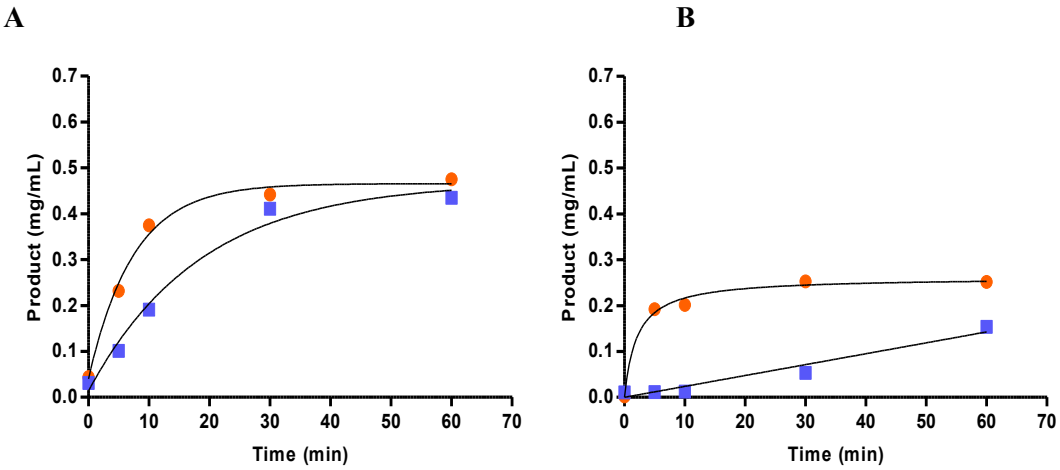


Figure 4.13 Rates of Wheat arabinoxylan (A) and Rye arabinoxylan (B) hydrolysis by OtXyn8A and Xyn8. OtXyn8A (0.2 µM) (■) and Xyn8 (0.2 µM) (●) were incubated with rye arabinoxylan **(A)** and wheat arabinoxylan **(B)** at a final substrate concentration of 8 mg/mL. Reactions were carried out in sodium acetate (0.05 M final concentration) pH 6 and BSA (1 mg/mL final concentration) at 25°C. At intervals 5 min, 10min, 30 min and 60 min samples were removed and rates were measured by reducing sugar assay using DNSA reagent (see section 2.2.12).

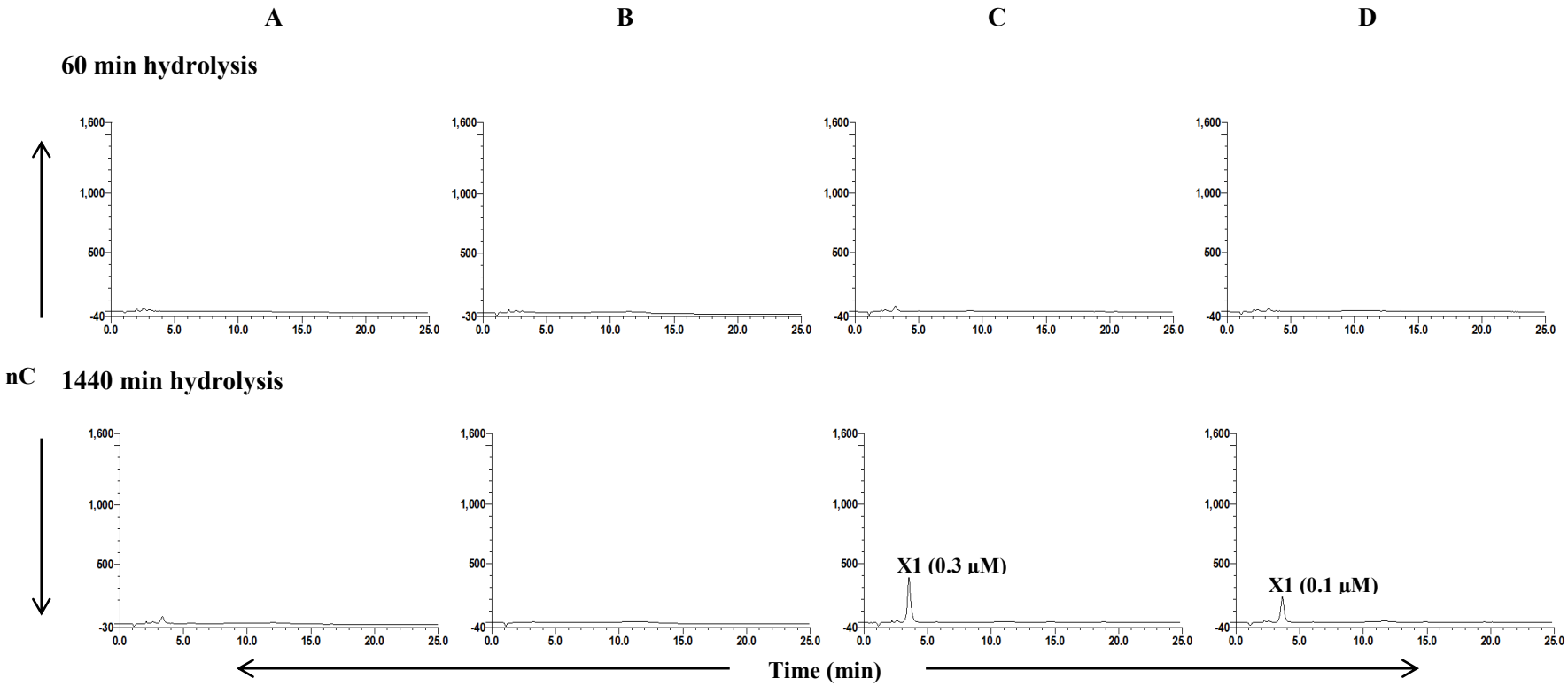


Figure 4.14 HPAEC-PAD analysis showing the hydrolysis products of wheat arabinoxylan, rye arabinoxylan, OSX and birchwood xylan by REX. REX (0.2 µM) was incubated with wheat arabinoxylan **(A)**, rye arabinoxylan **(B)**, OSX **(C)** and birchwood xylan **(D)** at a final substrate concentration of 8 mg/mL. Reactions were carried out in sodium acetate (0.05 M final concentration) pH 6 and BSA (1 mg/mL final concentration) at 25°C. At intervals 60 min and 1440 min samples were removed and subjected to HPAEC-PAD analysis (see section 2.2.12). Hydrolysis products at 24 h have been quantified (see section 2.2.12.1).

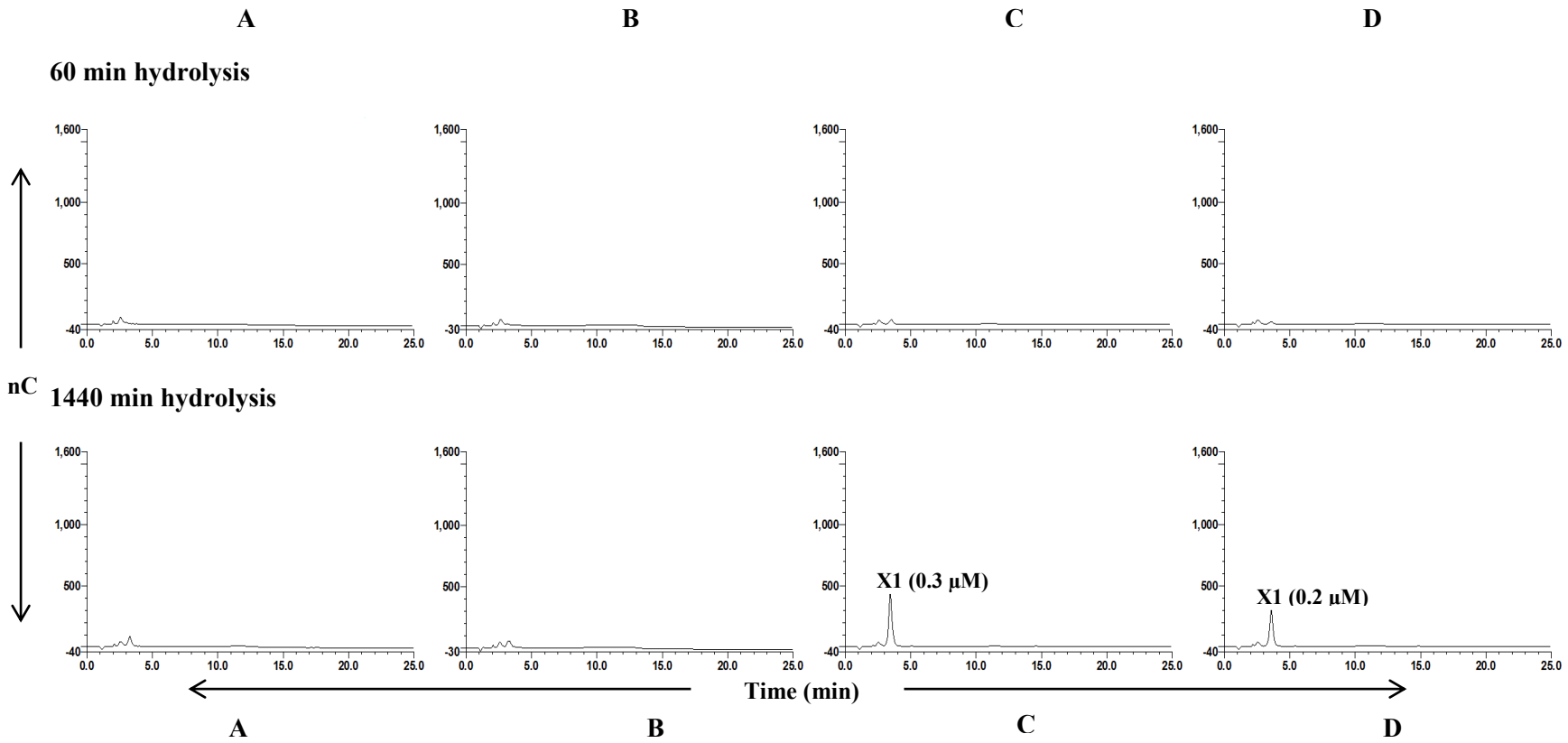


Figure 4.15 HPAEC-PAD analysis showing the hydrolysis products of wheat arabinoxylan, rye arabinoxylan, OSX and birchwood xylan by REXA. REXA (0.2 µM) was incubated with wheat arabinoxylan **(A)**, rye arabinoxylan **(B)**, OSX **(C)** and birchwood xylan **(D)** at a final substrate concentration of 8 mg/mL. Reactions were carried out in sodium acetate (0.05 M final concentration) pH 6 and BSA (1 mg/mL final concentration) at 25°C. At intervals 60 min and 1440 min samples were removed and subjected to HPAEC-PAD analysis (see section 2.2.12). Hydrolysis products at 24 h have been quantified (see section 2.2.12.1).

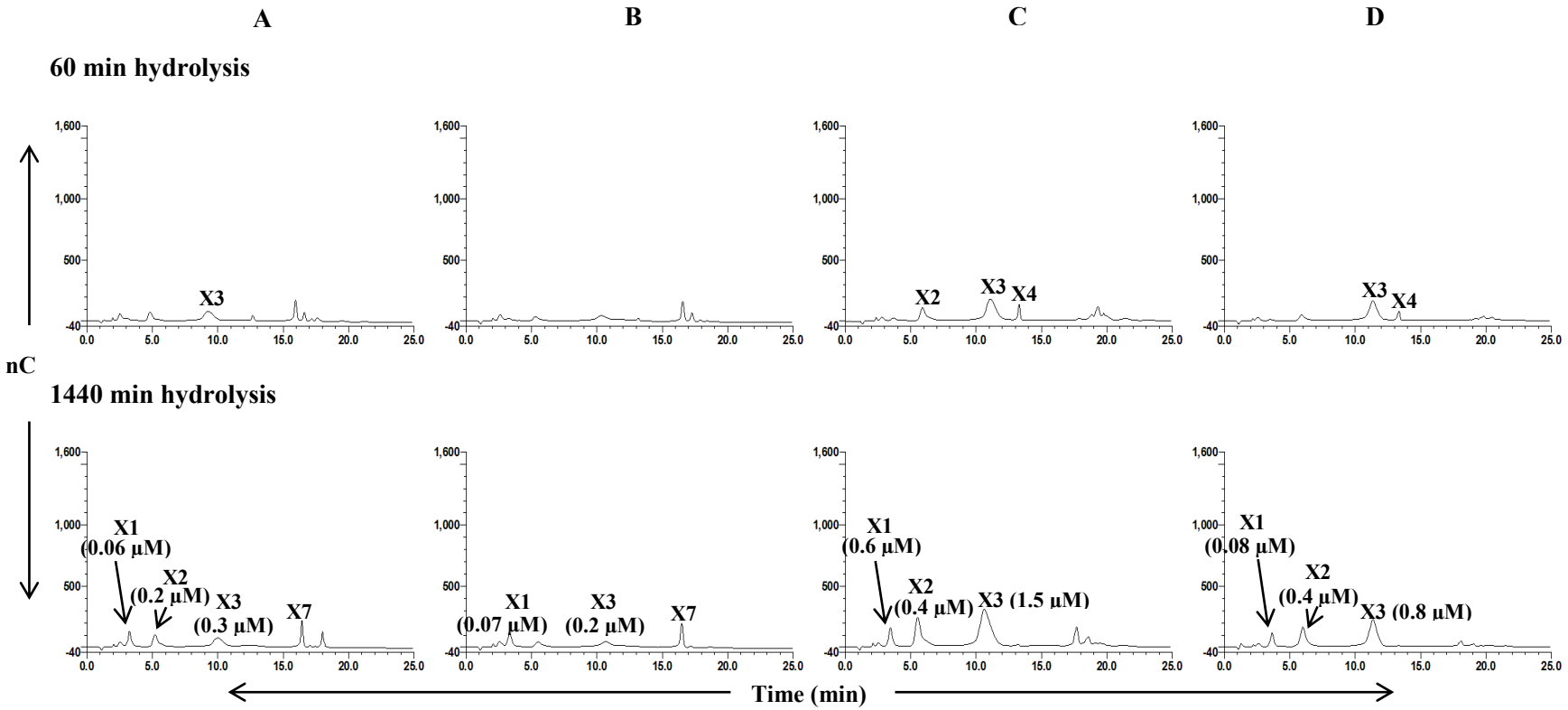


Figure 4.16 HPAEC-PAD analysis showing the hydrolysis products of wheat arabinoxylan, rye arabinoxylan, OSX and birchwood xylan by Xyn8. Xyn8 (0.2 µM) was incubated with wheat arabinoxylan **(A)**, rye arabinoxylan **(B)**, OSX **(C)** and birchwood xylan **(D)** at a final substrate concentration of 8 mg/mL. Reactions were carried out in sodium acetate (0.05 M final concentration) pH 6 and BSA (1 mg/mL final concentration) at 25°C. At intervals 60 min and 1440 min samples were removed and subjected to HPAEC-PAD analysis (see section 2.2.12). Hydrolysis products at 24 h have been quantified (see section 2.2.12.1).

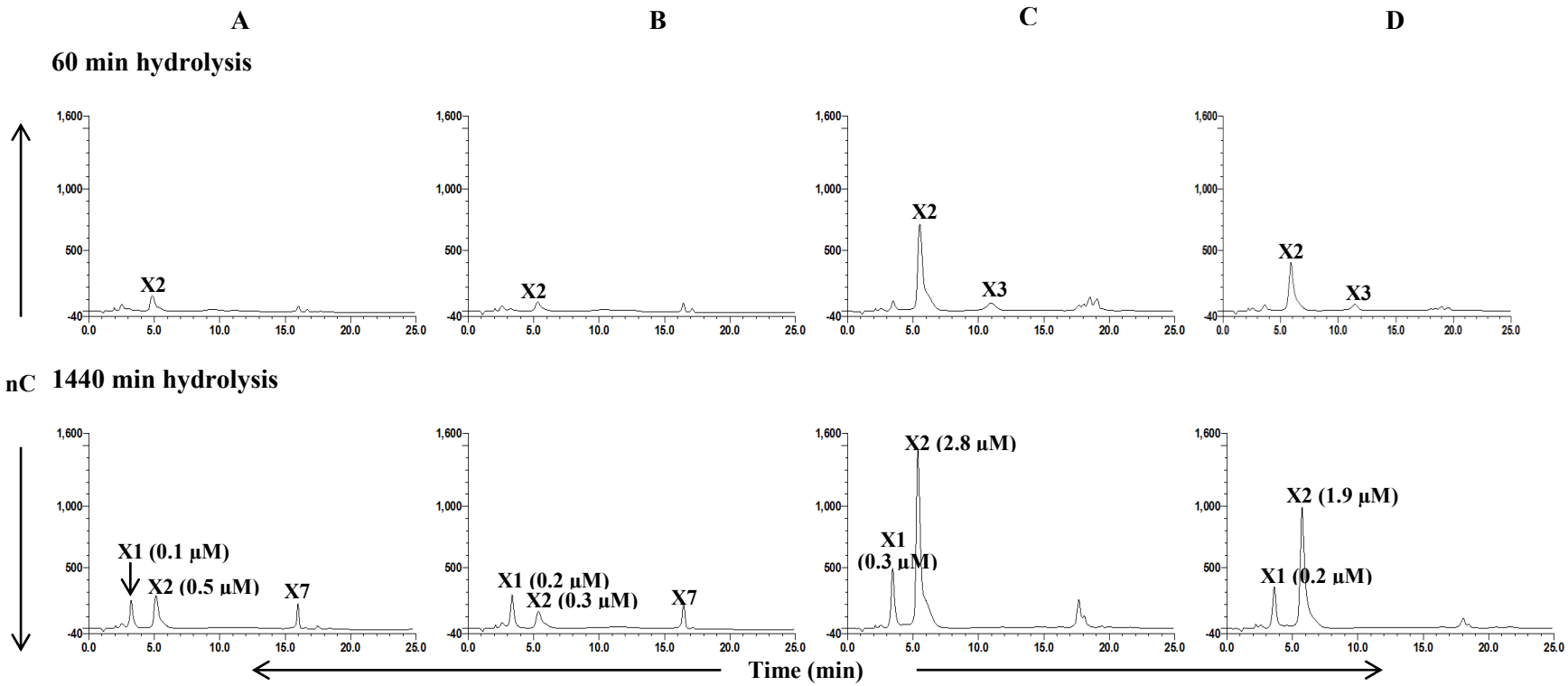


Figure 4.17 HPAEC-PAD analysis showing the hydrolysis products of wheat arabinoxylan, rye arabinoxylan, OSX and birchwood xylan by OtXyn8A. OtXyn8A (0.2 µM) was incubated with wheat arabinoxylan **(A)**, rye arabinoxylan **(B)**, OSX **(C)** and birchwood xylan **(D)** at a final substrate concentration of 8 mg/mL. Reactions were carried out in sodium acetate (0.05 M final concentration) pH 6 and BSA (1 mg/mL final concentration) at 25°C. At intervals 60 min and 1440 min samples were removed and subjected to HPAEC-PAD analysis (see section 2.2.12). Hydrolysis products at 24 h have been quantified (see section 2.2.12.1).

4.4.1.2 GH8 activities on xylooligosaccharides

The GH8 xylanases were tested against various xylooligosaccharides that were produced within this study and purified using gel chromatography. The reactions of each enzyme incubated with the individual xylooligosaccharides were monitored and analysed using HPAEC-PAD.

The data collected for xylooligosaccharide hydrolysis by Xyn8 revealed that the enzyme possessed no activity against X2 or X3 (Figures 4.18 and 4.19). Apparent activity was observed upon incubation of Xyn8 with X4, resulting in the accumulation of X3 and X1 (Figure 4.20). However, these results were obtained after an extended reaction of 24 h unlike OtXyn8A in which total hydrolysis of X4 into X2 was achieved in just 1 h of hydrolysis (Figure 4.20). This data generated was consistent with the data gathered from previous studies by Lee *et al*., (2006) and Pollet *et al*., (2010), who also described the activity of Xyn8 on X5 and X6 which was not investigated within this study. The activity of Xyn8 was reported to be ultimately higher against X6 than X5, in the release of X3. The difference in activity against X4 between Xyn8 and OtXyn8A indicates that OtXyn8A has a preference to this substrate while other studies suggest that Xyn8 prefers X6 to bind to. As a result of this and the main reaction product of Xyn8 being X3, it is presumed that this enzyme contains 6 subsites (-3, -2, -1, +1, +2 and +3) and similarly to this, PhXyl has also been reported to have six subsites (De Vos et al., 2006). However the varying affinities for both polymeric xylans and xylooligosaccharides makes their substrate specificities quite different as PhXyl has a higher activity against xylan substrates with a high degree of polymerisation while Xyn8 is more active against substrates with a lower degree of polymerisation (Pollet *et al*., 2010).

REX and REXA were active against xylooligosaccharides with a degree of polymerisation greater than 2 including X4 however for such reactions, the data was collected only at time points 1 h and 24 h. The hydrolysis patterns of X3 degradation by REX, REXA and

OtXyn8A appeared to be quite similar in that the products released by all three enzymes were X1 and X2 (Figure 4.19). However, the degradation of X3 by REX and REXA was relatively quicker as complete degradation was recorded after 1 h of hydrolysis. In comparison, the hydrolysis of X3 by OtXyn8A was completed after an extended incubation and such results were collected 24 h although since no data was collected between 1 h and 24 h, the end of the reaction could not be determined. REX and REXA also ultimately released X1 and X2 from X4 hydrolysis however data from previous studies by Honda *et al*., (2004) suggested that X3 was the preferred substrate of REX which coincides with the three catalytic subsites (-2, -1 and +1) of the enzyme. Similarly to REX, the preferred substrate of REXA was X3, as described by Lagaert *et al*., (2007). How the specificity of REX towards xylooligosaccharides differs from endo-xylanases such as PhXyl can be explained by structural features of the subsites of REX. Fushinobu *et al*., (2005) revealed how the +2 subsite of REX was blocked by a kink in one the structural loops not allowing sufficient binding to the substrate. Additionally, the His319 within the +1 subsite formed a specific hydrogen bond with the β-hydroxyl of the xylose unit therefore showing how REX is able to discriminate between the reducing and non-reducing end of a xylose residue for the release of xylose.

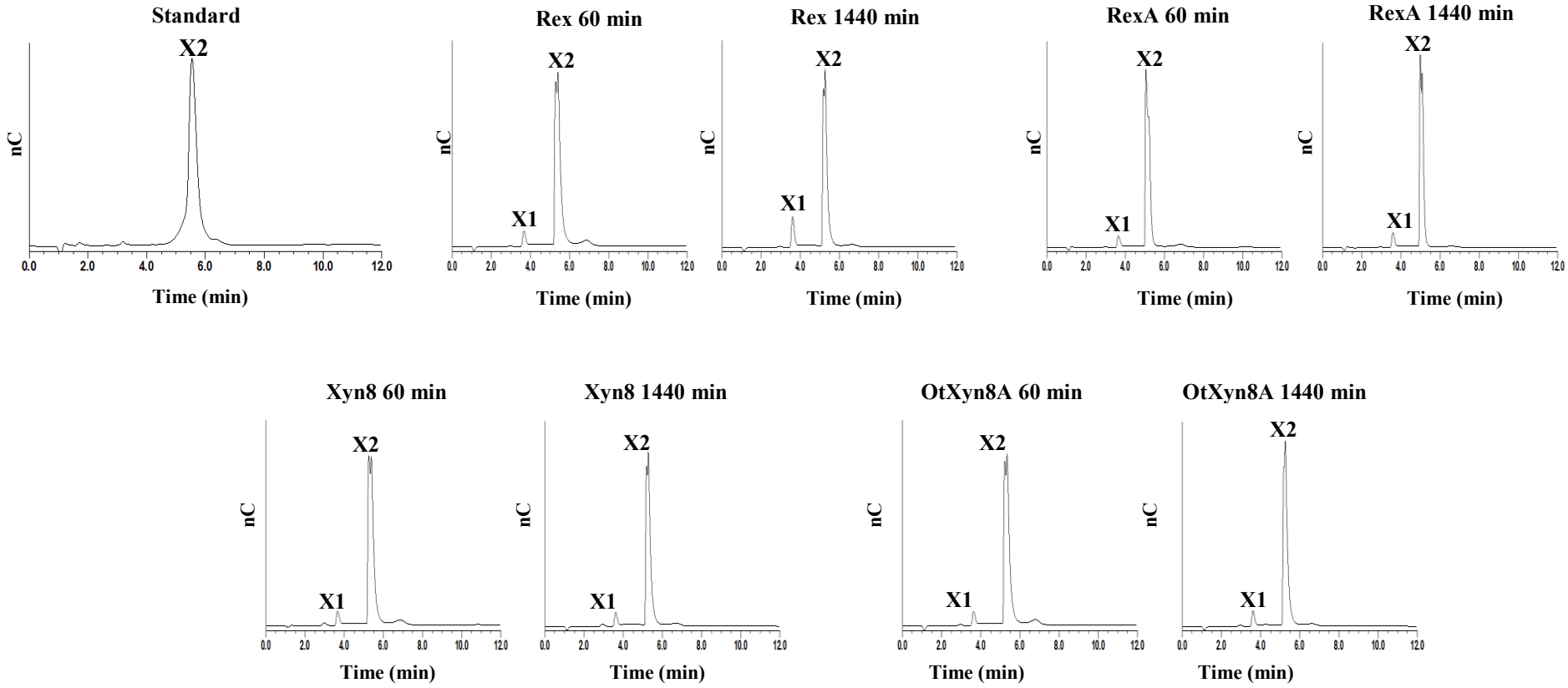


Figure 4.18 Activity against xylobiose (X2) by GH8 xylanases. REX (0.2 μM), REXA (0.2 μM), Xyn8 (0.2 μM) and OtXyn8A (0.2 μM) were incubated with X2 Reactions were carried out in sodium acetate (0.05 M final concentration) pH 6 and BSA (1 mg/mL final concentration) at 25°C. At intervals 60 min and 1440 min samples were removed and subjected to HPAEC-PAD analysis (see section 2.2.12).

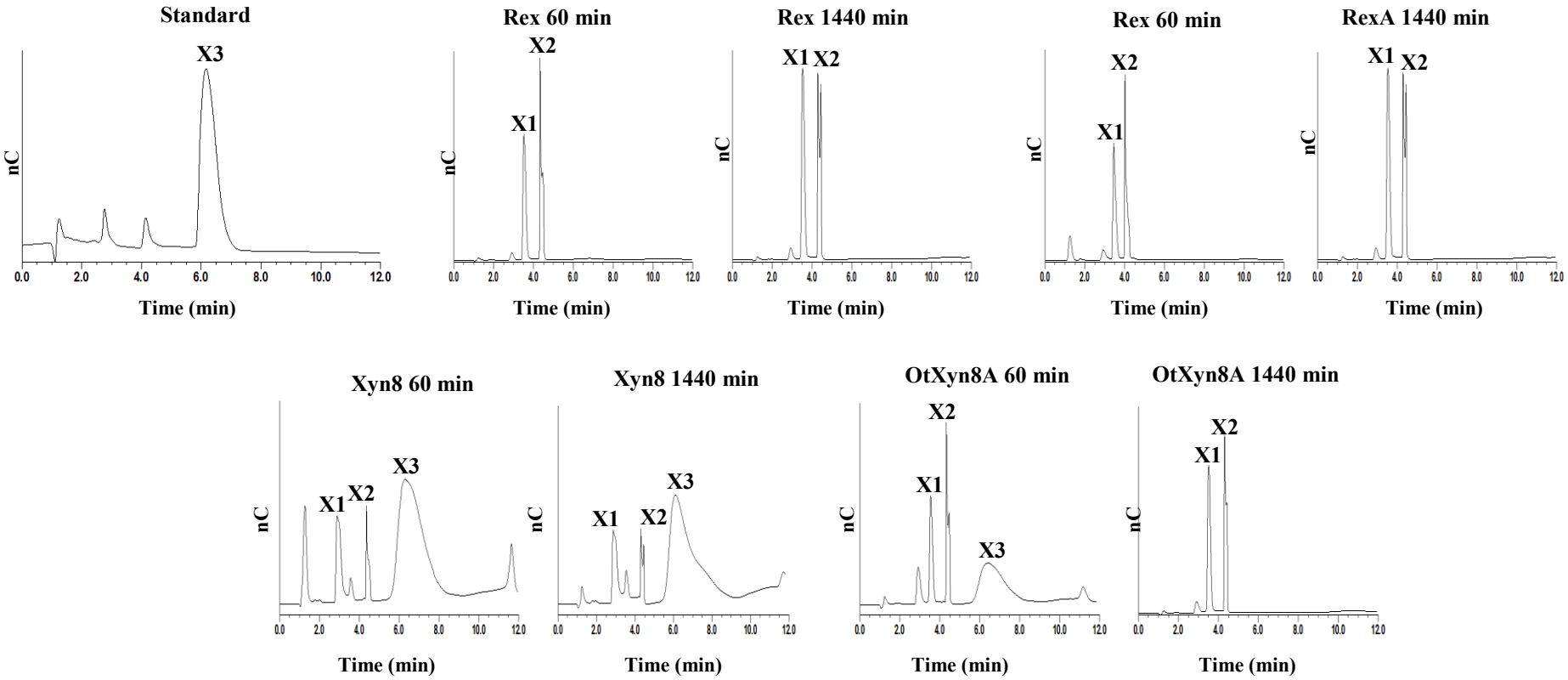


Figure 4.19 Activity against xylotriose (X3) by GH8 xylanases. REX (0.2 µM), REXA (0.2 µM), Xyn8 (0.2 µM) and OtXyn8A (0.2 µM) were incubated with X3 Reactions were carried out in sodium acetate (0.05 M final concentration) pH 6 and BSA (1 mg/mL final concentration) at 25°C. At intervals 60 min and 1440 min samples were removed and subjected to HPAEC-PAD analysis (see section 2.2.12).

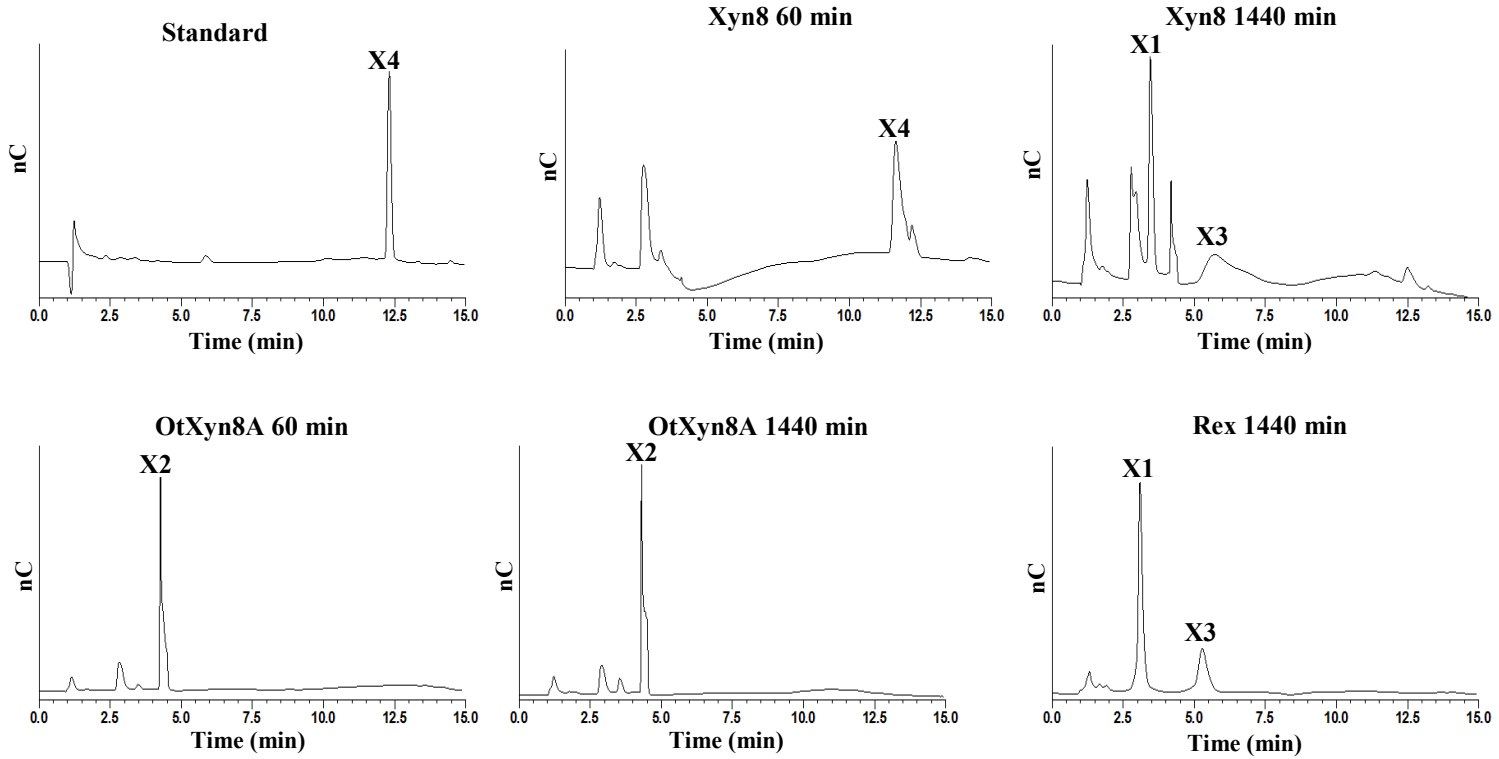


Figure 4.20 Activity against xylotetraose (X4) by GH8 xylanases. REX (0.2 µM), Xyn8 (0.2 µM) and OtXyn8A (0.2 µM) were incubated with X4 Reactions were carried out in sodium acetate (0.05 M final concentration) pH 6 and BSA (1 mg/mL final concentration) at 25°C. At intervals 60 min and 1440 min samples were removed and subjected to HPAEC-PAD analysis (see section 2.2.12).

4.4.1.3 GH8 activities on AZO-xylan

To confirm endo- activity of OtXyn8A and Xyn8, an AZO-OSX assay was carried out as described in section 2.2.11.4, in which REX and REXA were used as negative controls as essentially they do not possess any endo-activity. AZO-OSX is a chromogenic substrate containing one Remazolbrilliant Blue R molecule per 30 sugar residues in which endo-activity is measured by monitoring the release of soluble oligosaccharide fragments containing the dye. Reactions were undertaken at 45°C, resembling the optimum temperature for OtXyn8A, REX and REXA activity and at 25°C, the optimal temperature for Xyn8. For OtXyn8A and Xyn8, endo- activity was identified as an increase in absorbance was recorded over a 60 min time course as shown in Figure 4.21. REX and REXA showed no increase in absorbance but stayed at relatively negligible absorbance confirming no endo-action of the enzymes (Figure 4.21).

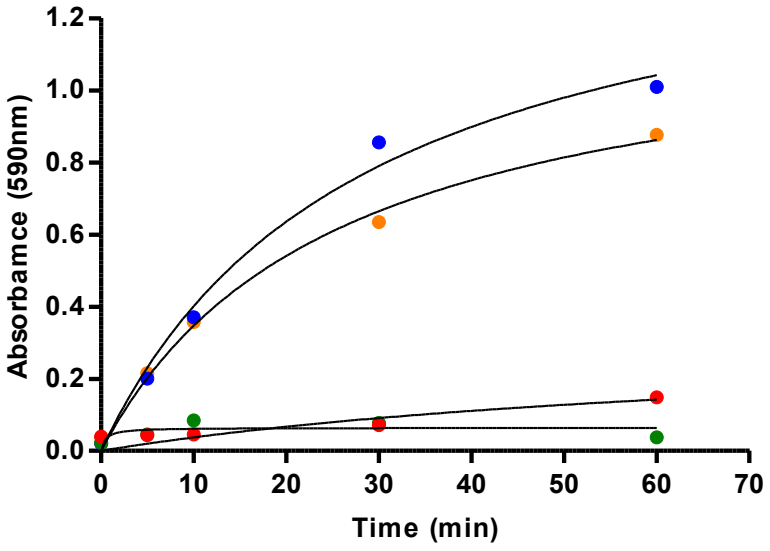


Figure 4.21 AZO-OSX assay for the determination of endo-xylanase activity by GH8 xylanases. OtXyn8A (0.2 µM) (●), Xyn8 (0.2 µM) (●), REX (0.2 µM) (●) and REXA (0.2 µM) (●) were incubated with AZO-OSX at a final concentration of 8 mg/mL and the assay was carried out as described in section 2.2.11.4 at their optimal temperatures. 1 mg/mL BSA (final concentration) and 0.05 M of appropriate buffer achieving optimal pH for individual GH8 xylanases were added.

4.4.1.4 Anomeric recognition analysis

A characteristic property of reducing sugars was first described by Dubrunfaut in 1846 (Finch, 199), whereby a change in optical rotation occurs resulting in the interconversion of anomers in which an equilibrium mixture of the two anomer forms is reached. This phenomenon is known as mutarotation and all reducing carbohydrates undergo mutarotation in an aqueous solution where the sugar ring takes the α- and β- forms. In water, D-xylose favours the β- form with the equilibrium concentrations approximately 65% for the β-pyranoid ring and 35% for the α-pyranoid ring (Schmidt *et al*., 1996). Knowledge of the anomeric configuration recognised by each enzyme gives an insight to the rates at which substrates are hydrolysed and if mutarotation is a rate-limiting step in such cleavage.

The anomeric analysis of the hydrolytic products produced from various xylooligosaccharides by REX demonstrated that the enzyme favoured the β-anomeric composition of X3 at the linkage closest to the reducing end. This suggested that the enzyme requires mutarotation to convert α-X3 into β-X3 and Honda *et al.,* (2004) hypothesised that the utilisation of a mutarotase would accelerate the action of REX on xylooligosaccharides. Therefore to establish whether mutarotation was favoured by the investigating GH8 xylanases, standard assays were carried out for the analysis of hydrolysis products released from OSX by each enzyme with the addition of mutarotase to catalyse mutarotation.

HPAEC-PAD results for OtXyn8A and Xyn8 hydrolysis reactions did not show an increased quantity of xylooligosaccharides released when mutarotase was introduced into the reaction (Figures 4.24 and 4.25). Therefore indicating that mutarotation was not a necessity and OtXyn8A and Xyn8 recognised both the α- and β-anomers of the reducing end xylose residue.

HPAEC-PAD analysis of REX and REXA hydrolysing xylooligosaccharides showed an elevated amount of xylose released as a consequence of accelerated mutarotation by

mutarotase in comparison to the reaction with no added mutarotase (Figures 4.22 and 4.23). This data was in agreement with the results founded previously for REX by Honda *et al.,* (2004) in that REX only recognised the β-anomer of the reducing end xylose residues. Therefore spontaneous mutarotation of the substrate is required for the successive hydrolysis by REX.

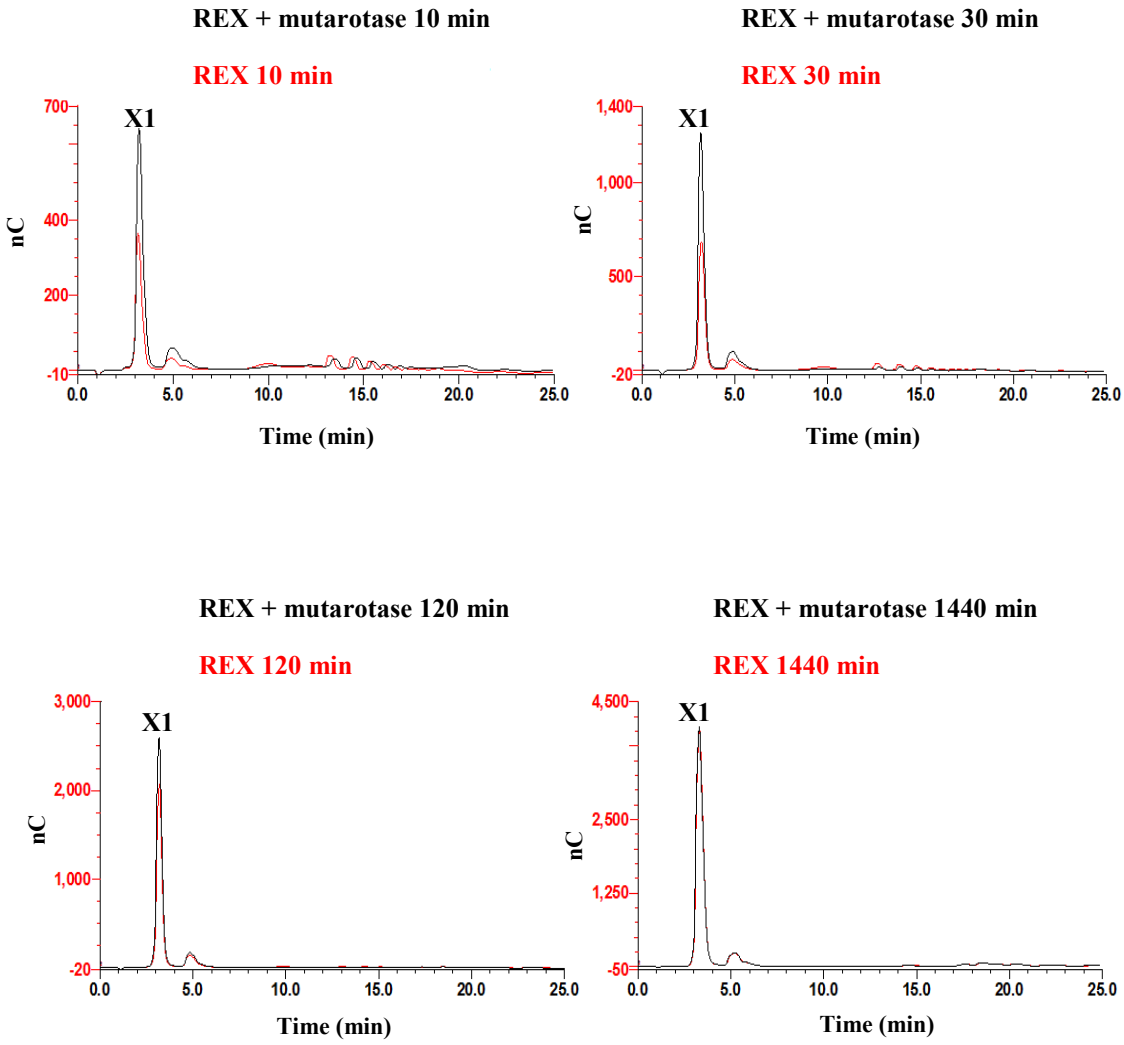


Figure 4.22 The effect of mutarotase on REX activity against xylooligosaccharides. A mixture of xylooligosaccharides with a substrate concentration of 8 mg/mL were incubated with REX (0.2 µM) and **with** / **without** mutarotase (4.1 µg/µL). Reactions were carried out in sodium acetate (0.05 M final concentration) pH 6 and BSA (1 mg/mL final concentration) at 25°C. At intervals 10 min, 30 min, 120 min and 1440 min, samples were removed and subjected to HPAEC-PAD analysis (see section 2.2.12). Red lines represent REX activity; black lines represent REX and mutarotase activity.

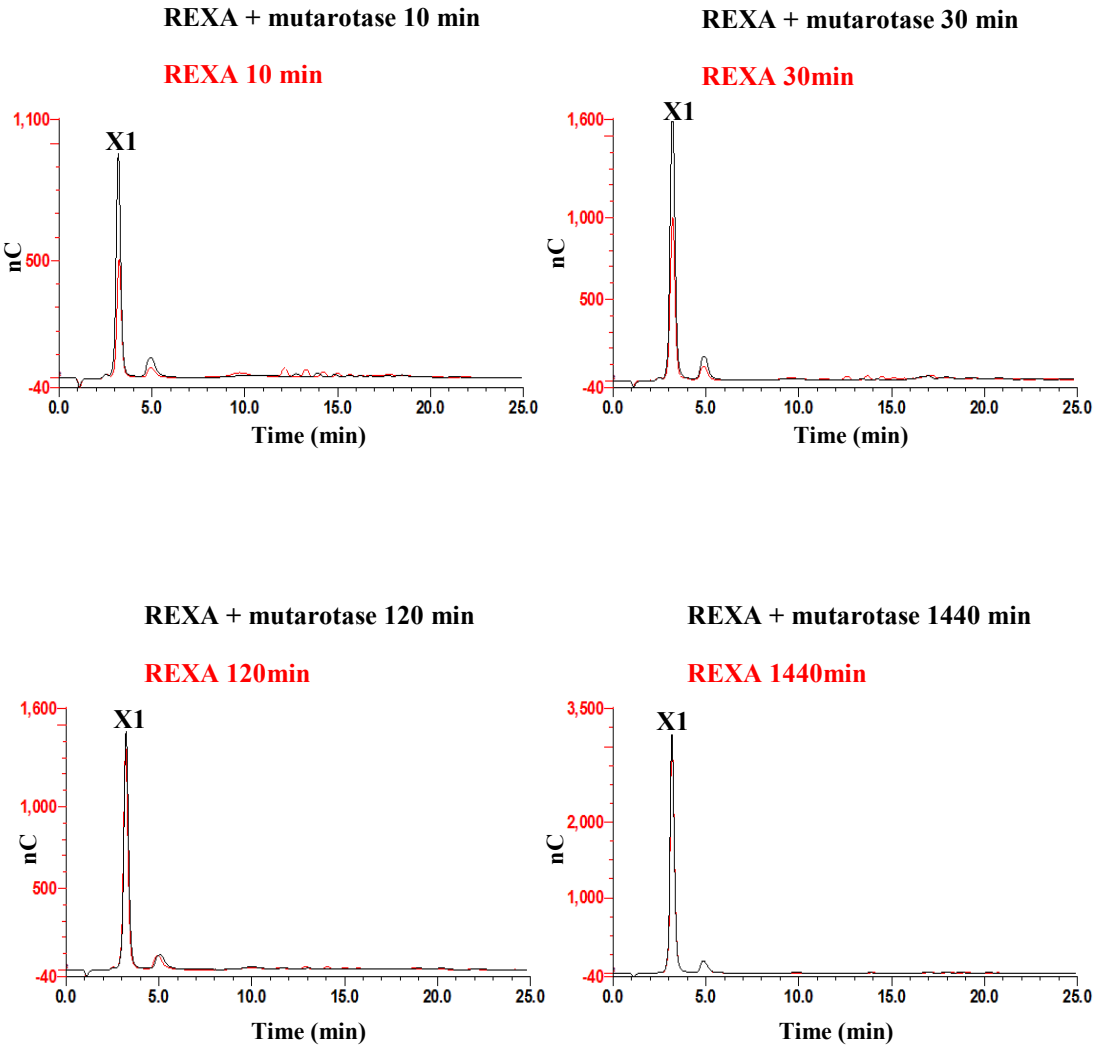


Figure 4.23 The effect of mutarotase on REXA activity against xylooligosaccharides. A mixture of xylooligosaccharides with a substrate concentration of 8 mg/mL were incubated with REXA (0.2 µM) and **with** / **without** mutarotase (4.1 µg/µL). Reactions were carried out in sodium acetate (0.05 M final concentration) pH 6 and BSA (1 mg/mL final concentration) at 25°C. At intervals 10 min, 30 min, 120 min and 1440 min, samples were removed and subjected to HPAEC-PAD analysis (see section 2.2.12). Red lines represent REXA activity; black lines represent REXA and mutarotase activity.

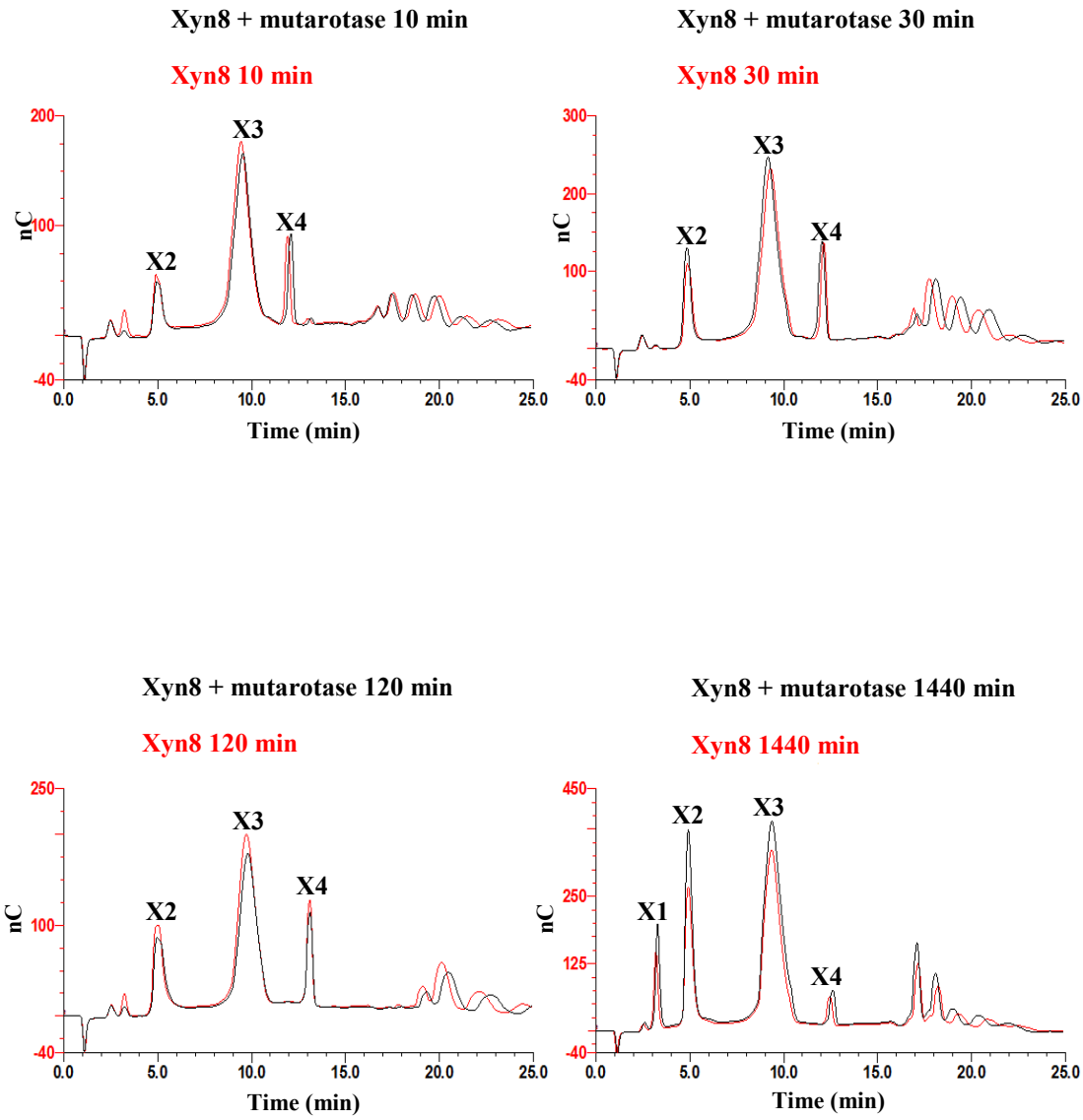


Figure 4.24 The effect of mutarotase on Xyn8 activity against OSX. Xyn8 (0.2 µM) was incubated with OSX at a final of 8 mg/mL and **with** / **without** mutarotase (4.1 µg/µL). Reactions were carried out in sodium acetate (0.05 M final concentration) pH 6 and BSA (1 mg/mL final concentration) at 25°C. At intervals 10 min, 30 min, 120 min and 1440 min, samples were removed and subjected to HPAEC-PAD analysis (see section 2.2.12). Red lines represent Xyn8 activity; black lines represent Xyn8 and mutarotase activity.

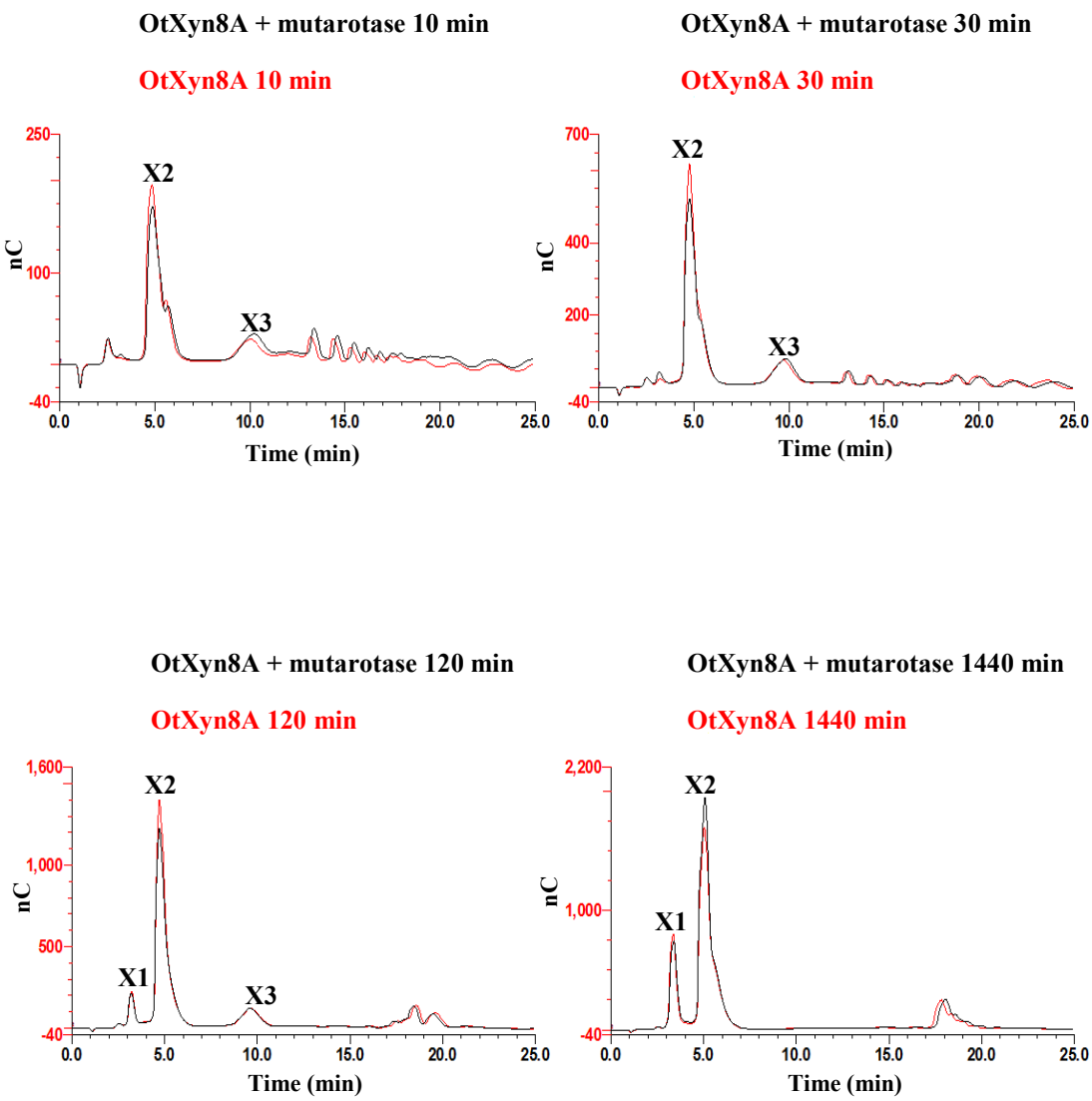


Figure 4.25 The effect of mutarotase on OtXyn8A activity against OSX. OtXyn8A (0.2 µM) was incubated with OSX at a final of 8 mg/mL and **with** / **without** mutarotase (4.1 µg/ µL). Reactions were carried out in sodium acetate (0.05 M final concentration) pH 6 and BSA (1 mg/mL final concentration) at 25°C. At intervals 10 min, 30 min, 120 min and 1440 min samples were removed and subjected to HPAEC-PAD analysis (see section 2.2.12). Red lines represent OtXyn8A activity; black lines represent OtXyn8A and mutarotase activity.

4.4.2 Biochemical characterisation comparison

From the data in Table 4.4 it can be observed that OtXyn8A and Xyn8 exhibit similar k_{cat} values, whilst the k_{cat} value for PhXyl is significantly higher. Furthermore, the k_M value was significantly lower for OtXy8A in comparison to values obtained for Xyn8 and PhXyl, against OSX. This lower k_m observed for OtXyn8A suggests the enzyme has an elevated affinity for OSX in comparison to the other GH8 endo-xylanases. The kinetic constants observed for OtXyn8A and Xyn8 may reflect the release of sugar from the initial and latter stages of xylan hydrolysis as initially Xyn8 and OtXyn8A released relative amounts of sugar products until after prolonged hydrolysis OtXyn8A accumulated greater quantities of its end products (Figures 4.16 and 4.17).

Enzyme	**k_{cat} (min^{-1}) (OSX)**	**K_M (mg mL^{-1}) (OSX)**	**k_{cat} / K_M (min^{-1} mg^{-1} mL)**
OtXyn8A	7661 ± 23	6.84 ± 0.66	1120 ± 9
Xyn8	8280*	14*	591*
REX	NA	NA	NA
REXA	NA	NA	NA
PhXyl	65940*	21*	3140*

Table 4.4 Michaelis-Menten kinetic parameters for the GH8 xylanases for the hydrolysis of OSX. The reaction of OtXyn8A on OSX was performed in triplicate at 45°C in 0.05 M sodium acetate buffer, pH6 and BSA was included at a final concentration of 1 mg/mL. DNSA analysis was carried out as described in section 2.2.11.2 to analyse kinetic values. *Kinetic parameters reproduced from Pollet *et al*., 2010.

4.4.4 Sequence comparison between OtXyn8A and characterised GH8 xylanases

Since the 3D structure of OtXyn8A has not yet been solved, it was important to compare the primary sequences of OtXyn8A and the characterised GH8 xylanases for investigation into structural determinants that may be responsible for different enzymes specificities. Therefore, amino acid alignments of OtXyn8A and the GH8 endo-xylanases, Xyn8 and PhXyl and the two exo-oligoxylanases, REX and REXA, were performed using ClustalW revealing significant catalytic residues and sequences involved in substrate binding.

From multiple sequence alignments, sequence identities between each of the GH8 xylanases could be established. The GH8 xylanases have been classified as GH8 enzymes due to their closest sequence homology to other members of family 8 despite the speculated identities to the other GH8 enzymes being moderately low. The primary sequence homologies between OtXyn8A and the remaining GH8 endo-xylanases can be described as relatively low (29 - 47%) with OtXyn8A showing highest sequence identity to Xyn8 and the lowest to PhXyl. Interestingly, OtXyn8A has a higher sequence identity to REX than to PhXyl which does not necessarily account for their individual modes of action since OtXyn8A and PhXyl are both endo-xylanases and REX is an exo-xylanase.

The crystal structures of PhXyl and REX revealed important catalytic residues with the general acid residing as a glutamic acid residue (Glu78 and Glu70 for PhXyl and REX, respectively) and the general base being an aspartic acid residue (Asp281 and Asp263 for PhXyl and REX, respectively) (Fushinobu *et al*., 2005; De Vos *et al*., 2006). With respect to this, these critical residues could be determined on a sequence level for the remaining GH8 xylanases, REXA, Xyn8 and OtXyn8A, in which such residues are highlighted in Figure 4.26. From probing the primary sequence alignments (Figure 4.26) the presented GH8 xylanases are conserved in their general acid and base residues, however a third conserved residue also exists in the form of an additional aspartic acid. In REX, the residue Asp128 was suggested to be involved in sugar-ring distortion (Fushinobu *et al*., 2005) and therefore

this may be performing the same role in all GH8 xylanases presented in Figure.. Furthermore, as these GH8 xylanases consist of an aspartic acid as their catalytic base, this means that they belong to GH8 subfamily GH8a, thus providing more information in the characterisation of OtXyn8A.

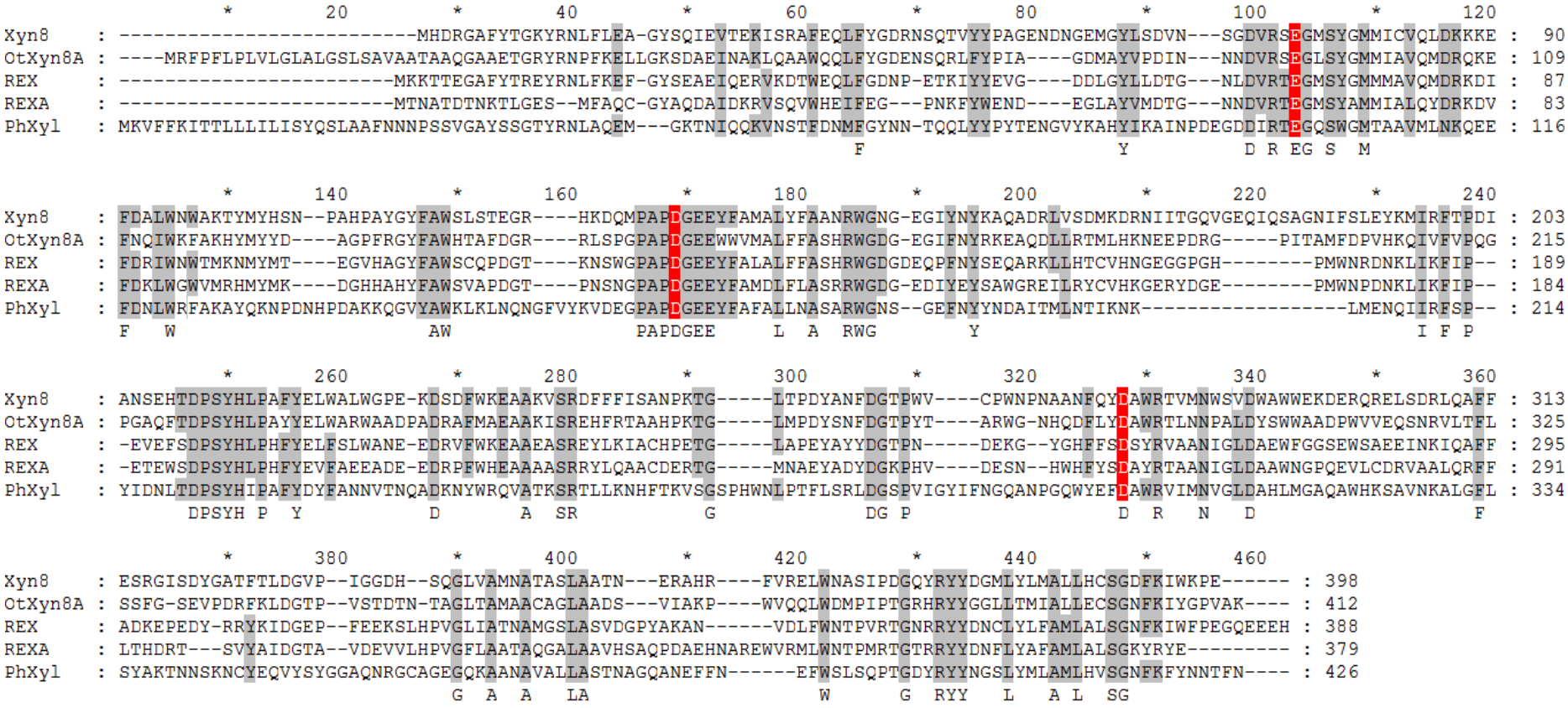


Figure 4.26 CLUSTALW Multiple sequence alignment of OtXyn8A, Xyn8A, REX, REXA and PhXyl. Similar amino acids are coloured in grey and catalytic residues are coloured in red.

4.4.5 Substrate specificity in GH8

Assessing the substrate preference and hydrolysis profiles of the GH8 xylanases has shed new light on the substrate specificity in family 8 as specificity is clearly diverse between that of the tested GH8 xylanases and PhXyl (described in literature). Not only the activities of such enzymes against polymeric xylans, but their activities against various xylooligosaccharides, have provided a greater understanding of preferred substrates of the GH8 enzymes. The GH8 enzymes, REX and REXA, release xylose from the reducing end of various xylooligosaccharides and are therefore termed as exo-xylanases and not β-xylosidases, due to their inability to cleave X2 and hydrolysis at the non-reducing end of the substrate. Furthermore, the GH8 endo-xylanases yield various products: OtXyn8A releases X2 as its main product, Xyn8 yields X3 and PhXyl produces predominant amounts of X3 and X4. Therefore it can be assumed that the catalytic domains of these enzymatic structures contain differing numbers of subsites giving rise to diversity in substrate binding. Although only a few GH8 endo-xylanases have been characterised to date, they have all shown high specificity in that they are only able to regulate productive binding with poorly substituted xylan substrates compared to highly decorated xylans.

The endo-xylanases, OtXyn8A and Xyn8 tested within this study both exhibited an endo-mode of action but to variable extents. Despite their 47% sequence identity to one another, OtXyn8A was more active on xylan and arabinoxylan polymeric substrates in comparison to Xyn8, possibly indicating that such polymeric substrates are possibly not the true substrates of Xyn8 in nature. This is quite consistent in terms of the cellular location of Xyn8, as it is an intracellular enzyme and therefore its preferred substrate is likely to be xylooligosaccharides entering the intracellular environment. Previous studies have indicated that Xyn8 is more active on substrates with a low degree of polymerisation compared to PhXyl which prefers substrates displaying a high degree of polymerisation (Pollet *et al*., 2010). This may be the case occurring within this current study and OtXyn8 and PhXyl are

quite comparable in that sense. However, the relatively low sequence identity of 27% between these two endo-xylanases does not justify the substrate similarity linking OtXyn8A and PhXyl and their structural features. Structural analysis of PhXyl in complex with substrate and product revealed that the enzyme's substrate specificity was determined by structural features at subsites +2, -1 and -2 despite the occurrence of six subsites (Petegem *et al*., 2003).

The substrate specificities of family GH8 enzymes can be compared to those of family GH5 in that the substrate specificity of GH5 xylanases is quite narrow but vary from one member to another. Similarly to family GH8, only a minority of xylanases are classified into GH5 even though this family is one of the larger families of glycoside hydrolases within the CAZy database, containing mainly cellulases, glucanases and mannanases. However, the GH5 xylanases including XYNA from *Erwinia chrysanthemi* (Vrsanská *et al*., 2007) and XYNC from *Bacillus subtilis* (St John *et al*., 2009), have shown activity against substituted xylans. Both of these enzymes have been thoroughly characterised and have a distinct action towards 4-*O*-methyl-D-glucuronoxylan as they require the presence of 4-*O*-methyl-D-glucuronosyl side residues for the hydrolytic action on the xylan backbone. Additionally, XYNIV from *Trichoderma reesei* has been reported to degrade substituted and unsubstituted xylans, releasing xylose as its main product (Parkkinen *et al*., 2004). Furthermore, a very recent GH5 xylanase was isolated from *Clostridium thermocellum* showing high activity against arabinoxylan with relatively limited activity against OSX, in comparison (Correia *et al*., 2011). This enzyme also did not have the capacity to degrade 4-*O*-methyl-D-glucuronoxylan or birchwood xylan, but very much displayed specificity for arabinoxylans and so was ultimately defined as an arabinoxylanase. This preference of highly substituted xylans that the GH5 xylanases acquire is different to that of the GH8 xylanases. Similarly, while none of the GH8 xylanases have been described to show activity against cellulose substrates, including that of OtXyn8A within this study, numerous GH5 xylanases have been reported to have both xylanase and cellulase activities (Collins *et al*., 2005). For example, a

xylanase from *Prevotella ruminicola* 23, with a high amino acid sequence identity to the catalytic domain of an endoglucanase exhibited endoglucanase, activity against the cellulose substrate, CMC, as well as xylan (Whitehead, 1993), although cellulase activity was relatively the lower activity of the two. Moreover, it was found that a gene encoding an endoglucanase from *Fibrobacter succinogenes* S85 encoded an enzyme with both endoglucanase and xylanase activities, as it hydrolysed both CMC and OSX (Cho *et al*., 2000).

It is interesting to note the high substrate specificities observed for both family 5 and family 8 xylanases, although enzyme functions are quite different to each other. The structural differences in catalytic folds of these enzyme families will account for such a substrate specificity distinction and further investigation into families GH5 and GH8 remains a significant challenge.

4.5 Comparison of the activity of OtXyn8A and characterised GH10 and GH11 xylanases

This section attempts to describe a comparative analysis of the xylanase properties between families GH8, GH10 and GH11 and to assess whether GH8 xylanases share any catalytic function or substrate specificity with the two major xylanase families. The GH8 xylanases established already demonstrate broad and diverse specificities within the family making this category of xylanases quite distinct from those members of GH10 and GH11. In the interest of comparing these xylanase families, this study focused upon the GH10 endo-xylanases, CmXyn10B, from *Cellvibrio mixtus* and CjXyn10A from *Cellvibrio japonicus and* the GH11 endo-xylanase, NpXyn11A, from *Neocallimastix patriciarum* alongside OtXyn8A. The GH10 and GH11 xylanases utilised in this study were kindly provided by Prozomix Ltd.

4.5.1 Structure/ function of family GH10 and GH11 xylanases

4.5.1.1 Structure/ function of CmXyn10B and CjXyn10A

The biochemical properties of CmXyn10B from *Cellvibrio mixtus* were first detailed by(Pell *et al*., 2004), describing this typical GH10 endo-xylanase as the most active GH10 xylanase reported to date. Pell *et al*., (2004), suggested that the activity of CmXyn10B against xylooligosaccharides increased with the degree of polymerisation and that the enzyme was considerably more active against smaller oligosaccharides in comparison to other GH10 xylanases. The activity of CmXyn10B against various xylooligosaccharides demonstrated that the enzyme had six subsites (+4, +3, +2, +1, -1 and -2). The increased catalytic activity observed for CmXyn8B was unlikely to be a result of interactions within the -1 and -2 subsites as the glycon region was described to be extremely similar to other GH10 xylanases. CmXyn10B also generally produced a range of oligosaccharides from poorly and highly decorated xylans. Through solvent exposure, the structure of CmXyn10B was

indicative of how side chains linked to O2 moieties of xylose are accommodated at the +1 subsite thus allowing xylan decoration to be tolerated by CmXyn10B and for subsequent degradation by accessory enzymes, such as α-glucuronidase.

CjXyn10A from *Cellvibrio japonicus* is a typical endo-xylanase randomly cleaving the β-1,4-glycosidic bonds of xylan (Charnock *et al*., 1997). CjXyn10A was reported to release a mixture of xylooligosaccharides at the initial stages of the reaction which were progressively degraded generating mostly X1, X2 and X3. CjXyn10A was originally thought to be an exo-xylanase by previous studies carried out by Harris *et al*., (1994) with findings were on crystallographic studies of CjXyn10A binding to xylopentose, suggesting that the xylanase contained the subsites -1 to +4. However, later studies by Charnock *et al*., (1997) predicted a sixth subsite (-2) as CjXyn10A more readily hydrolysed xylohexose than xylopentose. Furthermore, the -2 subsite was found to be essential for efficient hydrolysis of xylooligosaccharides but not polymeric xylan. This study suggested that subsites +4, +3, +2 and +1 bound to the polysaccharide while adjacent xylose residues filled the -1 subsite demonstrating that binding sites adjacent to the site of bond cleavage do not bind the substrate. Mutations of residues interacting with xylose in the -1 subsite lead to a major reduction in catalytic activity (Leggio *et al*., 2000).

The crystal structures of CmXyn10B and CjXyn10A were solved enabling the identification of catalytic residues conserved within the active site allowing investigation into the mechanism by which these enzymes function. The overall catalytic domain of both CmXyn10B and CjXyn10A was identified as a $(\beta/\alpha)_8$ barrel fold composed of 8 repeat units of α-helices and β-sheets. These enzymes are classical examples of the Clan GH-A enzymes, the clan in which the GH10 xylanases are categorised into. The catalytic acid-base and nucleophile residues of CmXyn10B were shown to be Glu157 and Glu262 (Pell *et al.,* 2004) respectively while the catalytic residues for CjXyn10A were Glu127 and Glu246 (Charnock *et al.,* 1997) which are both found at the carboxyl-terminal ends of β-sheets 4 and 7

respectively. These glutamic acid catalytic residues are usually conserved across family 10 xylanases.

4.5.1.2 Structure/ function of GH11 NpXyn11A

NpXyn11A was initially investigated by (Gilbert *et al*., (1992)who showed that the enzyme hydrolysed soluble xylan in a typical endo-manner. HPAEC-PAD analysis revealed that NpXyn11A released similar amounts of xylose and X2 while later studies carried out by (Fanutti *et al*., 1995), revealed that prolonged hydrolysis of both OSX and wheat arabinoxylan released X3 and X2 as the major hydrolysis products. The crystal structure of NpXyn11A revealed the presence of -3 to +3 subsites, explaining the high catalytic efficiency displayed by the enzyme in the generation of X3 (Vardakou *et al*., 2008).

On a structural basis, the enzyme displayed the β-jelly roll fold typical of GH11 xylanases (see section 1.9.2.1.2). Furthermore the two predicted catalytic residues, Glu113 (nucleophile) and Glu201 (acid-base) were located in the centre of the active site and separated by ~6 Å (Vardakou *et al.,* 2008) which is entirely consistent with the catalytic features required for the mechanism carried out by retaining glycoside hydrolases. Compared to other GH11 xylanases, NpXyn11A displayed unusually high activity which may possibly be a result of the extended substrate binding cleft of the enzyme (Vardakou *et al.,* 2008). The reported k_{cat} of NpXyn11A, 3.5 x 10^5 min^{-1}, on OSX is considerably higher than that of various GH11 xylanases previously studied such as NfXyn11A from *Nonomuraea flexuosa* and XynIII from *Acrophialophora nainiana* whose k_{cat} values were 2.3 min^{-1} and 5.95 min^{-1} respectively (Zhang *et al*., 2011; Cardoso *et al*., 2003).

4.5.2 Functional comparison of OtXyn8A, CmXyn10B, XYLA and NpXyn11A

To determine whether the hydrolysis behaviour varied between the GH8, GH10 and GH11 xylanases, the reaction products released by OtXyn8A, CmXyn10B, CjXyn10A and NpXyn11A from natural polymeric substrates, OSX and wheat arabinoxylan, were monitored over a prolonged period of time. Rates of hydrolysis were also measured (Figure 4.27). Reducing sugar analysis was carried out utilising the DNSA method with hydrolysis products identified by HPAEC-PAD. From the data shown in Figures 4.28 and 4.29, CmXyn10B, CjXyn10A and NpXyn11A showed typically endo-patterns of activity as was to be expected, with the early release of large xylooligosaccharides.

Although all xylanases, including OtXyn8A, shared the same mode of action, they produced different amounts of xylose and xylooligosaccharides at the earlier and prolonged stages of OSX hydrolysis. When comparing the reaction products of OtXyn8A with CmXyn10B, CjXyn10A and NpXyn11A, OtXyn8A appeared to be quite different in terms of the quantity of larger xylooligosaccharides generated.

Since CmXyn10B was cited as the most active GH10 xylanase, it was useful to compare the hydrolytic profile of such an established enzyme with OtXyn8A and to identify if the amount of sugar released by OtXyn8A was at all comparable to that of CmXyn10B. Also as CmXyn10B released X2 as one of its major product due to the -1 and -2 subsites it was important to monitor X2 release by CmXyn10B and OtXyn8A. CmXyn10B had a much higher rate of activity releasing larger quantities of xylooligosaccharides for the subsequent degradation into X2 (Figure 4.27). However OtXyn8A appeared to generate comparable quantities of X1 and X2 with minute traces of X3 to CmXyn10B after prolonged hydrolysis (Figure 4.29). This shows that OtXyn8A can possibly compete with the activity of CmXyn10B on xylan substrates resulting in purer X1 and X2 products which can be exploited in industrial biotechnological processes. On highly substituted xylans, however, CmXyn10B exhibited a more tolerant behaviour in comparison to OtXyn8A.

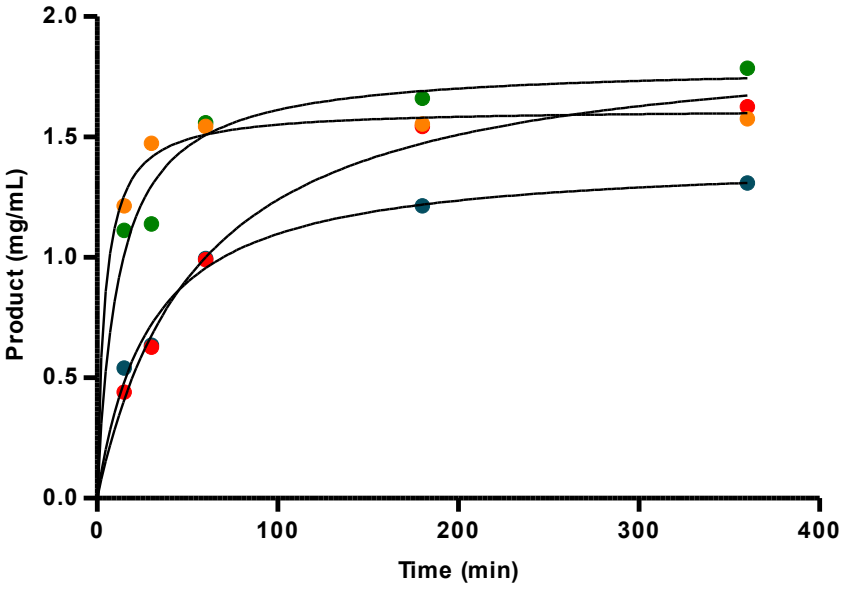


Figure 4.27 Rates of OSX hydrolysis by GH8, GH10 and GH11 endo-xylanases. OtXyn8A (0.2 µM) (●), CmXyn10B (0.2 µM), (●), CjXyn10A (0.2 µM) (●) and NpXyn11A (0.2 µM) (●) were incubated with OSX at a final substrate concentration of 8 mg/mL. Reactions were carried out in sodium acetate (0.05 M final concentration) pH 6 and BSA (1 mg/mL final concentration) at 25°C. At intervals 1 min, 2 min, 4 min, 5 min, 10 min, 20 min, 40 min, 60 min, 120 min and 60 min samples were removed and rates were measured by reducing sugar assay using DNSA reagent (see section 2.2.12).

The data present in Figures 4.28 and 4.30 revealed that CmXyn10B was highly active against wheat arabinoxylan and 4-*O*-methyl-D-glucuronoxylan respectively, unlike OtXyn8A which reduced levels of X1 and X2 were detected, suggesting differences between the GH10 and GH8 xylanases related to the hydrolysis of decorated xylans. The glycon subsites of solved xylanase structures are highly preserved within the GH10 xylanase family (Kaneko *et al*., 2004). Therefore the substrate specificity of the GH10 xylanases is most likely determined by the structure of the substrate in the aglycon region. The capacity to hydrolyse decorated xylans is conserved across the GH10 xylanases, as substituent side chains can be accommodated at the -2 and -3 subsites (Pell *et al*., 2004). The studies of CmXyn10B revealed that the identity of an aromatic residue, either tyrosine or

phenylalanine at the +1 subsite greatly influence the catalytic efficiency of the enzyme against glucuronoxylan which explains the results observed when comparing the activity of CmXyn10B against OtXyn8A. Additionally, substrate recognition and accommodation involving arabinose substituents seems to differ among the GH10 xylanases. The xylanase TaXyn10 from *Thermoascus aurantiacus* displayed a significant substrate specificity towards xylan decorated with arabinosyl units in which a higher activity was observed for the decorated substrate in comparison to the undecorated form of the substrate (Vardakou *et al.,* 2005). Specificity determinants contributing to productive substrate binding such as arabinose substituents are not a generic property of the GH10 xylanases however many members of this family are merely able to accommodate substituted xylans as observed with CmXyn10B within this study.

CjXyn10A displays properties typical of extracellular GH10 endo-xylanases and so an activity comparison this enzyme and OtXyn8A seemed to be quite reasonable. The data in Figure 4.29 shows that the OSX hydrolysis patterns of OtXyn8A and CjXyn10A were quite similar in that a mixture of xylooligosaccharides were generated at the initial stages of hydrolysis. However, while both enzymes exhibit a capacity to hydrolyse OSX, it was observed that CjXyn10A generated a rather substantial amount of xylooligosaccharides in comparison, while OtXyn8A yielded primarily X2 from its initial endo-products. The activity of CjXyn10A against highly decorated xylans is similar to other GH10 xylanases and the enzyme shows a higher activity towards arabinoxylan in comparison to OtXyn8A (Figure 4.28). The data implies the capacity of GH10 endo-xylanases, like CjXyn10A, in accommodating decorated xylans in subsites -2, -1 and +1 while sugar residues in the +2 and +3 subsites must be unsubstituted (Pell *et al*., 2004; Faulds *et al*., 2006).

When monitoring the OSX hydrolysis products released by NpXyn11A, the quantity of various xylooligosaccharides released throughout the reaction by NpXyn11A was very significant (Figure 4.29). Initially, a mixture of xylooligosaccharides was generated but X3

was predominantly the major product released by NpXyn11A demonstrating the importance of the +3 and -3 subsites with the end products being mainly X1 and X2 which is in agreement with findings reported by Gilbert *et al*., (1992). The data in Figures 4.28 and 4.29 illustrates that NpXyn11A had a higher activity against OSX in comparison to wheat arabinoxylan as far less of the end products were generated demonstrating that such a xylanase from GH11 preferably cleaves the unsubstituted regions of the xylan polymer. The substrate selectivity of decorated xylans by GH11 xylanases has been previously recorded from numerous observations indicating that xylans with a high degree of substitutions are not fully hydrolysed by such xylanases (Biely *et al*., 1997; Paës *et al*., 2011). Therefore, the results from this study are not surprising. Vardakou *et al*., (2008) suggested that the O3 of xylose interacting with residues in the -2 subsite of NpXyn11A could be linked to an arabinofuranoside substituent as it was found that xylotriose was the major decorated product released from cereal-derived arabinoxylan hydrolysis. This may be so, however from the data presented in Figure 4.28, it can be seen that larger decorated xylooligosaccharides accumulated from wheat arabinoxylan degradation by NpXyn11A and no xylotriose was detected after prolonged hydrolysis. Either way, it can be stated that even if GH11 xylanases can accommodate decorations linked to xylose in one subsite, the remaining subsites are unable to do this, limiting the action of such xylanases on highly substituted xylans. This incomplete hydrolysis on decorated xylan was similarly observed for OtXyn8A indicating a comparable functional property between the GH11 and GH8 families. However, when comparing the hydrolytic profiles of NpXyn11A and OtXyn8A it was observed that the GH11 xylanase was more active on arabinoxylan than the GH8 enzyme suggesting that different subsites accommodate decorated xylose residues to different extents between the two xylanase families. Solvent exposure revealed that the +1 and -1 subsites of NpXyn11A were not solvent accessible suggesting that side chains cannot be accommodated by these sites which also relates to the narrow shape of the catalytic cleft.

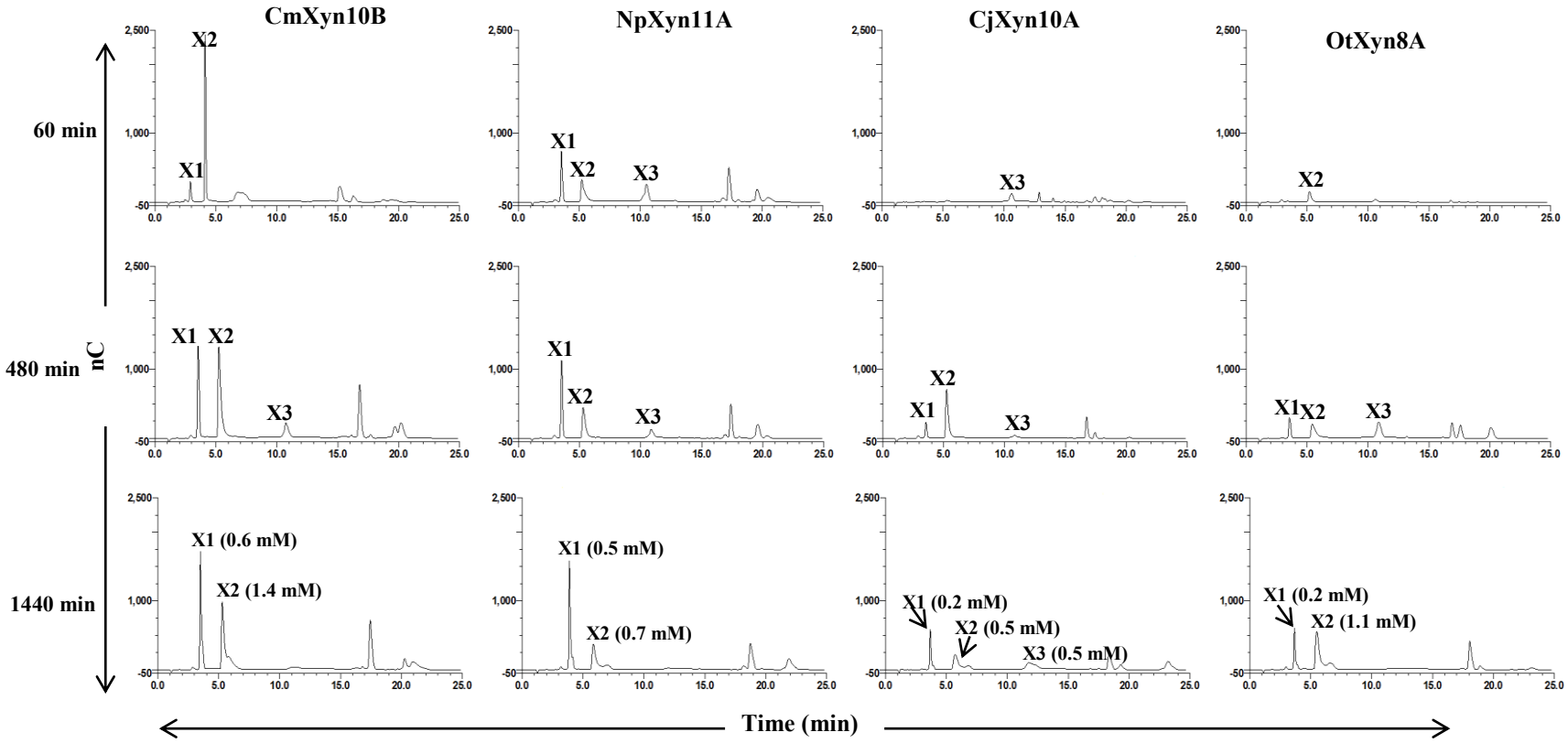


Figure 4.28 Hydrolysis of wheat arabinoxylan. CmXyn10B (0.2 µM), NpXyn11A (0.2 µM), CjXyn10A (0.2 µM) and OtXyn8A (0.2 µM) was incubated with wheat arabinoxylan at a final substrate concentration of 8 mg/mL. Reactions were carried out in sodium acetate (0.05 M final concentration) pH 6 and BSA (1 mg/mL finalconcentration) at 25°C. At intervals 15 min, 60 min, 180min, 480 min and 1440 min samples were removed and subjected to HPAEC-PAD analysis (see section 2.2.12). Hydrolysis products of last time point have been quantified (see section 2.2.12.1).

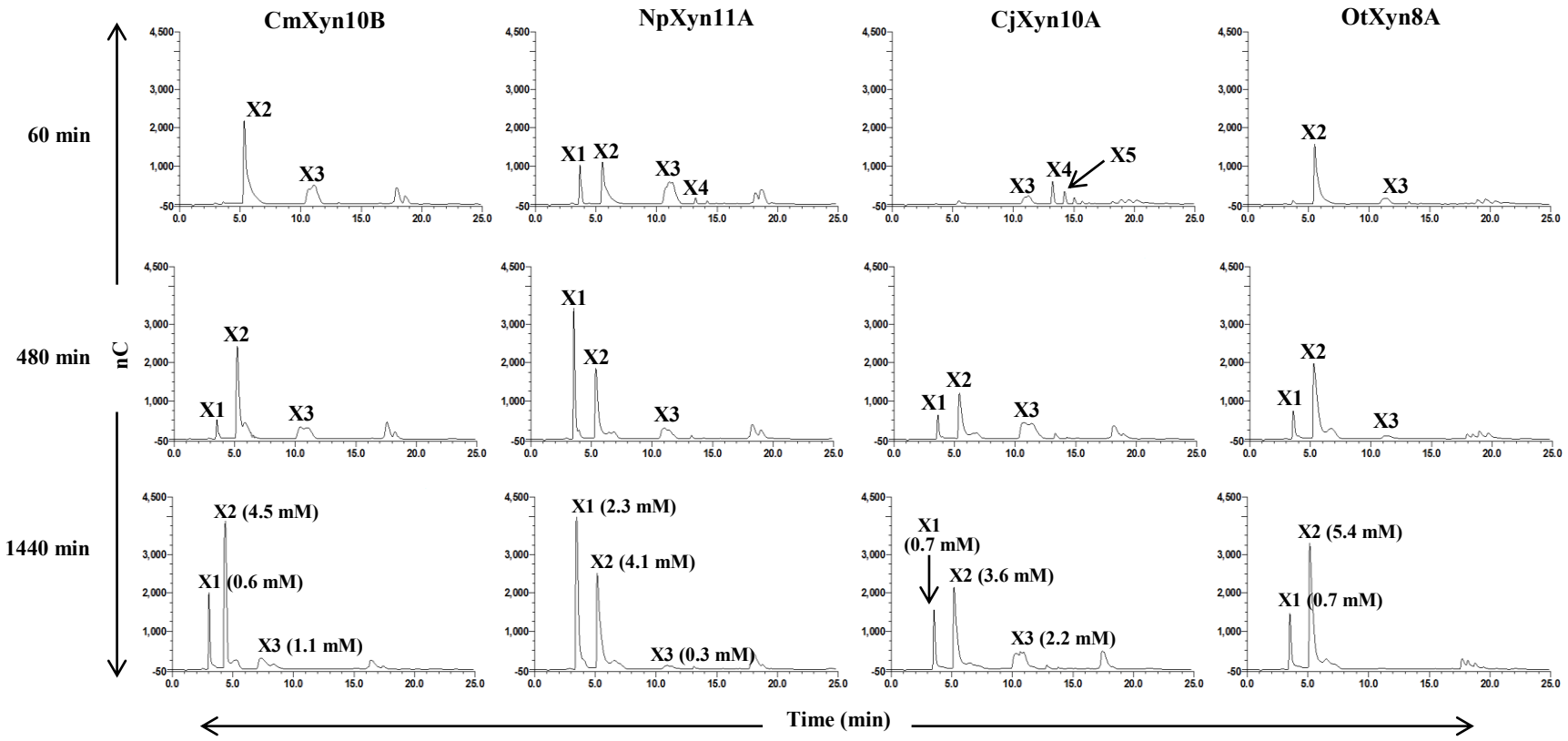


Figure 4.29 Hydrolysis of OSX. CmXyn10B (0.2 µM), NpXyn11A (0.2 µM), CjXyn10A (0.2 µM) and OtXyn8A (0.2 µM) was incubated with OSX at a final substrate concentration of 8 mg/mL. Reactions were carried out in sodium acetate (0.05 M final concentration) pH 6 and BSA (1 mg/mL final concentration) at 25°C. At intervals 15 min, 60 min, 180min, 480 min and 1440 min samples were removed and subjected to HPAEC-PAD analysis (see section 2.2.12). Hydrolysis products of last time point have been quantified (see section 2.2.12.1).

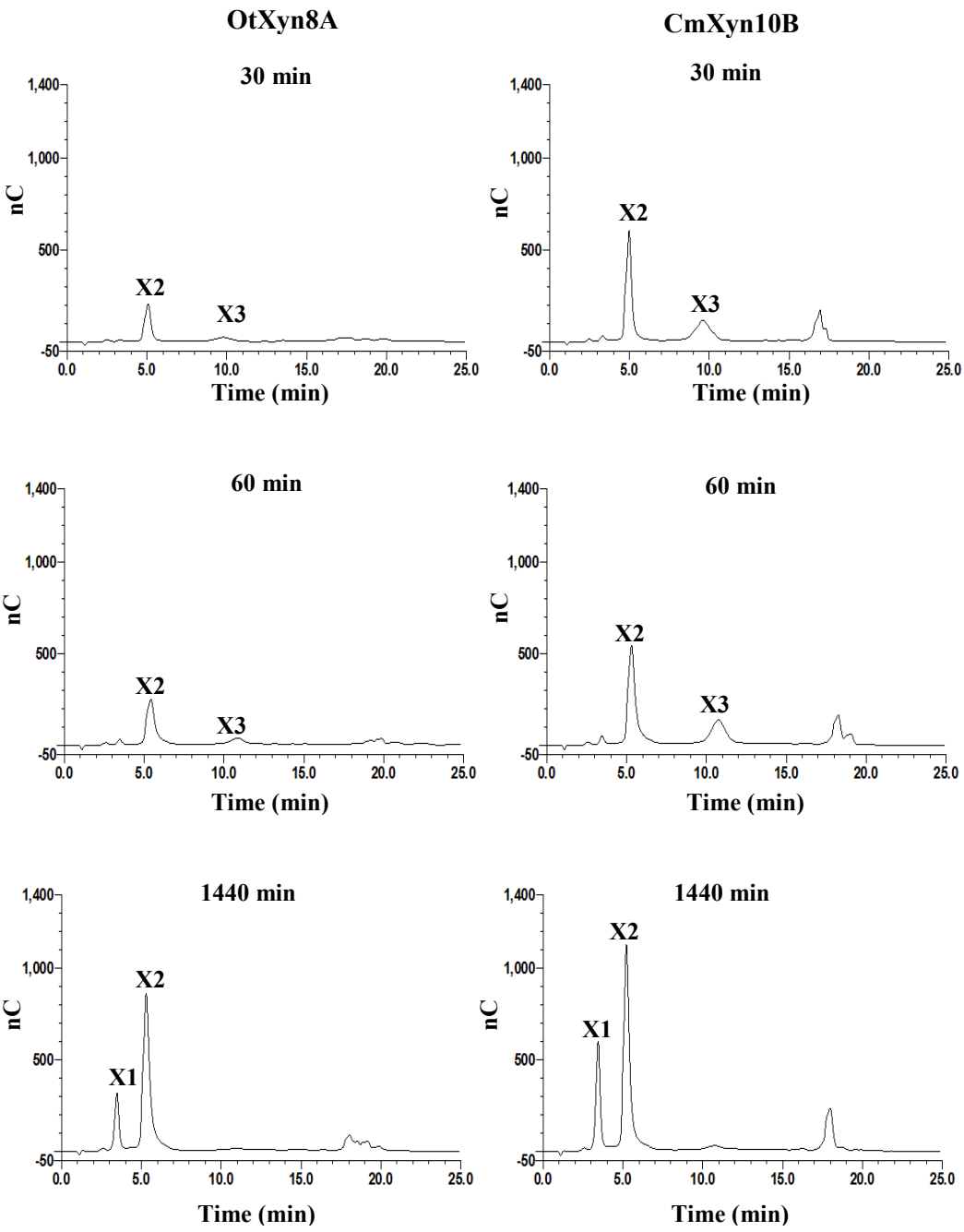


Figure 4.30 Hydrolysis of 4-*O*-methyl-D-glucuronoxylan. OtXyn8A (0.2 µM) and CmXyn10B (0.2 µM) were incubated with 4-*O*-methyl-D-glucuronoxylan at a final substrate concentration of 4 mg/mL. Reactions were carried out in sodium acetate (0.05 M final concentration) pH 6 and BSA (1 mg/mL final concentration) at 25°C. At intervals 30 min, 60 min and 1440 min samples were removed and subjected to HPAEC-PAD analysis (see section 2.2.12).

4.5.3 Biochemical comparison

The kinetic constants against OSX of the investigated enzymes were used to populate Table 4.5. The observation from Figure 4.30 revealed that sugar release in the earlier stages of xylan hydrolysis was quite different for OtXyn8A and the other GH10 and GH11 enzymes as OtXyn8A exhibited a lower catalytic efficiency (k_{cat} / K_M) in comparison to the other enzymes.

Enzyme	k_{cat} **(min^{-1})**	K_M **(mg mL^{-1})**	k_{cat} **/** K_M **(min^{-1} mg^{-1} mL)**
OtXyn8A	7661 $\pm$ 23	6.84 $\pm$ 0.66	11420 $\pm$ 9
CmXyn10B	19900*	6.17 *	3225*
CjXyn10A	41362*	1.29*	32063*
NpXyn11A	350000*	4.2*	84000*

Table 4.5 Michaelis-Menten kinetic parameters for the GH8, GH10 and GH11 xylanases for the hydrolysis of OSX. The reaction of OtXyn8A on OSX was performed in triplicate at 45°C in 0.05 M sodium acetate buffer, pH 6 and BSA was included at a final concentration of 1 mg/mL. DNSA analysis was carried out as described in section 2.2.11.2 to analyse kinetic values. *Kinetic parameters reproduced from Pell *et al*., (2004), Charnock *et al*., (1997) and Vardakou *et al*., (2008) respectively.

4.6 Conclusion

Evaluation of the xylanases representing members from families 8, 10 and 11 under the same conditions has shown significant variations in structure-function relations between these three groups of glycoside hydrolases. On a structural basis, these three xylanase families are unrelated with GH10 xylanases comprising a $(\beta/\alpha)_8$ barrel fold, GH11 xylanases comprising a β-jelly roll and GH8 xylanases displaying an $(\alpha/\alpha)_6$ barrel architecture. In addition, family GH8 xylanases catalyse hydrolysis with inversion of anomeric configuration whereas GH10 and GH11 enzymes catalyse hydrolysis via a double displacement mechanism resulting in retention of the anomeric configuration.

Characterisation of the GH8 xylanases OtXyn8A, Xyn8, REX and REXA has shown variation in substrate specificities within the GH8 family and such intra-family xylanase diversity is quite distinct from that of the GH10 and GH11 xylanases. It could be suggested however, that individual GH8 enzymes are highly specific differing from their GH10 counterparts as family 10 enzymes can accommodate both poorly and heavily substituted substrates. This was clearly demonstrated between OtXyn8A and CmXyn10B during this research as well as from previous studies (Collins *et al*., 2002; Pollet *et al*., 2010). In terms of function, the GH8 endo-xylanases are quite similar to GH11 endo-xylanases in that the degree of substrate substitution greatly influences the activity of the enzyme with larger products being released in comparison to the family GH10 enzymes. The GH10 xylanases have a lower specificity for heteroxylans with a high degree of substitution accommodating arabinose and methyl-D-glucuronic acid. Meanwhile, it has been shown that GH11 xylanases and to an extent GH8 xylanases, preferentially cleave unsubstituted regions of the xylan backbone. The functional property by which GH11 and GH8 xylanases are unable to cleave glycosidic linkages close to substituents on the xylan main chain was elucidated in the structure of the active site. A structural feature of the catalytic domain of GH11 enzymes is a long deep cleft, too narrow for +1 and -1 subsites to accommodate side chains. In

comparison, the structural feature hindering the GH8 endo-xylanase, PhXyl, is the occurrence of an extended loop folding over the glycon region decreasing the accessibility of the active site for decorated substrates (Pollet *et al*., 2010). In contrast, the active sites of GH10 xylanases exhibit shallow grooves and are able to accommodate side groups at the +1 subsite (Pell *et al*., 2004) as well as arabinose substituents linked to O3 in the distal -2 and -3 subsites (Fujimoto *et al*., 2004).

The examination of the relevant literature has revealed that GH8 xylanases are not characterised as well as their GH10 and GH11 counterparts. This current study provided more information about GH8 enzymes, demonstrating new insights into the varying substrate specificities exhibited by the family 8 xylanases. Thus, suggesting that predictions into catalytic and structural features of this newly identified enzyme cannot be solely based on the current characterisations of other family 8 members.

4.7 Further work

More data is necessary to achieve comprehensive structure-function analysis of OtXyn8A and so numerous opportunities for future studies exist in this area of research. The solved 3D structure of OtXyn8A would be a major asset contributing to the characterisation of the GH8 xylanases. Mutational analysis of specific catalytic residues within OtXyn8A active site in conjunction with the crystallisation complexes of xylan oligosaccharides or xylobiose would facilitate the interaction of amino acids with the substrate.

5 Results and Discussion

Synergy between enzymes from *Opitutus terrae* PB90-1

5.1 Introduction

The organisation of functionally related genes is a common feature within bacterial genomes in which such genes frequently reside in clusters. These clusters have been described to encode various functions that are essential to the microorganism, such as the function to catabolise particular substrates for energy production (Fischbach *et al*., 2010). Since early linkage studies of *E. coli* (Stahl *et al*., 1966), gene clustering has become a well documented feature of bacterial genomes and furthermore genes have been found to reside in such clusters in the order of the biochemical pathway they catalyse (Lawrence *et al*., 1996). Additionally, genes in clusters can be regulated under common control and are transcribed together via a shared promoter and operator sequence is utilised. Such a structured cluster is referred to as an operon, although not all clusters are recognised as operons as such genes can also be transcribed independently (Martin *et al*., 2009). This type of gene clustering usually occurs in eukaryotes including yeasts and filamentous fungi in which genes are not transcribed as a single mRNA (Osbourn *et al*., 2009). Notably, operons are not highly conserved with only a small number predicted within bacterial genomes. Most data suggests that in the instance of bacterial operons, there is an abundance of rare gene clusters with unique directions of genes that are separated by short intergenic spaces and are transcribed in the same direction (Koonin, 2009; Osbourn *et al*., 2009). Furthermore it has been suggested that the degree of operon distribution varies greatly among genomes (Koonin, 2009).

One of the major examples of an operon regulating the degradation of plant cell wall polysaccharides is the *cip-cel* operon found in *Clostridium cellulolyticum* in which genes responsible for the catabolism of cellulose reside (Abdou *et al*., 2008). The *cip-cel* cluster comprises 12 genes all of which are transcribed together producing cellulase enzymes,

mostly endoglucanases that act on crystalline cellulose. Additionally, this operon is described to be essential to the cellulosome complex (section 1.8) for the efficient degradation of crystalline cellulose. A newly identified *xyl-doc* operon from the same microorganism has also been described which is specialised for the degradation of hemicellulose (Blouzard *et al*., 2010).

This chapter will describe a gene cluster residing in the genome of *Opitutus terrae* PB90-1 in which genes encode proteins specific for xylan degradation. Enzyme cocktails that can be utilised to hydrolyse plant cell wall polysaccharides are becoming a widespread focus in the degradation of lignocellulosic biomass for the production of biofuel. Therefore, the identification of such a gene cluster can be beneficial to such an application as it may be a unique enzyme system for the efficient degradation of specific substrates.

5.2 *Opitutus terrae* PB90-1 gene cluster and approach taken for this study

The current research into OtXyn8A (Chapter 4), a novel GH8 endo-xylanase, lead to the discovery of a gene cluster composed of putative glycoside hydrolases within *O. terrae*. By identifying where the xylanase gene was located within the genome, it was observed that OtXyn8A resided next to a further 3 genes, all related to proteins utilised for xylan degradation. These included a gene encoding a putative GH52 β-xylosidase (EC 3.2.1.37), a GH67 α-glucuronidase (EC 3.2.1.131) and another GH8 glycoside hydrolase.

In keeping with the nomenclature proposed for plant cell wall degrading enzymes (Henrissat *et al*., 1998) the genes and their products of this study were designated with three letters in accordance to their substrate preference. Gene and enzyme names also provide information to which families each protein belongs and are presented in Table 5.1.

Putative glycoside hydrolase	**Gene**	**Enzyme name**	**Accession number**
GH8 endo-xylanase	*xyn8A*	OtXyn8A	ACB77586.1
GH8 endo-xylanase	*xyn8B*	OtXyn8B	ACB77857.1
GH52 β-xylosidase	*xyl52A*	OtXyl52A	ACB77584.1
GH67 α-glucuronidase	*glca67A*	OtGlcA67A	ACB77588.1

Table 5.1 Names for genes and encoded enzymes. Accession codes are given so genes and respective proteins can be identified in NCBI.

Within the gene cluster identified, *xyl52A* was located upstream of a sugar transporter which in turn was flanked by the remaining genes, *xyn8A*, *xyn8B* and, *glca67A* (Figure 5.1). The intergenic space between *xyn8A* and *xyn8B* was very short (only 3 nucleotides) with a larger space (79 nucleotides) residing between *xyn8B* and *glca67A*. Furthermore, two *lac*1 family transcription regulators were immediately upstream of the gene cluster with close proximity to *xyl52A* (~ 200 nucleotides).

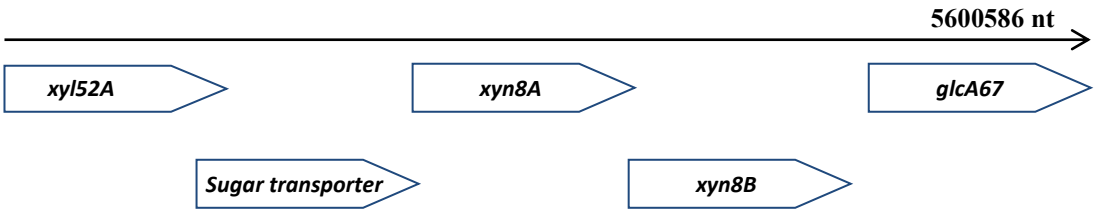


Figure 5.1 Organisation of genes within *O. terrae* cluster.

It has been suggested for rapid and reliable gene expression that transcription factor genes are usually located near the genes they regulate enabling coordinated binding and regulation (Osbourn *et al*., 2009). Therefore it is possible that only a single mRNA may transcribe these gene products. The closeness of the genes also suggests that they may interact synergistically in the efficient breakdown of xylan substrates; hence the reasoning for resolving the activities exhibited by the individual enzymes. A xylanolytic gene cluster was identified previously in *Prevotella ruminicola* 23 (Dodd *et al*., 2009) in which two GHs were isolated and found to function synergistically in releasing monomeric sugars from xylan. The gene encoding the product Xyl3A, a predicted β-xylosidase, was observed to improve the release of xylose from xylooligosaccharides produced by Xyn10D-Fae1A, a bifunctional enzyme exhibiting both endo-xylanase and esterase activities.

As the *O. terrae* genes had already been scoped for putative glycoside hydrolases, the genes within this cluster had already been targeted for this research. Therefore, in addition to OtXyn8A, the xylanolytic genes encoding OtXyl52A and OtGlcA67A had already been successfully cloned and expressed in *E. coli* BL21 (Section 3.4, Figures 3.4 - 3.6). However, while the cloning of gene product OtXyn8B was achieved, soluble expression of the protein was unsuccessful even though many attempts were carried out to obtain this GH8 enzyme. The study therefore aimed to determine synergism between the remaining three recombinant xylanolytic enzymes of this cluster in the hydrolysis of oat spelt xylan (OSX) and 4-*O*-methyl-D-glucuronoxylan. Sugar products were analysed by HPAEC-PAD analysis under standard conditions (see section 2.2.12).

5.2 Characterisation of OtXyl52A

5.2.1 Sequence analysis

Bioinformatics of the nucleotide and amino acid sequence was carried out as described in Chapter 3 for the subsequent cloning and expression of OtXyl52A. Amino acid sequence analysis of OtXyl52A revealed that the protein was mostly related to a GH52 β-xylosidase from *Thermoanaerobacterium xylanolyticum* LX-11 with an identity of 44 %. Furthermore, OtXyl52A was predicted to lack a signal peptide indicated that it was located intracellularly. The occurrence of intracellular β-xylosidases supports the concept that xylooligosaccharides produced from the action of extracellular endo-xylanases on xylan are transported directly into the cell where they are hydrolysed by β-xylosidases in an exo-fashion for the release of xylose (Tsujibo *et al*., 2004). Additionally, the gene encoding the sugar transporter within the same gene cluster could presumably regulate the transport of such xylooligosaccharides across the membrane.

5.2.2 Biochemical characterisation

OtXyl52A was assayed against the synthetic substrates 4-nitrophenyl-β-D-xylopyranoside and 4-nitrophenyl-α-L-arabinofuranoside to detect initial activity of the enzyme. Due to the structure of 4-nitrophenyl-β-D-xylopyranoside, this substrate is not hydrolysed by endo-xylanases and so sugar monomers released are usually attributed to xylosidase activity alone. Upon incubation with 4-nitrophenyl-β-D-xylopyranoside, there was a significant colour change of substrate indicating that OtXyl52A was active against the 4-nitrophenol substrate. This result was as expected, as the enzyme was originally annotated as a β-xylosidase and no activity was detected against 4-nitrophenyl-α-L-arabinofuranoside. Due to the activity of the OtXyl52A against 4-nitrophenyl-β-D-xylopyranoside, a thorough biochemical characterisation could be carried out with the determination of kinetic parameters. β-

Xylosidase activity exhibited by OtXyl52A was optimal at pH 5.5 and 55°C (Figures 5.2 and 5.3 respectively). The optimum temperature of activity was measured under standard conditions (see section 2.2.11.1) for 2 min with temperature ranging from 30 – 70°C and inactivated with 3 M Tris. The absorbance of each sample was then determined at 400 nm. Absorbances obtained were converted to concentrations using the published molar absorptivity coefficient of 18.3 mM^{-1} cm^{-1} for 4-nitrophenol (Kezdy *et al*., 1962).

The enzyme remained stable up to 30°C but no activity was detected at 50 – 70°C (Figure 5.4). The thermostability was determined by incubating OtXyl52A at various temperatures (25°C, 30°C, 35°C, 40°C, 45°C and 50°C) for 20 min before assaying with 5 mM 4-nitrophenyl-β-D-xylopyranoside for 2 min. The reaction was inactivated with 3 M Tris and absorbance of each sample read at 400 nm.

The pH and temperature optimum of β-xylosidases has been identified to be in the range of 3 – 5 and 50 – 70°C respectively (Knob *et al*., 2009). Numerous β-xylosidases have exhibited similar enzymatic properties to OtXyl52A including a GH52 β-xylosidase from *Aeromonas caviae* ME-1 which was stable up to 40°C and showed highest activity at 50°C (Suzuki *et al*., 2001). β-Xylosidases have also been found to display highest activity at elevated temperatures and have even retained stability at such temperatures. For example, a β-xylosidase from *Aspergillus sydowii* was found to have an optimum temperature and possess thermostability at 85°C (Ghosh *et al*., 1993).

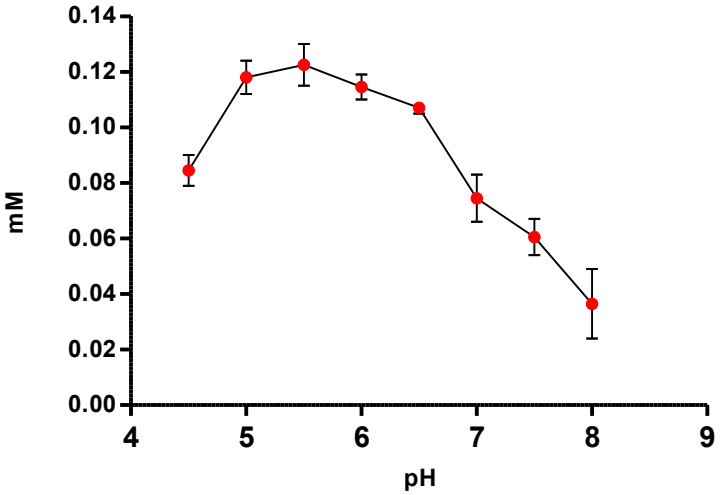


Figure 5.2 The effect of pH on the rate of OtXyl52A activity against 4-nitrophenyl-β-D-xylopyranoside. Reactions were carried out in duplicate with 0.6 µM of OtXyl52A at a final substrate concentration of 5 mM. Sodium phosphate and sodium acetate buffers (0.05 M final concentration) were used to cover pH range (4.5 - 8). BSA (1 mg/mL) was also added to reaction mix and reactions were carried out as described in section 2.2.11.1.

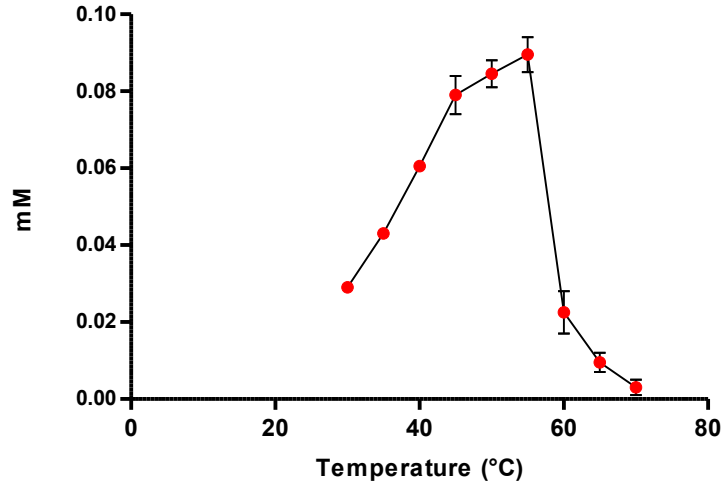


Figure 5.3 The effect of temperature on the rate of OtXyl52A activity against 4-nitrophenyl-β-D-xylopyranoside. Reactions were performed in duplicate with 0.6 µM of OtXyl52A at 30, 35, 40, 45, 50, 55, 60, 65 and 70°C at a final substrate concentration of 5 mM. Sodium phosphate buffer pH 6.5 (final concentration 0.05 M) and BSA (1 mg/mL) was added to the reactions as described in section 2.2.11.1.

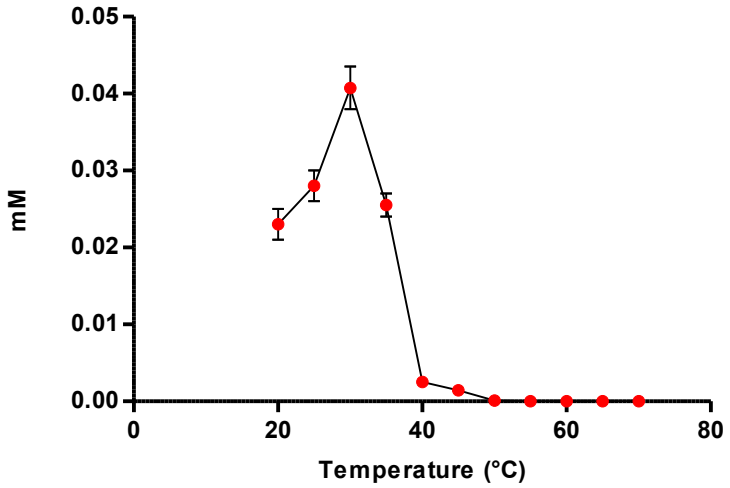


Figure 5.4 The effect of temperature on OtXyl52A stability as measured by activity against 4-nitrophenyl-β-D-xylopyranoside. Reactions were carried out in duplicate with 0.6 µM of OtXyl52A at 35°C, at a final substrate concentration of 5 mM after pre-incubation at 20, 25, 30, 35, 40, 45, 50, 55, 60, 65 and 70°C. Sodium phosphate buffer pH 6.5 (final concentration 0.05 M) and BSA (1 mg/mL) was added to the reactions as described in section 2.2.11.1.

5.2.3 Kinetic analysis

To determine apparent kinetic parameters, OtXyl52A was incubated with 4-nitrophenyl-β-D-xylopyranoside in 0.05 mM sodium phosphate buffer (pH 6.5) at 35°C. Substrate concentrations spanning the K_M were chosen. In order to ascertain if data reflected true Michaelis-Menten kinetics, a Lineweaver-Burk plot was constructed and used to determine values of K_M, k_{cat} and k_{cat} / K_M (Figure 5.5.). The apparent K_M and k_{cat} of OtXyl52A were found to be 0.9130 mM and 1417.16 min^{-1} respectively with a k_{cat} / K_M of 1552.7 min^{-1} mM^{-1}. The most catalytically active β-xylosidase to date was isolated from *Selenomonas ruminantum* by Jordan *et al*. (2007) which displayed the highest activity against xylobiose (X2) with a k_{cat} of 3030 min^{-1}. Furthermore, the k_{cat} value was found to successively decrease with increasing oligomierisation, demonstrating the catalytic importance of the -1 and +1 subsites of such an enzyme.

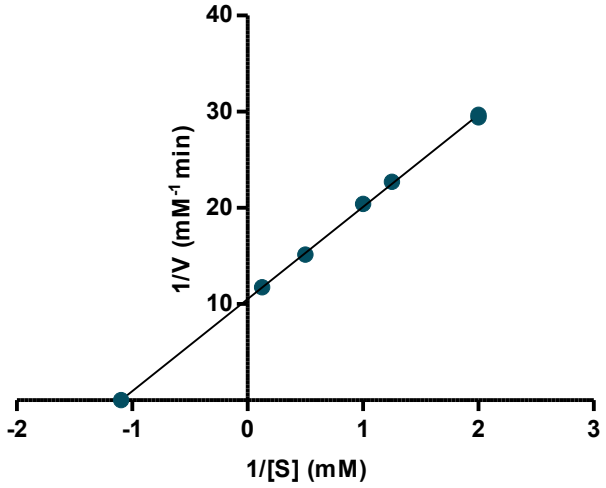


Figure 5.5 Lineweaver-Burk plots for OtXyl52A against 4-nitrophenyl-β-D-xylopyranoside. The reactions were performed with 0.6 µM of OtXyl52A at 35°C in 0.05 M sodium acetate buffer, pH 6.5 including BSA at a final concentration of 1 mg /mL.

5.2.4 Mode of action

β-Xylosidases ultimately cleave X2 and short xylooligosaccharides from the non-reducing end releasing xylose; making these accessory enzymes pivotal for the complete degradation of xylan (Shallom *et al*., 2003). The mode of action of OtXyl52A was essentially determined against natural substrates and through the release of reducing sugar. By DNSA assay, increased levels of reducing sugar were detected upon incubation of OtXyl52A with xylooligosaccharides produced from pre-treated OSX and wheat arabinoxylan by an endo-xylanase (OtXyn8A). HPAEC-PAD analysis of products generated from OSX digestion by OtXyn8A and OtXyl52A revealed a substantial amount of xylose after just 1 h of hydrolysis (Figure 5.6). This was clearly indicative of the β-xylosidase activity exhibited by OtXyl52A as the enzyme cleaves X2 which is preferentially released by OtXyn8A. After 24 h of hydrolysis, the only product detected, not including decorated residual sugars, was xylose indicating the end of the reaction and more importantly revealing the β-xylosidase activity of OtXyl52A against products released by OtXyn8A.

OtXyl52A was additionally compared to the activity of OtXyl43A, a different *O. terrae* β-xylosidase produced within this research belonging to family GH43. It should be noted that the gene encoding OtXyl43A was not located within the same organisation of genes surrounding OtXyl52A and OtXyn8A. The products released by both β-xylosidases from xylooligosaccharides produced by OtXyn8A were determined to give an insight into the possible synergy exhibited by enzymes produced from the investigated gene cluster. HPAEC-PAD analysis showed that OtXyl43A released only half the amount of xylose from OtXyn8A-treated OSX compared to OtXyl52A (Figure 5.6). Additionally, substantial amounts of X2 accumulated after 24 h hydrolysis by OtXyl43A (Figure 5.6). Therefore, it could be suggested that OtXyl52A showed greater synergy with OtXyn8A for the efficient degradation of OSX into xylose. Furthermore, the two β-xylosidases clearly demonstrated

various affinities for X2 produced by OtXyn8A which may affected by different enzymatic mechanisms and/or structural determinants.

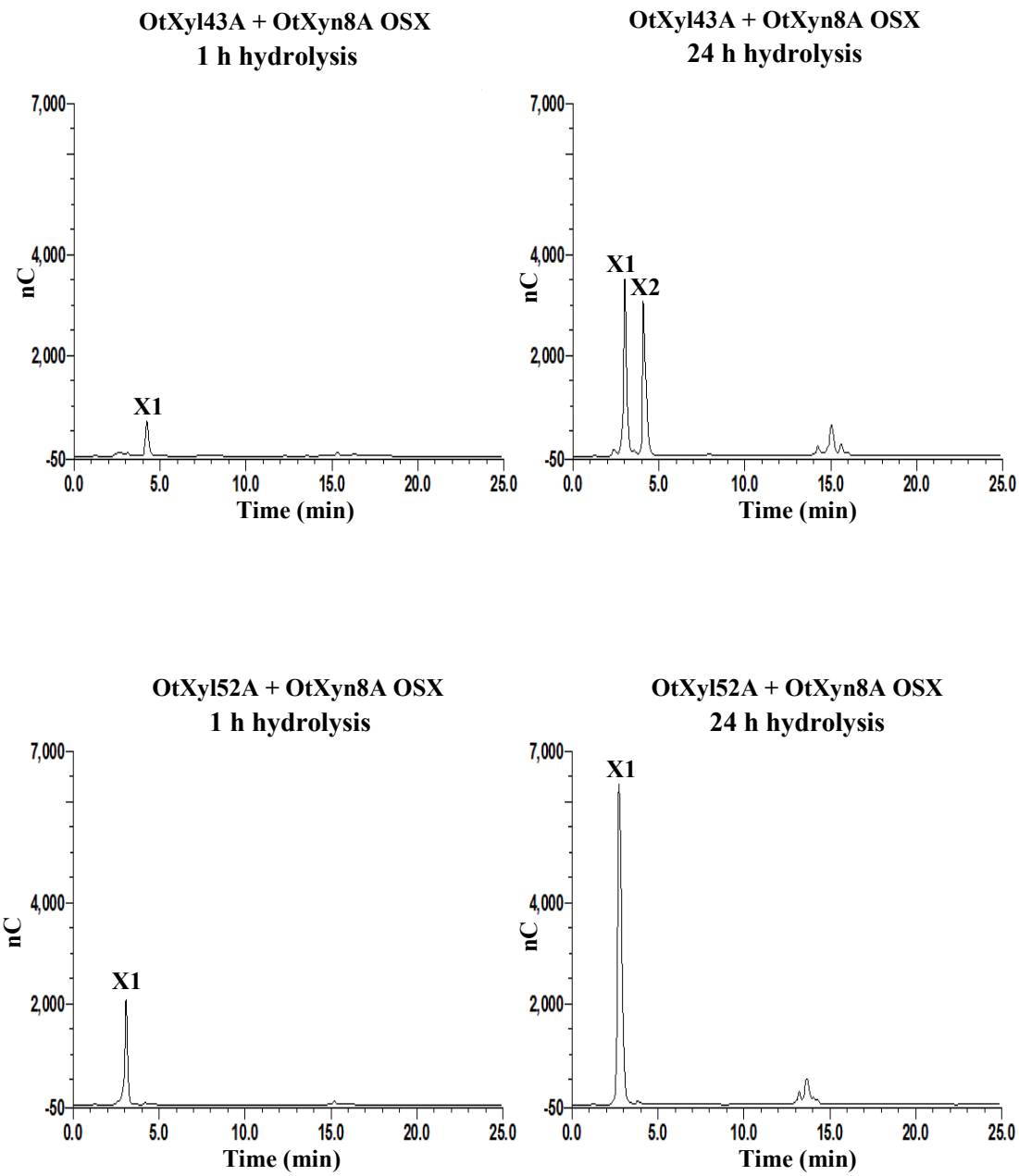


Figure 5.6 Analysis of synergy of OtXyn8A and OtXyl52A or OtXyl43A on OSX. OtXyl52A (0.6 µM) and OtXyl43A (0.6 µM) were incubated with OSX (final concentration 8 mg/mL). Reactions were carried out with OtXyn8A (0.2 µM) in sodium acetate (0.05 M final concentration) pH 6 and BSA (1 mg/mL final concentration) at 25°C. At 1 h and 24 h, a sample was removed and subjected to HPAEC-PAD (see section 2.2.12).

Based on primary sequence similarities, β-xylosidases, in general, have been classified into GH families 3, 39, 43, 52 and 54 although characterised β-xylosidases are quite limited in family GH52. To date there is only one three-dimensional structure available for family 52 β-xylosidases and this was solved by Contreras *et al*., (2008). This study reported that XynB2 from *Geobacillus stearothermophilus* was a dimeric protein whose active site was formed by two protomers. Furthermore, this study demonstrated how modifications to pH and temperature as well as chemical denaturants all affected the activity of XynB2. However, information on the mechanism by which this β-xylosidase or any other GH52 member cleaves the target substrate is still quite vague and the two corresponding catalytic residues are yet to be elucidated. Structural features of XynB2 include an active site comprising of an α- and β- helix to which have been suggested to be generally conserved among GH52 members.

Another property of the GH52 family members is their potential to catalyse hydrolysis via the retaining mechanism. This mechanism implies that GH52 β-xylosidases perform both hydrolysis and transglycosylation whereas the inverting GH43 β-xylosidases can only carry out hydrolysis (Smaali *et al*., 2006). A GH52 β-xylosidase, XysB, from *Aeromonas caviae* ME-1 showed significant transglycosylation activity against X3 where substantial amounts of X4 was produced (Suzuki *et al*., 2001). However, when a mixture of xylose, X2 and X3 were used as the substrate, larger products were not detected during incubation with the enzyme. No transglycosylation products were observed during the reaction with OtXyl52A against xylooligosaccharides within this study and this may be because when a mixture of xylooligosaccharides is present as the substrate the balance of the reaction is shifted towards xylose production. Additionally, there is evidence of β-xylosidases showing little or no transglycosylation against X2 which has been demonstrated by a GH39 retaining β-xylosidase from *Bacillus halodurans* C-125 (Smaali *et al*., 2006). This may explain the observation made for OtXyl52A in which it fully degrades X2 into xylose with no production of X3 in the earlier or prolonged stages of hydrolysis. On a physiological basis,

OtXyl52A is distributed intracellularly within the cell cytoplasm of *O. terrae*. Therefore, its main substrate could simply be X2 produced by the endo-xylanase OtXyn8A which is then transported across the cell membrane, emphasising the synergy between these enzymes.

Further emphasis that OtXyl52A and OtXyn8A work synergistically together is the reduced xylose production observed for the GH43 β-xylosidase, OtXyl43A, used within this study. Comparing the activities of the two different β-xylosidases against xylooligosaccharides generated by OtXyn8A revealed that OtXyl43A did not completely degrade X2 into xylose after prolonged incubation which was unexpected. Family GH43 β-xylosidases utilise the inverting mechanism for catalysis meaning they do not have the capacity to carry out transglycosylation. Furthermore, the most catalytically efficient xylosidase characterised to date falls under this family of enzymes (Knob *et al*., 2009; Jordan *et al*., 2007). There may be many reasons as to why OtXyl43A was unable to efficiently hydrolyse the xylooligosaccharides into xylose. For example, the catalytic rate may have been inhibited with high concentrations of xylose or the pH utilised within the experiment may have influenced the activity. Many β-xylosidases have been reported to be inhibited by xylose including that of highly active β-xylosidases such as the GH43 β-xylosidase from *Selenomonas ruminantium* in which high levels of accumulating xylose strongly inhibited the performance of the enzyme (Jordan *et al*., 2007). However, some β-xylosidases have demonstrated a high tolerance to high xylose levels including a β-xylosidase from *Scytalidium thermophilum* (Zanoelo *et al*., 2004).

5.3 Characterisation of OtGlcA67A

5.3.1 Sequence analysis

Bioinformatics of the nucleotide and amino acid sequence was carried out as described in Chapter 3 for the cloning and expression of OtGlcA67A. The deduced amino acid sequence was compared to other proteins and a BLAST search revealed that OtGlcA67A shared the closest identity to a family 67 α-glucuronidase from *Rhodothermus marinus* DSM 4252. SignalP analysis predicted an encoded peptide sequence. It was therefore assumed that OtGlcA67A was an extracellular enzyme. Many publications have reported the range of substrates against which α-glucuronidases are active . Some α-glucuronidases have been found to display specificity to polymeric substrates while others favour short xylooligosaccharides. Furthermore, it has been well established that fungal α-glucuronidases are generally expressed extracellularly while bacterial α-glucuronidases are generally intracellular enzymes (Nurizz *et al*., 2002; Heneghan *et al*., 2007).

5.3.2 Mode of action

All α-glucuronidases characterised to date have been classified into family GH67 and they hydrolyse the α-1,2-glycosidic bond between a 4-*O*-methyl-D-glucuronic acid and a xylose moiety at the reducing end of a xylooligosaccharide (Nagy *et al*., 2003). The activity of OtGlcA67A was determined against 4-*O*-methyl-D-glucuronoxylan in the presence and absence of the endo-xylanase, OtXyn8A. DNSA assay revealed that the enzyme was not active on the polysaccharide itself but rather active on the mixture of xylooligosaccharides substituted with 4-*O*-methyl-D-glucuronic acid produced by the original action of OtXyn8A. Reaction products of the reaction with OtGlcA67A and OtXyn8A were analysed by HPAEC-PAD which showed increased hydrolysis of larger xylooligosaccharides after 1 h

(Figure 5.7). This may suggest that the removal of side chains by OtGlcA67A enhanced the hydrolytic activity of OtXyn8A. After prolonged hydrolysis of 24 h, there was the accumulation of X2 but peaks observed in the presence and absence of OtGlcA67A were not identical, suggesting that this α-glucuronidase removes 4-*O*-methyl-D-glucuronic acid from this disaccharide (Figure 5.7). A peak was detected before the MeGlcA-xylose peak upon incubation with OtGlcA67A at earlier and later stages of hydrolysis which may have been 4-*O*-methyl-D-glucuronic acid but without substantial standards this could not be determined accurately. The activity of OtGlcA67A was compared to that of the characterised α-glucuronidase, CjGlcA67A, from *Cellvibrio japonicus* which was reported to have a substrate specificity for 4-*O*-methyl-D-glucuronoxylooligosaccharides rather than the polysaccharide (Nagy *et al*., 2002). HPAEC-PAD analysis showed a potential 4-*O*-methyl-D-glucuronic acid peak when incubated with the substrate in presence of OtXyn8A and a comparison of the data produced by both OtGlcA67A and CjGlcA67A revealed similar hydrolytic patterns (Figure 5.7).

From the results generated, it could be postulated that OtGlA67A released 4-*O*-methyl-D-glucuronic (MeGlc) acid from aldobiouronic acid (MeGlcAXyl) and aldotriouronic acid ($MeGlcAXyl_2$) which in turn, were the main products released by OtXyn8A on 4-*O*-methyl-D-glucuronoxylan as observed in Figure 5.7. Furthermore, as the accumulation of similar hydrolytic products with and without α-glucuronidase was observed after 24 h, then it could be suggested that OtGlA67A was possibly only able to interact with the sugar moiety that participated in the target glycosidic bond. This idea is supported by the crystal structure of CjGlcA67A which revealed that the enzyme could only accommodate 4-*O*-methyl-D-glucuronic acid attached to a xylose in which the 4'hydroxyl was not linked to another xylose moiety, i.e. the oligosaccharide terminus (Nurizzo *et al*., 2002).

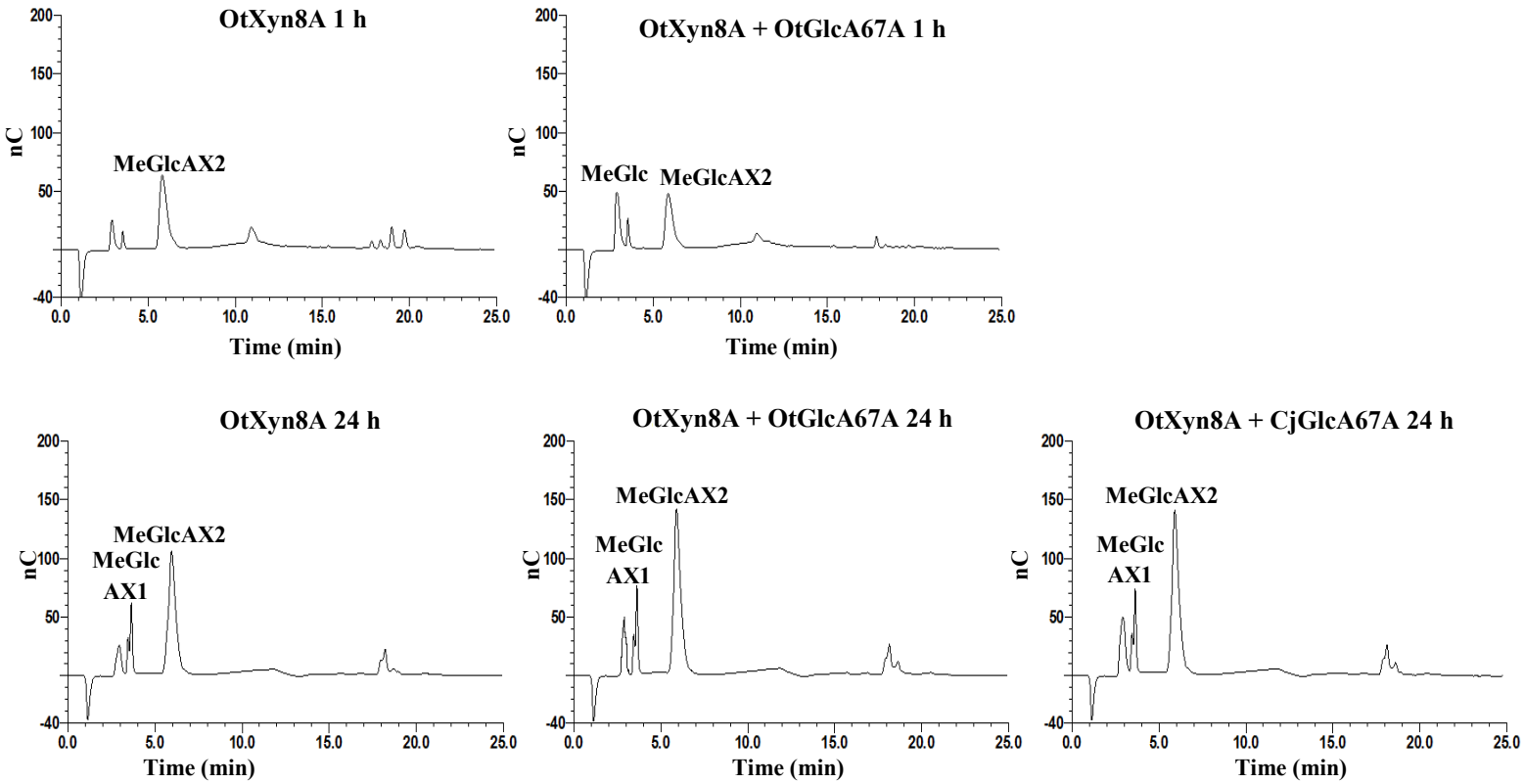


Figure 5.7 Analysis of OtGlcA67A on 4-*O*-D-methyl-glucronoxylan. OtGlcA67A (0.2 µM) plus OtXyn8A (0.2 µM) were incubated with 4-*O*-D-methyl-glucronoxylan (final concentration 4 mg/mL) and reactions were compare after 1 h hydrolysis and 24 h hydrolysis. Similarly the action on oligosaccharides of OtGlcA67A was compared to that of CjGlcA67A. Reactions were carried out in sodium acetate (0.05 M final concentration) pH 6 and BSA (1 mg/mL final concentration) at 25°C. At 1 h and 24 h, a sample was removed and subjected to HPAEC-PAD (see section 2.2.12).

Similarly to OtGlcA67A, CjGlcA67A contained a predicted signal peptide suggesting it was secreted out of the cell cytoplasm which is not necessarily consistent with the preferred activity of the enzyme against oligosaccharides rather than polysaccharides. Nagy *et al*., 2002 suggested the oligosaccharides were hydrolysed by CjGlcA67A on the outer membrane of *C. japonicus* which may be the apparent situation occurring with OtGlcA67A in *O. terrae*. Numerous extracellular α-glucuronidases have been isolated that do in fact attack the glucuronoxylan polymer. For example, the secretion of an α-glucuronidase from *Pichia stipitis* was so effective in debranching 4-*O*-methyl-D-glucuronoxylan that it aided hydrolysis of the polysaccharide by an endo-xylanase (Ryabova *et al*., 2009).

Structural analysis into CjGlcA67A showed that this α-glucuronidase was a dimeric protein with each monomer consisting of three domains (Nurizzo *et al*., 2002). The N-terminal domains formed a two layer β-sandwich whereas the C-terminal domain consisted of long α-helices. The central domain was found to be a classical $(\beta/\alpha)_8$ barrel fold that was ultimately the catalytic centre. It was elucidated from the structure of the enzyme in complex with its substrate, aldobiouronic acid, that the carboxylate and methyl moieties of 4-*O*-methyl-D-glucuronic acid were key determinants in substrate recognition (Nagy *et al*., 2003). To achieve glycosidic bond cleavage, the uronic acid had to be distorted to accommodate the oxocarbenium ion-like transition state required for the catalytic mechanism. The interaction between the enzyme and the carboxylate group was also found to play a crucial role in stabilising transition state stabilising. Nagy *et al*., (2003) observed that the removal of residues interacting with the carboxylate and methyl groups significantly reduced the k_{cat} of CjGlcA67A. This may have relevance to that of OtGlcA67A, since the residues focussed upon by Nagy *et al*., (2003) are completely conserved in family 67.

5.4 Analysis of potential synergy between enzymes on OSX and 4-*O*-D-methyl-glucronoxylan

Through HPAEC-PAD analysis, the collective activity of the recombinant enzymes described within this chapter was demonstrated, indicating the significant hydrolysis of 4-*O*-methyl-D-glucuronoxylan into xylose (Figure 5.8). This suggests that this cluster of genes may be critical for the degradation of xylan substrates by *O. terrae*. Not only that, but the β-xylosidase and endo-xylanase (OtXyn8A) produced from this gene cluster demonstrated a more efficient hydrolysis of OSX compared to an enzyme cocktail including OtXyn8A and a different *O. terrae* β-xylosidase not located elsewhere in the genome. This provides information on the physiological role of the enzymes and the element of synergy between the gene products from the same gene cluster of *O. terrae* rather than individual enzymes from the whole genome.

The complete hydrolysis of 4-*O*-D-methyl-glucronoxylan into xylose was observed with the incubation of all three recombinant enzymes, OtXyn8A, OtXyl52A and OtGlcA67A. Therefore, this was indicative of the synergistic effect of the addition of the β-xylosidase to the endo-xylanase and α-glucuronidase incubation mixture. However, due to the potentially low activity of OtGlcA67A upon the polymeric substrate in comparison to small oligomer substrates, this indicated the essential requirement of OtXyn8A. Many reports have found that α-glucuronidases act exclusively on the glycosidic bonds between terminal xylose at the non-reducing end (Ryabova *et al*., 2009). Therefore as β-xylosidases also cleave the non-reducing end, it was expected that a synergistic effect between OtXyl52A and OtGlcA67A would occur as the α-glucuronidase would remove 4-*O*-methyl-D-glucuronic acid for the terminal xylose to be accessible for the action of the β-xylosidase. This was confirmed further in this study as the reaction with OtXyn8A and OtXyl52A in the absence of OtGlcA67A showed decreased hydrolysis of the substrate (in terms of xylose released) suggesting that OtXyl52A was prevented in cleaving terminal xylose from small

oligosaccharides due to the location of 4-*O*-methyl-D-glucuronic acid (Figure 5.8). Therefore, it could be assumed that OtGlcA67A was essential for OtXyl52A to carry out efficient β-xylosidase activity.

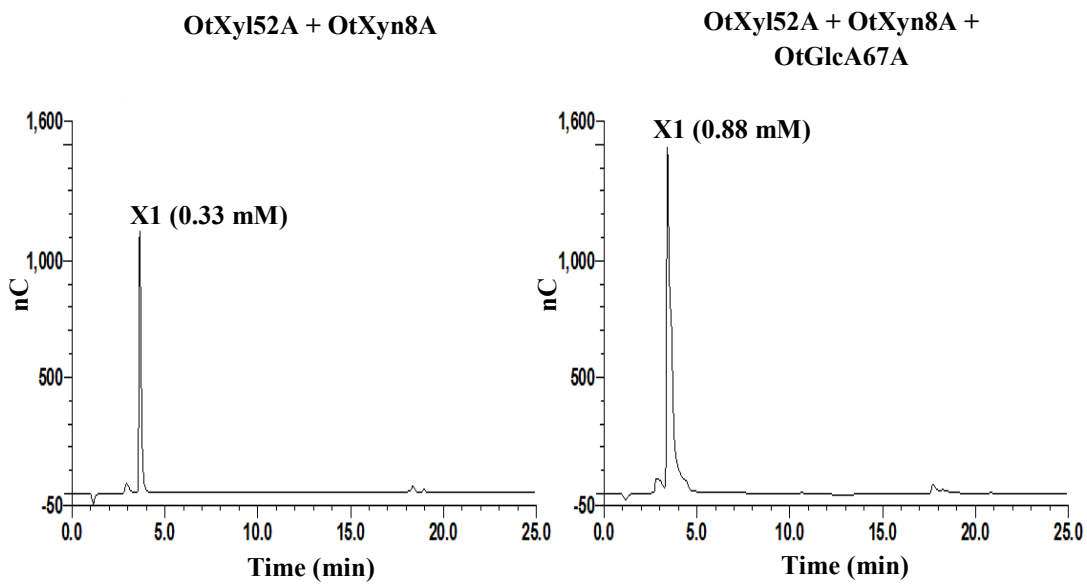


Figure 5.8 Analysis of potential synergy of enzymes from gene cluster on 4-*O*-D-methyl-glucronoxylan. OtXyl52A (0.7 µM) and OtXyn8A (0.2 µM) were incubated with 4-*O*-D-methyl-glucronoxylan (final concentration 5mg/mL) in the absence of OtGlcA67 and presence of OtGlcA67A (0.2 µM). Reactions were carried out in sodium acetate (0.05 M final concentration) pH 6 and BSA (1 mg/mL final concentration) at 25°C. At 24 h, a sample was removed and subjected to HPAEC-PAD (see section 2.2.12). Xylose released in each reaction has been quantified and is indicated in brackets.

Since the fourth gene, *xyn8B*, of the gene cluster could not be expressed it was difficult to determine the significance of this enzyme within the xylanolytic gene cluster. The gene product of *xyn8B* was annotated as an endo-xylanase and amino acid sequence analysis by NCBI BLAST showed that the OtXyn8B had the closest identity to OtXyn8A of 58%. However, substrate specificities of these two enzymes may differ somewhat according to their cellular distribution within the bacterium which may lead to their individual physiological roles, i.e., as pointed out in Chapter 4, OtXyn8A appears to be an extracellular enzyme, due to the existence of a signal peptide, while OtXyn8B appears to be an intracellular protein, due to the lack of a predicted signal peptide. Therefore, it could be hypothesised that even though these two GH8 enzymes share a relatively high identity, their cellular locations may support the suggestion that their activities are likely to differ. For example, while OtXyn8A acts on the polysaccharide substrate, OtXyn8B may be active against smaller xylooligosaccharide structures that enter the cell cytoplasm. This information is the first step into understanding the gene cluster as a whole and the assumption of xylanolytic synergism between all four gene products.

To understand this cluster of genes, an approach should be taken in which the regulation of xylanase activity is determined. For example, information would be required concerning the induction of the enzymes within *O. terrae* by certain substrates. A publication by Tsujibo *et al*., (2004) described the *bxl* operon of *Streptomyces thermoviolaceus* OPC-520 for xylooligosaccharide uptake and degradation. The operon comprised genes encoding a xylooligosaccharide binding protein, two permeases and an intracellular β-xylosidase all of which expression was induced by X2. In relation to this study it can be assumed, from the data provided in conjunction with the organisation and distribution of genes and their products, that the major inducer of such xylanase activity would be the xylan polysaccharide and possibly large xylooligosaccharides. Interestingly, Chin *et al*., (2001)who isolated *O. terrae* PB90-1 found that growth of the strain was supported with xylan but not xylose or methyl-α-glucopyranoside. A transcriptional and/or translational profile of the gene cluster

and maybe surrounding genes would be a useful approach to gain an insight into the physiological roles of these enzymes, especially OtXyn8B which has yet to be assigned a function. Identifying the transcription factors responsible for such genes would enable an understanding of expression regulation and the utilisation of xylose by the bacterium.

6 Results and Discussion

Characterisation of Glycoside Hydrolases Belonging to Family 43

6.1 Introduction

The expansion of family GH43 has gained much attention over recent years due its array of specificities but more importantly the occurrence of large numbers of GH43 enzymes encoded by individual microorganism genomes. Through the analysis of GH43 enzymes belonging to *Opitutus terrae* PB90-1 and other microorganisms, some definable activities were identified within this research although function could not be assigned for all of the isolated proteins even though protein expression and purification was successful. There are 21 putative GH43 sequences from *O. terrae* to which 15 were successfully expressed within this research (see Appendix B for SDS-PAGE images). Not all recombinant proteins are guaranteed to be active due to many reasons such as functional redundancy or lack of activity against the substrates tested. Furthermore, activities of investigated enzymes may only become apparent when acting in combination with other hydrolytic enzymes.

This chapter will describe the attempted preliminary characterisations of six family 43 enzymes that were expressed and subsequently exhibited activity against xylan and arabinan substrates. The majority of the GH43 enzymes investigated belonged to *O. terrae* with two of the isolated enzymes belonging to *Lactobacillus brevis* ATCC 367. As previously mentioned, *O. terrae* is an attractive target for this research due to the existence of twenty-one family 43 glycoside hydrolase encoding genes within its genome, which are now known to encode a variety of α-L-arabinofuranosidases and β-xylosidases. Family 43 is dominated by α-L-arabinanases, α-L-arabinofuranosidases and β-xylosidases. Furthermore, it is also recognised that this family contains β-1,3-galactanases and newly characterised xylanases. The GH43 group is an inverting family that utilises a triad of residues to perform catalysis. These residues include a Glu, acting as the general acid and two Asp residues, one as the

general base while the other modulates the pKa of the general acid (Nurizzo *et al*., 2002). Although the catalytic residues are conserved in all GH43 members, several studies have elucidated residues that are essential for substrate recognition which are only conserved in GH43 α-L-arabinofuranosidases and β-xylosidases (Cartmell *et al*., 2011 Brunzelle *et al*., 2008; Brüx *et al*., 2006; Fujimoto *et al*., 2010).

In total 29 genes encoding GH43 enzymes were amplified by PCR with 19 genes belonging to *O. terrae*. Furthermore, 24 of the amplified genes were successfully cloned in *E. coli* TOP10 with 17 expressed in *E. coli* BL21. Once solubility was achieved; recombinant proteins were purified as explained in section 2.2.10 and stored in ammonium sulphate for further use. SDS-PAGE gel images of the purified proteins investigated within this study can be observed in section 3.4, Figures 3.4 to 3.6. Family 43 enzymes were characterised by their activity exhibited on synthetic 4-nitrophenyl-glycosides and natural substrates including oat spelt xylan (OSX), wheat arabinoxylan, sugar beet arabinan and debranched arabinan. Determination of characterisation was made via DNSA assays and HPAEC-PAD analysis. As a control, characterised enzymes were utilised to assist the determination of new activities and these included the endo-1,5-α-arabinanase, ScAra43A, from *Streptomyces coelicolor* A3 and a GH43 β-xylosidase from *Bacillus subtilis* subsp. *subtilis* str. 168 respectively. Although, 17 GH43 enzymes were purified within this research (of which 15 belonged to *O. terrae*) the majority of enzymes either displayed no activity or trace activities that were difficult to interpret. Table 6.1 lists the enzymes that demonstrated sufficient activity against synthetic and natural substrates that could be monitored and subsequently analysed. Enzyme names are given with respect to the nomenclature proposed by Henrissat *et al*., (1998).

Putative glycoside hydrolase	**Organism**	**Enzyme name**	**Accession number**
α-L-arabinofuranosidase	*O. terrae*	OtAraf43A	ACB74677.1
α-L-arabinofuranosidase	*O. terrae*	OtAraf43B	ACB76503.1
α-L-arabinofuranosidase	*O. terrae*	OtAraf43C	ACB76451.1
α-L-arabinofuranosidase	*L. brevis*	LbAraf43A	ABJ64815.1
β-xylosidase	*O. terrae*	OtXyl43A	ACB73649.1
β-xylosidase	*L. brevis*	LbXyl43A	ABJ65333.1

Table 6.1 A list of the characterised GH43 enzymes described in this chapter and their activities.

6.2 GH43 α-L-arabinofuranosidases

The arabinan degrading enzymes are a well recognised group of glycoside hydrolases required for the hydrolysis of arabinofuranosyl residues from various substrates such as arabinan and arabinoxylans. More specifically endo-1,5-α-L-arabinanases catalyse the hydrolysis of the α-1,5- linked arabinan backbone whilst α-L-arabinofuranosidases attack α-1,5- linked arabinose residues of arabinan as well as α-L-arabinosyl side chains of arabinans and xylans which are usually α-1,2- and α-1,3- linked to the main backbone (Gilead *et al*., 1995). Within this study α-L-arabinofuranosidases were incubated with sugar beet arabinan and debranched arabinan substrates and activity was detected through the release of reducing sugar via DNSA before ultimately being characterised by HPAEC-PAD analysis. Debranched arabinan constitutes simply a α-1,5-linked arabinose backbone which is fully debranched of substituents. Sugar beet arabinan is a more complex polysaccharide consisting of α-1,5-linked arabinofuranose units which are highly substituted with single α-1,3-linked α-L-arabinose side chains and even less so, α-1,2-linked L-arabinofuranose (Caffall *et al*., 2009). Arabinoxylans have been reported to consist of mostly regions where singly (α-1,3-) and doubly (α-1,3-, α-1,2-) substituted xylose residues are separated by unsubstituted xylose residues (Izydorczyk *et al*., 1995).

6.2.1 Characterisation of OtAraf43A

6.2.1.1 Analysis of gene

Primary amino acid analysis of OtAraf43A suggested that the respective gene most probably encoded an α-L-arabinofuranosidase. The amino acid sequence of the encoded protein did not contain a putative signal peptide, so it is plausible to suggest that this enzyme may have been responsible for hydrolysing short oligosaccharide chains intracellularly. The amino acid sequence of OtAraf43A exhibited 74% and 69% sequence similarity with α-L-arabinofuranosidases from *Ricinus communis* and *Cellvibrio japonicus*, respectively. Furthermore, the *C. japonicus* enzyme was also an intracellular enzyme annotated as an exo-α-1,5-L-arabinofuranosidase.

6.2.1.2 Mode of action

The capacity of recombinant OtAraf43A to hydrolyse various natural substrates was tested upon sugar beet arabinan, debranched arabinan and rye arabinoxylan. HPAEC-PAD analysis revealed that OtAraf43A exhibited hydrolytic activity against sugar beet arabinooligosaccharides (Figure 6.1) and debranched arabinooligosaccharides (Figure 6.2), whereas no activity was detected against sugar beet arabinan or debranched arabinan. From such data it was apparent that the enzyme exhibited specificity for arabinooligosaccharides rather that polymeric substrates. The activity of OtAraf43A against debranched arabinooligosaccharides showed that the enzyme was able to hydrolyse α-1,5- linkages for the release of arabinose. Furthermore, as OtAraf43A displayed activity against branched arabinooligosaccharides generated from sugar beet, it indicates that the enzyme potentially has the ability to also cleave α-1,2- and / or α-1,3- linked side chains of arabinooligosaccharides. However, various arabinooligosaccharides accumulated after 24 h

hydrolysis, possibly meaning that OtAraf43A only attacked one of the types of side linkages present on sugar beet backbone as total hydrolysis was not achieved.

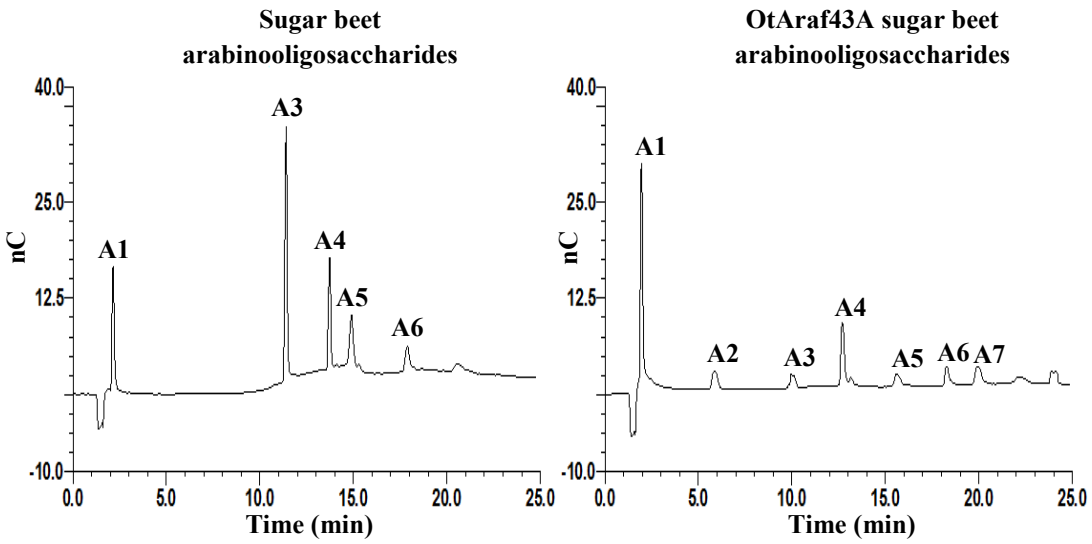


Figure 6.1 Activity of OtAraf43A on arabinooligosaccharides. OtArfa43A (5 µM) was incubated with sugar beet arabinooligosaccharides (final concentration 8mg/mL). Reactions were carried out in sodium phosphate (0.05 M final concentration) pH 6.5 and BSA (1 mg/mL final concentration) at 25°C. At 24 h, a sample was removed and subjected to HPAEC-PAD (see section 2.2.12).

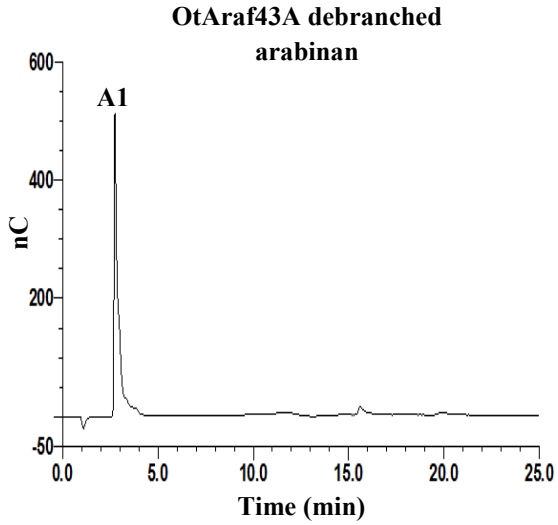


Figure 6.2 Activity of OtAraf43A on debranched arabinan. OtArfa43A (5 µM) was incubated with polymeric debranched arabinan (final concentration 2 mg/mL). Reactions were carried out in sodium phosphate (0.05 M final concentration) pH 6.5 and BSA (1 mg/mL final concentration) at 25°C. At 24 h, a sample was removed and subjected to HPAEC-PAD (see section 2.2.12).

Additionally, this α-L-arabinofuranosidase appeared to be arabinan specific as it did not possess the capacity to cleave side linkages present on rye arabinoxylan. The structure and function of an arabinan-specific enzyme cleaving side chains was recently elucidated from *C. japonicus* in which the α-L-arabinofuranosidase, CjAbf43A, was unable to hydrolyse arabinoxylan substrates (Cartmell *et al*., 2011). This study showed that the topology of the enzyme was complementary to the extended helical structure of the α-1,5-linked arabinan main chain of sugar beet arabinan indicating the enzyme could discriminate between arabinan and arabinoxylan backbones. The substrate binding cleft of Cjabf43A was identified as curved and therefore arabinoxylan would sterically clash with the surface of the binding site of the enzyme (Cartmell *et al*., 2011).

The functional analysis of OtAraf43A is in agreement with its cellular location as it could be suggested that the intracellular location of the α-L-arabinofuranosidase presumably targets arabinooligosaccharides with α-1,5-, α-1,3 and / or α-1,2- linkages that have been transported into the cell. Previous studies have shown this to be the case for numerous α-L-arabinofuranosidases. Inácio *et al*., (2008) were able to propose a pathway for the degradation of arabinose-containing polysaccharides by *Bacillus subtilis*. Their study indicated that this bacterium secreted two intracellular α-L-arabinofuranosidases specific for arabinooligosaccharides possibly suggesting that one of the other *O. terrae* α-L-arabinofuranosidases works synergistically with OtAraf43A for total degradation of substrates.

6.2.3 Characterisation of OtAraf43B

6.2.3.1 Analysis of gene

SignalP analysis of OtAraf43B revealed that the sequence at the N-terminus (amino acids 1 - 25) was predicted to be a signal peptide and therefore primers were designed to amplify the part of the gene that encodes the mature protein. OtAraf43B is therefore likely to be located extracellularly, which would relate to its function. The deduced amino acid sequence of OtAraf43B was compared to other protein sequences using BLAST. The sequence of OtAraf43B shared its highest identity (68 %) with a β-xylosidase from *Mucilaginibacter paludis*, hence the annotation of OtAraf43B as a putative β-xylosidase.

6.2.3.2 Mode of action

HPAEC-PAD analysis of the reaction products generated by OtAraf43B on sugar beet arabinan showed a significant release of arabinose from the polysaccharide after a prolonged reaction of 24 h (Figure 6.3). For the determination of which linkages were being hydrolysed, the enzyme was incubated with debranched arabinan and wheat arabinoxylan which represented arabinose linkages within main chain molecules and side chains, respectively. HPAEC-PAD results showed no release of arabinose from debranched arabinan but substantial release of arabinose from wheat arabino-xylooligosaccharides (Figure 6.5). These results indicated that OtAraf43B could not attack the α-1,5-linked arabinose chain of debranched arabinan but could hydrolyse 1,2- and / or 1,3- linked arabinose substituents that were present in wheat arabinoxylan and most likely sugar beet arabinan. It was not surprising that this α-L-arabinofuranosidase has the capacity to cleave arabinose side chains on arabinan and arabinoxylan backbones, as several residues have been found to be conserved in family 43 α-L-arabinofuranosidases and β-xylosidases which

are involved in substrate binding and hydrolysis (Moraïs *et al*., 2012). This activity is very complementary to that observed for OtAraf43A, which was arabinan-specific.

Moreover, when OtAraf43B was incubated with the endo-arabinanase, ScAra43A, on sugar beet arabinan, significantly more of the arabinooligosaccharides were produced possibly indicating that the effective removal of side chains by OtAraf43B enhanced the endo-activity of ScAra43A. The low activity observed against α-1,5-arabinooligosaccharides also suggested that OtAraf43B hydrolysed the side chains present in sugar beet arabinan, as a similar detection of arabinose released by the α-L-arabinofuranosidase was observed in the presence and absence of the endo-arabinanase.

The specificity of OtAraf43B could be further explored by determining exactly which linkage is hydrolysed by the enzyme, whether it be either of the α-1,2- or α-1,3- side chains or both. The α-L-arabinofuranosidase, CjAbf43A, was described to be specific for α-1,2-linkages present on sugar beet arabinan (Cartmell *et al*., 2011). Cartmell et al. (2011) demonstrated that the α-1,2-specificity of the enzyme could be determined by NMR spectroscopy in which arabinosyl residues were displayed in accordance to their chemical shifts. NMR analysis revealed resonances diagnostic of arabinosyl residues substituted with α-1,2- side chains which disappeared after the treatment with the α-L-arabinofuranosidase. The structure of CjAbf43A was determined within this same study showing the five-bladed β-propeller fold typical of GH43 enzymes. Furthermore, insights into the topology of the enzyme revealed a deep pocket situated in the centre of a curved cleft which was able to accommodate the arabinan backbone.

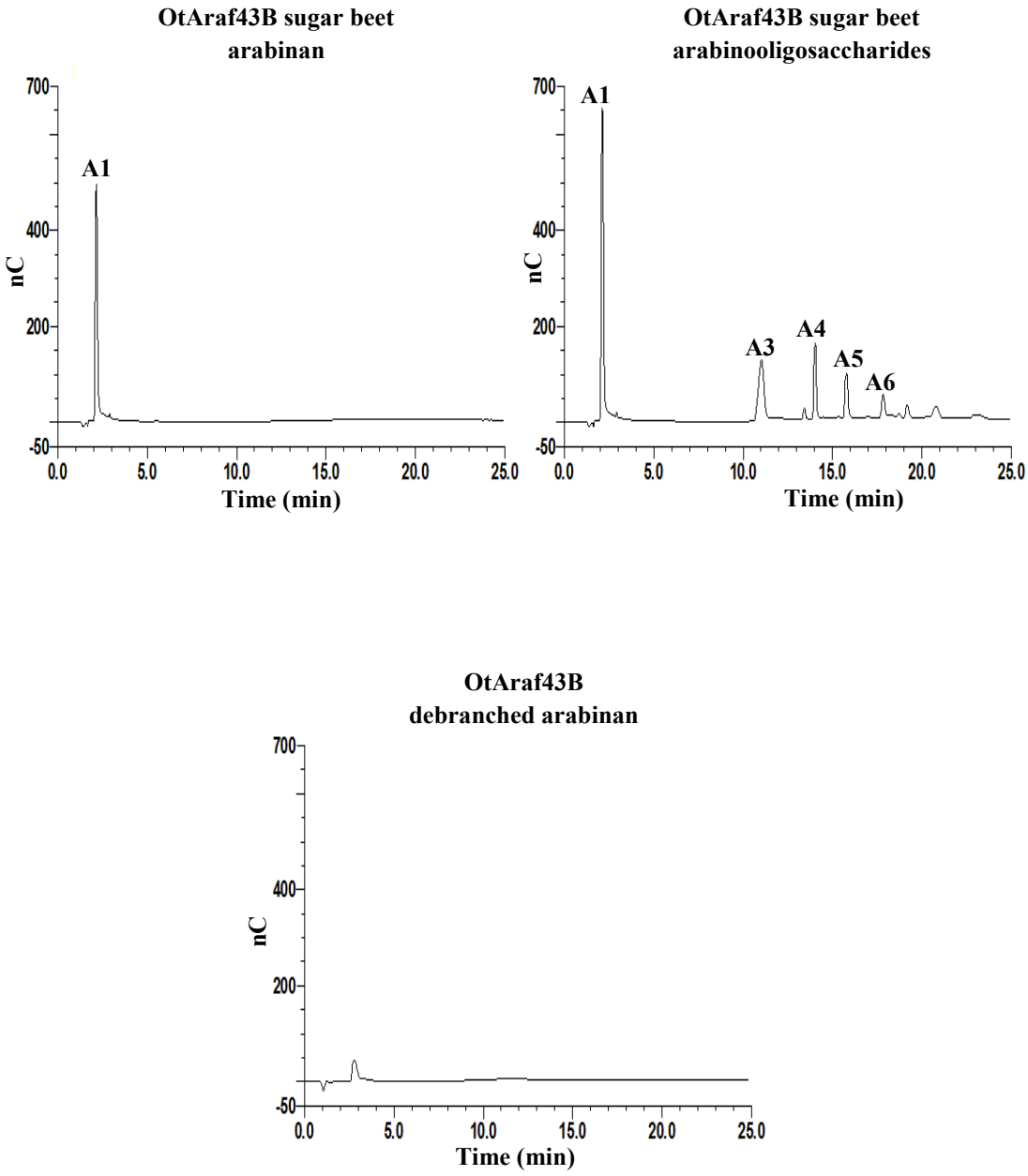


Figure 6.3 Activity of OtAraf43B on sugar beet arabinan, arabinooligosaccharides and debranched arabinan (24 h hydrolysis). OtArfa43B (5 µM) was incubated with sugar beet arabinan (final concentration of 8 mg/mL), arabinooligosaccharides (final concentration of 8 mg/mL) and debranched arabinan (final concentration of 2 mg/mL). Reactions were carried out in sodium phosphate (0.05 M final concentration) pH 6.5 and BSA (1 mg/mL final concentration) at 25°C. At 24 h, a sample was removed and subjected to HPAEC-PAD (see section 2.2.12).

6.2.4 Characterisation of OtAraf43C

6.2.4.1 Analysis of gene

SignalP analysis of the deduced amino acid sequence of OtAraf43C revealed a signal peptide at the N-terminus. This suggested that the glycoside hydrolase was an extracellular enzyme, indicating its primary target substrate was a polysaccharide. Furthermore, when a BLAST search was performed with the sequence, the results showed that OtAraf43C shared highest identity (64%) to an α-L-arabinofuranosidase from *Paenibacillus lactis* 154.

6.2.4.2 Mode of action

To determine the activity of OtArbf43C, the enzyme was incubated with 4-nitrophenol-glycosides, debranched arabinan, sugar beet arabinan and wheat arabinoxylan. The enzyme only showed activity against 4-nitrophenyl-α-L-arabinofuranoside among the tested 4-nitrophenyl-glycosides which may be indicative of important interactions with arabinose in the +1 subsite. HPAEC-PAD analysis demonstrated that when the enzyme was incubated with polysaccharides, OtAraf43C exhibited activity against debranched arabinan and to a much lesser extent, sugar beet arabinan, with the release of arabinose identifying the hydrolysis of α-1,5- linked arabinose units (Figure 6.4). It was observed that OtAraf43C was not active against wheat arabinoxylan, which indicated that this α-L-arabinofuranosidase did not hydrolyse α-1,2- and α-1,3- arabinose linkages present on the xylan backbone. Figure 6.5 compares the reaction products generated from wheat arabinoxylan oligosaccharide hydrolysis by OtAraf43B and OtAraf43C. OtAraf43B was able to attack the α-1,2- and α-1,3- linkages present on the xylan main chain releasing arabinose, while OtAraf43C was unable cleave such substituents.

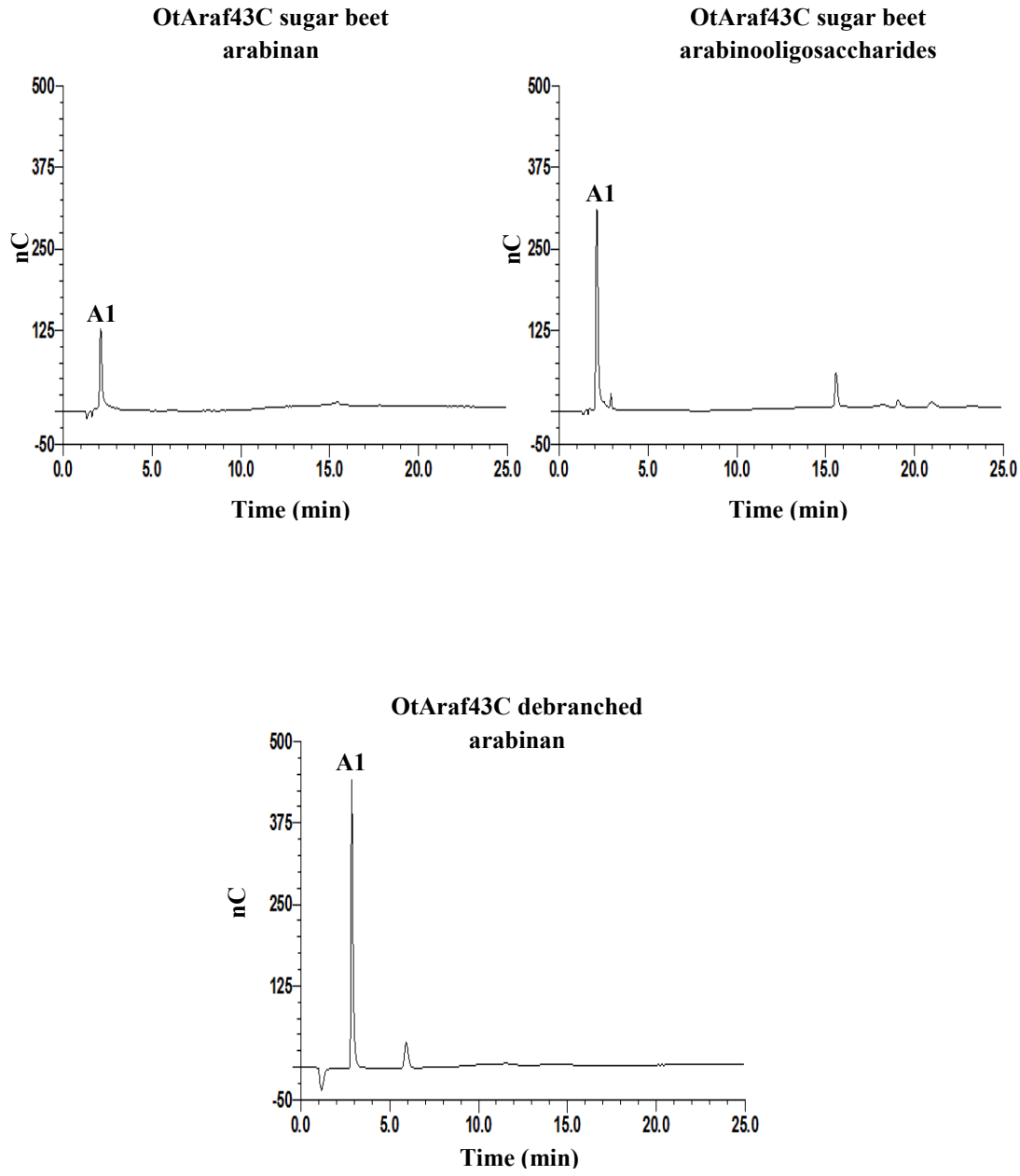


Figure 6.4 Activity of OtAraf43C on sugar beet arabinan, arabinooligosaccharides and debranched arabinan (24 h hydrolysis). OtArfa43C (5 µM) was incubated with sugar beet arabinan (final concentration 8 mg/mL), arabinooligosaccharides (final concentration 8 mg/mL) and debranched arabinan (final concentration 8 mg/mL). Reactions were carried out in sodium phosphate (0.05 M final concentration) pH 6.5 and BSA (1 mg/mL final concentration) at 25°C. At 24 h, a sample was removed and subjected to HPAEC-PAD (see section 2.2.12).

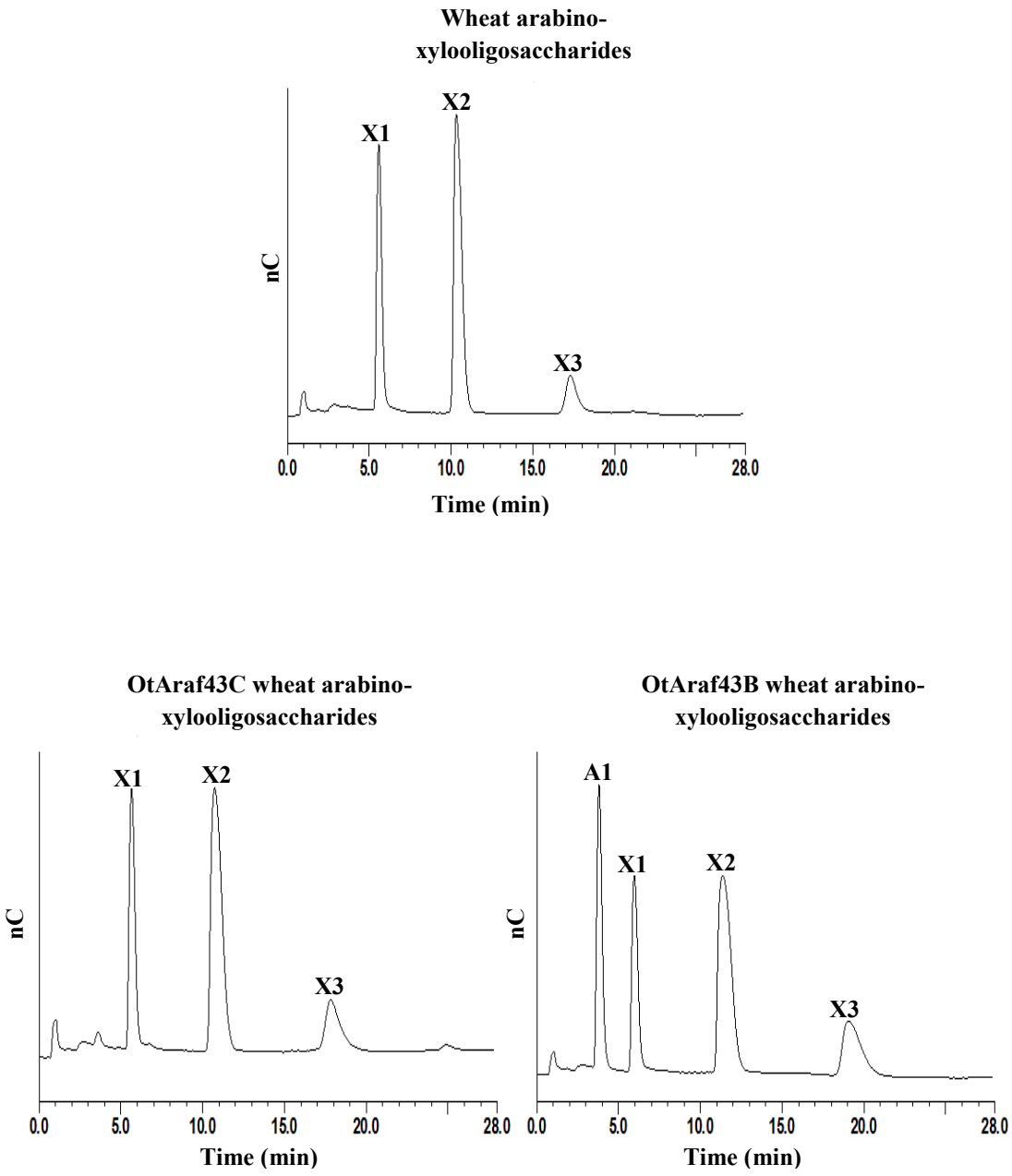


Figure 6.5 Activity of OtAraf43B and OtAraf43C on wheat arabino-xylooligosaccharides. OtAraf43B (5 µM) was incubated with wheat arabino-xylooligosaccharides and OtArfa43C (5 µM) was incubated with wheat arabino-xylooligosaccharides (final concentration 8 mg/mL). Reactions were carried out in sodium phosphate (0.05 M final concentration) pH 6.5 and BSA (1 mg/mL final concentration) at 25°C. At 24 h, a sample was removed and subjected to HPAEC-PAD (see section 2.2.12).

Collectively this data suggested that OtAraf43C could discriminate between the different linkages present within arabinose-related substrates and that it was specific for α-1,5- linked arabinose residues. OtAraf43C only released arabinose from the investigated arabinan polysaccharide suggesting it was strictly an exo-acting enzyme in comparison to ScAra43A, the control endo-arabinanase, which released arabinooligosaccharides with various degrees of polymerisation from arabinan. In comparison to OtAraf43B, OtAraf43C released less arabinose from sugar beet arabinan which may have been due to the steric hindrance caused to the enzyme by 1,2- and / or 1,3- linked arabinofuranosyl side chains present on the arabinan backbone. Also the release of arabinose did not significantly increase with the addition of an endo-arabinanase again indicating the hindrance caused by side chains to the -1 and +1 subsites.

Therefore, OtAraf43C could potentially be classed as an exo-1,5-α-L-arabinofuranosidase with the release of arabinose, most likely from the non-reducing end of the arabinose backbone which has been described previously for other GH43 enzymes (Brüx *et al*., 2006; Proctor *et al*., 2005). The mode of action exhibited by the enzyme can also be explained by the cellular location of OtAraf43C , as it occurs as an extracellular enzyme therefore suggesting its main target substrates are polysaccharide molecules. Additionally, the fact that OtAraf43C is a member of GH43 is consistent with the other activities present in this family, for example, endo-arabinanases, which also preferentially hydrolyse the same α-1,5-linked bond.

Numerous GH43 exo-acting arabinofuranosidases that strictly attack α-1,5- linkages of arabinan substrates have been isolated from *Streptomyces chartreusis* (Matsuo *et al*., 2000) and *Streptomyces avermitilis* (Ichinose *et al*., 2008). The crystal structure of the exo-1,5-α-L-arabinofuranosidase from *S. avermitilis*, SaAraf43A, has been recently solved (Fujimoto *et al*., 2010) and the structural features relating to its α-1,5- linked exo-mode of action elucidated. SaAraf43A was found to have a catalytic domain composed of the common

GH43 5-bladed β-propeller which is also typical for members belonging to families 32 and 68. Furthermore, this common domain is assembled from five radially orientated β-sheet repeats which are composed of four anti-parallel β-strands. Fujimoto *et al*., (2010) reported that the active site of SaAraf43A was located within the central space of the β-propeller fold forming the binding pocket. Conversely, the active-site of GH43 endo-arabinanases is described as residing across the surface of the β-propeller fold with the cleft long enough for the binding of linear arabinan substrate (Nurizzo *et al*., 2002; Alhassid *et al*., 2009). Structural analysis of SaAraf43A also revealed a steric hindrance due to a longer loop region (Tyr-281-Arg-294) close to the active site compared to that of the endo-arabinanase, BsArb43A, which meant that linear arabinan could not effectively bind. Fujimoto *et al*., (2010) were able to determine these structural loops encoded within the sequences of the exo-1,5-α-L-arabinofuranosidases, SaAraf43A and ScAraf43A. This is indicative of the exo-action exhibited by SaAraf43A for the single release of arabinose and therefore a similar active site arrangement may be observed within the structure of OtAraf43C.

The amino acid sequence of OtAraf43C was aligned with that of the exo-1,5-α-L-arabinofuranosidases, SaAraf43A and ScAraf43A and the endo-arabinanase BsArb43A from *Bacillus subtilis* subsp. *subtilis* str. 168, to determine if any structural elements could be identified in OtAraf43C (Figure 6.6). The endo-arabinanase BsArb43A contains a five residue loop, LTEER (highlighted in Figure 6.6). This loop was found to be responsible for the endo-action of BsArb43A, as the insertion of this loop into the exo-enzyme CjArb43A from *Cellvibrio japonicus* influenced the -3 subsite; diminishing chain end recognition (Proctor *et al*., 2005). This loop was not present within the structures of SaAraf43A and ScAraf43A and is not represented within the primary sequence of OtAraf43C which can be observed in Figure 6.6.

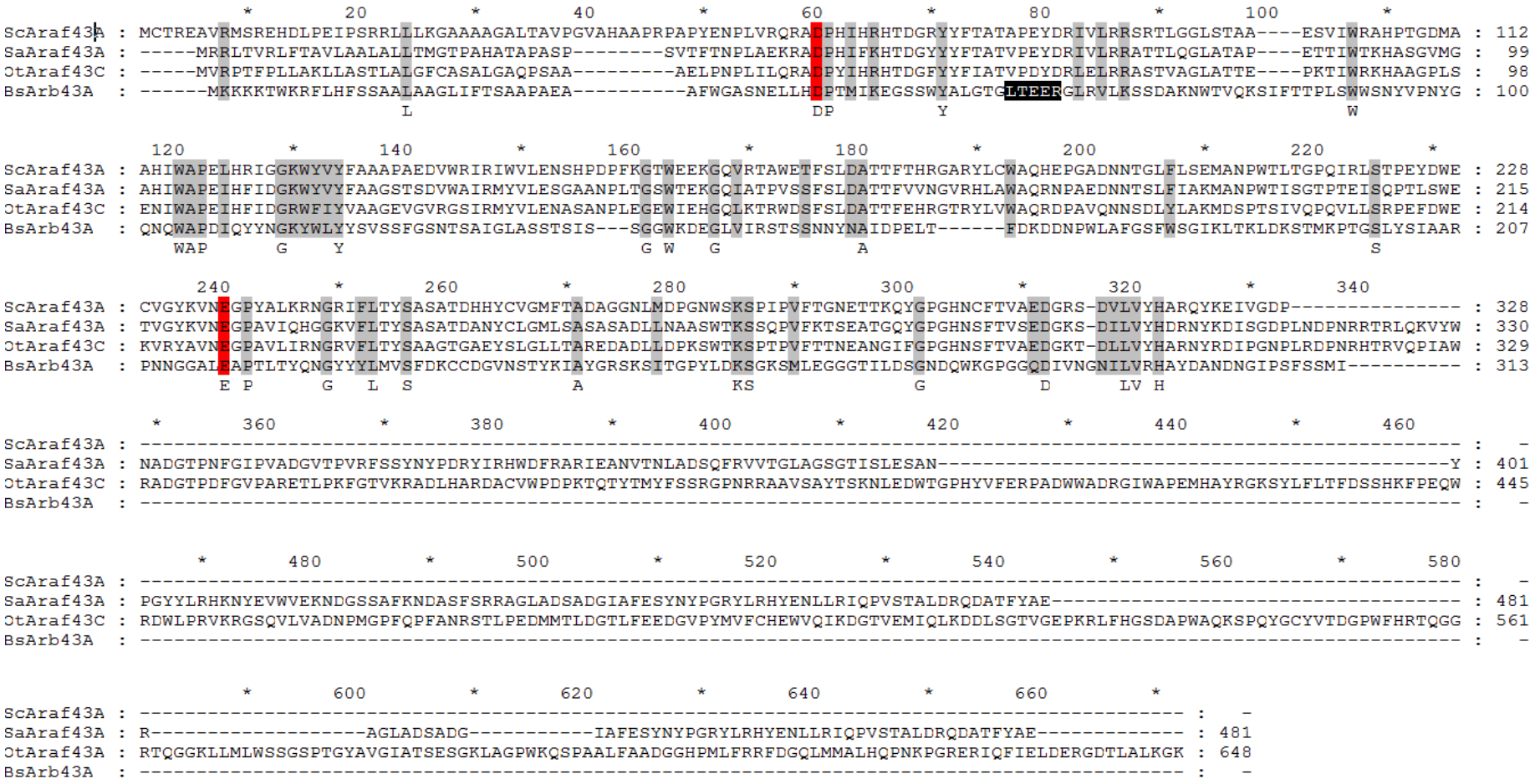


Figure 6.6 CLUSTALW Multiple sequence alignment of ScAraf43A, SaAraf43A, OtAraf43A and BsArb43A. Similar amino acids are coloured in grey and catalytic residues are coloured in red. The LTEER loop of BsArb43A is highlighted in black.

6.2.4 Characterisation of LbAraf43A

6.2.4.1 Analysis of gene

The deduced amino acid sequence of LbAraf43A was analysed using SignalP. No putative signal peptide sequence was identified, thus suggesting the probability that the encoded α-L-arabinofuranosidase is an intracellular enzyme. A BLAST search revealed that LbAraf43A had the highest sequence identity (87%) to an α-L-arabinofuranosidase from *Pediococcus acidilactici.*

6.2.4.2 Mode of action

The capacity of purified LbAraf43A to hydrolyse a range of different substrates including sugar beet arabinan, debranched arabinan and rye arabinoxylan was evaluated. HPAEC-PAD analysis showed that the enzyme hydrolysed arabinooligosaccharides from sugar beet arabinan (Figure 6.7) and rye arabinoxylan (Figure 6.8) as well as significant activity upon the polymeric substrate, debranched arabinan (Figure 6.7). Thus demonstrating that LbAraf43A could hydrolyse α-1,5-, α-1,3-, and /or α-1,2- linkages of arabinose containing substrates. However, the accumulation of a few arabinooligosaccharides after prolonged hydrolysis (24 h) of sugar beet arabinan with the an endo-arabinanase, ScAra43A, indicated that LbAraf43A was only potentially releasing arabinose from either α-1,2- or α-1,3- but not both since total hydrolysis of the substrate was not achieved. Furthermore, the activity on rye arabinoxylan could be considered quite low as the amount of arabinose released was significantly less than, for example, that released by OtAraf43B (Figure 6.5).

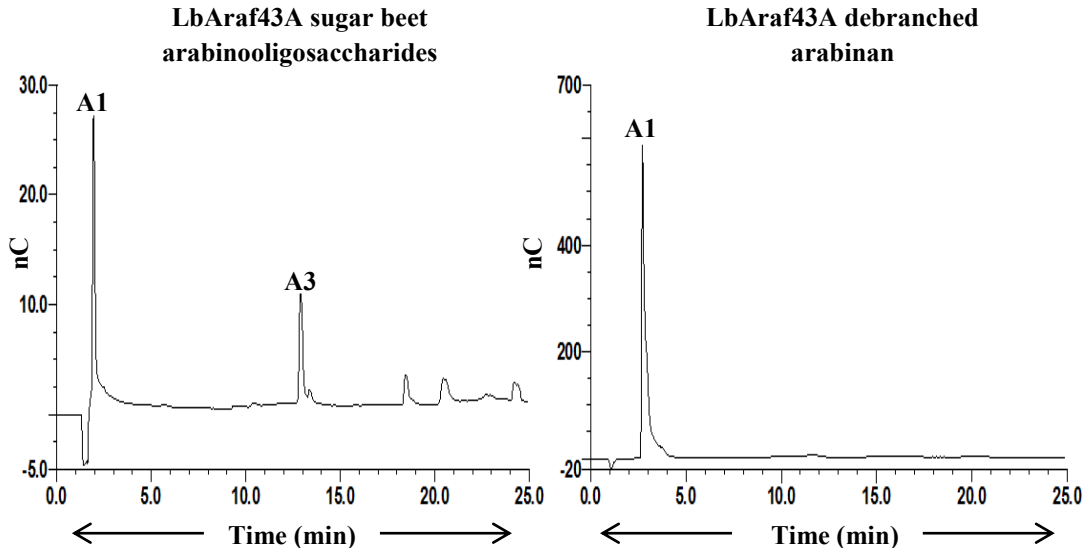


Figure 6.7 Activity of LbAraf43A on arabinan substrates. LbAraf43A (5 µM) was incubated with sugar beet arabinooligosaccharides (final concentration of 8 mg/mL) and debranched arabinan (final concentration 2 mg/mL). Reactions were carried out in sodium phosphate (0.05 M final concentration) pH 6.5 and BSA (1 mg/mL final concentration) at 25°C. At 24 h, a sample was removed and subjected to HPAEC-PAD (see section 2.2.12).

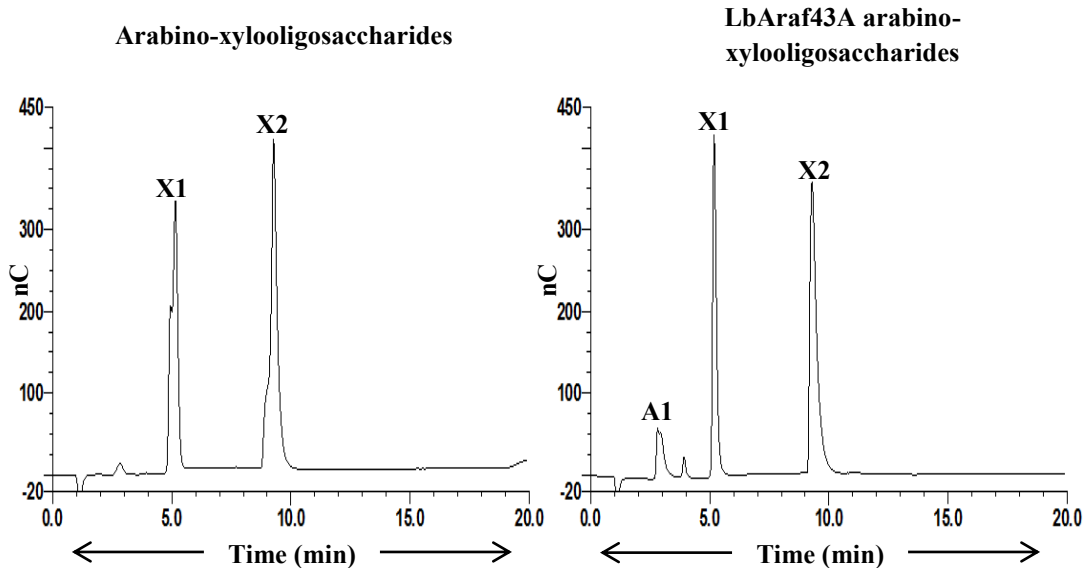


Figure 6.8 Activity of LbAraf43A on rye arabinoxylan. LbAraf43A (5 µM) was incubated with rye arabino-xylooligosaccharides (final concentration of 8 mg/mL). Reactions were carried out in sodium phosphate (0.05 M final concentration) pH 6.5 and BSA (1 mg/mL final concentration) at 25°C. At 24 h, a sample was removed and subjected to HPAEC-PAD (see section 2.2.12).

A better understanding as to why such an α-L-arabinofuranosidase like LbAraf43A displays high activity against polymeric debranched arabinan and sugar beet arabinan and not wheat arabinoxylan is required. However, this observation could indicate that *L. brevis* must secrete other enzymes capable of releasing arabinose from arabinoxylan substrates. Interestingly, a GH51 α-L-arabinofuranosidase from *L. brevis* has been isolated and the activity of the recombinant enzyme assessed (Michlmayr *et al*., 2011). Conversely, this study revealed that the GH51 enzyme exhibited relatively poor activity on both branched and debranched arabinan. Therefore such divergent activities between the α-L-arabinofuranosidases belonging to *L. brevis* could potentially indicate different specificities against arabinose-containing polysaccharides and furthermore a potential for synergistic interactions for the efficient hydrolysis of such substrates.

Previous studies have reported differences in the substrate specificities of α-L-arabinofuranosidases belonging to GH43 and GH51 in their action towards singly and doubly substituted side chains of arabinoxylans. α -L-Arabinofuranosidases belonging to family 43 have been demonstrated to liberate α-1,3- linked arabinofuranosyl residues from doubly substituted xylose units (Sørensen *et al*., 2006) as well as single α-1,2- and α-1,3-linked arabinose (Bourgois *et al*., 2007). In family 51, linkage specificity has been determined in which α-L-arabinofuranosidases show strong preference for singly linked arabinose and cleaving doubly residues to a much lesser extent (Sørensen *et al*., 2006; Beylot *et al*., 2001). However, a recent study by Lagaert *et al*., (2010) reported the existence of a GH51 α-L-arabinofuranosidase that attacked doubly linked α-1,3- arabinose residues, indicating that such an action was not confined to GH43. Therefore at this stage of characterisation of LbAraf43A, it is impossible to assume linkage specificity of α-L-arabinofuranosidases, maybe more so in GH51, but rather suggest that respective enzymes are synergistically required for total hydrolysis.

6.3 GH43 β-xylosidases

β-Xylosidases have been classified into numerous families including GH30, GH39, GH51, GH52 and GH54, however the most catalytically active β-xylosidases belong to GH43 (Knob *et al*., 2009; Jordan *et al*., 2007). Through the structural analysis of GH43 β-xylosidases from *Geobacillus stearothermophilus* and *Selenomonas ruminantium*, the identification of the catalytic base and catalytic acid required for the inverting mechanism have been elucidated for GH43 β-xylosidases (Brüx *et al*., 2006; Brunzelle *et al*., 2008). Furthermore, such enzymes have been found to possess a pocket topology comprising two subsites for the binding capacity of two xylose moieties which is consistent with their exo mode of action.

Within this research two GH43 β-xylosidases, OtXyl43A and LbXyl43A, were identified from *O. terrae* and *L. brevis* respectively. They both exhibited xylosidase activity with the release of xylose (X1) from xylooligosaccharides derived from OSX and wheat arabinoxylan which can be observed in Figures 6.9 and 6.10 respectively. The hydrolysis pattern obtained for OtXyl43A suggested that the enzyme may not be as active as LbXyl43A, as the β-xylosidase from *O. terrae* failed to completely degrade xylobiose (X2) into X1 even after a prolonged hydrolysis of 24 h, whereas Lbxyl43A completely degraded X2 into X1. Moreover, both OtXyl43A and LbXyl43A showed activity against the synthetic substrates, 4-nitrophenyl-β-D-xylopyranoside and 4-nitrophenyl-α-L-arabinofuranoside suggesting that the enzymes potentially possessed both xylosidase and arabinofuranosidase activities. Such enzymes that exhibit such dual function are known as bifunctional xylosidases / arabinofuranosidases. To assess the capability of these enzymes in releasing arabinose, both enzymes were incubated with rye arabinoxylan and sugar beet arabinan in the presence and absence of respective endo-enzymes. HPAEC-PAD analysis revealed after 24 h hydrolysis, both OtXyl43A and LbXyl43A released arabinose from rye arabinoxylan (Figure 6.10) but

with no significant release of arabinose from sugar beet arabinooligosaccharides. Therefore, it could be suggested that these two β-xylosidases are bifunctional enzymes with the release of arabinose as well as X1 from arabinoxylan. However, they are potentially arabinoxylan specific, since no activity was detected on arabinan substrates.

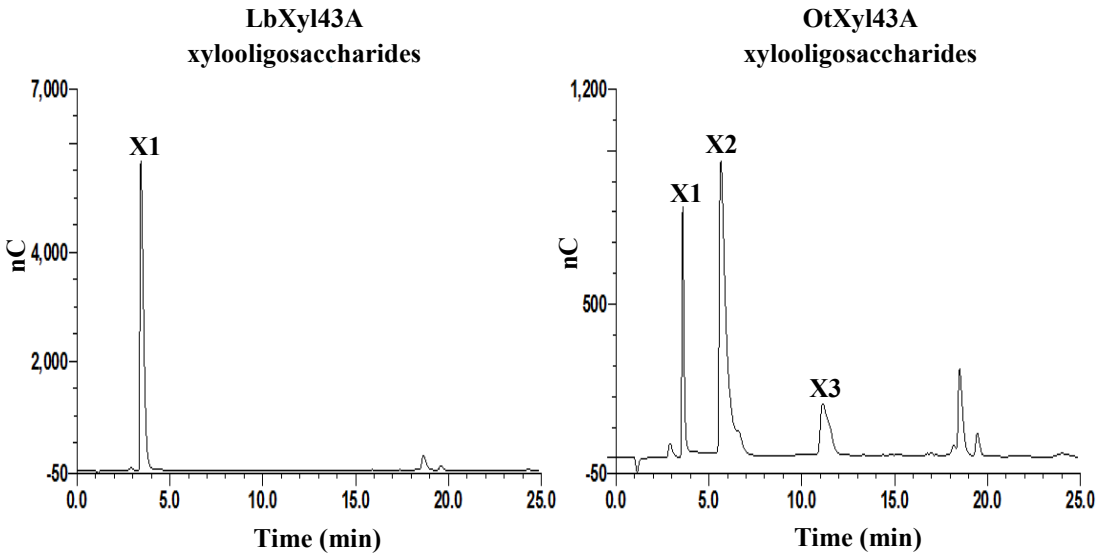


Figure 6.9 Activity of LbXyl43A and OtXyl43A on OSX. LbXyl43A (5 µM) and OtXyl43A (5 µM) were incubated with oat spelt xylooligosaccharides (final concentration of 8 mg/mL). Reactions were carried out in sodium acetate (0.05 M final concentration) pH 6 and BSA (1 mg/mL final concentration) at 25°C. At 24 h, a sample was removed and subjected to HPAEC-PAD (see section 2.2.12).

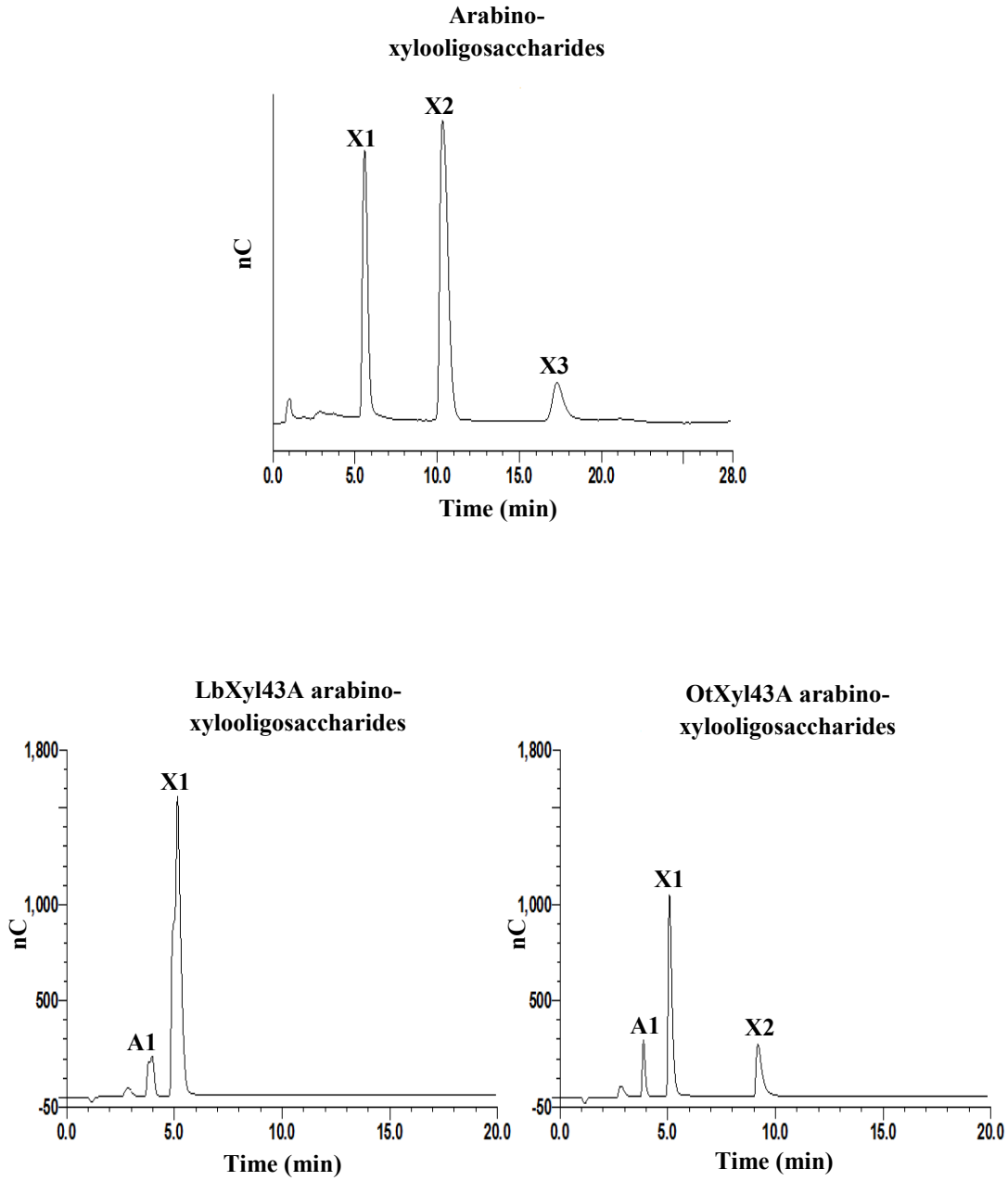


Figure 6.10 Activity of LbXyl43A and OtXyl43A on rye arabinoxylan. LbXyl43A (5 µM) and OtXyl43A (5 µM) were incubated with rye arabino-xylooligosaccharides (final concentration of 8 mg/mL). Reactions were carried out in sodium acetate (0.05 M final concentration) pH 6 and BSA (1 mg/mL final concentration) at 25 °C. At 24 h, a sample was removed and subjected to HPAEC-PAD (see section 2.2.12).

Dual-functional β-xylosidases enzymes are widely recognised and have been isolated from families GH43, GH30 and GH54 (Zhou *et al*., 2012), however GH43 β-xylosidases are quite unique in that they operate via an inversion mechanism. The rationale behind dual-functionality may lie in the fact that D-xylose and L-arabinose are spatially similar in that their hydroxyl groups and glycosidic bonds can be overlaid. However, a common feature of bifunctional xylosidase / arabinofuranosidase enzymes is that some possess a greater xylosidase activity than arabinofuranosidase activity and vice versa which has mainly been determined with 4-nitrophenol-glycoside substrates. Zhou *et al*., (2012) isolated two bifunctional β-xylosidases, RuXyn1 and RuXyn2, from GH43 and GH30 respectively in which RuXyn1 exhibited activity against both 4-nitrophenyl-β-D-xylopyranoside and 4-nitrophenyl-α-L-arabinofuranoside, whereas RuXyn2 only hydrolysed 4-nitrophenyl-β-D-xylopyranoside However against natural substrates, RuXyn1 was able to degrade xylooligosaccharides into X1 but unable to cleave arabinobiose while RuXyn2 exhibited the opposite action by hydrolysing arabinobiose and not xylooligosaccharides. Therefore, the basis of the dual-functionality of these enzymes was solely on their hydrolytic actions on 4-nitrophenyl-glycosides. There have been a few studies reporting dual-functionality on natural substrates as well as synthetic, including a study by Wagschal *et al*., 2009 who observed that a GH43 enzyme, deAX, exhibited arabinofuranosidase activity on arabinose-containing substrates, releasing arabinose and xylosidase activity against X2, X3 and arabinoxylan, releasing X1. Such activity is similar to the activities demonstrated by OtXyl43A and LbXyl43A within this study.

In an attempt to better understand the relative substrate specificity of bifunctional xylosidases, Jordan *et al*., (2007) searched for residues of the -1 subsite that potentially favour the xylosidase reaction over the arabinofuranosidase reaction using 4-nitrophenyl-glycosyl substrates. By mutating the residues of the glycone binding pocket of a GH43 β-xylosidase / α-L-arabinofuranosidase from *S. ruminantium* to alanine, the catalytic activity

towards both 4-nitrophenyl-β-D-xylopyranoside and 4-nitrophenyl-α-L-arabinofuranoside was lowered. Therefore, indicating that the native residues of the active site were important for catalysis but not essentially substrate specificity. However, several residues are conserved across both xylosidases and arabinofuranosidases within GH43, therefore lending more support to the existence of bifunctional xylosidases and arabinofuranosidases.

6.4 Conclusion

Despite that not all GH43 enzymes were functionally allocated within this research, the data presented here gives an insight into the diversity of family 43 glycoside hydrolases within *O. terrae*. This bacterium has an extensive library of GH43 enzymes targeted towards arabinan and xylan degradation which sequence analysis predicted to be mostly extracellularly located. When *O. terrae* PB90-1 was first isolated, growth was shown to be supported by xylan, pectin and arabinose (Chin *et al*., 2001) therefore demonstrating the importance of GH43 enzymes for the degradation of complex polysaccharides as a source of nutrients. Similar to *O. terrae*, *C. japonicus* is well known to secrete a large number of family 43 enzymes which are also predicted to act in an exo-fashion.

Within this study, two α-L-arabinofuranosidases, OtAraf43B and OtAraf43C, were isolated showing quite different specificities. It was demonstrated that while OtAraf43B was removing side chains from arabinan and arabinoxylan, OtAraf43C was solely attacking α-1,5- linkages of the arabinan main chain in an exo-manner. Such diversity may possibly result in a degree of synergy between the two enzymes for efficient degradation of polysaccharides by the bacterium as their actions reflect their ability to make target substrates more accessible to other enzymes. For example, the removal of arabinose side chains from doubly substituted units by OtAraf43B may leave single arabinose decorations which would then be more accessible to GH51 α-L-arabinofuranosidases also expressed by

O. terrae. However, only via a thorough analysis of linkage specificity could such determinations be made. Furthermore, the elucidation of these GH43 enzyme structures would confirm the specificity displayed by these family 43 members.

7 Final Discussion

The plant cell wall constitutes great potential in providing a sustainable energy resource for the industrial conversion of polysaccharides into biofuel. However, while a biochemical conversion of hemicellulose and cellulose in theory appears to be feasible, the industrial processes associated are still far from being economically viable as well as being impeded by the recalcitrant nature of lignocellulose. Microorganisms in nature, such as *Opitutus terrae*, utilise the plant cell wall as their primary source of energy by the secretion carbohydrate-active enzymes to degrade the polysaccharide molecules. The knowledge of such microbial systems has led to a further understanding of enzymes, such as the glycoside hydrolases, which could ultimately pave the way for economically efficient biomass conversion into bioethanol.

The focus of this research was to identify and characterise novel biomass degrading enzymes. The degradation of plant cell wall associated polysaccharides has been a widespread research area for the past few decades contributing greatly to knowledge and the significance of carbohydrate-active enzymes. However, it could be argued that much more is yet to be discovered with the ever expansion of GH families resulting from the increasing numbers of newly sequenced microbial genomes.

Chapter 4 reports the characterisation of a novel endo-xylanase, OtXyn8A, from GH8, a family to which there is only a limited amount of data on xylanase specificity. Previous reports have described that the GH8 xylanases constitute diverse substrate specificities with differences in major hydrolysis products. The study detailed in Chapter 4 aimed at evaluating such xylanases for their specificities against various substrates. A GH8 from an uncultured bacterium (Xyn8) demonstrated an endo-mode of action but with the ultimate release of xylotriose (Pollet *et al*., 2010; Lee *et al*., 2006). In contrast, OtXyn8A, proved to exhibit a higher activity against polysaccharide linear xylans from oat spelt and birchwood

with selectively releasing xylobiose from its substrates. Furthermore, these endo-activities differed completely from GH8 xylanases, REX and REXA, enzymes from *Bacillus halodurans* and *Bifidobacterium adolescentis*, respectively (Lagaert *et al*., 2007; Honda *et al*., 2004). These xylanases selectively released xylose from xylooligosaccharides but also partially from polymeric unsubstituted xylan after prolonged hydrolysis, which had not been reported by Lagaert *et al*., (2007) or Honda *et al*., (2004). Comparing the activities of the GH8 xylanases ultimately established that OtXyn8A exhibits novel specificity and furthermore contributing to the diversity of the GH8 family.

Additionally, Chapter 4 presented a study in which the specificity of OtXyn8A was compared to GH10 and GH11 xylanase members. Xylanases belonging to families 10 and 11 are well established and the understanding of xylan substrate specificity has mainly been deduced on the characterisation of enzymes from these two families. Therefore, against this background it seemed rational to survey the potential diversities between these two well recognised xylanase families and GH8 xylanases. OtXyn8A displayed endo-activity similar to the typical endo-modes of action of CmXyn10B (*Cellvibrio mixtus*), CjXyn10A (*Cellvibrio japonicus*), NpXyn11A (*Neocallimastix patriciarum*), however to varying degrees. CmXyn10B was active against highly decorated xylan polysaccharides, a property not observed among the GH11 and GH8 members. Even so, products of linear substrate hydrolysis, for example, oat spelt xylan (OSX), generated by OtXyn8A and CmXyn10B, were significantly similar with comparable amounts of xylose and xylobiose being released. Active-site characteristics of GH10 and GH11 xylanases, such as subsite properties, are highly conserved within each family but it is harder to make a generalisation for the lesser characterised GH8 xylanases. For example, the xylanase hydrolytic of the GH8 enzymes, REX and REXA, show that the enzymes have 3 subsites, while so far it can be assumed that OtXyn8A has at least 4. Furthermore, Collins *et al*., 2002 suggested that the GH8 xylanase, PhXyl, has at least 6 subsites. This variation in the number of subsites, along with possible

varying binding affinities, highlights the diverse substrate specificities among the GH8 xylanases.

Chapter 5 presented the identification of a gene cluster within *O. terrae* which encoded products utilised in xylan hydrolysis. Gene clusters are common within bacterial genomes but the identification of a gene organisation directed towards the breakdown of xylan substrates is potentially beneficial to biomass degradation and biorefining applications. Such a gene cluster would suggest the enzymes working synergistically and ultimately attacking substrates efficiently for total hydrolysis. The study in Chapter 5 showed that the enzymes produced from this gene cluster indeed generated more xylose when working together rather than the action of individual activities or in conjunction with other enzymes from the same organism. Furthermore, the understanding of such synergy between enzymes can be further assessed and exploited for the breakdown of complex substrates.

Chapter 6 attempted to explore various GH43 activities occurring in *O. terrae* to understand the multiple number of genes encoding GH43 enzymes within this organism and to propose the idea that each enzyme performs a particular role. There are 21 putative GH43 sequences within the *O. terrae* genome. While 15 GH43 proteins were successfully expressed, most of these either displayed no activity or trace activity against arabinan and arabinoxylan polysaccharides, suggesting that their natural substrates are more complex that those utilised within this research. However, Chapter 6 described the GH43 of *O. terrae* that exhibited significant activities against the tested substrates as well as two GH43 enzymes from *Lactobacillus brevis*. The data revealed that the GH43 α-L-arabinofuranosidases were functionally similar with the release of arabinose from their substrates. However, linkage specificity between enzymes appeared quite different as OtAraf43C was only active against α-1,5-linked arabinose while OtAraf43B exhibited activity against α-1,2- and / or α-1,3-linked arabinose units. Furthermore, the GH43 β-xylosidase, OtXyl43A displayed dual functionality as it was able to release arabinose as well as xylose from arabino-

xylooligosaccharides. Similar activity was identified for LbXyl43A, suggesting it too possessed a dual functional property. Although efforts into finding novel specificity within *O. terrae* GH43 were not a complete success, this study elucidated various activities and has given insights into the degree of diversity within enzyme specificities.

Ultimately, understanding the reasoning behind the evolutionary expansion of GH43 within *O. terrae* and other genomes such as *C. japonicus*, remains elusive. This however goes beyond just GH43 and into other families. For example, the *O. terrae* genome encodes 6 putative GH10 xylanases but the organism still requires the activity of a GH8 xylanase for xylan hydrolysis. Similarly, why genomes like *Clostridium thermocellum* host to a large number of cellulose degrading enzymes is also unclear. An approach into looking at gene clusters, like that described in Chapter 5, rather than looking at individual gene products, may provide insights in to how and why particular enzymes act synergistically to degrade substrates. It is only when an understanding of how glycoside hydrolases function within the context of the endogenous microbe as a whole is elucidated, that such enzymes can be applied to efficient degradation of the plant cell wall.

References

Abdou, L., Boileau, C., de Philip, P., Pagès, S., Fiérobe, H.-P., & Tardif, C. (2008). Transcriptional regulation of the Clostridium cellulolyticum cip-cel operon: a complex mechanism involving a catabolite-responsive element. *Journal of bacteriology*, *190*(5), 1499-506.

Achyuthan, K. E., Achyuthan, A. M., Adams, P. D., Dirk, S. M., Harper, J. C., Simmons, B. A., & Singh, A. K. (2010). Supramolecular self-assembled chaos: polyphenolic lignin's barrier to cost-effective lignocellulosic biofuels. *Molecules (Basel, Switzerland)*, *15*(12), 8641-88.

Adachi, W., Sakihama, Y., Shimizu, S., Sunami, T., Fukazawa, T., Suzuki, M., Yatsunami, R., et al. (2004). Crystal Structure of Family GH-8 Chitosanase with Subclass II Specificity from Bacillus sp . K17. *Journal of Molecular Biology*, *343*, 785-795.

Agrawal, P., Verma, D., & Daniell, H. (2011). Expression of Trichoderma reesei β-Mannanase in Tobacco Chloroplasts and Its Utilization in Lignocellulosic Woody Biomass Hydrolysis. *PloS one*, *6*(12), e29302.

Al-bari, M. A. A., Rahman, M. M. S., Islam, M. A. U., Flores, M. E., Bhuiyan, M. S. A., & Microbiology, P. (2007). Purification and Characterization of a â- (1 , 4) - Endoxylanase of Streptomyces bangladeshiensis sp . *Research Journal of CEell and Molecular Biology*, *1*(1), 31-36.

Alalouf, O., Balazs, Y., Volkinshtein, M., Grimpel, Y., Shoham, G., & Shoham, Y. (2011). A new family of carbohydrate esterases is represented by a GDSL/acetylxylan esterase from Geobacillus stearothermophilus. *The Journal of biological chemistry*, *3*, 1-18.

Alhassid, A., Ben-David, A., Tabachnikov, O., Libster, D., Naveh, E., Zolotnitsky, G., Shoham, Y., et al. (2009). Crystal structure of an inverting GH 43 1,5-alpha-L-arabinanase from Geobacillus stearothermophilus complexed with its substrate. *The Biochemical journal*, *422*(1), 73-82.

Alzari, P. M., Souchon, H., & Dominguez, R. (1996). The crystal structure of endoglucanase CelA, a family 8 glycosyl hydrolase from Clostridium thermocellum. *Structure (London, England : 1993)*, *4*(3), 265-75.

Antelmann, H., Tjalsma, H., Voigt, B., Ohlmeier, S., Bron, S., van Dijl, J. M., & Hecker, M. (2001). A proteomic view on genome-based signal peptide predictions. *Genome research*, *11*(9), 1484-502.

Bertran, M. S., & Dale, B. E. (1985). Enzymatic Hydrolysis and Recrystallization Behavior of Initially Amorphous Cellulose. *Biotechnology*, *XXVII*, 177-181.

Beylot, M., McKie, V., Voragen, A., Doeswijk-Voragen, C., Gilbert, H. (2001). The Pseudomonas cellulosa glycoside hydrolase family 51 arabinofuranosidase exhibits wide substrate specificity. *Biochemical Journal*, *358*, 607-614.

Biely, P., Vrsanska, M., Tenkanan, M., Kluepfel, D. (1997). Endo-P - 1,4-xylanase families: differences in catalytic properties. *Journal of biotechnology*, *57*, 151-166.

Blouzard, J.-C., Coutinho, P. M., Fierobe, H.-P., Henrissat, B., Lignon, S., Tardif, C., Pagès, S., et al. (2010). Modulation of cellulosome composition in Clostridium cellulolyticum: adaptation to the polysaccharide environment revealed by proteomic and carbohydrate-active enzyme analyses. *Proteomics*, *10*(3), 541-54.

Boerjan, W., Ralph, J., & Baucher, M. (2003). Lignin biosynthesis. *Annual review of plant biology*, *54*, 519-46.

Boisset, C., Fraschini, C., Schülein, M., Chanzy, H., Henrissat, B., & Schu, M. (2000). Imaging the Enzymatic Digestion of Bacterial Cellulose Ribbons Reveals the Endo Character of the Cellobiohydrolase Cel6A from Humicola insolens and Its Mode of Synergy with Cellobiohydrolase Imaging the Enzymatic Digestion of Bacterial Cellulose Ribbons R. *Applied and environmental microbiology*, *66*(4), 1444-1452.

Boraston, a B., Creagh, a L., Alam, M. M., Kormos, J. M., Tomme, P., Haynes, C. a, Warren, R. a, et al. (2001). Binding specificity and thermodynamics of a family 9 carbohydrate-binding module from Thermotoga maritima xylanase 10A. *Biochemistry*, *40*(21), 6240-7.

Boraston, A. B., Bolam, D. N., Gilbert, H. J., & Davies, G. J. (2004). Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *The Biochemical journal*, *382*(Pt 3), 769-81.

Bourgois, T. M., Van Craeyveld, V., Van Campenhout, S., Courtin, C. M., Delcour, J. a, Robben, J., & Volckaert, G. (2007). Recombinant expression and characterization of XynD from Bacillus subtilis subsp. subtilis ATCC 6051: a GH 43 arabinoxylan arabinofuranohydrolase. *Applied microbiology and biotechnology*, *75*(6), 1309-17.

Breton, C., Snajdrová, L., Jeanneau, C., Koca, J., & Imberty, A. (2006). Structures and mechanisms of glycosyltransferases. *Glycobiology*, *16*(2), 29R-37R.

Brown, R. M. (1994). Assembly of synthetic cellulose I. *Biochemistry*, *91*(August), 7425-7429.

Brown, R. M. (2004). Cellulose structure and biosynthesis: What is in store for the 21st century? *Journal of Polymer Science Part A: Polymer Chemistry*, *42*(3), 487-495.

Brunzelle, J. S., Jordan, D. B., McCaslin, D. R., Olczak, A., & Wawrzak, Z. (2008). Structure of the two-subsite beta-d-xylosidase from Selenomonas ruminantium in complex with 1,3-bis[tris(hydroxymethyl)methylamino]propane. *Archives of biochemistry and biophysics*, *474*(1), 157-66.

Brüx, C., Ben-David, A., Shallom-Shezifi, D., Leon, M., Niefind, K., Shoham, G., Shoham, Y., et al. (2006). The structure of an inverting GH43 beta-xylosidase from Geobacillus stearothermophilus with its substrate reveals the role of the three catalytic residues. *Journal of molecular biology*, *359*(1), 97-109.

Caffall, K. H., & Mohnen, D. (2009). The structure, function and biosynthesis of plant cell wall pectic polysaccharides. *Carbohydrate research*, *344*(14), 1879-900.

Cardoso, O. A. V., & Filho, E. X. F. (2003). Purification and characterization of a novel cellulase-free xylanase from Acrophialophora nainiana. *FEMS Microbiology Letters*, *223*(2), 309-314.

Cartmell, A., McKee, L. S., Peña, M. J., Larsbrink, J., Brumer, H., Kaneko, S., Ichinose, H., et al. (2011). The structure and function of an arabinan-specific alpha-1,2-arabinofuranosidase identified from screening the activities of bacterial GH43 glycoside hydrolases. *The Journal of biological chemistry*, *286*(17), 15483-95.

Carvalheiro, F., Duarte, L.C., Girio, F. M. (2008). Hemicellulose biorefineries: a review on biomass pretreatments. *Journal of Scientific &Industrial Research*, *67*, 849-864.

Charnock, S J, Bolam, D. N., Turkenburg, J. P., Gilbert, H. J., Ferreira, L. M., Davies, G. J., & Fontes, C. M. (2000). The X6 "thermostabilizing" domains of xylanases are carbohydrate-binding modules: structure and biochemistry of the Clostridium thermocellum X6b domain. *Biochemistry*, *39*(17), 5013-21.

Charnock, S J, Lakey, J. H., Virden, R., Hughes, N., Sinnott, M. L., Hazlewood, G. P., Pickersgill, R., et al. (1997). Key residues in subsite F play a critical role in the activity of Pseudomonas fluorescens subspecies cellulosa xylanase A against xylooligosaccharides but not against highly polymeric substrates such as xylan. *The Journal of biological chemistry*, *272*(5), 2942-51.

Charnock, S J, Spurway, T. D., Xie, H., Beylot, M. H., Virden, R., Warren, R. a, Hazlewood, G. P., et al. (1998). The topology of the substrate binding clefts of glycosyl hydrolase family 10 xylanases are not conserved. *The Journal of biological chemistry*, *273*(48), 32187-99.

Charnock, Simon J, Brown, I. E., Turkenburg, J. P., Black, G. W., & Davies, G. J. (2002). Convergent evolution sheds light on the anti-beta -elimination mechanism common to family 1 and 10 polysaccharide lyases. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(19), 12067-72.

Chin, K. J., Liesack, W., & Janssen, P. H. (2001). Opitutus terrae gen. nov., sp. nov., to accommodate novel strains of the division "Verrucomicrobia" isolated from rice paddy soil. *International journal of systematic and evolutionary microbiology*, *51*(Pt 6), 1965-8.

Cho, K. K., Kim, S. C., Woo, J. H., Bok, J. D., & Choi, Y. J. (2000). Molecular cloning and expression of a novel family A endoglucanase gene from Fibrobacter succinogenes S85 in Escherichia coli. *Enzyme*, *27*, 475-481.

Collins, T., Gerday, C., & Feller, G. (2005a). Xylanases, xylanase families and extremophilic xylanases. *FEMS microbiology reviews*, *29*(1), 3-23.

Collins, T., Gerday, C., & Feller, G. (2005b). Xylanases, xylanase families and extremophilic xylanases. *FEMS microbiology reviews*, *29*(1), 3-23.

Collins, T., Meuwis, M.-A., Stals, I., Claeyssens, M., Feller, G., & Gerday, C. (2002). A novel family 8 xylanase, functional and physicochemical characterization. *The Journal of biological chemistry*, *277*(38), 35133-9.

Contreras, L. M., Gómez, J., Prieto, J., Clemente-Jiménez, J. M., Las Heras-Vázquez, F. J., Rodríguez-Vico, F., Blanco, F. J., et al. (2008). The family 52 beta-xylosidase from Geobacillus stearothermophilus is a dimer: structural and biophysical characterization of a glycoside hydrolase. *Biochimica et biophysica acta*, *1784*(12), 1924-34.

Correia, M. a S., Mazumder, K., Brás, J. L. a, Firbank, S. J., Zhu, Y., Lewis, R. J., York, W. S., et al. (2011). Structure and function of an arabinoxylan-specific xylanase. *The Journal of biological chemistry*, *286*(25), 22510-20.

Correia, M. a S., Prates, J. a M., Brás, J., Fontes, C. M. G. a, Newman, J. a, Lewis, R. J., Gilbert, H. J., et al. (2008). Crystal structure of a cellulosomal family 3 carbohydrate esterase from Clostridium thermocellum provides insights into the mechanism of substrate recognition. *Journal of molecular biology*, *379*(1), 64-72.

Crombie, H. J., Chengappa, S., Hellyer, a, & Reid, J. S. (1998). A xyloglucan oligosaccharide-active, transglycosylating beta-D-glucosidase from the cotyledons of nasturtium (Tropaeolum majus L) seedlings--purification, properties and characterization of a cDNA clone. *The Plant journal : for cell and molecular biology*, *15*(1), 27-38.

Czjzek, M., Ben David, A., Bravman, T., Shoham, G., Henrissat, B., & Shoham, Y. (2005). Enzyme-substrate complex structures of a GH39 beta-xylosidase from Geobacillus stearothermophilus. *Journal of molecular biology*, *353*(4), 838-46.

Dashtban, M., Schraft, H., & Qin, W. (2009). Fungal bioconversion of lignocellulosic residues; opportunities & perspectives. *International Journal of Biological Sciences*, *5*(6), 578.

Davies, G, & Henrissat, B. (1995). Structures and mechanisms of glycosyl hydrolases. *Structure (London, England : 1993)*, *3*(9), 853-9.

Davies, GJ, & Wilson, K. (1997). Nomenclature for sugar-binding subsites in glycosyl hydrolases. *Biochemical Journal*, *559*, 557-559.

De Vos, D., Collins, T., Nerinckx, W., Savvides, S. N., Claeyssens, M., Gerday, C., Feller, G., et al. (2006). Oligosaccharide binding in family 8 glycosidases: crystal structures of active-site mutants of the beta-1,4-xylanase pXyl from Pseudoaltermonas haloplanktis TAH3a in complex with substrate and product. *Biochemistry*, *45*(15), 4797-807.

DeBoy, R. T., Mongodin, E. F., Fouts, D. E., Tailford, L. E., Khouri, H., Emerson, J. B., Mohamoud, Y., et al. (2008). Insights into plant cell wall degradation from the genome sequence of the soil bacterium Cellvibrio japonicus. *Journal of bacteriology*, *190*(15), 5455-63.

Divne, C., Ståhlberg, J., Teeri, T. T., & Jones, T. a. (1998). High-resolution crystal structures reveal how a cellulose chain is bound in the 50 A long tunnel of cellobiohydrolase I from Trichoderma reesei. *Journal of molecular biology*, *275*(2), 309-25.

Do, B.-C., Dang, T.-T., Berrin, J.-G., Haltrich, D., To, K.-A., Sigoillot, J.-C., & Yamabhai, M. (2009). Cloning, expression in Pichia pastoris and characterization of a

thermostable GH5 mannan endo-1,4-beta-mannosidase from Aspergillus niger BK01. *Microbial cell factories*, *8*, 59.

Dodd, D., & Cann, I. K. O. (2009). Enzymatic deconstruction of xylan for biofuel production. *Global change biology. Bioenergy*, *1*(1), 2-17.

Dodd, D., Kocherginskaya, S. a, Spies, M. A., Beery, K. E., Abbas, C. a, Mackie, R. I., & Cann, I. K. O. (2009). Biochemical analysis of a beta-D-xylosidase and a bifunctional xylanase-ferulic acid esterase from a xylanolytic gene cluster in Prevotella ruminicola 23. *Journal of bacteriology*, *191*(10), 3328-38.

Dror, T. W., Morag, E., Rolider, A., Bayer, E. A., Lamed, R., & Shoham, Y. (2003). Regulation of the Cellulosomal celS (cel48A) Gene of Clostridium thermocellum Is Growth Rate Dependent. *Society*, *185*(10), 3042-3048.

Durand, A., Hughes, R., Roussel, A., Flatman, R., Henrissat, B., & Juge, N. (2005). Emergence of a subfamily of xylanase inhibitors within glycoside hydrolase family 18. *The FEBS journal*, *272*(7), 1745-55.

Ebringerová, A. (2005). Structural Diversity and Application Potential of Hemicelluloses. *Macromolecular Symposia*, *232*(1), 1-12.

Elisabetta Sabini, Gerlind Sulzenbacher, Miroslava Dauter, Zbigniew Dauter, Per Lina Jorgensen, Martin Schulein, Claude Dupont, G. J. D. and K. S. W. (1999). Catalysis and specificity in enzymatic glycoside hydrolysis: a 2,5B confomation for the glycosy-enzyme intermediate revealed by the structure of the Bacillus agaradhaerens family 11 xylanase. *Chemistry & Biology*, *6*, 483-492.

Eneyskaya, E. V., Brumer, H., Backinowsky, L. V., Ivanen, D. R., Kulminskaya, A. a, Shabalin, K. a, & Neustroev, K. N. (2003). Enzymatic synthesis of beta-xylanase substrates: transglycosylation reactions of the beta-xylosidase from Aspergillus sp. *Carbohydrate research*, *338*(4), 313-25.

Fanutti, C., Ponyi, T., Black, G. W., Hazlewood, G. P., & Gilbert, H. J. (1995). The conserved noncatalytic 40-residue sequence in cellulases and hemicellulases from anaerobic fungi functions as a protein docking domain. *The Journal of biological chemistry*, *270*(49), 29314-22.

Faulds, C. B., Mandalari, G., Lo Curto, R. B., Bisignano, G., Christakopoulos, P., & Waldron, K. W. (2006). Synergy between xylanases from glycoside hydrolase family 10 and family 11 and a feruloyl esterase in the release of phenolic acids from cereal arabinoxylan. *Applied microbiology and biotechnology*, *71*(5), 622-9.

Fernandez-Abalos, J. M. Reviejo, V., Diaz, M., Rodriguez, S., Leal, F., Santamaria, R. I. (2003). Posttranslational processing of the xylanase Xys1L from Streptomyces halstedii JM8 is carried out by secreted serine proteases. *Microbiology*, *149*(7), 1623-1632.

Ficko-Blean, E., Stuart, C. P., & Boraston, A. B. (2011). Structural analysis of CPF_2247, a novel α-amylase from Clostridium perfringens. *Proteins*, *79*(10), 2771-7.

Finch, P. (1999). Carbohydrate: Structures, Syntheses and Dynamics. Kluwer Academic Press. pp 37-38.

Fischbach, M., & Voigt, C. a. (2010). Prokaryotic gene clusters: a rich toolbox for synthetic biology. *Biotechnology journal*, *5*(12), 1277-96.

Flipphi, M. J., Visser, J., van der Veen, P., & de Graaff, L. H. (1994). Arabinase gene expression in Aspergillus niger: indications for coordinated regulation. *Microbiology*, *140*, 2673-82.

Fort, S., Varrot, a, Schülein, M., Cottaz, S., Driguez, H., & Davies, G. J. (2001). Mixed-linkage cellooligosaccharides: a new class of glycoside hydrolase inhibitors. *Chembiochem : a European journal of chemical biology*, *2*(5), 319-25.

Fujimoto, Z., Ichinose, H., Maehara, T., Honda, M., Kitaoka, M., & Kaneko, S. (2010). Crystal structure of an Exo-1,5-{alpha}-L-arabinofuranosidase from Streptomyces avermitilis provides insights into the mechanism of substrate discrimination between exo- and endo-type enzymes in glycoside hydrolase family 43. *The Journal of biological chemistry*, *285*(44), 34134-43.

Fujimoto, Z., Kaneko, S., Kuno, A., Kobayashi, H., Kusakabe, I., & Mizuno, H. (2004). Crystal structures of decorated xylooligosaccharides bound to a family 10 xylanase from Streptomyces olivaceoviridis E-86. *The Journal of biological chemistry*, *279*(10), 9606-14.

Fushinobu, S., Hidaka, M., Honda, Y., Wakagi, T., Shoun, H., & Kitaoka, M. (2005). Structural basis for the specificity of the reducing end xylose-releasing exo-oligoxylanase from Bacillus halodurans C-125. *The Journal of biological chemistry*, *280*(17), 17180-6.

Garron, M.-L., & Cygler, M. (2010). Structural and mechanistic classification of uronic acid-containing polysaccharide lyases. *Glycobiology*, *20*(12), 1547-73.

Gasparic, a, Martin, J., Daniel, a S., & Flint, H. J. (1995). A xylan hydrolase gene cluster in Prevotella ruminicola B(1)4: sequence relationships, synergistic interactions and oxygen sensitivity of a novel enzyme with exoxylanase and beta-(1,4)-xylosidase activities. *Applied and environmental microbiology*, *61*(8), 2958-64.

Geib, S. M., Filley, T. R., Hatcher, P. G., Hoover, K., Carlson, J. E., Jimenez-Gasco, M. D. M., Nakagawa-Izumi, A., et al. (2008). Lignin degradation in wood-feeding insects. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(35), 12932-7.

Ghosh, M., & Nanda, G. (1993). Thermostability of P-xylosidase from Aspergillus sydowi MG49. *FEBS Letters*, *330*(3), 275-278.

Gielkens, M., González-Candelas, L., Sánchez-Torres, P., van de Vondervoort, P., de Graaff, L., Visser, J., & Ramón, D. (1999). The abfB gene encoding the major alpha-L-arabinofuranosidase of Aspergillus nidulans: nucleotide sequence, regulation and construction of a disrupted strain. *Microbiology (Reading, England)*, *145*, 735-741.

Gilbert, H. J., Hazlewood, G. P., Laurie, J. I., Xue, G. P., & Ne, T. (1992). Homologous catalytic domains in a rumen fungal xylanase : evidence for gene duplication and prokaryotic origin. *Molecular Microbiology*, *6*(15), 2065-2072.

Gilead, S., & Shoham, Y. (1995). Purification and characterization of alpha-L-arabinofuranosidase from Bacillus stearothermophilus T-6. *Applied and environmental microbiology*, *61*(1), 170-174.

Gilkes, N. R., Warren, R. a, Miller, R. C., & Kilburn, D. G. (1988). Precise excision of the cellulose binding domains from two Cellulomonas fimi cellulases by a homologous protease and the effect on catalysis. *The Journal of biological chemistry*, *263*(21), 10401-7.

Gloster, T. M., & Davies, G. J. (2010). Glycosidase inhibition : assessing mimicry of the transition state. *Org Biomol Chem*, *8*(2), 305.

Gröndahl, M., Teleman, A., & Gatenholm, P. (2003). Effect of acetylation on the material properties of glucuronoxylan from aspen wood. *Carbohydrate Polymers*, *52*(4), 359-366.

Guimarães, B. G., Souchon, H., Lytle, B. L., David Wu, J. H., & Alzari, P. M. (2002). The Crystal Structure and Catalytic Mechanism of Cellobiohydrolase CelS, the Major Enzymatic Component of the Clostridium thermocellum Cellulosome. *Journal of Molecular Biology*, *320*(3), 587-596.

Gupta, S., Bhushan, B., & Hoondal, G. S. (2000). Isolation , purification and characterization of xylanase from Staphylococcus sp . SG-13 and its application in biobleaching of kraft pulp. *Journal of Applied Microbiology*, *88*, 325-334.

Hatfield, R., & Vermerris, W. (2001). Update on Lignification Lignin Formation in Plants . The Dilemma of Linkage Specificity. *Plant physiology*, *126*, 1351-1357.

Heinze, T., Barsett, H., Ebringerová, A. (2005) Polysaccharides I: structure, characterisation and use. *Adv. Polym. Sci*. 186, 1-67.

Heneghan, M. N., McLoughlin, L., Murray, P. G., & Tuohy, M. G. (2007). Cloning, characterisation and expression analysis of α-glucuronidase from the thermophilic fungus Talaromyces emersonii. *Enzyme and Microbial Technology*, *41*(6-7), 677-682.

Henrissat, B, Teeri, T. T., & Warren, R. a. (1998). A scheme for designating enzymes that hydrolyse the polysaccharides in the cell walls of plants. *FEBS letters*, *425*(2), 352-4.

Henrissat, Bernard. (1991). A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochemical Journal*, *280*, 309-316.

Hervé, C., Rogowski, A., Blake, A. W., Marcus, S. E., Gilbert, H. J., & Knox, J. P. (2010). Carbohydrate-binding modules promote the enzymatic deconstruction of intact plant cell walls by targeting and proximity effects. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(34), 15293-8.

Honda, Y., Fushinobu, S., Hidaka, M., Wakagi, T., Shoun, H., Taniguchi, H., & Kitaoka, M. (2008). Alternative strategy for converting an inverting glycoside hydrolase into a glycosynthase. *Glycobiology*, *18*(4), 325-30.

Honda, Y., & Kitaoka, M. (2004). A family 8 glycoside hydrolase from Bacillus halodurans C-125 (BH2105) is a reducing end xylose-releasing exo-oligoxylanase. *The Journal of biological chemistry*, *279*(53), 55097-103.

Hövel, K., Shallom, D., Niefind, K., Belakhov, V., Shoham, G., Baasov, T., Shoham, Y., et al. (2003). Crystal structure and snapshots along the reaction pathway of a family 51 alpha-L-arabinofuranosidase. *The EMBO journal*, *22*(19), 4922-32.

Ichinose, H., Yoshida, M., Fujimoto, Z., & Kaneko, S. (2008). Characterization of a modular enzyme of exo-1,5-alpha-L-arabinofuranosidase and arabinan binding module from Streptomyces avermitilis NBRC14893. *Applied microbiology and biotechnology*, *80*(3), 399-408.

Inácio, J. M., Correia, I. L., & de Sá-Nogueira, I. (2008). Two distinct arabinofuranosidases contribute to arabino-oligosaccharide degradation in Bacillus subtilis. *Microbiology (Reading, England)*, *154*(Pt 9), 2719-29.

Izydorczyk, M. S., & Biliaderi, C. G. (1995). Cereal arabinoxylans : advances in structure and physicochemical properties, *Carbohydrate Polymers, 28,* 33-48.

Jitonnom, J., Lee, V. S., Nimmanpipug, P., Rowlands, H. A., & Mulholland, A. J. (2011). Quantum Mechanics/Molecular Mechanics Modeling of Substrate-Assisted Catalysis in Family 18 Chitinases: Conformational Changes and the Role of Asp142 in Catalysis in ChiB. *Biochemistry*, *50*(21), 4697-4711.

Jones, P. G., VanBogelen, R. a, & Neidhardt, F. C. (1987). Induction of proteins in response to low temperature in Escherichia coli. *Journal of bacteriology*, *169*(5), 2092-5.

Jordan, D. B., Li, X.-L., Dunlap, C. a, Whitehead, T. R., & Cotta, M. a. (2007). Structure-function relationships of a catalytically efficient beta-D-xylosidase. *Applied biochemistry and biotechnology*, *141*(1), 51-76.

Juge, N., Payan, F., & Williamson, G. (2004). XIP-I, a xylanase inhibitor protein from wheat: a novel protein function. *Biochimica et biophysica acta*, *1696*(2), 203-11.

K Usui., K, Ibata., T, Suzuki., K, K. (1999). XynX, a possible Exo-xylanase of Aeromonas caviae ME-1 that Produces Exclusively Xylobiose and Xylotetraose from Xylan. *Biosci. Biotechnol. Biochem.*, *63*(8), 1346-1352.

KOSHLAND, D. E. (1953). STEREOCHEMISTRY AND THE MECHANISM OF ENZYMATIC REACTIONS. *Biological Reviews*, *28*(4), 416-436.

Kaneko, S., Ichinose, H., Fujimoto, Z., Kuno, A., Yura, K., Go, M., Mizuno, H., et al. (2004). Structure and function of a family 10 beta-xylanase chimera of Streptomyces olivaceoviridis E-86 FXYN and Cellulomonas fimi Cex. *The Journal of biological chemistry*, *279*(25), 26619-26.

Kawai, R., Igarashi, K., Kitaoka, M., Ishii, T., & Samejima, M. (2004). Kinetics of substrate transglycosylation by glycoside hydrolase family 3 glucan (1-->3)-beta-glucosidase from the white-rot fungus Phanerochaete chrysosporium. *Carbohydrate research*, *339*(18), 2851-7.

Ketudat Cairns, J. R., & Esen, A. (2010). β-Glucosidases. *Cellular and molecular life sciences : CMLS*, *67*(20), 3389-405.

Kezdy, F.J., Bender, M. L. (1962). The Kinetics of the a-Chymotrpsin-Catalyzed Hydrolysis of p-Nitrophenyl Acetate. *Biochemistry*, *1*, 1097-1106.

Khasin, a, Alchanati, I., & Shoham, Y. (1993). Purification and characterization of a thermostable xylanase from Bacillus stearothermophilus T-6. *Applied and environmental microbiology*, *59*(6), 1725-30.

Kleman-leyer, K. M., Siika-aho, M., & Teeri, T. T. (1996). The Cellulases Endoglucanase I and Cellobiohydrolase II of Trichoderma reesei Act Synergistically To Solubilize Native Cotton Cellulose but Not To Decrease Its Molecular Size . *Applied and environmental microbiology*, *62*(8), 2883-2887.

Knob, a., Terrasan, C. R. F., & Carmona, E. C. (2009). β-Xylosidases from filamentous fungi: an overview. *World Journal of Microbiology and Biotechnology*, *26*(3), 389-407.

Kolenová, K., Vrsanská, M., & Biely, P. (2006). Mode of action of endo-beta-1,4-xylanases of families 10 and 11 on acidic xylooligosaccharides. *Journal of biotechnology*, *121*(3), 338-45.

Koonin, E. V. (2009). Evolution of genome architecture. *The international journal of biochemistry & cell biology*, *41*(2), 298-306.

Kruus, K., Wang, W. K., Ching, J., & Wu, J. H. (1995). Exoglucanase activities of the recombinant Clostridium thermocellum CelS, a major cellulosome component. *Journal of bacteriology*, *177*(6), 1641-4.

Kubata, B K, Takamizawa, K., Kawai, K., Suzuki, T., & Horitsu, H. (1995). Xylanase IV, an Exoxylanase of Aeromonas caviae ME-1 Which Produces Xylotetraose as the Only Low-Molecular-Weight Oligosaccharide from Xylan. *Applied and environmental microbiology*, *61*(4), 1666-8.

Kubata, Bruno Kilunga, Suzuki, T., & Horitsu, H. (1994). Purification and Characterization of Aeromonas caviae ME-1 Xylanase V , Which Produces Exclusively Xylobiose from Xylan. *Applied and Environmental Microbiology*, *60*(2), 531-535.

Kuhad, R. C., Gupta, R., & Singh, A. (2011). Microbial cellulases and their industrial applications. *Enzyme research*, *2011*, 280696.

Kumar, P., Barrett, D. M., Delwiche, M. J., & Stroeve, P. (2009). Methods for Pretreatment of Lignocellulosic Biomass for Efficient Hydrolysis and Biofuel Production. *Society*, 3713-3729.

Kurasin, M., & Väljamäe, P. (2011). Processivity of cellobiohydrolases is limited by the substrate. *The Journal of biological chemistry*, *286*(1), 169-77.

Lagaert, S., Pollet, A., Delcour, J. a, Lavigne, R., Courtin, C. M., & Volckaert, G. (2010). Substrate specificity of three recombinant α-L-arabinofuranosidases from Bifidobacterium adolescentis and their divergent action on arabinoxylan and arabinoxylan oligosaccharides. *Biochemical and biophysical research communications*, *402*(4), 644-50.

Lagaert, S., Van Campenhout, S., Pollet, A., Bourgois, T. M., Delcour, J. a, Courtin, C. M., & Volckaert, G. (2007). Recombinant expression and characterization of a reducing-end xylose-releasing exo-oligoxylanase from Bifidobacterium adolescentis. *Applied and environmental microbiology*, *73*(16), 5374-7.

Lamed, R., Setiter, E. V. A., & Bayer, E. A. (1983). Characterization of a complex in Clostridium Characterization of a Cellulose-Binding , Cellulase-Containing Complex in Clostridium thermocellum. *American Society for Microbiology*, *156*, 828-836.

Lawrence, J. G., & Roth, J. R. (1996). Selfish operons: horizontal transfer may drive the evolution of gene clusters. *Genetics*, *143*(4), 1843-60.

Lee, C. C., Kibblewhite-Accinelli, R. E., Wagschal, K., Robertson, G. H., & Wong, D. W. S. (2006). Cloning and characterization of a cold-active xylanase enzyme from an environmental DNA library. *Extremophiles : life under extreme conditions*, *10*(4), 295-300.

Leggio, L. L., Jenkins, J., Harris, G. W., & Pickersgill, R. W. (2000). X-ray crystallographic study of xylopentaose binding to Pseudomonas fluorescens xylanase A. *Proteins*, *41*(3), 362-73.

Mahalingeshwara Bhat, Joel Solomon Gaikwad, R. M. (1993). Purification and characterization of an extracellular P-glucosidase from the thermophilic fungus Sporotrichum thermophile and its influence on cellulase activity. *Journal of General Microbiology*, *139*, 2825-2832.

Mai, V., Wiegel, J., & Lorenz, W. W. (2000). Cloning, sequencing and characterization of the bifunctional xylosidase-arabinosidase from the anaerobic thermophile thermoanaerobacter ethanolicus. *Gene*, *247*(1-2), 137-43.

Martin, F. J., & McInerney, J. O. (2009). Recurring cluster and operon assembly for Phenylacetate degradation genes. *BMC evolutionary biology*, *9*, 36. doi:10.1186/1471-2148-9-36.

Matsuo, N., Kaneko, S., Kuno, A., Kobayashi, H., & Kusakaber, I. (2000). Purification, characterisation and gene cloning of two α- L -arabinofuranosidases from Streptomyces chartreusis GS901. *Biochemical Journal*, *346*, 9-15.

Mazur, O., & Zimmer, J. (2011). Apo- and cellopentaose-bound structures of the bacterial cellulose synthase subunit BcsZ. *The Journal of biological chemistry*, *286*(20), 17601-6.

McKie, V. A. M. C., Black, G. W., Millward-sadler, S. J., Hazlewood, G. P., Laurie, J. I., & Gilbert, H. J. (1997). Arabinase A from Pseudomonase fluorescens subsp. cellulsa exhibits both an endo- and exo- mode of action. *Biochemical Journal*, *323*, 547-555.

Michlmayr, H., Schümann, C., Kulbe, K. D., & del Hierro, A. M. (2011). Heterologously expressed family 51 alpha-L-arabinofuranosidases from Oenococcus oeni and Lactobacillus brevis. *Applied and environmental microbiology*, *77*(4), 1528-31.

Miller, G. (1959). Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Analytical chemistry*, *31*(3), 426–428. ACS Publications.

Moraïs, S., Salama-Alber, O., Barak, Y., Hadar, Y., Wilson, D. B., Lamed, R., Shoham, Y., et al. (2012). Functional Association of Catalytic and Ancillary Modules Dictates Enzymatic Activity in Glycoside Hydrolase Family 43 β-Xylosidase. *The Journal of biological chemistry*, *287*(12), 9213-21.

Murata, T., & Usui, T. (2002). Efficient Synthetic Method of Obtaining Oligosaccharide Units and Deriva- tives Utilizing Endoglycosidases. *Biotechnol. Bioprocess Eng.*, *7*, 263-267.

Nagy, T., Emami, K., Fontes, C. M. G. A., Luis, M., Ferreira, A., Humphry, D. R., Gilbert, H. J., et al. (2002). The Membrane-Bound α -Glucuronidase from Pseudomonas cellulosa Hydrolyzes 4- but Not 4- O- Methyl-d-Glucuronoxylan. *Journal of Bacteriology*, *184*(17), 4925-4929.

Nagy, T., Nurizzo, D., Davies, G. J., Biely, P., Lakey, J. H., Bolam, D. N., & Gilbert, H. J. (2003). The alpha-Glucuronidase , GlcA67A , of Cellvibrio japonicus Utilizes the Carboxylate and Methyl Groups of Aldobiouronic Acid as Important Substrate Recognition Determinants *. *The Journal of Biological Chemistry*, *278*(22), 20286-20292.

Nakajima, M., Nishimoto, M., & Kitaoka, M. (2010). Characterization of d-galactosyl-β1→4-l-rhamnose phosphorylase from Opitutus terrae. *Enzyme and Microbial Technology*, *46*(3-4), 315-319.

Nataf, Y., Bahari, L., Kahel-Raifer, H., Borovok, I., Lamed, R., Bayer, E. a, Sonenshein, A. L., et al. (2010). Clostridium thermocellum cellulosomal genes are regulated by extracytoplasmic polysaccharides via alternative sigma factors. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(43), 18646-51.

Ng, I.-S., Tsai, S.-W., Ju, Y.-M., Yu, S.-M., & Ho, T.-hua D. (2011). Dynamic synergistic effect on Trichoderma reesei cellulases by novel β-glucosidases from Taiwanese fungi. *Bioresource technology*, *102*(10), 6073-81. Elsevier Ltd.

Nurizzo, D., Nagy, T., Gilbert, H. J., & Davies, G. J. (2002). The structural basis for catalysis and specificity of the Pseudomonas cellulosa alpha-glucuronidase, GlcA67A. *Structure (London, England : 1993)*, *10*(4), 547-56.

Nurizzo, D., Turkenburg, J. P., Charnock, S. J., Roberts, S. M., Dodson, E. J., McKie, V. a, Taylor, E. J., et al. (2002). Cellvibrio japonicus alpha-L-arabinanase 43A has a novel five-blade beta-propeller fold. *Nature structural biology*, *9*(9), 665-8.

Ochiai, A., Itoh, T., Maruyama, Y., Kawamata, A., Mikami, B., Hashimoto, W., & Murata, K. (2007). A novel structural fold in polysaccharide lyases: Bacillus subtilis family 11 rhamnogalacturonan lyase YesW with an eight-bladed beta-propeller. *The Journal of biological chemistry*, *282*(51), 37134-45.

Osbourn, A. E., & Field, B. (2009). Operons. *Cellular and molecular life sciences : CMLS*, *66*(23), 3755-75.

Ospina-Giraldo, M. D., McWalters, J., & Seyer, L. (2010). Structural and functional profile of the carbohydrate esterase gene complement in Phytophthora infestans. *Current genetics*, *56*(6), 495-506.

Park, S., Baker, J. O., Himmel, M. E., Parilla, P. a, & Johnson, D. K. (2010). Cellulose crystallinity index: measurement techniques and their impact on interpreting cellulase performance. *Biotechnology for Biofuels*, *3*(1), 10.

Parkkinen, T., Hakulinen, N., Tenkanen, M., Siika-aho, M., & Rouvinen, J. (2004). Crystallization and preliminary X-ray analysis of a novel Trichoderma reesei xylanase IV belonging to glycoside hydrolase family 5. *Acta crystallographica. Section D, Biological crystallography*, *60*(Pt 3), 542-4.

Paës, G., Berrin, J.-G., & Beaugrand, J. (2011). GH11 xylanases: Structure/function/properties relationships and applications. *Biotechnology advances*. Elsevier Inc.

Pell, G., Szabo, L., Charnock, S. J., Xie, H., Gloster, T. M., Davies, G. J., & Gilbert, H. J. (2004a). Structural and biochemical analysis of Cellvibrio japonicus xylanase 10C: how variation in substrate-binding cleft influences the catalytic profile of family GH-10 xylanases. *The Journal of biological chemistry*, *279*(12), 11777-88.

Pell, G., Szabo, L., Charnock, S. J., Xie, H., Gloster, T. M., Davies, G. J., & Gilbert, H. J. (2004b). Structural and biochemical analysis of Cellvibrio japonicus xylanase 10C: how variation in substrate-binding cleft influences the catalytic profile of family GH-10 xylanases. *The Journal of biological chemistry*, *279*(12), 11777-88.

Pell, G., Taylor, E. J., Gloster, T. M., Turkenburg, J. P., Fontes, C. M. G. A., Ferreira, L. M. A., Nagy, T., et al. (2004). The Mechanisms by Which Family 10 Glycoside Hydrolases Bind Decorated Substrates *. *The Journal of Biological Chemistry*, *279*(10), 9597-9605.

Petegem, F. V., Collins, T., Meuwis, M.-alice, Gerday, C., Feller, G., & Beeumen, J. V. (2003). The Structure of a Cold-adapted Family 8 Xylanase at 1 . 3 Å Resolution. *The Journal of Biological Chemistry*, *278*(9), 7531-7539.

Pollet, A., Schoepe, J., Dornez, E., Strelkov, S. V., Delcour, J. a, & Courtin, C. M. (2010). Functional analysis of glycoside hydrolase family 8 xylanases shows narrow but distinct substrate specificities and biotechnological potential. *Applied microbiology and biotechnology*, *87*(6), 2125-35.

Pons, T., Naumoff, D. G., Martínez-Fleites, C., & Hernández, L. (2004). Three acidic residues are at the active site of a beta-propeller architecture in glycoside hydrolase families 32, 43, 62 and 68. *Proteins*, *54*(3), 424-32.

Proctor, M. R., Taylor, E. J., Nurizzo, D., Turkenburg, J. P., Lloyd, R. M., Vardakou, M., Davies, G. J., et al. (2005). Tailored catalysts for plant cell-wall degradation: redesigning the exo/endo preference of Cellvibrio japonicus arabinanase 43A. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(8), 2697-702.

Punja, Z. K., & Zhang, Y. Y. (1993). Plant chitinases and their roles in resistance to fungal diseases. *Journal of nematology*, *25*(4), 526-40.

Pérez-Gonzalez, J. a, De Graaff, L. H., Visser, J., & Ramón, D. (1996). Molecular cloning and expression in Saccharomyces cerevisiae of two Aspergillus nidulans xylanase genes. *Applied and environmental microbiology*, *62*(6), 2179-82.

Quay, D. H. X., Bakar, F. D. A., Rabu, A., Said, M., Illias, R. M., Mahadi, N. M., & Hassan, O. (2011). Overexpression , purification and characterization of the Aspergillus niger endoglucanase , EglA , in Pichia pastoris. *African Journal of Biotechnology*, *10*(11), 2101-2111.

Rasmussen, L. E., Sørensen, J. F., & Meyer, A. S. (2010). Kinetics and substrate selectivity of a Triticum aestivum xylanase inhibitor (TAXI) resistant D11F/R122D variant of Bacillus subtilis XynA xylanase. *Journal of biotechnology*, *146*(4), 207-14.

Rouvinen, J., Bergfors, T., Teeri, T., Knowles, J. K., & Jones, T. A. (1990). Three-dimensional structure of cellobiohydrolase II from Trichoderma reesei . *Science* , *249* (4967), 380-386.

Ryabova, O., Vrsanská, M., Kaneko, S., van Zyl, W. H., & Biely, P. (2009). A novel family of hemicellulolytic alpha-glucuronidase. *FEBS letters*, *583*(9), 1457-62. Federation of European Biochemical Societies.

Saha, B. C. (2003). Hemicellulose bioconversion. *Journal of industrial microbiology & biotechnology*, *30*(5), 279-91.

Sambrook, J., Russell, D.J. (2001) Molecular Cloning: A Laboratory Manual,*Volume 1* Cold Spring Harbor Laboratory Press.

Saxena, I. M., Brown, R. M., & Dandekar, T. (2001). Structure–function characterization of cellulose synthase: relationship to other glycosyltransferases. *Phytochemistry*, *57*(7), 1135-1148.

Scavetta, R. D., Herron, S. R., Hotchkiss, a T., Kita, N., Keen, N. T., Benen, J. a, Kester, H. C., et al. (1999). Structure of a plant cell wall fragment complexed to pectate lyase C. *The Plant cell*, *11*(6), 1081-92.

Schmidt, R. K., Karplus, M., & Brady, J. W. (1996). The Anomeric Equilibrium in D - Xylose : Free Energy and the Role of Solvent Structuring. *Journal of the American Chemical Society*, (5), 541-546.

Seyedarabi, A., To, T. T., Ali, S., Hussain, S., Fries, M., Madsen, R., Clausen, M. H., et al. (2010). Structural insights into substrate specificity and the anti beta-elimination mechanism of pectate lyase. *Biochemistry*, *49*(3), 539-46.

Shallom, D., & Shoham, Y. (2003). Microbial hemicellulases. *Current Opinion in Microbiology*, *6*(3), 219-228.

Shoham, Y., Lamed, R., & Bayer, E. a. (1999). The cellulosome concept as an efficient microbial strategy for the degradation of insoluble polysaccharides. *Trends in microbiology*, *7*(7), 275-81.

Sidhu, G., Withers, S. G., Nguyen, N. T., McIntosh, L. P., Ziser, L., & Brayer, G. D. (1999). Sugar ring distortion in the glycosyl-enzyme intermediate of a family G/11 xylanase. *Biochemistry*, *38*(17), 5346-54.

Simpson, D. J., Fincher, G. B., Huang, A. H. ., & Cameron-Mills, V. (2003). Structure and Function of Cereal and Related Higher Plant (1→4)-β-Xylan Endohydrolases. *Journal of Cereal Science*, *37*(2), 111-127.

Sinnott, M. L. (1990). Catalytic Mechanisms of Enzymic Glycosyl Transfer. *Chemical Reviews*, *90*(7), 1171-1202.

Sjöström, E. (1993). Wood Chemistry: Fundamentals and applications, Academic Press, 2nd Edition.

Smaali, I., Rémond, C., & O'Donohue, M. J. (2006). Expression in Escherichia coli and characterization of beta-xylosidases GH39 and GH-43 from Bacillus halodurans C-125. *Applied microbiology and biotechnology*, *73*(3), 582-90.

Sofie A. Frederix., Klaartje E. Van Hoeymissen., Christophe M. Courtin., J. A. D. (2004). Water-Extractable and Water-Unextractable Arabinoxylans Affect Gluten Agglomeration Behavior during Wheat Flour Gluten − Starch Separation. *Journal of Agricultural And Food Chemistry*, *52*, 7950-7956.

St Leger, R. J., Joshi, L., & Roberts, D. W. (1997). Adaptation of proteases and carbohydrates of saprophytic, phytopathogenic and entomopathogenic fungi to the requirements of their ecological niches. *Microbiology (Reading, England)*, *143 (Pt 6*(1 997), 1983-92.

Stahl, Franklin, W., Murray, N. E. (1966). THE EVOLUTION OF GENE CLUSTERS AND GENETIC CIRCULARITY IN MICROORGANSIMS. *Genetics*, *53*, 569-576.

Subramaniyan, S., & Prema, P. (2002). Biotechnology of microbial xylanases: enzymology, molecular biology and application. *Critical reviews in biotechnology*, *22*(1), 33-64.

Sugiyama, J., Gustavsson, M., Fransson, L., Linder, M., & Teeri, T. T. (2002). The binding specificity and affinity determinants of family 1 and family 3 cellulose binding modules. *PNAS*, *100*(2), 484-489.

Sulzenbacher, G., Driguez, H., Henrissat, B., Schülein, M., & Davies, G. J. (1996). Structure of the Fusarium oxysporum endoglucanase I with a nonhydrolyzable substrate

analogue: substrate distortion gives rise to the preferred axial orientation for the leaving group. *Biochemistry*, *35*(48), 15280-7.

Suzuki, T., Kitagawa, E., Sakakibara, F., Ibata, K., Usui, K., Kawai, K. (2001). Cloning, Expression and Charcaterization of a Family 52 B-Xylosidase Gene (XysB) of a Multiple-xylanase-producing Bacterium, Aeromonas caviae ME-1. *Biosci. Biotechnol. Biochem.*, *65*, 487-494.

Sørensen, H. R., Jørgensen, C. T., Hansen, C. H., Jørgensen, C. I., Pedersen, S., & Meyer, A. S. (2006). A novel GH43 alpha-L-arabinofuranosidase from Humicola insolens: mode of action and synergy with GH51 alpha-L-arabinofuranosidases on wheat arabinoxylan. *Applied microbiology and biotechnology*, *73*(4), 850-61.

Taherzadeh, M., Karimi, K. (2007). Acid-based hydrolysis processes for ethanol from lignocellulosic materials: a review. *BioResources*, *2*, 472-499.

Taylor, E. J., Smith, N. L., Turkenburg, J. P., Souza, D., Gilbert, H. J., & Davies, G. J. (2005). Structural insight into the ligand specificity of a thermostable family 51 arabinofuranosidase , Araf51 , from Clostridium thermocellum. *Biochemical Journal*, 1-22.

Tokunaga, T., & Esaka, M. (2007). Induction of a novel XIP-type xylanase inhibitor by external ascorbic acid treatment and differential expression of XIP-family genes in rice. *Plant & cell physiology*, *48*(5), 700-14.

Tsujibo, H., Kosaka, M., Ikenishi, S., Sato, T., & Miyamoto, K. (2004). Molecular Characterization of a High-Affinity Xylobiose Transporter of Streptomyces thermoviolaceus OPC-520 and Its Transcriptional Regulation. *Journal of Bacteriology*, *186*, 1029-1037.

Turner, P., Mamo, G., & Karlsson, E. N. (2007). Potential and utilization of thermophiles and thermostable enzymes in biorefining. *Microbial cell factories*, *6*, 9.

Vanholme, R., Demedts, B., Morreel, K., Ralph, J., & Boerjan, W. (2010). Lignin biosynthesis and structure. *Plant physiology*, *153*(3), 895-905.

Vardakou, M., Dumon, C., Murray, J. W., Christakopoulos, P., Weiner, D. P., Juge, N., Lewis, R. J., et al. (2008). Understanding the structural basis for substrate and inhibitor recognition in eukaryotic GH11 xylanases. *Journal of molecular biology*, *375*(5), 1293-305.

Vardakou, M., Flint, J., Christakopoulos, P., Lewis, R. J., Gilbert, H. J., & Murray, J. W. (2005). A family 10 Thermoascus aurantiacus xylanase utilizes arabinose decorations of xylan as significant substrate specificity determinants. *Journal of molecular biology*, *352*(5), 1060-7.

Varrot, a, Schülein, M., & Davies, G. J. (1999). Structural changes of the active site tunnel of Humicola insolens cellobiohydrolase, Cel6A, upon oligosaccharide binding. *Biochemistry*, *38*(28), 8884-91.

Vasconcelos, E. a R., Santana, C. G., Godoy, C. V., Seixas, C. D. S., Silva, M. S., Moreira, L. R. S., Oliveira-Neto, O. B., et al. (2011). A new chitinase-like xylanase inhibitor protein (XIP) from coffee (Coffea arabica) affects Soybean Asian rust (Phakopsora pachyrhizi) spore germination. *BMC biotechnology*, *11*(1), 14. BioMed Central Ltd.

Volynets, B. (2011). Assessment of pretreantments and enzymatic hydrolysis of wheat straw as a sugar source for bioprocess industry. *I NTERNATIONAL JOURNAL Of Energy and Environment*, *2*(3), 427-446.

Vorwerk, S., Somerville, S., & Somerville, C. (2004). The role of plant cell wall polysaccharide composition in disease resistance. *Trends in plant science*, *9*(4), 203-9.

Vrsanská, M., Kolenová, K., Puchart, V., & Biely, P. (2007). Mode of action of glycoside hydrolase family 5 glucuronoxylan xylanohydrolase from Erwinia chrysanthemi. *The FEBS journal*, *274*(7), 1666-77.

Wagschal, K., Heng, C., Lee, C. C., & Wong, D. W. S. (2009). Biochemical characterization of a novel dual-function arabinofuranosidase/xylosidase isolated from a compost starter mixture. *Applied microbiology and biotechnology*, *81*(5), 855-63.

Watanabe, S., Viet, D. N., Kaneko, J., Kamio, Y., & Yoshida, S. (2008). Cloning, Expression and Transglycosylation Reaction of Paenibacillus sp. Strain W-61 Xylanase 1. *Bioscience, Biotechnology and Biochemistry*, *72*(4), 951-958.

Weng, X.-Y., Huang, Y.-Y., Gao, H., & Sun, J.-Y. (2010). Characterization of a xylanase inhibitor TAXI-I from wheat. *Biologia Plantarum*, *54*(1), 154-158.

White, a, & Rose, D. R. (1997). Mechanism of catalysis by retaining beta-glycosyl hydrolases. *Current opinion in structural biology*, *7*(5), 645-51.

Whitehead, T. R. (1993). Analyses of the gene and amino acid sequence of the Prevotella (Bacteroides) ruminicola 23 xylanase reveals unexpected homology with endoglucanases from other genera of bacteria. *Current microbiology*, *27*(1), 27-33.

Winterhalter, C., & Liebl, W. (1995). Two Extremely Thermostable Xylanases of the Hyperthermophilic Bacterium Thermotoga maritima MSB8. *Applied and environmental microbiology*, *61*(5), 1810-5.

Withers, S. (2001). Mechanisms of glycosyl transferases and hydrolases. *Carbohydrate Polymers*, *44*(4), 325-337.

Xie, H., Gilbert, H. J., Charnock, S. J., Davies, G. J., Williamson, M. P., Simpson, P. J., Raghothama, S., et al. (2001). Clostridium thermocellum Xyn10B Carbohydrate-Binding Module 22-2 : The Role of Conserved Amino Acids in Ligand Binding. *Biochemistry*, *40*, 9167-9176.

Yasutake, Y., Kawano, S., Tajima, K., Yao, M., Satoh, Y., Munekata, M., & Tanaka, I. (2006). Structural Characterization of the Acetobacter xylinum Endo- ᴺᴸ -1 , 4-Glucanase CMCax Required for Cellulose Biosynthesis. *Bioinformatics*, *1077*(February), 1069-1077.

Yoon, K.-hong, Yun, H. N., & I, K. H. J. (1998). M O L E C U L A R CLONING OF A BACILLUS SP . KK-1 XYLANASE GENE AND CHARACTERIZATION OF THE GENE PRODUCT. *Biochemistry and Molecular Biology International*, *45*(2), 337-347.

Zanoelo, F. F., Polizeli Md, M. D. L. T. D. M., Terenzi, H. F., & Jorge, J. A. (2004). Purification and biochemical properties of a thermostable xylose-tolerant beta- D-xylosidase from Scytalidium thermophilum. *Journal of industrial microbiology & biotechnology*, *31*(4), 170-6.

Zechel, D. L., & Withers, S. G. (2000). Glycosidase mechanisms: anatomy of a finely tuned catalyst. *Accounts of chemical research*, *33*(1), 11-8.

Zhang, J., Siika-Aho, M., Puranen, T., Tang, M., Tenkanen, M., & Viikari, L. (2011). Thermostable recombinant xylanases from Nonomuraea flexuosa and Thermoascus aurantiacus show distinct properties in the hydrolysis of xylans and pretreated wheat straw. *Biotechnology for biofuels*, *4*(1), 12. BioMed Central Ltd.

Zheng, Y., Pan, Z., & Zhang, R. (2009). Overview of biomass pretreatment for cellulosic ethanol production. *International Journal of Agricultural and Biological Engineering*, *2*(3), 51-68.

Zhou, J., Bao, L., Chang, L., Zhou, Y., & Lu, H. (2012). Biochemical and kinetic characterization of GH43 β-D: -xylosidase/α-L: -arabinofuranosidase and GH30 α-L: -arabinofuranosidase/β-D: -xylosidase from rumen metagenome. *Journal of industrial microbiology & biotechnology*, 143-152.

Zou, J. Y., Kleywegt, G. J., Ståhlberg, J., Driguez, H., Nerinckx, W., Claeyssens, M., Koivula, a, et al. (1999). Crystallographic evidence for substrate ring distortion and protein conformational changes during catalysis in cellobiohydrolase Ce16A from trichoderma reesei. *Structure (London, England : 1993)*, *7*(9), 1035-45.

v.d. Veen, Peter, Flipphi, M. J. A., Voragen, A. G. J., & Visser, J. (1991). Induction, purification and cahracterisation of arabinases produced by Aspergillus niger. *Archives of Microbiology*, *157*, 23-28.

van Passel, M. W. J., Kant, R., Palva, A., Copeland, A., Lucas, S., Lapidus, A., Glavina del Rio, T., et al. (2011). Genome sequence of the verrucomicrobium Opitutus terrae PB90-1, an abundant inhabitant of rice paddy soil ecosystems. *Journal of bacteriology*, *193*(9), 2367-8.

von Ossowski, I., Ståhlberg, J., Koivula, A., Piens, K., Becker, D., Boer, H., Harle, R., et al. (2003). Engineering the Exo-loop of Trichoderma reesei Cellobiohydrolase, Cel7A. A comparison with Phanerochaete chrysosporium Cel7D. *Journal of Molecular Biology*, *333*(4), 817-829.

Appendices

Appendix A. Primer list and pET-28a vector map

A list of the primers designed within this research for target gene fragments are presented here. Primers are designated by the GenBank accession code of each gene/protein. Restriction sites are underlined.

1) ACB77586.1

TGCCATAG GGATCC ATGCAAGGCGCCGCTGAAA
TGCCATAG AAGCTT TCATTTCGCAACGGGACC

2) ACB77587.1

TGCCATAG GAATTC ATGACGTTTGTCCTGCGAAA
TGCCATAG CTCGAG CTATTTGGAAGTGGTGGGTGC

3) ACB75510.1

TGCCATAG GAATTC ATGGAGGCGGCCGCGGCG
TGCCATAG AAGCTT CTATCGCTGCGCGGTTTGCA

4) ACB76979.1

TGCCATAG AAGCTT TG ATGGACCGGGAATCGATC
TGCCATAG CTCGAG TTAGCGAAAGTCCACGGTCG

5) ACB76489.1

TGCCATAG GGATCC ATGCAGACTTCCTTGCCC
TGCCATAG AAGCTT TCAGGGAAGCTGGAGCTCGA

6) ACB77881.1

TGCCATAG GAGCTC ATGGCGACGACGCTGCCCA

TGCCATAG AAGCTT TCAGTTCGACGGATTCACCA

7) ACB74389.1

TGCCATAG GAATTC ATGCAAGTCGTTTCCAACACC
TGCCATAG AAGCTT TTATCGGAGGATGATCGCG

8) ACB76939.1

TGCCATAG GAATTC ATGTCCGATGCGGTGGAG
TGCCATAG AAGCTT TCATTGCAGCCGCACCGT

9) ACB75400.1

TGCCATAG GAATTC ATGCTCGCCCTCGGCTCCGTT
TGCCATAG CTCGAG TCACCACACCGTGGCGTT

10) ACB76447.1

TGCCATAG GAGCTC ATGGTTTCCGTCACCGTC
TGCCATAG AAGCTT TCACTCCCATTCGAGCACGA

11) ACB74536.1

TGCCATAG GAATTC ATGAATGGCACGTTCACCAA
TGCCATAG CTCGAG TCACTCGGCGTAGCGCCACA

12) ACB76503.1

TGCCATAG GAATTC ATGGCCGCCGTCGCGTCCA
TGCCATAG CTCGAG TTACTCGTGCGCGAGCTCCA

13) ACB76389.1

TGCCATAG GAATTC ATGCTGATCACGCTCCCC
TGCCATAG CTCGAG TCAGCTTCGCTTTCGTGG

14) ACB74678.1

TGCCATAG GAATTC ATGCTTGGCGCTTTGATG

TGCCATAG CTCGAG CTATGGTCCCGTGACGATC

15) ACB74639.1

TGCCATAG CATATG ATGACGGCGTTTCTCTTCA
TGCCATAG AAGCTT TCACGGGGTGATCGGTTT

16) ACB76993.1

TGCCATAG GAATTC ATGTGGGGCGACCAGGGCGAT
TGCCATAG CTCGAG TCAGCGAGTTGGATTCGCG

17) ACB73649.1

TGCCATAG GAATTC ATGCCGGCCCAGCCCGCG
TGCCATAG AAGCTT TCAGTCTGCCGCGGCGGT

18) ACB77562.1

TGCCATAG CATATG ATGCTCCGCCCCTTTAGC
TGCCATAG GTCGAC TCAACGCGCACTCACGGG

19) ACB77584.1

TGCCATAG GAATTC ATGGCTCAAGAAAGTTACCA
TGCCATAG AAGCTT TTACTCCGAAAGCCACAGGTA

20) ACB75541.1

TGCCATAG AAGCTT TG ATGGAGGCGCCAAGGGCGCT
TGCCATAG CTCGAG TTACTTTAAATCTCGATAG

21) ACB77588.1

TGCCATAG GAATTC ATGGCCGACGACGGTTAT
TGCCATAG CTCGAG TCAGTGCGCGCCGGGATG

22) ACB76502.1

TGCCATAG GAATTC ATGCCGCCGCCGATCGACT
TGCCATAG CTCGAG TCATTTGAAGTACAACTGGAG

23) ACB73635.1

TGCCATAG GAATTC ATGGAAAAGCCCGAAGACGTT
TGCCATAG AAGCTT TCAGCGGAAGATCCGCTG

24) ACB73399.1

TGCCATAG GAATTC ATGCCCGATCCCGTTTCG
TGCCATAG AAGCTT TCAGCGGAACAGCAGCGG

25) ACB77599.1

TGCCATAG GAATTC ATGTCGTTTACGAGGGCCGGT
TGCCATAG CTCGAG CTATCGGATCGCCGCCGGCTT

26) ACB73776.1

TGCCATAG GAATTC ATGCCGGCAGTGCTGCAG
TGCCATAG CTCGAG TCAAGGCTCCTGGCCTTT

27) ACB75584.1

TGCCATAG GAATTC ATGGCCACTCCGATCACG
TGCCATAG AAGCTT TTACTTCACTCCGATCGCCT

28) ACB75560.1

TGCCATAG GAATTC ATGGGACTGATCGGGCCG
TGCCATAG AAGCTT CTACCGCTCGAGGTGCTCC

29) ACB73407.1

TGCCATAG GAATTC ATGCTCGATGACACGCCA
TGCCATAG CTCGAG CTAGCGCTGCGGTTCGCG

30) ACB77600.1

TGCCATAG CATATG ATGGCCCGCTCCCGCTTG
TGCCATAG AAGCTT TCAGTTGGCGTTACGTGAA

31) ACB75551.1

TGCCATAG GAGCTC ATGGCACACTCGGGTCGT
TGCCATAG AAGCTT CTACGGCGTCTCGTCGCGCA

32) ACB77177.1

TGCCATAG GAATTC ATGCTACGTCCCCGCTTG
TGCCATAG CTCGAG CTACTTCGCTCCGCCAACTTT

33) ACB74675.1

TGCCATAG GAATTC ATGAAGGAAGGCGACACGTT
TGCCATAG CTCGAG TCACCGCACCAGTTCGCT

34) ACB76508.1

TGCCATAG GAATTC ATGGCGCCCGCGAGCTTC
TGCCATAG AAGCTT TCAATCCGCCGGCGCCGTC

35) ACB76517.1

TGCCATAG GAGCTC ATGCAGTCCGACAATGGC
TGCCATAG AAGCTT CTACTGTTGCGAGACCAGACG

36) ACB74677.1

TGCCATAG GAATTC ATGCAGCCGCTCAAACCT
TGCCATAG AAGCTT TCACTCCACGCACGACGG

37) ACB73665.1

TGCCATAG GAATTC ATGAGCACGAAGAAACCGTT
TGCCATAG CTCGAG TCAGGGCTCCATCGGCCGGAT

38) ACB77594.1

TGCCATAG GAATTC ATGGACTCCTTCGAATACACC
TGCCATAG CTCGAG CTACTCGGTCCGCGCGTGCAT

39) ACB76822.1

TGCCATAG GAATTC ATGCCCTCCACCATCCGCAA
TGCCATAG AAGCTT CTACGCGCGGACGTCGGG

40) ACB75811.1

TGCCATAG GAATTC ATGGAGCATCGGGCATTTT
TGCCATAG AAGCTT TTAGGCGCTGCGGGCGAG

41) ACB76504.1

TGCCATAG GAATTC ATGAGCGGAGCCCGCATC
TGCCATAG CTCGAG TCACTCGAACCGGAGCCA

42) ACB76487.1

TGCCATAG GAATTC ATGACCACGCACCGGTCG
TGCCATAG CTCGAG TCAGGTCCCGCCGATCTCA

43) ACB74651.1

TGCCATAG GAATTC ATGGAGCTGCCAAACCCA
TGCCATAG AAGCTT TCATTTGCCTTTCAGCGCGA

44) ACB76488.1

TGCCATAG GAATTC ATGGCAACGGCCGAGCCG
TGCCATAG CTCGAG CTAGCGCGCTTTCGCCGGGAG

45) ACB73620.1

TGCCATAG GAGCTC ATGGAGCCCCGAACCCTCA
TGCCATAG AAGCTT CTATCTGGTGGCGATTCGCA

46) ACB76492.1

TGCCATAG GAATTC ATGCTTCCTTGCCTCCCCGC
TGCCATAG AAGCTT CTACGGCCGCGCCGAGCC

47) ACB76505.1

TGCCATAG GAATTC ATGACGGCCACCATCAAG
TGCCATAG AAGCTT CTATCGGTTCACGGCGAC

48) ACB76398.1

TGCCATAG GAGCTC ATGTCTGACATCCCTGTTTAC
TGCCATAG AAGCTT TCAATCGCCGCCATCGTC

49) ACB74659.1

TGCCATAG GAATTC ATGAACGCCACGATCGTGCTT
TGCCATAG AAGCTT TCAGAGCTCGAGCACGACCA

50) ACB75101.1

TGCCATAG GAGCTC ATGTGGCTGCGAAGCATC
TGCCATAG AAGCTT TCAGGCCGTCGCAAACTG

51) ACB77608.1

TGCCATAG GGATCC ATGAAGCAAGGCGCGCCGAT
TGCCATAG CTCGAG TCACTTCACCGGCAGCGCGAA

52) ACB73617.1

TGCCATAG CATATG ATGCTGACGGGCCCGTTTTT
TGCCATAG AAGCTT TCAGTAGTCGACGATCACCAC

53) CAH06053.1

TGCCATAG CATATG AAAGCAACTACCTCCTTG
TGCCATAG AAGCTT TTAACGTGAAATCCCGACCG

54) CAH06512.1

TGCCATAG CATATG AACACATTCGGGAAGAAA
TGCCATAG CTCGAG TTACATTTTAAAAGTATAT

55) CAH06110.1

TGCCATAG CATATG CAAAAACCTCCTCAGGATATG
TGCCATAG CTCGAG TCAATGCAGCGTAAACCGT

56) NP_768005.1

TGCCATAG CTCGAG TCAAACCCCGATCGCGTTC
TGCCATAG CATATG GAGAGCGAACGCCAGCAT

57) NP_772817.1

TGCCATAG CATATG TCCGACCTTCCTCAAAGCTC
TGCCATAG CTCGAG CTACACCCCCGCCGCGTT

58) NP_771297.1

TGCCATAG GAATTC ATGCCGACAACGCAACTG
TGCCATAG AAGCTT TCATGGCTGCACCGCCCT

59) ABC35066.1

TGCCATAG CATATG GAAAACGGCGCCCGTCGAA
TGCCATAG CTCGAG TCACGGCTGCTGAAGCCGAAT

60) ABC92395.1

TGCCATAG CATATG ATCGATGCGAATACGCT
TGCCATAG CTCGAG TCAACCCGGCTTGTGATT

61) ABC94363.1

TGCCATAG CATATG ATCGATTCCATTCTCG
TGCCATAG AAGCTT TCATGCAGTCGACGGCGG

62) AAM55007.1

TGCCATAG CATATG ATTGATAAGAAAGCATTG
TGCCATAG CTCGAG TCAGAGCGCCAAGACAGC

63) CAE27177.1

TGCCATAGTGCCATAG CATATG GCCTCATCGCTCACCCC
TGCCATAG CTCGAG TCAGGCGGCTGTGCCGCGG

64) BAE17525.1

TGCCATAG CATATG TGGGACCGTTATAATAAACCG
TGCCATAG CTCGAG TTATAAATCACGTGCATTCG

65) BAE17642.1

TGCCATAG CATATG ATCATTTTTAAAAAT
TGCCATAG CTCGAG TTAATTGAGGGACTCACCA

66) BAE17270.1

TGCCATAG CATATG TTGTTAAATCAAACATCAA
TGCCATAG CTCGAG TTATTTGCTTATTAAGCTA

67) BAC69512.1

TGCCATAG CATATG ACCATCGACCTCGCC
TGCCATAG CTCGAG TCAGGCGGCCCGGCGCGC

68) BAC72965.1

TGCCATAG CATATG ACCATCGACCTCGCC
TGCCATAG CTCGAG TCAGGCGGCCCGGCGCGC

69) BAC73310.1

TGCCATAG CATATG GACGAGCGCGAGCCCTCC
TGCCATAG CTCGAG TCATCTCCGCGCCCGCAGC

70) AAM79614.1

TGCCATAG CATATG AAACAAAGAAAACACCG
TGCCATAG CTCGAG CTATTTATCAAAACCATTTGA

71) AAM79900.1

TGCCATAG CATATG TTAGCATTTCCAAAAGAG
TGCCATAG CTCGAG TCAATCACTGAGGTGAGAAGC

72) CAB15962.1

TGCCATAG CATATG AGTTCAAATGAAAAACGA
TGCCATAG CTCGAG TCAGAGACTCTCTCCGTTTGT

73) CAH06510.1

TGCCATAG CATATG AAAAGAATTGGAATA
TGCCATAG CTCGAG TCAGTGCCAAGTATAAAAA

74) CAH07509.1

TGCCATAG CATATG TATGGACTTGTATGTTGTTGC
TGCCATAG CTCGAG TTAGTTATAGGTATCTCGTAC

75) ACB66572.1

TGCCATAG CATATG CTCGCAACGGCCGCG
TGCCATAG CTCGAG TCAGCCCGGCGCGGCGCGCA

76) ABC89301.1

TGCCATAG CATATG ATCGAGAAGACCGTGCT
TGCCATAG CTCGAG TTATTGCTGACCTGCCACAA

77) ABC91234.1

TGCCATAG CATATG ACTCAACAACCGCCCCT
TGCCATAG CTCGAG TCAATCCGCTGATTGGCA

78) BAE17243.1

TGCCATAG CATATG AATGTAATTGACTTA
TGCCATAG CTCGAG CTAAAATACCGTATTGAT

79) BAE17242.1

TGCCATAG CATATG TGTCGTGCAAATATCAA
TGCCATAG CTCGAG TTACGCTTTCCATCGACCAT

80) CAH09839.1

TGCCATAG CATATG GACATGCGGCAAGATACTTA
TGCCATAG CTCGAG TTATTCTACATATTTAGTTTC

81) CAH06920.1

TGCCATAG CATATG AGTGGGAACCCTGTATTTCA
TGCCATAG CTCGAG CTATTTCTTTTTAGACTTCTT

82) ABJ61956.1

TGCCATAG CATATG TTAAAAATAAAAAACCCG
TGCCATAG CTCGAG TCAAAGACTCGTCACTTT

83) ABC91230.1

TGCCATAG CATATG ATCCGCAATCCTGTCCT
TGCCATAG CTCGAG TCACAATTCATCATAATTGAA

84) BAC74280.1

TGCCATAG CATATG CACGGCACGCCCGCC
TGCCATAG CTCGAG TCAGCGCCGCAGCCACTT

85) BAC69163.1

TGCCATAG CATATG ATCAGCGTCCCCGCCGAG
TGCCATAG CTCGAG TCATCGCGGCTCCACCGGG

86) BAC74083.1

TGCCATAG CATATG ATCCTCACCTCCGACCT
TGCCATAG AAGCTT TCAGGCCGAGGTGGAGGC

87) BAC69820.1

TGCCATAG CATATG GGCGGCGGGGTCATCAAG
TGCCATAG CTCGAG TCATGTGCCACGCGTCCACT

88) BAC72694.1

TGCCATAG GAATTC ATGACCCACCCCTGGCAG
TGCCATAG CTCGAG TCACTTCACCCGCACCTC

89) ABJ62166.1

TGCCATAG CATATG GATGTTACTGTTAAAAAAA
TGCCATAG CTCGAG TTAATGACCAGTGTCTATAT

90) ABC36653.1

TGCCATAG CATATG CCGAGCGAACTCGGCGAC
TGCCATAG CTCGAG TCAATTTTCTATCCTCGGGC

91) ABQ36152.1

TGCCATAG GGATCC GAGGCTGCGGTGACGGAT
TGCCATAG AAGCTT TCACGCACTCACGGCATT

92) ABQ38386.1

TGCCATAG CATATG ATCGAAGGCGCGGCGCAT
TGCCATAG CTCGAG TCACTCTTCCGCCCGGTT

92) CAD55266.1

TGCCATAG CTCGAG TCACGGCGTGGCGCAGAC
TGCCATAG CATATG TCGCCCGCGGCGCGGTCC

93) CAA20610.1

TGCCATAG CATATG GCCGGTGGCATCCACGTCAG
TGCCATAG AAGCTT TCAGTGGTCCGCCGCGCTGC

94) CAB56687.1

TGCCATAG CATATG CGACGCCACAGCGCC
TGCCATAG CTCGAG TCATCGCAGCTCCAGCAC

95) CAB55731.1

TGCCATAG CATATG ACGACCGGCACCATCACC
TGCCATAG CTCGAG CTAGCGGCCGCGTTCCAT

96) CAB53318.1

TGCCATAG CATATG GTGCATAGGATCGTGAAGCG
TGCCATAG AAGCTT TCAAGCGTCGCCCGTGAACT

97) ABJ65333.1

TGCCATAG CATATG AAGATTCAAAATCCAGTT
TGCCATAG CTCGAG TTAATCTGGCAGTTCTTGAT

98) ABJ63538.1

TGCCATAG CATATG ACTTTAATTCAAAATCCA
TGCCATAG AAGCTT TTATTCATGAGTTAAGTCCTC

99) ABJ64815.1

TGCCATAG CATATG ATGGTCCAAGCTCAAGACTA
TGCCATAG CTCGAG CTATTTATAATTGTAAGGAAC

100) BAC69034.1

TGCCATAG CATATG TTCCACGCGAGACGC
TGCCATAG CTCGAG TCAGCGGTGGCTGAACAA

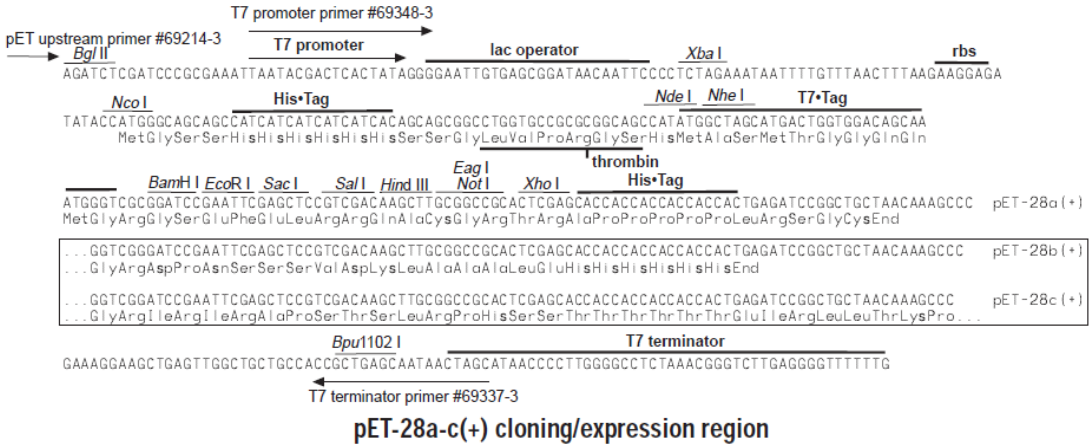


Figure A.1 Cloning region for the pET-28a vector (Novagen, 2007). Target gene fragments were cloned with the inclusion of the stop codon to allow expression of an N-terminally his-tagged protein.

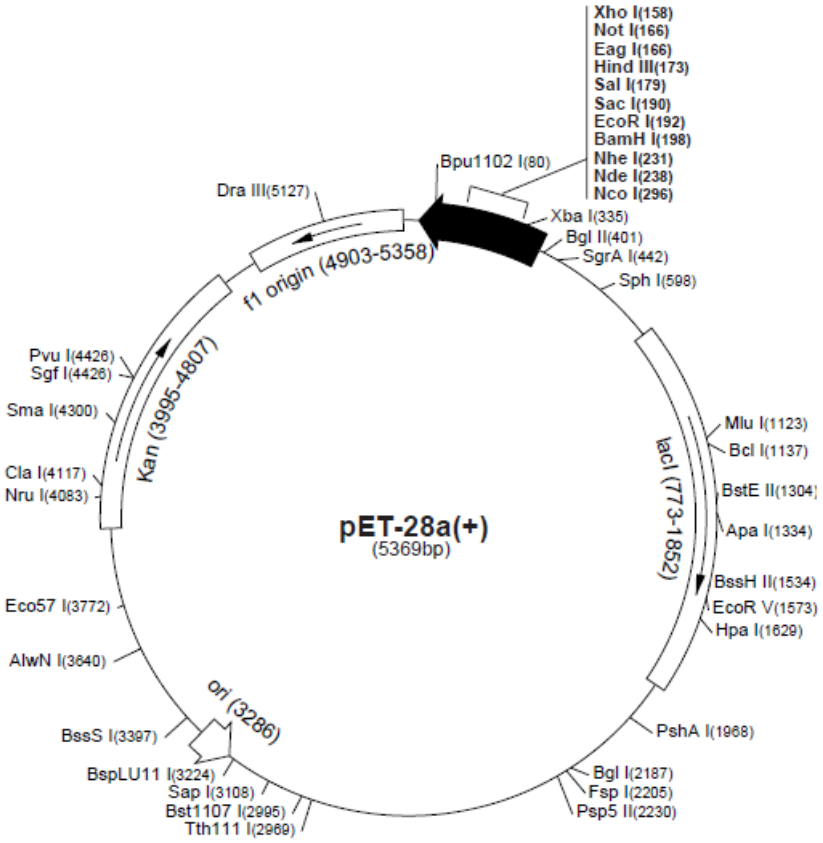
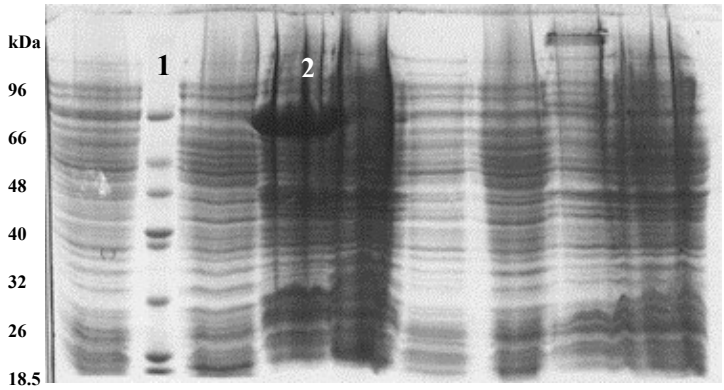


Figure A.2 Map of pET-28a vector.

Appendix B. SDS-PAGE Images

Expression of *Opitutus terrae* GH43 proteins

ACB74651.1 (OtAraf43C)

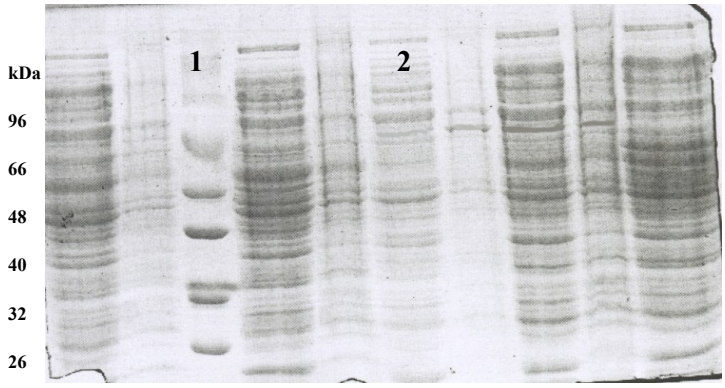


Lane 1 = size std (NZYT)

Lane 2 = ACB74651.1

(73360.6 Da)

ACB76822.1

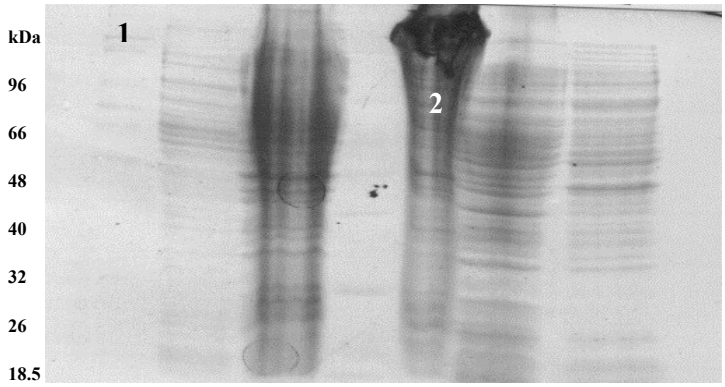


Lane 1 = size std (NZYT)

Lane 2 = ACB76822.1

(100645 Da)

ACB76517.1

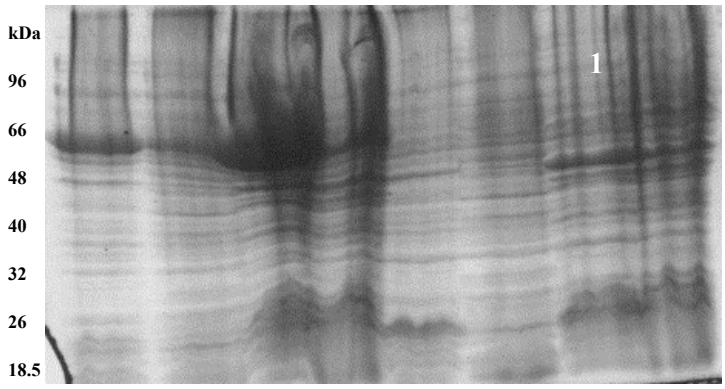


Lane 1 = size std (NZYT)

Lane 2 = ACB76517.1

(130528.9 Da)

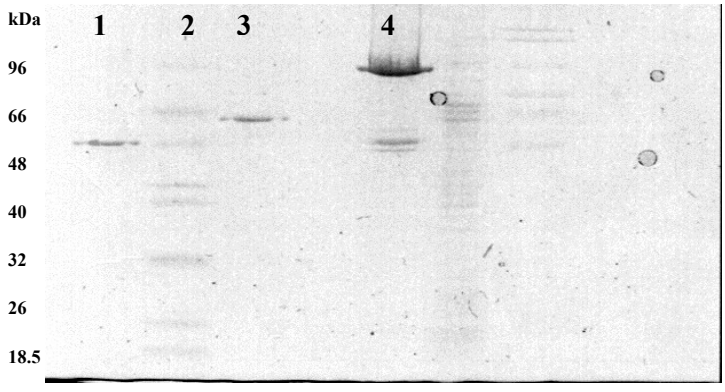
ACB76508.1



Lane 1 = ACB76508.1

(63043.8 Da)

ACB74675.1, ACB73665.1, ACB74651.1



Lane 1 = ACB74675.1

(40298.1 Da)

Lane 2 = size std (NZYT)

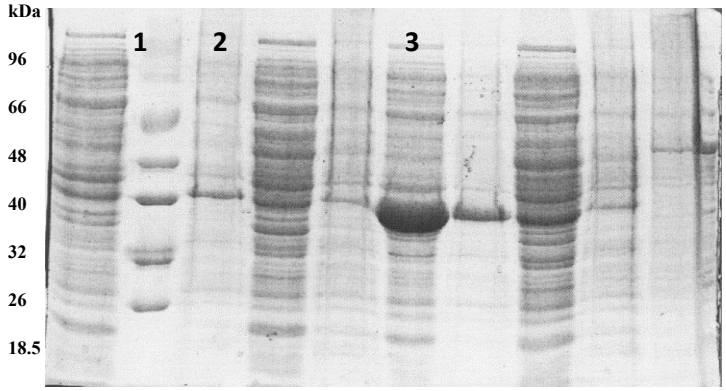
Lane 3 = ACB73665.1

(39853.8 Da)

Lane 4 = ACB74651.1

(73360.6 Da)

ACB77177.1 and ACB74677.1 (OtAraf43A)



Lane 1 = size std (NZYT)

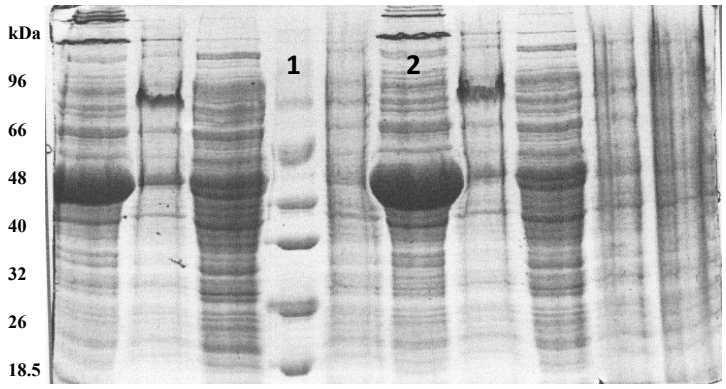
Lane 2 = ACB77177.1

(41644.5 Da)

Lane 3 = ACB74677.1

(39588.7 Da)

ACB73649.1

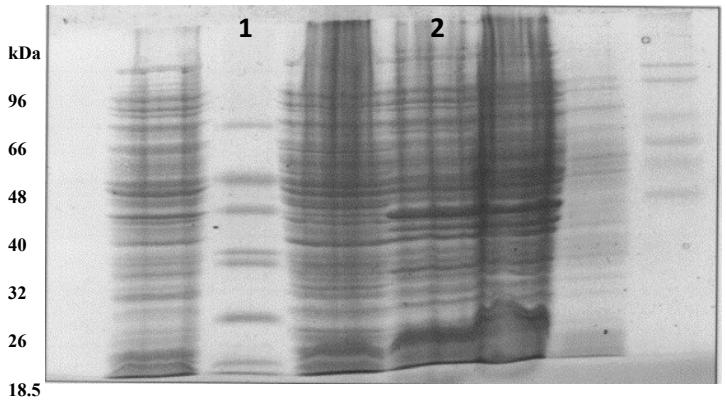


Lane 1 = size std (NZYT)

Lane 2 = ACB73649.1

(56365.1 Da)

ACB76993.1

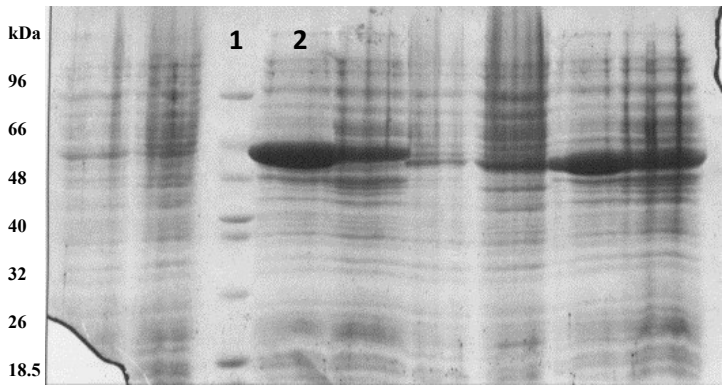


Lane 1 = size std (NZYT)

Lane 2 = ACB76993.1

(56657.4 Da)

ACB76389.1

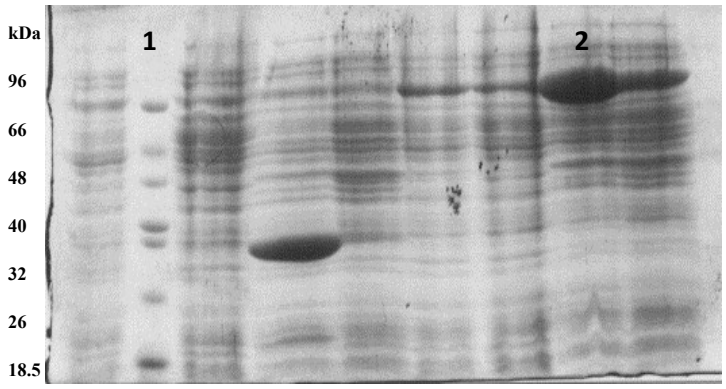


Lane 1 = size std (NZYT)

Lane 2 = ACB76389.1

(42740 Da)

ACB74536.1

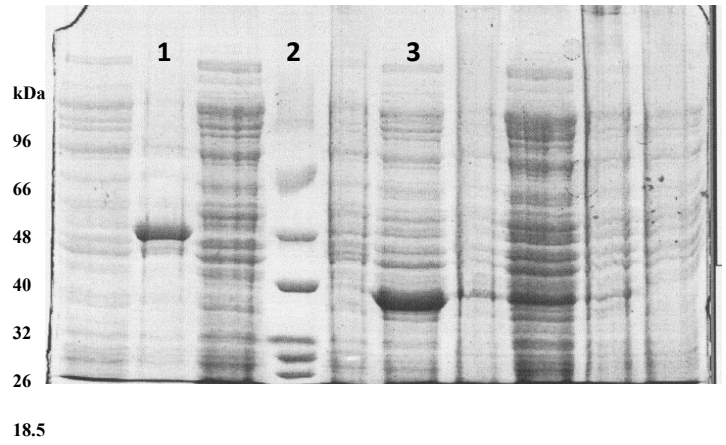


Lane 1 = size std (NZYT)

Lane 2 = ACB74536.1

(75705.2 Da)

ACB76504.1 and ACB76487.1



Lane 1 = ACB76504.1

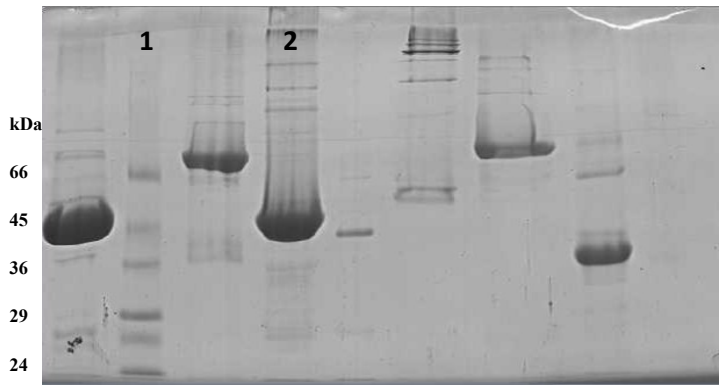
(52665.5 Da)

Lane 2 = size std (NZYT)

Lane 3 = ACB76487.1

(39829.9 Da)

ACB76503.1 (OtAraf43B)



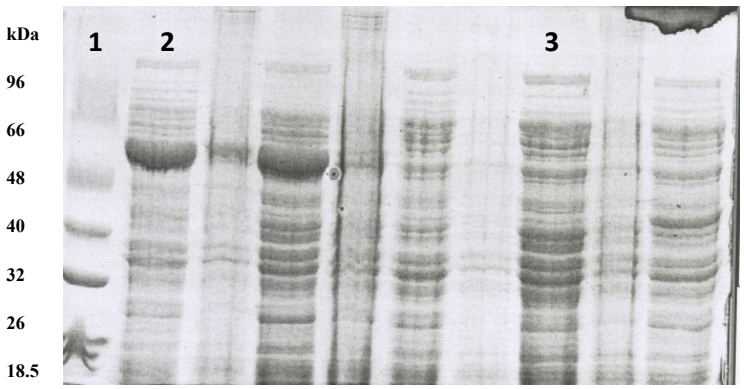
Lane 1 = size std (SIG)

Lane 2 = ACB76503.1

(42959 Da)

Expression of OtXyl52A and OtGlcA67A

ACB77584.1 (OtXyl52A) and ACB77588.1 (OtGlcA67A)



Lane 1 = size std (NZYT)

Lane 2 = ACB77584.1

(81101.4 Da)

Lane 3 = ACB77588.1

(82357.1 Da)

Appendix C. General use equipment

Sterilisation

Autoclave sterilisation was achieved at 121°C for 30 min using a bench-top Prestige Medical 2100 Classic autoclave or 121°C for 40 min in a Priorclave 200 L front loading autoclave. Filter sterilisation was performed through a 0.2 μm Sartorius, Ministart® filter unit.

Incubators

For the growth of bacteria in liquid cultures an Innova 44 orbital shaker Brunswick Scientific (BRS) was used. Growth of bacteria on solid media was performed in a static GallenKamp incubator.

pH meter

All adjustments to the pH of solutions and media were achieved using a Jenway Ion Meter 3340 calibrated with buffers at pH 4.0, 7.0 and 9.2.

Centrifugation:

Centrifugation of volumes up to 1.5ml was achieved using a small Sigma (SIG) 1-15 bench top micro-centrifuge. MSE Microcentaur bench-top centrifuge was used to centrifuge PCR tubes.

Agarose gel electrophoresis (AGE)

Electrophoresis of DNA was performed using an electrophoresis system (BR) powered by a power PAC 300 apparatus (BR).

Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) gel kit

Electrophoresis of proteins was achieved using A Mini-PROTEAN™ 3 Cell kit powered (BR) by a power PAC 300 apparatus (BR).

Sonication

Cell lysis was achieved using a MSE Soniprep 150 ultra-sonication machine (Sanyo).

Spectrophotometer

Spectrophotometric measurements were made using a UV-visible Hελios α Spectronic Unicam spectrophotometer (TS).

PCR

All PCR reactions were performed using a Mastercycler® gradient thermocycler (Eppendorf).

Vortexing

Fisher Scientific (FIS) whirlmixer was used to vortex samples.

Waterbath

Grant (SUB series) incubator water bath used to incubate samples at digitally controlled temperatures.

Pippetting

Gilson pipettes P2-P5000 used to pipette samples and reagents from 0.01-5000 μl accurately.

Weighing Reagents

Mettler AT250 analytical balance used to measure micro gram quantities of chemicals and reagents.