Characterization of termite *Trinervitermes trinervoides* metagenome-derived glycoside hydrolases, the formulation of synergistic core enzyme sets for effective sweet sorghum and corncob saccharification, and their potential industrial applications

A thesis submitted in fulfilment of the requirements for the

the degree of

DOCTOR OF PHILOSOPHY

at

RHODES UNIVERSITY

by

Mpho Stephen Mafa

November 2018

“For wisdom will enter your heart, and knowledge will be pleasant to your soul. Discretion will protect you, and Understanding will guard you.”

Proverbs 2: verses: 10-11

Abstract

The current study investigated the biochemical properties of endo-glucanase (GH5E), exo-glucanase (GH5D), xylanase (GH5H) and endo-glucanase/xylanase (GH45), derived from the hindgut bacterial symbionts of a termite (*Trinervitermes trinervoides*) for their potential role in the biotechnology industry. All these enzymes, except GH5D, exhibited activities on cellulosic and xylan-rich polymeric substrates, which only displayed activity on *p*-nitrophenyl cellobioside. GH5D, GH5E, GH5H and GH45 enzymes retained more than 80% of their activities at pH 5.5 and also retained more than 80% of their activities at 40°C. Furthermore, these enzymes were thermostable at 37°C for 72 hours. GH5E, GH5H and GH45 were generally stable over a range of metal-ion. The kinetic parameters for GH5E were 5.68 mg/ml (K_M) and 34.36 U/mg protein (V_{max}). GH5D activity did not follow classical Michaelis-Menten kinetics, suggesting product inhibition. GH5H displayed K_M values of 5.53, 95.03 and 2.10 mg/ml and V_{max} values of 112.36, 144.45 and 180.32 U/mg protein on beechwood xylan, CMC, and xyloglucan, respectively. GH45 displayed a K_M of 6.94 mg/ml and a V_{max} of 12.30 U/mg protein on CMC. GH5D [cellobiohydrolase (CBH)] and a commercial CBHII (GH6) enzyme outperformed a commercial CBHI (GH7) enzyme when these enzymes hydrolysed β-glucan. GH5D and CBHII also displayed a higher degree of synergy on β-glucan but failed to show synergy on Avicel. We therefore concluded that GH5D and CBHII are β-glucan-specific cellobiohydrolases. The corncob (CC) and sweet sorghum bagasse (SSB) substrates were pretreated with lime, NaOH and $NaClO_2$. Subsequent to pretreatment, these substrates were used to investigate if GH5D, GH5E, GH5H and GH45 could operate in synergy. Results revealed that out of 12 possible core enzyme sets constructed, only two (referred to as CES-E and CES-H) displayed higher activities on pretreated CC or SSB. Simultaneous synergy was generally the most effective mode of synergy during hydrolysis of alkaline pretreated SSB and CC samples by both CES-E and CES-H. Both core enzyme sets did not display synergy on oxidative pretreated substrates. These findings suggest that lime and NaOH are more effective pretreatments for CC and SSB substrates. We used PRotein Interactive MOdeling (PRIMO) software to demonstrate that GH5D protein structure is an $(\alpha/\beta)_8$ barrel with a tunnel-like active site. Enzymes with this type of protein structure are able to perform transglycosylation, a process in which GH5D produced methyl, ethyl and propyl cellobiosides. We concluded that the GH5D, GH5E, GH5H and GH45 enzymes possess novel biochemical properties and that they form synergy

during the hydrolysis of complex substrates (SSB and CC). GH5D transglycosylation could be used to produce novel biodegradable chemicals with special properties (e.g. anti-microbial properties). In conclusion, our findings suggest that GH5D, GH5E, GH5H and GH45 can potentially be used to improve biorefinery processes.

Declaration

I declare that this thesis is my own, Unaided work. It is being submitted for the degree of Doctor of Philosophy at Rhodes University. It has not been submitted before for any degree or examination at any other university.

M.S. MAFA

Signed: Mpho Stephen Mafa

On this day 14 of November 2018

Table of Contents

Abbreviations

°C	Degree(s) Celsius
BWX	Beech wood xylan
CAZy	Carbohydrate-Active enzymes
CBH	Cellobiohydrolase
CBM	Carbohydrate-binding module
CC	Corncob
CDSs	Coding domain sequences
CMC	Carboxymethyl cellulose
DNS	Dinitrosalicylic acid
DP	Degree of polymerization
DS	Degrees of synergy
EC	Enzyme commission number
EDTA	Ethylenediaminetetraacetic acid
FTIR	Fourier-transform infrared spectroscopy
g	Gram
g	Gravity
GH	Glycoside hydrolase
GH45	Glycosyl hydrolase family 45
GH5D	Glycosyl hydrolase family 5 clan D
GH5E	Glycosyl hydrolase family 5 clan E
GH5H	Glycosyl hydrolase family 5 clan H
GOPOD	Glucose oxidase/peroxidase

h	Hour
HBI	Hydrogen bond intensity
kDa	Kilo Daltons
L	Litre
LOI	Lateral order index
mg	Milligram
min	Minute
mL	Millilitre
mM	Millimolar
$NaClO_2$	Sodium chlorite
NaOH	Sodium hydroxide
nm	Nanometer
NREL	National Renewable Energy Laboratory
ORF	Open reading frame
PASC	Phosphoric acid pretreated cellulose
*p*NP	*p*-Nitrophenol
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SEM	Scanning electron microscopy
SSB	Sweet sorghum bagasse
TCI	Total crystallinity index
TEMED	N, N, N', N'-tetramethylethylenediamine
TLC	Thin layer chromatography
U	Units of enzyme activity

w	Weight
µg	Microgram
µL	Microlitre
µM	Micromolar
µmol	Micromole

List of Figures

List of Tables

Dedication and Acknowledgements

I dedicate my work to my mother Neo Mafa who has supported and encouraged me through my years of study, and for giving me the opportunity to further my education

I would like to take this opportunity to thank the following people, without whom this study would not have been possible:

Prof Brett I. Pletschke and Dr Konanani Rashamuse, whose supervision, commitment and mentorship facilitated the completion of my PhD degree.

My family, especially my siblings for their continued support and encouragement throughout my years of study.

I would also like to extend my thanks to the Enzyme Science Programme (ESP) Research group for their technical assistance, with special thanks to Mariska Thoresen and Glynis Oree and Samkelo Malgas for their words of encouragement and extra support when necessary.

All my friends, in particular, Noko Pele and my fiancé Lindiwe Dhlamini for their love, support and prayers which kept me focused during the years of my study.

I also wish to acknowledge the CSIR and DST for providing me with the funding which enabled me to complete my research project.

Research outputs emanating directly from this study

PUBLICATIONS:

1. Bhattacharya A., Mafa M.S., Rashamuse K and Pletschke. B.I. (2016). A bi-functional xylanase/ xyloglucanase, GH5H, from a termite metagenome improve the hydrolysis of pretreated biomass. International Journal of Advances in Science Engineering and Technology, Vol: 4, Iss: 3.

MANUSCRIPTS READY FOR PUBLICATION:

1. Mafa M.S., Malgas S., Bhattacharya A., Rashamuse K and Pletschke B.I. Characterization and formation of a core enzyme set consisting of termite metagenome derived multifunctional and processive enzymes with potential application in the biofuel or feed industry. To be submitted to Enzyme and Microbial Technology.

2. Mafa M.S., Malgas S., Bhattacharya A., Rashamuse K and Pletschke B.I.
An investigation on the effects of alkaline and oxidative pretreatments on agricultural biomass and its hydrolysis by a termite-derived enzyme core set. To be submitted to Enzyme and Microbial Technology.

3. Mafa M.S., Malgas S., Rashamuse K and Pletschke B.I.
Chemical and morphological changes in biomass crystallinity are necessary for GH5D activity: A comparative study between a termite metagenome derived cellobiohydrolase (GH5D), Cel7 and Cel6. To be submitted to Bioresource Technology.

BOOK CHAPTER TO BE SUBMITTED FOR PUBLICATION:

1. Malgas S., Mafa M.S and Pletschke B.I. A review of xylanolytic enzymes with regards to their synergistic interactions during hetero-xylan degradation.

INTERNATIONAL CONFERENCE (PUBLISHED) PROCEEDINGS:

1. Shrivastava Bhattacharya A., Bhattacharya A., Rashamuse K and Pletschke B.I. Differential adsorption and inhibition of endoglucanase from different sources on lignin and its derivatives. IRF International Conference. Nagpur, India, 19th June 2016. (Talk). Proceedings: ISBN: 978-93-85832-29-1

2. Bhattacharya A., Mafa M.S., Rashamuse K and Pletschke B.I. A bi-functional xylanase/xyloglucanase, GH5H, from a termite metagenome improves the hydrolysis of pretreated biomass. IRF International Conference. Nagpur, India, 19th June 2016. (Talk). Proceedings: ISBN: 978-93-85832-29-1 pages 1-8.

3. Pletschke B.I., Malgas S., Bhattacharya A., Bhattacharya-Shrivastava A., Clarke M.D., Mafa M.S., Morake S and Thoresen M. Enzyme Synergism: A Powerful Tool for Decreasing Enzyme Loading for Efficient Biomass Conversion. EUBCE 2016. 24th European Biomass Conference & Exhibition, RAI Amsterdam Exhibition and Convention Centre, Amsterdam, The Netherlands, June 6-9, 2016. (Talk). European Biomass Conference and Exhibition Proceedings ISSN 2282-5819 pages 68-82.

LOCAL CONFERENCES:

1. Mafa M.S., Bhattacharya A., Rashamuse K and Pletschke, B.I. Synergistic associations between termite metagenome derived plant cell wall degrading enzymes during the hydrolysis of complex substrates. 26th Annual Conference of the Catalysis Society Of South Africa (CATSA) 2015. Arabella Hotel & Spa, Kleinmond, 15-18 November 2015 (Poster).

2. Mafa M., Rashamuse K and Pletschke B.I. Synergism: Development of a termite derived core enzyme set for sweet sorghum and corncob depolymerisation. 25th South African Society for Biochemistry and Molecular Biology (SASBMB) Congress. 40th Anniversary-2016. East London ICC, East London, 10- 14th July 2016 (Talk).

3. Mafa M.S., Rashamuse K and Pletschke B.I. Termite metagenome derived novel multifunctional bio-catalysts with potential applications in biofuel and biochemical production. 27th Annual Conference of the Catalysis Society Of South Africa (CATSA) 2016. Champagne Sports Resort, Central Drakensberg region of KwaZulu-Natal of South Africa 6-9 December 2016 (Talk).

4. Mafa. M.S., Rashamuse K and Pletschke B.I. Enzymatic synthesis of alkyl-cellobiosides and their application as antibacterial agents. 8th The Interdisciplinary Postgraduate Conference (IPGC) 2017: Rhodes University, Grahamstown Eastern Cape 29 -30 September 2017 (Talk).

Chapter 1: Literature Review

1.1. Termites: Reservoir of glycosyl hydrolase enzymes

1.1.1. Introduction

The currently known termite classes are made up of more than 2600 species (Ni and Tokuda 2013). Termites are generally classified into seven families: *Mastotermitidae, Kalotermitidae, Termopsidae, Hodotermitidae, Rhinotermitidae, Serritermitidae* and *Termitidae* (Eggleton 2001; Bignell and Jones 2011; Ni and Tokuda 2013). The *Mastotermitidae* are the most primitive family, containing only one wood-feeding species in Australia. Termites of the *Kalotermitidae* feed on the materials from dry wood, and *Termopsidae* nest in and feed on wet dead logs. Termites of the *Hodotermitidae* are grass-feeders and *Rhinotermitidae* consume wood, primarily in temperate zones. The *Serritermitidae* group is endemic to South America and contains a few species, suggesting that it may be better classified within the Rhinotermitidae family (Ni and Tokuda 2013).

All termites feed on lignocellulosic biomass (Watanabe and Tokuda 2010; Ni and Tokuda 2013; Brune 2014). Lignocellulose is the principal component of the plant cell wall, and termites consume it either in the form of wood or grass or in different stages of decomposing plant biomass. Lignocellulose is composed of a linked complex of cellulose, hemicelluloses and lignin. These complexes are highly recalcitrant to enzymatic degradation, but termites degrade and extract necessary nutrients from lignocellulose efficiently (Brune 2014). Termites are reported to degrade about 74–99% of the cellulose and 65–97% of the hemicellulosic components (Prins and Kreulen 1991). These reports demonstrate that termites are a potential source of the novel glycosyl hydrolase (GH) enzymes that can degrade lignocellulosic biomass.

Bignell and Jones (2011) published the taxonomic index of 494 termite entries and showed that some of these termites are in a symbiotic association with fungi (nine were identified), protist (60 were identified), and 64 prokaryotes (64 prokaryotes were identified). Furthermore, termites that have flagellated protistan symbionts in the hindgut are grouped into “lower termites”, while those that lack protistan symbionts in their hindguts are referred to as “higher termites” (Watanabe and Tokuda 2010). Lower termites survive/feed on lignocellulosic based biomass with the assistance of the prokaryotic and eukaryotic

symbionts, whereas in higher termites nutritional mutualism has been reduced to host-prokaryote symbiosis.

Lignocellulose digestion in termites gut is best described using the lower termite digestive system (Ni and Tokuda 2013; Brune 2014). Figure 1.1 shows a simplified termite digestive system. In the foregut, the small particles of lignocellulosic biomass that are produced by the mandibles are mixed with enzymes of the salivary glands and further crushed to smaller particles by the muscular gizzard. Any glucose that is released in the midgut is resorbed via the epithelium. However, the partially digested lignocellulosic biomass particles pass through the enteric valve into the voluminous hindgut paunch. Once in the hindgut paunch, the biomass particles are immediately hydrolysed by cellulolytic flagellates, which use cellulytic and hemicellulytic enzymes that are secreted into their digestive vacuoles. The microbial fermentation products (which include short-chain fatty acids) are resorbed by the termites, and the lignin-rich residues are voided as faeces. In higher termites, hindgut bacteria have a similar role to that of flagellates in cellulose degradation (Ni and Tokuda 2013; Brune 2014).

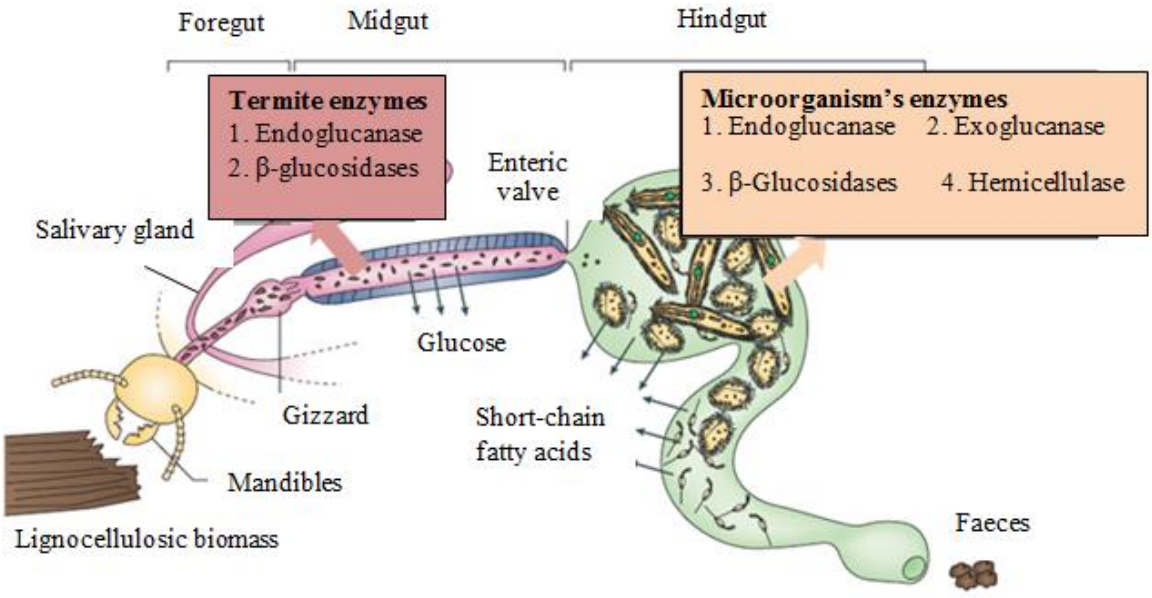


Figure 1.1. Simplified termite digestive system, [modified from Brune (2014)].

Termites are therefore a reservoir or source of enzymes that are able to degrade lignocellulosic biomass. Thus, it is no surprise that termites are an attractive potential source of novel enzymes that can enhance hydrolysis of the lignocellulosic biomass. Lignocellulosic biomass is used in the bio-refinery and bio-energy industries, hence, we propose that lignocellulosic enzymes derived from termites or their symbionts can be employed to degrade biomass.

1.1.2. Metagenomic studies of termites and their symbionts for mining lignocellulases

Metagenomic studies combine biology and chemistry because these studies explore the production of natural products from uncultured soil, marine water, insect and vertebrate gut microorganisms (Handelsmanl *et al*. 1998; Sleator *et al*. 2008). The approach involves extraction of genomic DNA of organisms that cannot be or have not been cultured by isolating their DNA. The isolated DNA is cloned into organisms that can be cultured and the resultant clones products (enzymes) are screened for the production of new chemicals.

Metagenomics can be divided into classes: function-driven and sequence-based analysis of the DNA sequence (genes) or entire genome of uncultured microorganisms (Sleator *et al*. 2008; Batista-García *et al*. 2016). Functional metagenomics involves screening metagenomic libraries for a particular phenotype, which include, but are not limited to salt tolerance, antibiotic production or lignocellulose degrading enzyme activity. In addition, It is important to identify the origin of the uncultured DNA using phylogenetic approches. Secondly, sequence-based methods involve screening clones for the highly conserved gene sequences for identification purposes (e.g. the 16S rRNA gene). Sequence based approaches also include sequencing the entire clone to identify genes of interest, or large-scale sequencing of the complete metagenome to search for marker genes in the reconstructed genomes (Warnecke *et al*. 2007; Sleator *et al*. 2008; Uchima *et al*. 2012; Bastista-Garcia *et al*. 2016). Our study will focus on metagenomic studies that have been carried out on the gut of termites to produce/clone glycosyl hydrolase (GH) enzymes by (Rashamuse *et al*. 2016).

Most reports on termite metagenomic research have focused on termite cellulytic enzymes (Bastista-Garcia *et al*. 2016). These enzymes were expressed in heterologous host systems such as *Escherichia coli*, *Pichia pastoris* and *Aspergillus oryzae*. Through functional metagenomic screening of the hindgut bacterial symbionts of a termite, *Trinervitermes*

trinervoides, Rashamuse *et al*. (2016) identified open reading frames (ORF) for 25 glycoside hydrolase (GH). The ORFs were classified into 14 different glycoside hydrolase (GH) families: eight from GH family 5, four from GH family 9 and two from GH family 13. In addition, one ORF was identified for each of following GH families: 2, 10, 11, 26, 29, 43, 44, 45, 67, and 94. Only eight of these GH families were heterologously overexpressed in an *E. coli* system and partially characterized. The results revealed that enzymes GH5C, GH5E, GH5F, and GH5G were classified as endo-cellulases, GH5D was classified as an exo-cellulase, GH5H and GH11 were classified as endo-xylanases, and GH29 was classified as a α-fucosidase. Prior to that study Warnecke *et al*. (2007) used metagenomic analysis of the bacterial symbionts in the hind gut paunch of a wood-feeding *Nasutitermes* species and shown the presence of a large, diverse set of bacterial genes (45 in total) encoding glycoside hydrolase.

Termites and their symbionts therefore possess a number of genes encoding proteins which digest lignocellulosic biomass and support the hypothesis that termites are a reservoir for GH enzymes. In the past decade, megenomic research has been at the forefront of discovering these enzymes from the gut of wood and grass feeding termites. In the following sections, we will demonstrate how termite enzymes work in synergy to hydrolyze lignocellulose. All the enzymes, which include cellulases, hemicellulases and accessory enzymes, derived from termites and their symbionts, will be comprehensively discussed.

1.2. Termite glycoside hydrolase (GH) classification

Glycoside hydrolases (GH) are enzymes that catalyze the hydrolysis of the glycosidic linkage of glycosides (http://www.cazy.org). Glycoside hydrolases are also referred to as glycosidases. The GH enzymes can also catalyze the hydrolysis of O-, N- and S-linked glycosides (Naumoff 2011). In addition, N-glycoside hydrolases [Enzyme Commission (EC 3.2.2)] are mainly enzymes of nucleic acid metabolism and only S-glycoside hydrolase (EC 3.2.3) thioglucosidase (EC 3.2.3.1) was reclassified as an O-glycoside hydrolase with the EC number 3.2.1.147.

GH enzymes have a variety of activities on a vast number of natural substrates, which include oligosaccharides, polysaccharides and carbohydrate derivatives (http://www.cazy.org). Traditional biochemical classification of GH enzymes is based exclusively on their substrate specificity. In addition, for the amylases with EC number 3.2.1.1 and EC number 3.2.1.2, classification takes into account molecular mechanism of the catalyzed reaction. In the case

of traditional biochemical classification, only the main or a preferred substrate of any GH-enzyme is considered. This form of classification has some advantages, because of its relative simplicity and long term experience in its use.

The accumulation of data on GH enzyme polypeptide or amino acid sequences (primary structures) has made it possible for the establishment of a new system for classification of GH enzymes (Naumoff 2011). This classification is based on a comparison of GH enzyme catalytic domain and amino acid sequences. The mechanisms employed by the GH enzymes during substrate hydrolysis are also used and they are important when classifying the GH enzymes into different GH families. There are two major mechanisms currently known for GH enzymes, namely, 1) retention or 2) inversion of the stereochemistry at the anomeric carbon of the hydrolyzed carbohydrate (Henrissat and Bairoch 1996; Naumoff 2011). The mechanisms are generally conserved within each family. For instance, GH families1, 2, 5, 7, 10, 11, 12, 13, 16, 17, 22, 30, 31, 32, 33, 34, 35, 39 and 42 display a retaining mechanism, while the inverting mechanism prevails in GH families 6, 8, 9, 14, 15, 19, 24, 37, 43, 44, 45, 46, 47 and 48 (Henrissat and Bairoch 1996). The additional information on GH enzyme classification on CAZy (www.cazy.org) indicates the type of mechanism for each of the 145 GH families where it is known. In addition, the grouping of GH enzymes into clans is also indicated in the CAZy (www.cazy.org) classification.

A "clan" is a group of families that have a common evolutionary origin, which is recognized by significant similarities in their three dimensional structures, together with conservation of the catalytic residues and catalytic mechanism (Henrissat and Bairoch 1996). The growing number of three dimensional structures solved for glycosyl hydrolases and/or improved sequence comparison strategies have revealed the relationships between some glycosyl hydrolases families which can be grouped in clans.

Interestingly, when we searched for the GH families of GH enzymes derived from termites and their symbionts in CAZy (www.cazy.org) we could only obtain 21 hits, which included GH families 1, 3, 4, 5, 8,10, 11, 13, 23, 36, 43, 51, 78, 88, 94, 97, 99, 109, 115 and 116. In addition, the search also showed carbohydrate esterase (CE) families 1 and 2; Carbohydrate-binding modules (CBM) 4 and 32; and glycosyltransferase (GT) families 2, 4, 28, 51 and 84. However, our focus in this review was only on the enzymes from GH families. According to Warnecke *et al*. (2007) and Rashamuse *et al*. (2016), there are more than 21 GH families of GH enzymes derived from termites documented in the literature. These observations suggest

that there is a need to update the CAZy database with GH families of termite derived carbohydrate degrading enzymes. Thus, the next section focuses only on the currently documented GH families of GH enzymes derived from termite or their symbionts.

1.3. Cellulases and hemicellulases GH families

The information provided above shows that there are currently two forms of classifying the GH enzymes, which are mainly based on amino acid sequence (www.CAZy.org) and substrate specificity (IUBMB enzyme nomenclature). The section discusses the functions of GH enzymes (cellulases and hemicellulases) with regard to their families. In addition, a summary of the characteristics of the polysaccharides hydrolyzed by the GH enzymes will be provided.

1.4. Lignocellulosic biomass

The strategies that termites use for the breakdown of lignified plant cell walls are unique and efficient. They include the mechanical milling of the lignocellulosic biomass by mandibles and gizzards. The midgut is an enzymatic loaded chamber and degradation of the small particles of lignocellulosic biomass is initiated in this chamber. In the hindgut paunch, the remains of the lignocellulose are digested under anaerobic conditions (fermentation) by the termite symbionts which convert lignocellulose to microbial products (Ni and Tokuda 2013; Brune 2014). The reason for the complex termite digestive system lies in the nature of the lignified plant cell wall, which consists mainly of lignin, hemicellulose and cellulose.

1.4.1. Cellulosic component of lignocellulose

Cellulose is the main component of the plant cell wall (Saini *et al*. 2014). Cellulose makes up about 20–40% of dry weight in the primary cell wall and it generally increases up to 50% in the secondary cell wall, except for a few cases such as cotton seed hairs, which consist of 100% cellulose (Watanabe and Tokuda 2010). It is a linear homo-polysaccharide consisting of about 500–15000 glucose units that are linked by β-1,4-glycosidic bonds. The β-1,4 orientation of the glycosidic bonds results in the potential formation of intramolecular and intermolecular hydrogen bonds. The hydrogen bonds make native cellulose highly crystalline, insoluble, and recalcitrant to enzyme hydrolysis. The highly crystalline regions of cellulose in the plant cell wall are separated by less ordered amorphous regions. Cellulases are a group of GH enzymes required to completely hydrolyze the cellulosic component of the lignocellulose.

1.4.2. Hemicellulosic component of lignocellulose

Hemicellulose is a general term for the major non-cellulosic polysaccharides in plant cell walls, excluding pectin and starch (Watanabe and Tokuda 2010). Hemicellulose is a branched polymer of pentoses such as xylose and arabinose or hexoses such as mannose, galactose, and glucose (Watanabe and Tokuda 2010; Malgas *et al*. 2015a). Generally, the hemicellulose polymer is made up of about 50–200 units of sugar residues. Some of the branches of the hemicellulose polymer consist of ferulate or acetate groups that randomly attached with ester linkages to the hydroxyl groups of the sugar rings. The hemicellulose plays an important role by linking lignin and cellulose. Given the complexity of the hemicellulose polymer, a number of enzymes are required to completely hydrolyze the hemicellulose. These enzymes include debranching enzymes which are also referred to as accessory enzymes, the polysaccharide backbone cleaving enzymes and the oligosaccharide (e.g. xylobiose or mannobiose) cleaving enzymes (Malgas *et al*. 2015a; Van Dyk and Pletschke 2012).

1.5. Cellulolytic enzymes derived from termites

Two separate cellulolytic systems, endogenous cellulases and cellulases of symbiotic microbe origin, are well established to coexist in termite species (Brune 2014). Cellulase is a general term for cellulolytic enzymes, of which three classes are documented based on the mode of enzymatic action and the substrate specificity (Watanabe and Tokuda 2010; Van Dyk and Pletschke 2012). The cellulolytic enzymes include endoglucanases (EC 3.2.1.4), exoglucanases (EC 3.2.1.74 and 3.2.1.91), and β-glucosidases (EC 3.2.1.21). All of these enzymes hydrolyze the β-1,4-glycosidic bonds in cellulose. The endoglucanases hydrolyze amorphous sites of cellulose chains randomly, while exoglucanases, cellobiohydrolases (EC 3.2.1.91) and glucohydrolases (EC 3.2.1.74), hydrolyze the cellulose fibers from the non-reducing or reducing ends and release either cellobiose (cellobiohydrolases) or glucose (glucohydrolases). The β-glucosidases hydrolyze cellobiose or cello-oligomers to glucose from the non-reducing ends. Figure 1.2 shows a simplified enzymatic degradation process for cellulose (Van Dyk and Pletschke 2012).

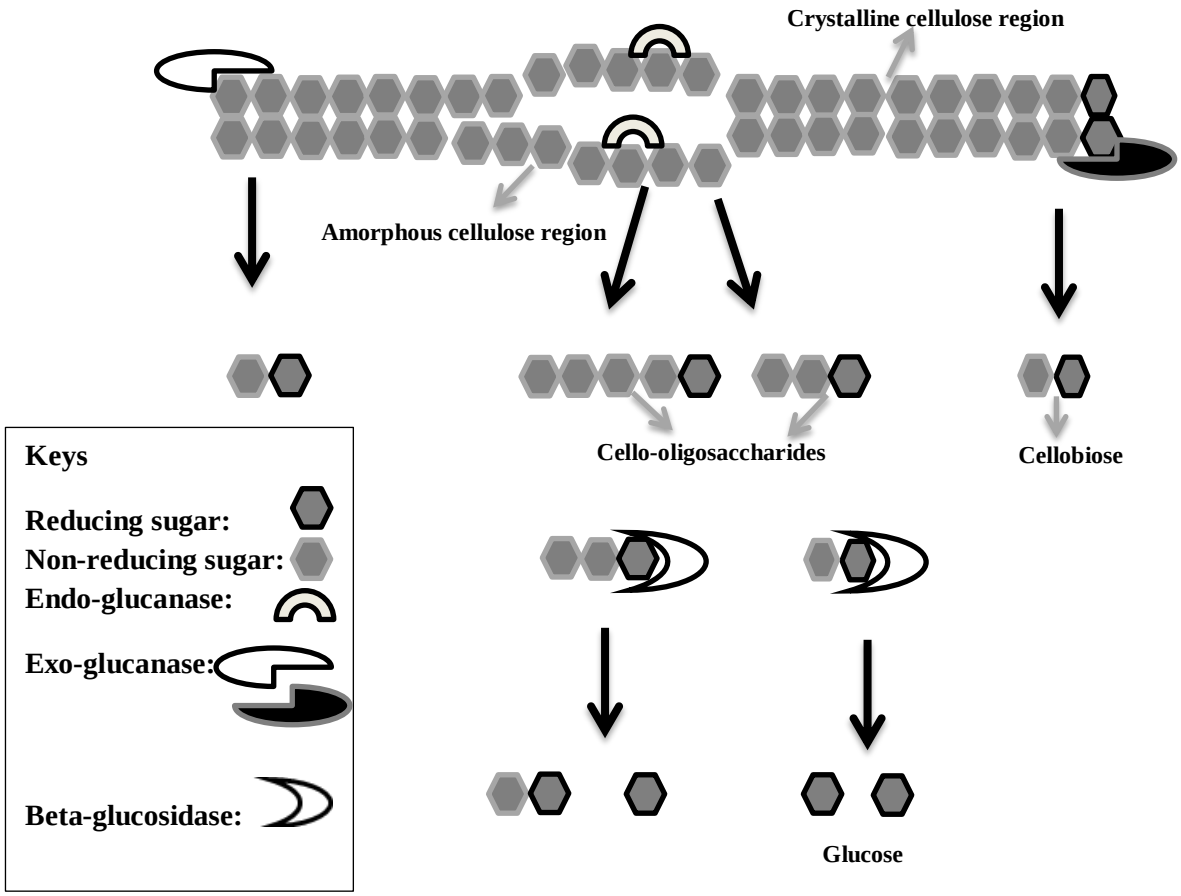


Figure 1.2. Schematic representing the enzymatic hydrolysis of cellulose. Cellobiohydrolases act on the non-reducing or reducing termini of cellulose fibers to processively release cellobiose. Endo-β-1,4-glucanases randomly cleave cellulose chains. β-glucosidases hydrolyze cellobiose or cello-oligomers to glucose from the non-reducing ends.

1.5.1. Endoglucanases

Termite endo-glucanase enzymes can hydrolyze the β-1,3- and β-1,4-glycosidic bonds of cellulose/glucan. Watanabe *et al.* (1998) reported that a notable feature of the termite endo-β-1,4-glucanase is the absence of a cellulose binding domain which is generally found in most bacterial and fungal cellulase components, particularly the endoglucanase from the salivary gland of the lower termite *Reticulitermes speratus*. The endo-glucanases from this termite

appears to consist only of a single catalytic domain and are classified as endo-β-1,4-glucanases from GH family 9. Interestingly, the termite endo-β-1,4-glucanases exhibit low or no activity towards crystalline cellulose, but attack the amorphous region of the biomass, producing oligosaccharides (Figure 1.2). An endo-glucanase from a symbiotic protist of *R. speratus* was expressed in *Aspergillus oryzae* (Todaka *et al*. 2010). It was then purified and characterized. The results showed that the enzyme was endoglucanse-I from GH family 7 and it hydrolyzed cello-oligosaccharide (e.g. cellohexaose) to glucose, cellobiose and cellotriose.

A comprehensive analysis of the *Nasutitermes* P3 luminal community metagenome revealed that this termite hindgut lumen and its symbionts have a range of endo-glucanase enzymes, which derive from different GH families (Warnecke *et al*. 2007). Some of these endo-glucanase enzymes displayed bifunctional activity. For instance, endo-glucanases from GH family 51 displayed β-1,4-endoglucanase and arabinofuranosidase activities. The endoglucanases from other GH families include families 5, 8, 9, 44, 45 and 74 and were also reported to be present in the *Nasutitermes* P3 luminal community. In contrast, gene modules corresponding to the catalytic domains of GH family 6 and family 48 family endoglucanases (which are well known endoglucanses from the microbial cellulosic system) were not observed in the *Nasutitermes* P3 luminal community metagenome. Scharf (2015) also supports this observation that endoglucanases from families 6 and 48 are absent from the termite cellulase system. However, endo-glucanases, from families 5, 9 and 45 seem to be prevalent in different termite species.

Rashamuse *et al*. (2016) showed that the higher termites group also host bacteria in the hindgut which produce endoglucanases. Through functional metagenomic screening of the hindgut bacterial symbionts of a termite, *Trinervitermes trinervoides*, Rashamuse and co-workers expressed and purified endoglucanases which belong to GH families, 5, 9, 44 and 45. The endoglucanases that are produced by the lower and the higher termites share similarities because: 1) most of them belong to GH family 5 or 9, 2) they have more than one substrate specificity, 3) they are produced by the termites or termite symbionts and 4) there are no reports of endoglucanases from GH families 6 and 48.

1.5.2. Exoglucanases

The termite's midgut is proposed to abundantly produce the endo-β-1,4-glucanase (EC 3.2.1.4), which preferably attacks amorphous regions of the cellulose (Tokuda and Watanabe 2007; Tokuda *et al*. 2012). Hence, it is likely that cellulosic food particles reaching the

termite's hindgut primarily comprise non-amorphous crystalline regions. For instance, the hindguts of *Nasutitermes takasagoensis* and *Nasutitermes walkeri* showed cellulase activities against crystalline cellulose (Tokuda and Watanabe 2007). In addition, Sethi *et al*. (2013) showed that the cellobiohydrolase (CBH)/exoglucanase genes from GH family 7 were highly expressed in the protist symbionts compared to the *Reticulitermes flavipes* termite hindgut, midgut and the foregut tissue, as well as the salivary gland.

Three *R. flavipes* termite CBHs from GH family 7 were highly active against *p*NPC, a well-known model substrate for testing CBH activity (Sethi *et al*. 2013). The exoglucanases derived from *Pseudotrichonympha grassii* which are classified into GH family 7 was also reported by Liu *et al*. (2017). *P. grassii* is a flagellate found in the hindgut of *Coptotermes formosanus*. Functional genomic screens confirm that novel exo-glucanases from GH family 5 clades showed activity on acid solubilized and microcrystalline cellulose (Warnecke *et al*. 2007; Scharf 2015). In addition, Rashamuse *et al*. (2016) used functional metagenomics and purified a putative exo-glucanase enzyme that is classified as GH5D. The GH5D enzyme displayed activity on *p*NPC. Ni and Tokuda (2013) proposed that further studies on more diverse enzyme overexpression techniques using different systems, especially overexpression of CBHs, are required. Even though this suggestion was made five years ago, there is still little information available on the termite exo-glucanase enzymes. Thus, we suggest that more studies are required to address the lack of CBH/exoglucanase enzymes reported from termite cellulose degrading systems. Lastly, the suggested studies should also include the GH family classification of these enzymes.

1.5.3. β-glucosidases

The termite foregut primarily secretes β-glucosidase, while the midgut secretes both endoglucanase and beta-glucosidase that make essential contributions to cellulose degradation, by primarily hydrolyzing the amorphous regions of cellulose (Tokuda *et al*. 1997; Ni and Tokuda 2013). Figure 1.2 showed that beta-glucosidase is involved in the hydrolysis of short oligosaccharides that are produced by endoglucanase or exoglucanase. Scharf (2015) reported beta-glucosidase from seven different termite species (*R. flavipes, C. formosonus, Neotermes koshunensis, Globitermes sulphureus, Neotermes koshunensis, N. takasagoensis* and *C. gestroi*) which were classified to GH family 1. Feng et al. (2015) purified three beta-glucosidase homologs (CfGlu1C, CfGlu1B, and CfGlu1D) from the termite *C. formosanus* and studied their effect on cellulose degradation process. The results showed that CfGlu1B favors cellobiose, while CfGlu1C favors sucrose as a substrate. In

addition, the metagenomic study suggests that the hydrolysis of oligosaccharides into simple sugars in the termites is performed by glucosidase enzymes from GH family 3, or is performed by a cellobiose phosphorylase enzyme which belongs to GH family 94 (Warnecke *et al*. 2007).

Some termite β-glucosidase enzymes have multifunctional properties - for instance, Uchima *et al*. (2012) expressed, purified and characterized the endogenous β-glucosidase from the midgut of the higher termite *N. takasagoensis*. The β-glucosidase belongs to GH family 1 and proved to be multifunctional. It hydrolyzed *p*NP-fucopyranoside, *p*NP-glucopyranoside, *p*NP- lactoside and *p*NP- galactopyranoside. This β-glucosidase was also able to hydrolyze a range of oligosaccharides, including the disaccharide laminaribiose (β-1,3-linked glucose) which was the most preferred substrate, while gentiobiose (β-1,6-linked glucose) was the least preferred. However, β-glucosidase showed moderate activity on salicin (β-glucose), lactose (β-1,4-glucose-linked galactose), and sophorose (β-1,2-linked glucose). The diversity of β-glucosidase enzymes derived from the termite was reported by Do *et al*. (2014), who indicated that the *C. gestroi* termite gut has β-glucosidases, 6-phospho-β-glucosidases, licheninases, glucan endo-1,3-β-D-glucosidases, cellulose 1,4-β-cellobiosidases, glucan 1,3-β-glucosidases, and cellobiose phosphorylases, which belong to GH family 1, 1,4, 16, 5, 17 and 94. Finally, the β-glucosidase enzyme was able to hydrolyze cellobiose, cellotriose, cellotetraose, cellopentaose and cellohexaose. It was observed by thin-layer chromatography that β-glucosidase also possesses transglycosylation activity. A β-glucosidase with multifunctional activity was also reported by Scharf *et al*. (2010). These observations support the proposal that termites may be reservoirs for novel GH-enzymes.

1.6. Hemicellulolytic enzymes derived from termites

The main chains of hemicellulose are composed of xylose for xylan or glucose for xyloglucan or mannose for mannan. Malgas *et al*. (2015a) reviewed in detail the types and composition of the mannan polysaccharides; while xylan composition is reviewed in detail by Van Dyk and Pletschke (2012) and details of xyloglucan composition are provided by Gilbert *et al*. (2008). In summary, chemical compositions of hemicelluloses vary across plant species. Large differences are documented to exist between gymnosperms (wood) and angiosperms (grasses or agricultural cereal crops). Generally, hemicelluloses of gymnosperms consist primarily of acetylated glucomannan, arabinoglucuronoxylan, and galactoglucomannan. In addition, the hemicelluloses of some gymnosperms include xyloglucan. In contrast,

hemicelluloses of angiosperms primarily contain non-acetylated glucuronoxylan and glucomannan.

Given the complexity of the hemicellulose structure, the digestion of hemicellulose requires a large set of enzymes that degrade the polysaccharides or oligosaccharides. These enzymes include xylanases (EC 3.2.1.8), β-xylosidases (EC 3.2.1.37), α-glucuronidases (EC 3.2.1.131), α-arabinofuranosidases (EC 3.2.1.39) and acetylxylan esterases (EC 3.1.1.72) (Van Dyk and Pletschke 2012; Walia *et al*. 2017). Xylan, a linear polymer made of xylose units linked by β-1,4-glycosidic bonds, is the major component of hemicellulose. Xylanases and β-xylosidases cleave the backbone of xylan to release xylose (Walia *et al*. 2017). The glucuronidase, arabinofuranosidase and esterases are required to hydrolyze the branching sugars (Figure 1.3). Mannanases (EC 3.2.1.78) are enzymes which initiate the degradation of mannans by randomly cleaving the backbone to liberate short β-1,4-mannooligomers (Figure 1.4). The next key group of enzymes in mannan hydrolysis are β-mannosidases (EC 3.2.1.25), enzymes which catalyze the hydrolysis of β-1,4-linked mannosides and releasing mannose from the non-reducing end of mannans and mannan oligosaccharides. The α-galactosidases (EC 3.2.1.22) are enzymes that hydrolyze α-galactoside substituents from simple oligosaccharides and polymeric substrates (Figure 1.4). The enzymes that cleave the xyloglucan backbone at unbranched glucose residues are called xyloglucanases (EC 3.2.1.151). Although multiple hemicellulolytic enzymes could collaboratively act against lignocellulose in conjunction with cellulases, studies on hemicellulose degradation in termites are limited compared to what is available on cellulose digestion. Thus, in this section, we will focus on xylanases and mannanases that are produced by the termite.

1.6.1. Xylanases

Ni and Tokuda (2013) reported that, based on most of the research conducted on xylanases activities in the lower termite digestive system, it appears that xylanase activity is mainly found in the hindguts of lower termites. These suggest that xylanase activity from lower termites is mostly associated with the symbionts found in the hindgut. For instance, 25 strains of 13 yeast species found in a lower termite (*Reticulitermes chinensis*) hindgut were identified to be xylanase producers (Ali *et al*. 2017). In addition, novel yeast species isolated from wood-feeding termite (*R. chinensis*) guts, such as *Candida gotoi, Candida pseudorhagii, Hamamotoa lignophila*, *Meyerozyma guilliermondii, Sugiyamaella* sp. 1, *Sugiyamaella* sp. 2 and *Sugiyamaella* sp. 3 were also confirmed to produce xylanases. Even though this study by Ali *et al*. (2017) provides a detailed classification of yeast species found

in the gut of a lower termite, it does not give details of the type of xylanases produced by these yeast species or classifies them. Using biochemical characterization and a proteomic approach, Cairo *et al.* (2011) demonstrated that the lower termite *C. gotoi* gut extract contained xylanases which were classified into GH families 10, 11 and 43.

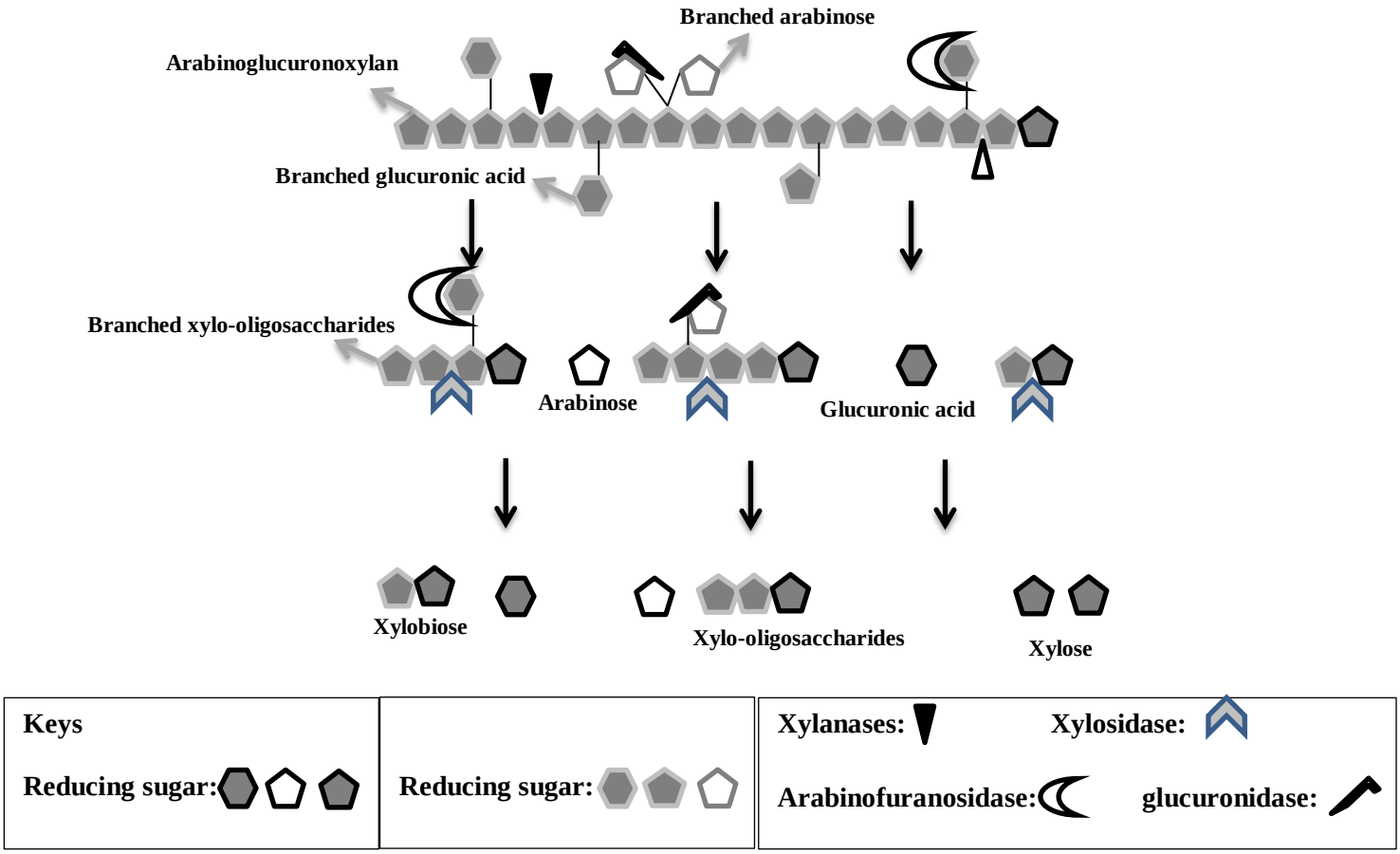


Figure 1.3. Schematic representation of xylan enzymatic hydrolysis. Xylanase enzymes act on the xylan backbone and produce xylo-oligosaccharides or xylobiose. Arabinofuranosidase cleaves arabinose branching from the xylan backbone. α-Glucuronidase cleaves glucuronic acid branching from the xylan backbone. Arabinoglucuronoxylan was chosen to show the enzyme cleaving sites because it consists of the sugar molecules that are found in arabino- and glucurono-xylan.

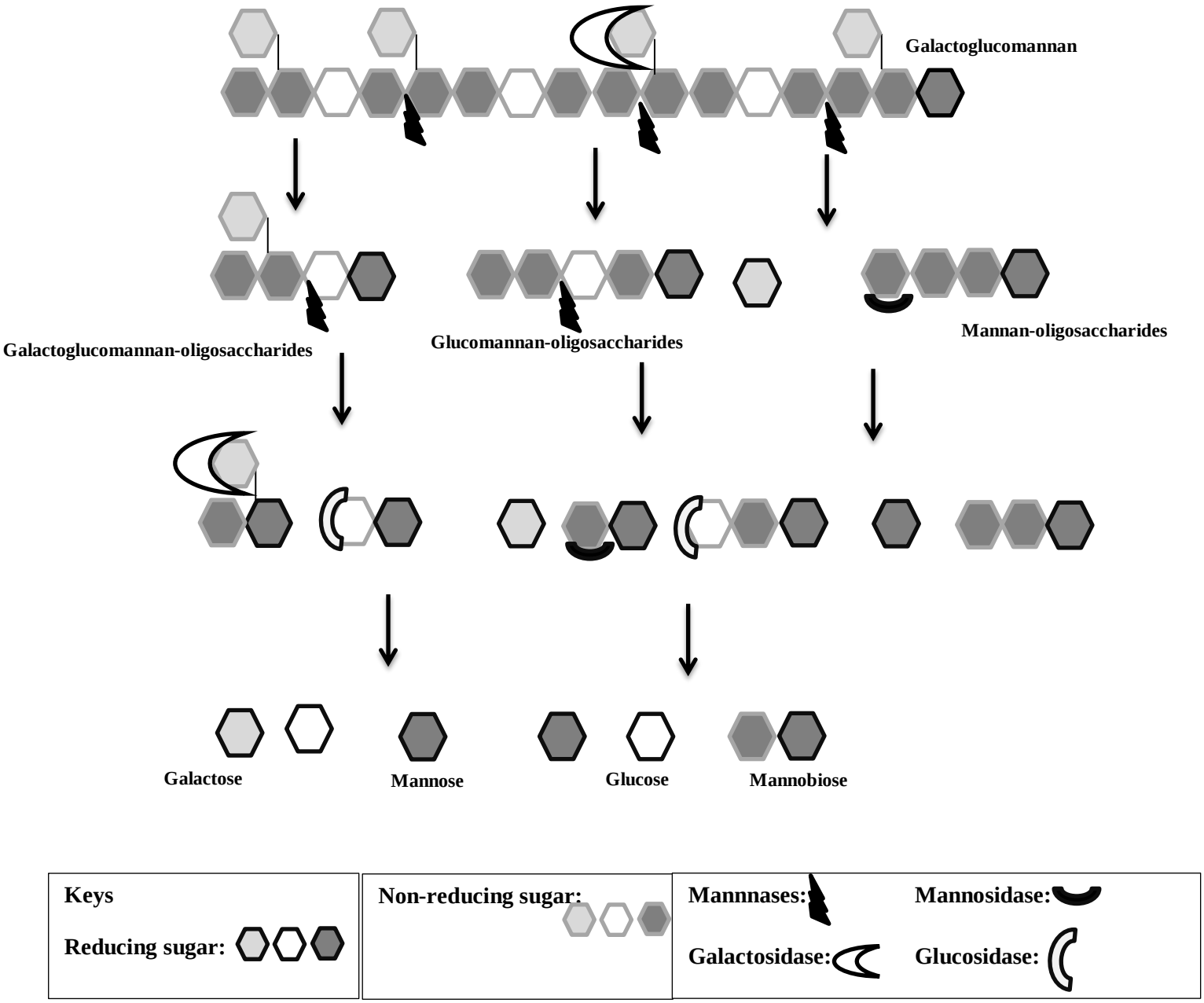


Figure 1.4. Schematic representation of mannan enzymatic hydrolysis. Mannanase enzymes act on mannan backbone and produce mannan-oligosaccharide or mannobiose. Mannosidase cleaves mannan-oligosaccharides or mannobioses. α-Galactosidase cleaves galactose branching from the mannan backbone. The glucosidase cleaves the glucose units on the backbone of glucomannan oligosaccharides. The galactoglucomannan was chosen to show the enzyme cleaving sites because it consists of the sugar molecules that are found in galactomannan and glucomannan.

In contrast, the fungus-growing higher termites, or wood and litter feeding higher termites, display xylanase activity in their midguts (Ni and Tokuda 2013). These observations suggest that higher termites and their symbionts can produce xylanase enzymes. Rashamuse *et al.* (2016) demonstrated that a higher termite (*Trinervitermes trinervoides*) hindgut possessed the multi modular β-1,4-xylanase (from GH family 11) which included two putative carbohydrate-binding modules (CBMs). This endoxylanase was able to release xylooligosacharides from wheat arabinoxylan. Substrate specifity profiles of this termite metagenome derived enzyme showed an additional endoxylanase from GH family 5. The enzyme was further classified to GH clan H and it was referred to as GH5H.

A termite gut bacterial xylanase, Xyl-ORF19, derived from gut bacteria of a wood-feeding termite (*Globitermes brachycerastes*) was cloned and characterized (Han *et al*. 2013). The protein structure of Xyl-ORF19 has two domains, a C-terminal GH10 catalytic domain and an N-terminal Big_2 noncatalytic domain. The catalytic domain was shown to be a $(\alpha/\beta)_8$ barrel as in GH10 xylanases but it has two extra β-strands. The non-catalytic domain is structurally similar to an immunoglobulin-like domain of intimins. The recombinant enzyme without the non-catalytic domain has fairly low catalytic activity, suggesting that the non-catalytic domain could affect the enzyme performance. Three endoxylanases from the flagella of *C. formosanus* was assigned to GH family 11 (Arakawa *et al*. 2009). GH family 11 xylanases seem to be more prevalent in termite symbionts that feed on wood or grasses, including cereal crops (Mattéotti et al. 2012; Ni and Tokuda 2013; Rashamuse *et al*. 2016).

The novel strain *Paenibacillus macerans* IIPSP3 isolated from wood-feeding termites, produced high levels of the pH- and thermo-stable xylanase. The IIPSP3 xylanase is a novel enzyme that can be used in the biofuel industry as it efficiently operates at pH 4.5 and in the paper and pulp industry as it can efficiently operate at pH 9.0 (Dheeran *et al*. 2012). Figure 1.3 shows that xylanases are important xylan degrading enzymes because they produce oligosaccharides that are hydrolyzed by xylosidase enzymes. In addition to xylanases that belong to GH family 10 or 11 shown above, Ni and Tokuda (2013) also reported xylanase enzymes which belong to GH family 8 and 43, respectively. CAZy classification of termite enzymes still requires attention as many researchers publish work on xylanase enzymes without assigning them to a specific GH family. Recently, a study was conducted with about

30 xylanase active bacterial isolates, but not even a single xylanase was assigned to the relevant GH family (Mathew *et al*. 2017).

1.6.2. Xylosidases

Generally, the xylo-oligosaccharides produced by endoxylanases are further hydrolyzed to xylose by the enzymes called xylosidases. The digestive enzymes in the gut of *N. corniger* worker and soldiers termites included β-xylosidases (Lima *et al*. 2014). The gut of the termite *Reticulitermes santonensis* contains a unique diversity of prokaryotic and eukaryotic microorganisms. A novel β-glucosidase, which belongs to GH family 1, was identified from the gut bacteria of *R. santonensis*. However, this enzyme displayed higher β-xylosidase activity compared to β-glucosidase (Matt *et al*. 2017). A study by Liu *et al*. (2016) demonstrated that the metatranscriptome of the protistan community in the gut of lower termites, *R. flaviceps*, contained an abundance of lignocellulases, which play a vital role in termite nutrition. In addition, β-xylosidase identified among the lignocellulases was derived from the protistan community. Even though Ni and Tokuda (2013) indicated that there is a shortage of research on the hemicellulose degrading enzymes five years ago, the situation remains the same when it relates to the β-xylosidases derived from termites or their gut symbionts.

1.6.3. Mannanases and mannosidases

The lower termite *Coptotermes gestroi* crude extract was used to hydrolyze hemicellulosic substrates such as mannan and galactomannan (Cairo *et al*. 2011). Using proteomic techniques, Cairo and co-workers identified two mannan degrading enzymes which belong to GH family 2 and 38, respectively. Six genes encoding for mannan endo-1,4-β- mannanases were identified in *C. gestroi* (Do *et al*. 2014). All the endomannanases were assigned to GH family 26. This family of enzymes is well known to contain the endomannanases (Malgas *et al*. 2015a, 2015b). Glycoside hydrolase family 26 mannanase C (RsMan26C) from a symbiotic protist of the termite *Reticulitermes speratus* was characterized after it was heterologously expressed in an *E. coli* system (Tsakagoshi *et al*. 2014). Specific activities of RsMan26C on mannotetraose, mannopentaose, and mannohexaose produced mannose, mannobiose, and mannotriose. In addition, the enzyme displayed partial activity against mannotriose but failed to cleave mannobiose after 24 h. RsMan26C is a typical endo-mode mannanase that cleaves the internal β-1,4-linkage of manno-oligosaccharides with a degree of polymerization higher than three. The substrate specificity of RsMan26C was analyzed to

elucidate the hydrolysis profiles of glucomannan, galactomannan from locust bean gum and mannan, respectively. RsMan26C, in addition, degraded and released reducing sugars from crystalline ivory nut mannan and galactomannan from guar gum, but activities against these substrates were lower.

Heterologous expression, purification, and characterization of a termite *Reticulitermes speratus* symbiotic protist β-mannanase (belonging to GH family 26) was performed using a *Pichia pastoris* strain KM71H (Tsukagoshi *et al*. 2014). This β-mannanase was referred to as RsMan26H and biochemical analysis demonstrated that RsMan26H was an endo-processive mannobiohydrolase. The endo-processive manner of RsMan26H was characterized by a prominent production of mannobiose from the onset of the reaction against manno-oligosaccharide (mannohexaose – mannotriose) and β-mannan, ivory nut mannan and locust bean gum. *Bacteroidales* are among the major constituents in the termite gut microbial community (Yuki *et al*. 2014). Strain Rs-03T was isolated from the gut of the wood-feeding subterranean termite *R. speratus* and documented in the Japanese Collection of Microorganisms as JCM 10512T, a novel bacterial species, *Bacteroides reticulotermitis*. The genome sequence of *B. reticulotermitis* strain JCM 10512T showed that this bacterium has a number of the enzymes that hydrolyze lignocellulose. Among these enzymes were α- and β-mannanases and β-mannosidases, however these enzymes were not assigned to GH families. Finally, metagenomic fosmids and their functional screening showed a number of lignocellulose degrading enzymes activities, including the mannanase enzymes derived from the gut of termite *N. corniger* and classified into GH families 31, 74, 43, 95 and 42 (Abot *et al*. 2016). Figure 1.4 illustrates the mannan degradation by the different GH enzymes, which includes the backbone cleaving enzymes and the oligomannan cleaving enzymes. The enzymes breaking down the manno-oligosaccharides are called mannosidases.

The α- and β-mannosidase enzyme activities were shown in the termite as early as 1984 (Azuma *et al*. 1984). Pine wood is rich in mannan and the termite (*R. flavipes*) gut extracted β-mannosidase displayed high activity on pine wood (Karl and Scharf 2015; Rajarapu and Scharf 2017). The single-cell genomics approach was used to clarify the functions of an ectosymbiotic relationship of bacterium and protist found in the termite hindgut (Yuki *et al*. 2015). The authors focused on *Candidatus symbiothrix dinenymphae*, which is attached to the cell surface of the cellulolytic protist *Dinenympha* spp. and accounts for 2.5% of the bacterial cells in the gut of the lower termite *R. speratus*. Genome screening of the coding domain

sequences (CDSs) against the CAZy database (http://www.cazy.org/) revealed that the bacterium genome possessed at least 58 genes coding for cellulolytic or hemicellulolytic enzymes. The identified hemicellulytic enzymes included two α-mannosidases which were classified into GH families 38 and 92. α-Mannosidase hydrolyzes the non-reducing α-D-mannose residues of α-D-mannosides.

1.7. Hemicellulolytic branch cleaving enzymes derived from termites

Hemicellulose has a complex composition compared to cellulose and therefore requires a large number of enzymes for complete hydrolysis. Enzymes degrading hemicellulose can be divided into depolymerizing enzymes which cleave the backbone and enzymes that remove substituents (debranching enzymes) which may pose steric hindrances to the hemicellulose depolymerizing enzymes (Van Dyk and Pletschke 2012). Xylans and mannans generally have several different substituents linked to their backbone. The substituents sugar units include arabinose, galactose and glucose and their removal from the backbone of xylans or mannans require a number of different debranching enzymes. Some of these accessory enzymes are α-L-arabinofuranosidases, α-glucuronidase and α-galactosidase.

The genes from the termite hindgut were shown to encode α-L-arabinofuranosidases and arabinan endo-1,5-α-L-arabinosidases which belong to GH families 43 and 51, respectively (Yuki et al. 2015). These enzymes cleave α-L-arabinofuranoside residues from arabinoxylan and arabinogalactan. Arabinoxylan is a major component of hemicellulose in plant cell walls and contains β-1,4-linkages between xylose and arabinose. Yuki et al. (2014) reported that *Bacteroides reticulotermitis* strain JCM 10512 displayed arabinofuranosidase activity. In addition, the *C. gestroi* termite gut microbiota hosts two different arabinofuranosidases, namely α-N-Arabinofuranosidase and Arabinan endo-1,5-α-L-arabinosidase belonging to GH family 43 (Do et al. 2014). An α-L-arabinofuranosidase belonging to GH family 43 was also identified in *C. gestroi* (Cairo et al. 2011). Functional metagenomic screening of termite gut *N. corniger* was used to identify and confirm the enzymatic activities of arabinofuranosidases which belong to GH families 2, 51 and 88. The α-N-arabinofuranosidase was also found in the protistan community of *R. flaviceps* using the metatranscriptomics technique. The data on the metagenomics and metatranscriptomics of arabinofuranosidase enzymes is rapidly accumulating. However, the techniques and properties of GHs are not well characterised or given the necessary importance. The biochemical reactions can help elucidate the novel arabinofuranosidases based on their physico-chemical properties.

Glucuronoxylans (GXs) consist of the main chain of 1,4-linked β-D-xylopyranosyl residues to which α-1,2-D-linked glucopyranosyluronic acid residues are attached (Togashi *et al*. 2009). This form of the xylan is the most abundant in hardwoods, for example, *Eucalyptus globulus*. Glucuronidases are enzymes that remove the glucuronic acid side chains from the xylan backbones. However, only a few studies have paid attention to the presence of the glucuronidases in the termite digestive system (and its symbionts). Warnecke *et al*. (2007) used a metagenomic analysis of the bacterial community resident in the hindgut paunch of a wood-feeding higher termite *Nasutitermes* species and revealed that the bacterial symbionts produced some glucosidases including α-glucuronidases (belonging to GH families 4 and 67) and a D-4,5-unsaturated β-glucuronyl hydrolase belonging to GH family 88.

Galactomannans are composed of linear chains of β-1,4-linked D-mannose residues with α-1,6-linked D-galactose side groups (Malgas *et al*. 2015a). α-Galactosidases are enzymes that hydrolyze α-galactoside substituents from galactomanno-oligosaccharides and galactomannans (Malgas *et al*. 2015a, 2015b). The *Nasutitermes* species hosts a symbiotic bacterial community that produces β-galactosidases and α-galactosidases which belong to GH families 1, 2, 4, 27, 35, 36, 42 and 98 (Warnecke *et al*. 2007). The genome of the ectosymbionts of protists in the termite gut also encoded for the galactosidase enzymes, which belong to GH family 2 and 42. β-galactosidase cleaves non-reducing β-D-galactose residues in β-D-galactosides. However, GH family 27 and family 36 α-galactosidase cleave non-reducing α-D-galactose residues in α-D-galactosides (Yuki *et al*. 2015). An α-galactosidase and a β-galactosidase belonging to GH family 4 and 42, respectively were also reported in the gut microbiota of the termite, *C. gestroi* (Cairo *et al*. 2011; Do et al. 2014). These studies suggest that galactosidase is one of the most abundant substituents cleaving enzymes (after arabinofuranosidase) found in termites and their gut symbiotic microbiota.

1.8. Termite enzyme synergy studies

Synergy studies have been used to elucidate the specific mechanism of action and interaction of cellulolytic and hemicellulolytic enzymes (Van Dyk and Pletschke 2012). Synergy studies are generally carried out using pure enzymes on defined substrates. To elucidate the specific mechanisms of action of the cellulolytic or hemicellulolytic enzymes, a number of defined substrates are used, which include colorimetric substrates or polymeric substrates, such as *p*-nitrophenol substrates and carboxymethyl cellulose. However, to elucidate the synergism between the enzymes more complex substrates like crystalline cellulose, wood derived

substrates or agricultural residues are required (Tokuda *et al*. 2005; Van Dyk and Pletschke 2012).

Natural feeding behavior of termites generally occurs in a sequentially synergistic manner (using both hemicellulolytic and cellulolytic enzymes) and this phenomenon is well established in termite studies (Azuma *et al*. 1984; Tokuda *et al*. 2005; Ni and Tokuda 2013; Brune 2014). Interestingly, Tokuda *et al*. (2005) reported that crystalline cellulose hydrolysis in the higher termite digestive system was mainly located in the midgut and hindgut compared to the lower termites, which significantly hydrolyzed the crystalline cellulose in the foregut (salivary glands) and hindgut. The activities of the crystalline cellulose hydrolyzing enzymes were measured by quantifying the amount of glucose and total reducing sugar produced by these enzymes. The hydrolysis of the crystalline cellulose by termites suggest that the termite digestive system possesses either the individual action of cellobiohydrolases, or synergistic enzymatic reactions catalysed by cellobiohydrolases, endo-β-1,4-glucanases, and possibly β-glucosidases.

Scharf *et al*. (2011) showed that in the case of the termites and their symbiotic microbiota, synergistic activities occur not only between the lignocellulolytic enzymes but also between the enzymes that are extracted from the termite (host of symbionts) tissue and the enzymes derived from termite symbionts. The authors used a combination of native gut tissue preparations (from *R. flavipes*) and recombinant enzymes derived from the host gut transcriptome to identify synergistic collaborations between the host and the symbionts. Pinewood sawdust was used as the substrate of interest. Results of synergy studies between enzymes prepared from termite gut tissue and those produced by the symbionts revealed that there was a significant increase in the amounts of glucose and xylose released. There was a 2 and a 1.25 fold increase in glucose and xylose produced in the synergy studies, respectively.

Synergy between purified cellulolytic or hemicellulolytic enzymes or combinations of these enzyme systems will be discussed below. However, it is important to note that there is a lack sufficient studies on the characterisation of recombinant pure GH enzymes derived from termites tissue or derived from termite uncultured microbiota metagenomes (Saadeddin, 2014).

1.8.1. Synergy studies with cellulolytic enzymes derived from termites

Synergy studies with the recombinant enzymes endoglucanase from GH family 9 (referred to as Cell-1) and β-glucosidase from GH family 1 showed more than 300-fold and 70-fold increases in amounts of glucose produced from pine sawdust and beechwood xylan,

respectively, relative to each enzyme alone (Scharf *et al*. 2011). Furthermore, *C. formosanus* termite β-glucosidase (referred to as CfGlu1C, CfGlu1B, and CfGlu1D) synergy studies were performed using three different natural substrates, which included D-lactosum, cellobiose, and sucrose (Feng *et al*. 2015). The combination of CfGlu1C and CfGlu1B enhanced the hydrolysis of D-lactosum with about 2-fold, but not that of cellobiose or sucrose. Addition of the CfGlu1D enzyme to the mixture of CfGlu1C and CfGlu1B enhanced the hydrolysis activity with about 2.5-fold. The synergy between termite β-glucosidases (G1mgNtBG1) and Celluclast 1.5 L was compared to synergy between Novozyme 188 and Celluclast 1.5.L (commercial enzymes), in a study which aimed to hydrolyze Avicel (Uchima *et al*. 2012). After 48 h of reaction mixture incubation, the combination of G1mgNtBG1 to Celluclast 1.5 L resulted in the release of higher amounts of reducing sugars (9 mg/ml) compared to Celluclast 1.5 L alone, which released about 7 mg/ml total reducing sugar. The combination of Novozyme 188 and Celluclast 1.5 L only produced about 7.5 mg/ml total reducing sugar. These results show that *N. takasagoensis* termite G1mgNtBG1 was more effective than the commercial β-glucosidase, Novozyme 188, and was also able to form synergy with commercial Celluclast 1.5 L. Three recombinant cellobiohydrolases which belong to GH family 7 (referred to as GHF7-3, GHF7-5, and GHF7-6) were used to hydrolyze pinewood (Sethi *et al*. 2013). The synergistic combinations of each of the three GHF7 enzymes with Cell-1 and β-glucosidase reported by Scharf *et al*. (2011), Showed that synergy was formed between GHF7-3, Cell-1, and β-glucosidase. The synergy assays released about 2.5-fold more glucose compared to glucose produced by individual enzymes.

Even though efforts were made to document the presence of the cellulolytic enzymes in the lower and higher termites through the use of "omics' studies, purified cellulolytic enzymes from termites are still required if researchers want to understand the cellulolytic enzymes in termites. We share the same view with Watanabe and Tokuda (2010) and Ni and Tokuda (2013) that there are many cellulolytic gene sequences currently available. Thus, we suggest that heterologous expression systems of these genes will help produce pure cellulolytic enzymes, which will contribute to our understanding of how termite cellulolytic enzymes act. In addition, a synergistic application of heterologously expressed cellulolytic enzymes (cellobiohydrolase/ exoglucanases, endoglucosidases and β-glucosidase) during hydrolysis of lignocellulosic biomass could explain the mechanisms these enzymes employ. Understanding these enzyme mechanisms will assist with regards to the appropriate industrial application of termite cellulolytic enzymes.

1.8.2. Synergy studies with hemicellulolytic enzymes derived from termites

Functional metagenomic screening for hemicellulases in the fungus growing termite *Pseudacanthotermes militaris* revealed a large number of hemicellulose degrading enzymes (Bastien *et al*. 2013). These enzymes included arabinofuranosidases, xylosidases and endoxylanases. This study and other studies by Watanabe and Tokuda (2010), Ni and Tokuda (2013), Brune (2014) and Scharf (2015) demonstrate that termites and their symbionts possess hemicellulase enzymes, which could be used to hydrolyze the hemicellulosic biomass in a synergistic manner. However, it is surprising to note that to date there are no synergy study reports with the recombinant xylanolytic enzymes derived from termites.

It has been established that application of lignocellulose degrading enzymes in the bio-refinery industry requires pH and temperature-stable enzymes. In addition, it has been shown that xylanases derived from termites and their symbionts are pH-stable and some are thermo-stable. Dheeran et al. (2012) reported a novel xylanase from *Paenibacillus macerans* IIPSP3 isolated from a termite hindgut. Xylanase was found to be thermostable over a broad temperature which ranged between 40 and 90°C; however, the optimum temperature was 60°C. Its activity was observed to increase sharply with gradual increases in temperature up to 60°C (specific activity of 16.6 U/mg). When the enzyme was incubated at temperatures ranging from 30 to 100°C, it maintained 100% activity of up to 60°C, but thereafter xylanase activity decreased by 5, 12, 30 and 49% when the temperature was further increased to 70, 80, 90, and 100°C, respectively. Rashamuse et al. (2016) also showed that xylanases from a termite gut possesses special features, for instance, xylanase from GH family 11 (Xyl1) was of particular interest as it was shown to consist of a catalytic domain and two carbohydrate-binding modules (CBMs) which function to selectively bind insoluble xylan and increase the rate of hydrolysis. With such features of xylanase and much more reported by others cited in the four reviews mentioned above, there is no doubt that the synergistic activities of the termite hemicellulolytic enzymes could improve our knowledge of hemicellulose degradation. Models for studying enzyme synergy for bioconversion of hemicellulosic biomass are well explained by Van Dyk and Pletschke (2012) and we suggest that using this approach can help improve our knowledge of hemicellulose degradation.

Chapter 2: Research Motivation and Hypothesis

2.1. Problem statement

Glycoside hydrolases (GH) are important enzymes because are used different biotechnological industrial applications, which include the biofuel, textile or pulp, and paper industries. However, the GH enzymes that are currently commercially available are expensive and some of the commercial enzymes are not sufficiently effective during hydrolysis of the biomass. This necessitates a search for sources of effective novel GH enzymes which, would hydrolyse lignocellulose effectively at a low enzyme concentration dosage. Van Dyk & Pletschke (2012) argued that the use of GH enzymes in synergy stratergies can improve the hydrolysis of lignocellulosic biomass. However, we have demonstrated (in Chapter 1 section 1.8) that there are very few synergy studies performed with cellulolytic enzymes and there are no reports on the synergy studies with hemicellulytic enzymes.

Most of the commercially available cellulases, hemicellulases or enzyme cocktails are produced from fungal or bacterial sources. However, termites are suggested to be novel GH enzyme reservoirs because termites' digestive systems degrade more that 90% of lignocellulosic biomass. Rashamuse et al. (2016) heterologously overexpressed and partially characterised endocellulases (GH5E), an exocellulase (GH5D), endoxylanases (GH5H) and endoxylanase/endocellulase (GH43) derived from the hindgut bacterial symbionts of a termite (*Trinervitermes trinervoides*). Motivations to perform the current study were developed: 1) The partial characterisation of novel GH5E, GH5D, GH5H and GH43 is insufficient for undestanding the functions and applications of these novel enzymes. Thus, we proposed that comprehensive characterisation of these enzymes could assist us in elucidating GH5E, GH5D, GH5H and GH43 functions, classifications (based on their activities) and their potential applications in the biorefinery industry. 2) Studies on GH enzymes derived from the gut microbiota of the temites show that there is not much information available regarding heterologously expressed GH enzymes and their applications. 3) Only a few GH enzyme synergy studies have been performed using termite derived cellulases; however, there is no documentation of synergy studies performed with hemicellulase enzymes derived from termite gut microbiota. 4) Our studies will mitigate the knowledge gap regarding the functions of purified GH enzymes derived from the gut microbiota of temites. These motivations led to the research questions below.

2.2. Research questions

1) Can cellulases and xylanases (purified from a *Trinervitermes trinervoides* termite metagenome) work in a synergistic manner for effective hydrolysis of lignocellulosic biomass?; 2) Can these termite metagenome derived enzymes be combined in a cocktail with commercial enzymes that will increase sugar yields from sweet sorghum bagasse and corncob?; 3) How can we best explain the mechanisms that are employed by these termite derived enzymes, based on their behavior during the synergistic hydrolysis of the biomass?; 4) Which type of enzyme synergy strategy (between simultaneous, sequential or successive synergy) can hydrolyse sweet sorghum and corncob most effectively? and 5) Can the products of the reaction that was conducted using termite derive GH enzymes produce any other valuable products that can be applied in other industries besides the biofuel industry?

2.3. Aims and Objectives

In order to address the research questions above, the following aims and objectives were formulated:

1) To pretreate the corncob and sweet sorghum substrates with lime, NaOH and $NaClO_2$ and analyse their major components (lignin and holocellulose)
2) To characterise the biochemical properties of the purified recombinant enzymes and to determine the kinetic parameters of these enzymes (using assays with model substrates).
3) To investigate the effect of product inhibition and metal ions on the activities of the purified enzymes.
4) To determiner if the enzymes (GH5D, GH5E, GH5H and GH45) can operate in synergy and to determine if the enzymes act/behave in a simultaneous, sequential or successive manner during the reaction.
5) To optimise enzyme synergy by adding commercial enzymes (in particular β-glucosidase and xylosidase) to our enzyme combinations, and to construct enzyme cocktails that will increase the yields of sugar monomers that can subsequently be fermented to produce biofuels.
6) To investigate if exo-glucanase (GH5D) can perform transglycosylation, which can be used to produce biochemicals with special properties.

2.4. Overview of thesis

Alkaline and oxidative pretreatments of sweet sorghum bagasse and corncob were performed and the biomasses were characterized after pretreatment as discussed in Chapter 3. In Chapter 4 we successfully characterized the termite derive enzymes, namely endoglucanase GH5E, exoglucanase GH5D, a multifunctional GH5H and a bifunctional GH45. The effects of metal ions and hydrolysis products (glucose and xylose) on the performance of these enzymes were investigated. GH5D and CBHI hydrolysed *p*NPC suggesting that they hydrolyse the substrates with the same mechanism. Based on this observation a comparative study between GH5D, CBHI and CBHII was conducted in Chapter 5. The results showed that GH5D was not able to hydrolyze the crystalline cellulose like CBHI and CBHII. However, the GH5D was highly active on insoluble β-glucan from barley. In Chapter 6 we used synergy studies to optimise and produce core enzyme sets made up of GH5E, GH5E, GH5H, GH45, β-glucosidase, and xylosidase. The model substrates CMC and insoluble arabinoxylan were used to optimise cellulolysic and xylanolytic core enzyme sets. Simultaneous application of enzymes in a synergy study (referred to as simultaneous synergy) was more effective than the sequential or successive applications, referred to as sequential or successive synergy. GH5D was also shown to possess transglycosylation properties in Chapters 3 and 4. We addressed this observation by using bioinformatics techniques and modeling the structure of the GH5D enzyme in Chapter 7. The structure of GH5D was shown to be a $(\alpha/\beta)_8$ barrel with a long tunnel at the active site, which supports our early observations in the characterisation Chapter 4. Furthermore, the transglycosylation reactions performed using GH5D, in the presence of alcohols, were shown to produce alkyl cellobiosides that have antimicrobial activities.

Chapter 3: Alkaline and Oxidative Pretreatments of Corncob and Sweet Sorghum Bagasse for Lignin Removal

3.1. Introduction

Lignocellulosic biomass is classified into four different groups (woody biomass, agricultural residues, energy crops and cellulosic wastes like municipal wastes) based on its source (Saha 2003; Hoon and Hyun 2014; Saini *et al*. 2014; Kim *et al*. 2016). In the current study we focused agricultural residues, which include rice/wheat/barley straws, corn stover and sugar cane bagasse. Agricultural residues are an important substrate in the biorefinery industry. The agricultural residues consists of carbohydrate (cellulose and hemicellulose) and lignin as the major components (Van Dyk and Pletschke 2012; Kim *et al*. 2016).

Cellulose is a homogeneous, crystalline and water insoluble polymer that is made up of glucose units which are linked by β-1,4-glycosidic linkages (Saini *et al*. 2014; Lee *et al*. 2016). The crystalline nature of the cellulose makes it resistant to enzymatic attacks. In contrast, hemicellulose is a heterogeneous polymer (Saha 2003). It is made up of pentose (xylose and arabinose) and hexose (mannose, glucose and galactose) sugars. In most plant species, the major component of hemicellulose is xylan. The third major component of lignocellulosic biomass is lignin and it covers both cellulose and hemicellulose (Hendriks and Zeeman 2009; Zhao et al. 2012). Thus, lignin removal is important for increasing surface area biomass and increasing accessibility of the holocellulose to GH enzymes (Zhao et al. 2012).

Alkaline pretreatment is one of the best methods to delignify lignocellulosic biomass because of its desirable benefits (Kim et al. 2016). These benefits include the use of non-polluting chemicals, less-corrosive chemicals, use of mild conditions compared to acids and low cost of production.

Sodium hydroxide pretreatment is one viable method to remove lignin from biomass because of its low cost compared to dilute acid pretreatment (Modenbach 2013). Unlike ammonia fiber explosion (AFEX), sodium hydroxide pretreatment is performed at ambient pressure and temperature. Wu *et al.* (2011) pretreated sweet sorghum bagasse (SSB) with sodium hydroxide. Pretreatment removes lignin from the biomass and increases holocellulose content compared to untreated biomass. The removal of lignin and hemicellulose from the biomass was directly proportional to sodium hydroxide concentration, increase in temperature and period (time) of treatment (Wu *et al*. 2011; Wang *et al*. 2012). Sodium hydroxide not only removed the lignin but also reduced the crystallinity of the biomass. In addition, Wang *et al.*

(2012) suggested that removal of lignin and a reduction biomass crystallinity could enhance the accessibility of cellulose and hemicellulose to enzymes.

Lime pretreatment is also one of the low cost methods used to remove lignin from lignocellulosic biomass. It has a less of an environmental impact and is available in many countries (Bensah and Mensah 2013). Pretreatment of lignocellulosic biomass with lime is mainly performed using energy crops and cereal crops such as sugar cane bagasse and corn stover. Rabelo *et al.* (2013) pretreated sugar cane bagasse with lime. The pretreatment removed lignin from the biomass and decreased the crystallinity. The changes suggest a larger surface area of biomass is accessible to enzymatic hydrolysis, and changes in the structure (increase in amorphous regions) could also improve the enzyme digestibility of the biomass.

In addition, sodium chlorite has been used to delignify woody and agricultural biomass (Yu *et al.* 2011; Siqueira *et al.* 2013; Wu *et al.* 2014). Yu *et al.* (2011) showed that the biomass crystallinity index increased with the decreasing lignin content after $NaClO_2$ pretreatment. The increase of crystallinity index was an interesting observation because it has been reported that biomass pretreated with alkaline (NaOH and lime) change its structure to a more amorphous/cellulose-II structure, which means crystallinity index decreases with the removal of the lignin (Jungnikl et al. 2008; Beukes and Pletschke 2010).

3.2. Aims and Objectives

3.2.1. Aim

To remove lignin from sweet sorghum bagasse (SSB) and corncob (CC) by pretreating these biomasses with alkaline and oxidative chemicals

3.2.2. Objectives

1. To de-lignify sweet sorghum bagasse (SSB) and corncob (CC) by pretreating them with lime, sodium hydroxide (NaOH) and sodium chlorite ($NaClO_2$).
2. To biochemically characterise the SSB and CC biomass after each pretreatment by using the National Renewable Energy Laboratory (NREL) methods.

3. To analyse the topology of the SSB and CC biomass with scanning electron microscopy (SEM) and analyse their chemical changes with FTIR subsequent to pretreatments.
4. To determine which of the two pretreatment (alkaline and oxidative) effectively de-lignify the SSB and CC biomass.

3.3. Materials and Methods

3.3.1. Substrate pretreatment

Sweet sorghum (*Sorghum bicolor* L. Moench) was provided by Prof J. Dames and sweet corn was purchased at the local store. The corn kernels were removed and the corncob(CC) was used as a substrate. The CC and sweet sorghum were dried overnight at 60°C. Dry biomass was milled to a fine powder (less than 2 mm) using an industrial bench top blender/homogenizer. The powdered sweet sorghum was referred to as sweet sorghum bagasse (SSB). The powdered biomass was chemically pretreated with alkaline (lime and NaOH) and oxidative ($NaClO_2$) pretreatments. Alkaline pretreatment was conducted using modified NaOH and lime pretreatment according to Panagiotopoulos *et al.* (2010) and Beukes & Pletschke (2010), respectively. Oxidative pretreatment was performed according to Siqueira *et al.* (2013). For lime pretreatment, 5 g (w/v) of powdered SSB and CC were pretreated with 2 g (w/v) lime and then distilled water was added to make a final volume of 50 ml. For NaOH pretreatment, about 1.25 g (w/v) was added to 5 g (w/v) of SSB and CC and distilled water was added to make a final volume of 100 ml. One and a half grams (w/v) of $NaClO_2$ and 0.3% (v/v) acetic acid were used to pretreat 5 g (w/v) SSB and CC, distilled water was added to a final volume of 160 ml. The NaOH and $NaClO_2$ reactions were incubated in the oven at 70°C for 4 h, while the lime pretreatment was incubated for 24 h at the same temperature. Samples were filtered and washed with distilled water until pH 7.0 was reached. The samples were air dried and stored in an airtight container for future use.

3.3.2. Microscopy studies

3.3.2.1. Scanning electron microscopy (SEM)

To determine the SSB or CC biomass topological changes (due to lignin removal) subsequent to pretreatment with lime, NaOH, and $NaClO_2$, the dried SSB and CC samples were mounted on SEM stubs using double sided graphite tape, sputter coated with gold using a Balzers Union sputtering device. The gold coated SSB and CC samples were viewed under a Tescan

Vega SEM at 20kV. Digital images were captured using a Vega imaging system (Pinchuck 2009).

3.3.2.2. Light microscopy

The removal of lignin on the SSB and CC biomass subsequent to pretreatments was qualitatively assessed through histochemical techniques using a light microscope. The SSB and CC samples were stained with 10% (w/v) phlorogucinol in 95% (v/v) ethanol (Wiesner reagent) prior to light microscope analysis (Botha 2008; Tao et al. 2009). The SSB and CC samples were covered with the phlorogucinol solution and incubated at room temperature for 3 min. After the samples were drained, a few drops of concentrated HCl were added to the samples and incubated for another 3 min. The excess HCl was removed from the samples. A few drops of paraffin were used to preserve the colour that formed. The samples were visualized using an Olympus BX40 light microscope (magnification at X20) and photographed using an Olympus DP72 digital camera.

3.3.3. FTIR measurements

The functional group analysis of the untreated or pretreated SSB and CC biomass samples were analysed with Fourier Transform Infrared (FTIR). The FTIR spectra of Avicel (control), untreated, lime, NaOH and $NaClO_2$ pretreated SSB and CC samples were recorded at room temperature using an UATR-FTIR instrument (Perkin-Elmer, USA). The UATR-FTIR was equipped with a ZnSe-Diamond crystal composite that allows collection of FTIR spectra directly on a dry sample without any special preparation. The "pressure arm" or electronic force gauge was set to 80% on the SSB and CC samples positioned on top of the ZnSe-Diamond crystal to ensure a good contact between the sample and the incident IR beam, thereby minimizing loss of the IR beam. All FTIR spectra were collected at a spectrum resolution of 4 cm^{-1}, with four co-added scans per sample over the range of 4000 cm^{-1} to 650 cm^{-1}. A background scan of the clean ZnSe-Diamond crystal was acquired by scanning without the sample. The Perkin-Elmer software (Spectrum V. 6.3.5, 2009) was used to perform spectra normalization, baseline corrections and peak integration. The spectra of the Avicel, and untreated and pretreated SSB and CC samples, were presented as absorbance values, and each value represented the means of the four scans.

3.3.4. Biomass composition characterisation

Biomass composition of untreated, lime, NaOH and $NaClO_2$ treated SSB and CC were determined according to a modified method of Moxley & Zhang (2007) and the National Renewable Energy Lab (NREL) (Sluiter *et al.* 2010). The pretreated SSB and CC biomass

[300 mg (w/v)] was transferred into a test-tube. Three milliliters of strong sulfuric acid (72% w/v) was gently added to the tube containing sample. The sample was mixed with a glass rod and incubated at 37°C for 60 min on a bench top shaker (set at 170 rpm). Following the incubation, the samples were placed on ice for 10 min. The sample's hydrolysates were diluted with 84 ml of distilled water to achieve 4% (v/v) sulfuric acid concentration. One milliliter was further diluted to 1% (v/v) sulfuric acid concentration by adding 3 ml of distilled water in a separate tube. Both the hydrolysates containing 4% and 1% sulfuric acid were autoclaved at 121°C for 1 h. After autoclaving the samples, calcium carbonate was added upon cooling to neutralize the samples to pH 6.0. The samples hydrolyzed with 4% and 1% (v/v) sulfuric acid were used to determine total reducing sugar (glucose and xylose) and lignin content in the biomass.

3.3.5. Saccharide content of SSB and CC

3.3.5.1. Total reducing sugar quantification

After sulfuric acid hydrolysis of the untreated, lime, NaOH and $NaClO_2$ pretreated SSB and CC biomasses, the samples were centrifuged at 13 000 rpm for 5 min. The supernatant was used to measure total reducing sugars using a modified dinitrosalicylic acid (DNS) method (Miller 1959). The hydrolysate supernatants (150 µl) of pretreated SSB and CC were transferred to Eppendorf tubes. Subsequently, 300 µl of DNS reagent was added to the tube and the mixture was incubated at 100°C for 5 min and then cooled on ice for 10 min. Cooled samples (250 µl) were transferred into microtiter plates and readings were taken at 540 nm on a Powerwave-X microplate reader (Bio-Tek Instruments) using KC Junior software. Glucose and xylose were used as suitable standards (Appendix: Figures A4.1 B and C).

3.3.5.2. Glucose and xylose quantification

Glucose and xylose quantities in untreated, lime, NaOH and $NaClO_2$ pretreated SSB and CC biomasses were quantified using Megazyme kits and calculations were performed according to the instructions in the kits. Glucose and xylose were measured using a glucose oxidase/peroxidase (GOPOD) kit and a xylan and an arabinoxylan kit (using the microplate assay), respectively.

3.2.6. Total phenolic compounds

3.2.6.1. Total phenolic compounds in the acid hydrolysate

The amounts of total phenolics were determined using the Folin–Ciocalteu reagent. After acid hydrolysis of untreated, lime, NaOH and $NaClO_2$ pretreated SSB and CC biomass, the

samples were centrifuged and the hydrolysates were used to measure the total phenolic compounds. The samples (100 µl) were transferred into eppendorf tubes, then 500 µl of 10% (v/v) Folin-Ciocalteu reagent was added to the tubes and the samples were mixed properly. The mixtures were incubated for 3 min at room temperature and 400 µl of sodium carbonate was added. The mixture was left at the room temperature for 60 min. After 60 min, the samples (250 µl) were transferred to a microtiter plate and the absorbance values were measured at 765 nm on a Powerwave-X microplate reader (Bio-Tek Instruments) using KC Junior software. The total phenolic compounds were quantified using a standard curve. Gallic acid was prepared and used as a suitable standard (Turkmen *et al*. 2006).

3.3.6.2. Insoluble lignin

After sulfuric acid hydrolysis of the untreated, lime, NaOH, $NaClO_2$ pretreated biomass SSB and CC biomass, samples were filtered and the residues were collected to determine the insoluble lignin content according to NREL protocols (Sluiter et al. 2010). The collected residues were transferred into crucibles. The crucibles containing the biomass were weighed and were referred to as W1 (weight one). The crucibles were incubated at 105°C for 60 min, followed by measuring the crucible containing the dry biomass residues (W2=weight two) of untreated, lime, NaOH, $NaClO_2$ pretreated SSB and CC. The crucibles containing dry biomass residues were placed in the furnace set at 575°C for 24 h. The crucibles were then cooled in a desiccator and were weighed for the third time (W3). Insoluble lignin was measured using the simple equation: IL= W2-W3

3.3.6.3. Data Analysis

One way analysis of variance (ANOVA) was used to analyse the biomass composition of alkaline or oxidative pretreated SSB and CC biomass. Evaluation was conducted to elucidate significant increases or decrease of lignin, cellulose and hemicellulose content of SSB or CC biomass post alkaline or oxidative pretreatment. All pairwise comparison procedures were conducted using the Data analysis feature in Microsoft Excel.

3.4. Results and discussion

3.4.1. Biomass morphological changes analysis

SEM was used to analyze the morphological changes on the surface of both SSB and CC samples after pretreatments. Figure 3.1 shows that lignin was removed from the biomass after pretreatment. Control/untreated samples were covered with the layer, which histochemical analysis detected as lignin (Appendix: Figure A3.1).

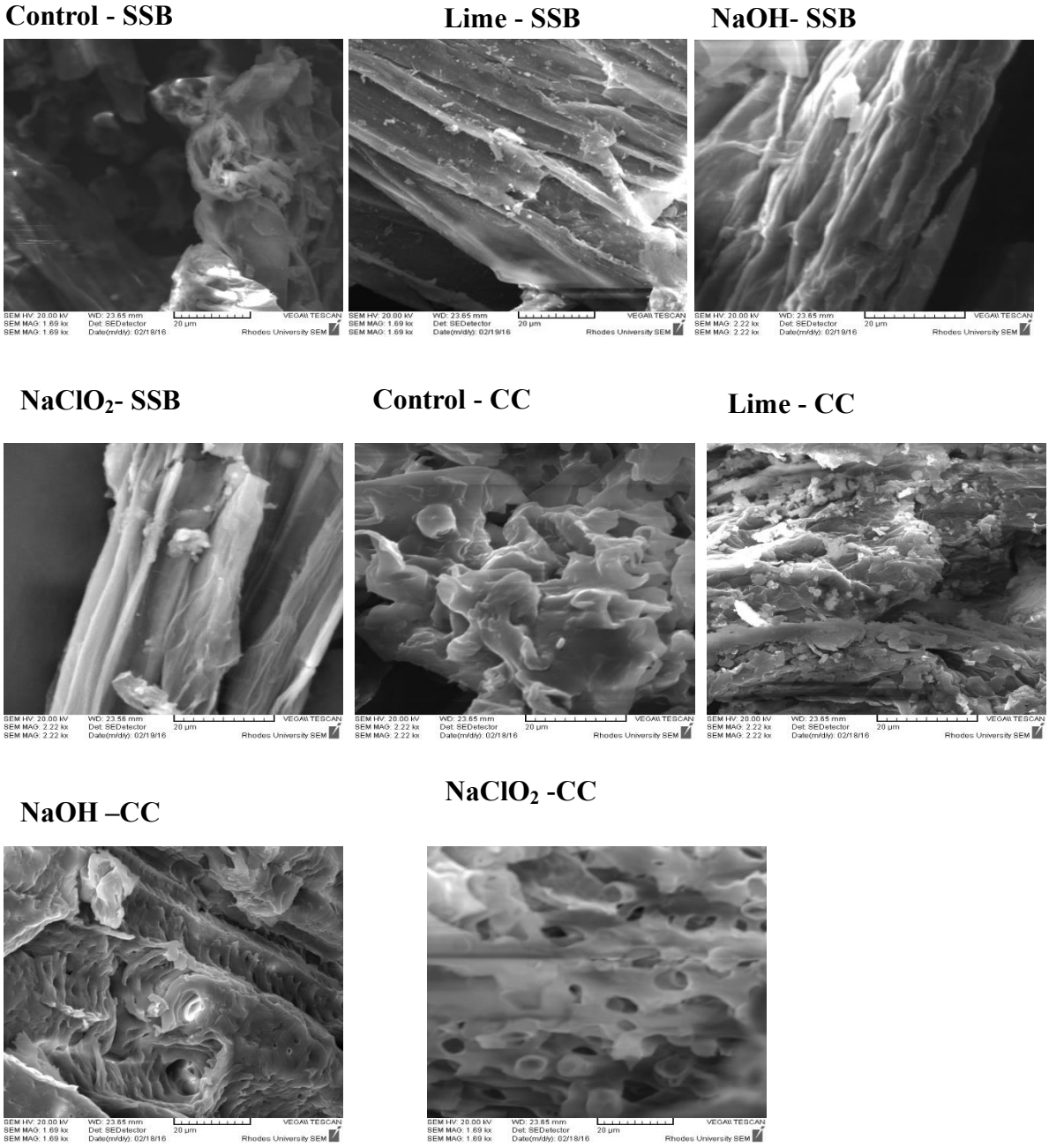


Figure 3.1. Scanning electron microscope (SEM) images that show changes in the morphology of sweet sorghum and corncobbiomass. SSB represents sweet sorghum bagasse that was pretreated with lime, sodium hydroxide (NaOH) and sodium chlorite ($NaClO_2$). CC represents corncobthat was pretreated with lime, sodium hydroxide (NaOH) and sodium chlorite ($NaClO_2$). Controls were not treated with chemicals. Scale bar is set at 20 µm.

Subsequent to lime pretreatment, the layer of lignin was broken and small droplets of the lignin were formed on both SSB and CC samples. Beukes & Pletschke (2010) suggested that lime pretreatment solubilized the lignin, which could explain the small droplets formed on both SSB and CC biomass after pretreatment. NaOH pretreatment removed lignin from both SSB and CC biomass. After NaOH pretreatment, the biomass changed in structure and pore openings appeared on the surface of the biomass. These findings were consistent with other reports (Lima *et al*. 2013; Gabhane *et al*. 2014; Kamarullah *et al*. 2015). The oxidative pretreatment with $NaClO_2$/acetic acid removed the lignin from the biomass, such that the pores appeared on the surface of the biomass after pretreatment. Siqueira *et al.* (2013) also reported that $NaClO_2$ solubilized parts of the lignin and degraded most of the lignin, except for the lignin in the middle lamella.

Our findings coincide with the proposed model of lignin migration by Donohoe et al. (2008). Donohoe and co-workers proposed that lignin re-localized during pretreatment of the biomass. The movement of lignin starts from inside the cell-wall and moves outwards, leaving pores on the surface of the pretreated biomass. It is suggested that changes in the biomass structure and the resulting pore openings should give enzymes more access to the biomass during hydrolysis (Beukes and Pletschke 2010; Kim and Han 2012; Kim et al. 2016).

3.4.2. Biomass structural and chemical changes - FTIR spectra analysis

ATR-FTIR was used to analyze the structural and chemical changes that occurred on SSB and CC biomass after alkaline or oxidative pretreatment. The distinct bands on the FTIR spectra and their positions allow the identification of polysaccharide structures and their composition (Kacuráková et al. 2000). In addition, O'Connor et al. (1958) used cotton fiber as a model biomass to show that FTIR can also be used to predict the crystallinity of the biomass. Table 3.1 shows five important IR spectra regions used to qualitatively characterise the structure of untreated, lime, NaOH and $NaClO_2$ pretreated CC and SSB biomass.The FTIR spectra regions between 4000 and 650 cm^{-1} were analyzed for all the samples, except for the regions between 1800 and 2800 cm^{-1}, because this region is not significant for the analysis of the polysaccharides (Abidi et al. 2014). The IR spectra of the untreated, alkaline and oxidative pretreated CC and SSB biomass were compared to that of Avicel (crystalline cellulose polymer), which was used as a model biomass because it has a well known structural and chemical composition.

The IR spectra ranging between 3700 and 3000 cm^{-1} forms a broad band because of the OH-stretching vibration coming from the hydrogen bonds in cellulose (Siroky et al. 2010; Abidi et al. 2014). We observed a broad band at approximately 3333 cm^{-1} in all our samples, coinciding to this OH-stretching vibration coming from the hydrogen binding the cellulose. The 3333 cm^{-1} band was also seen on the Avicel FTIR spectra. Interestingly, the FTIR spectra of the pretreated CC samples were different compared to those of pretreated SSB biomass at the 3333 cm^{-1} band (Appendix Figure A3.2). Figure A3.2-A shows that $NaClO_2$/acetic acid pretreated CC had the highest peak at 3333 cm^{-1}, followed by NaOH pretreated CC, Avicel, untreated and lime pretreated CC.

Table 3.1: The five important IR spectral regions for alkaline and oxidative pretreated corncob (CC) and sweet sorghum (SSB) samples

Vibrations (cm^{-1})										
IR region (cm^{-1}) and polymers	Avicel	CC-C	CC-lime	CC-NaOH	CC-$NaClO_2$	SSB-C	SSB-lime	SSB-NaOH	SSB-$NaClO_2$	Assignments
650-1200 (cm^{-1})										
	665.61	666.43	666.96	666.78	666.71	666.62	667.17	666.90	661.75	OH out-of-plane bending
	708.40	697.41	712.08	697.10	698.33				799.99	CH_2 rocking
Cellulose II	897.84	897.95	872.84	897.60	898.46		863.37	898.40		β-Linkage of cellulose
Cellulose, hemicellulose	1029.56	1031.75	1032.15	1027.70	1032.05	1030.99	1032.15	1029.52	1031.57	C-O and ring stretching modes
Cellulose, hemicellulose	1053.80									C-O stretch
cellulose I	1104.67									Anti-symmetric in-plane stretching band
Cellulose II	1160.54	1157.46	1158.17	1157.23	1158.53	1159.66	1159.06	1158.03	1157.66	Anti-symmetrical bridge C-O-C Stretching

1200-1500 (cm^{-1})										
Cellulose, hemicellulose	1202.85							1202.43		C-O stretching
Guaicyl lignin		1243.86			1243.26	1242.71			1234.06	C-O stretching or NH_2 deformation
Cellulose, hemicellulose	1315.68			1317.34	1316.45			1316.60	1318.58	CH_2 rocking
Cellulose II	1369.02	1371.15		1370.77	1371.86	1368.27		1370.80	1370.79	C-H bending
Cellulose II	1428.66	1428.08	1410.45	1421.90	1428.98		1416.72	1423.33	1434.71	CH_2 scissoring
1500-1800 (cm^{-1})										
Lignin						1519.97			1533.16	NH_2 deformation
Water	1640.43	1624.40		1643.19	1634.54	1634.69		1634.98	1627.06	O-H bending of adsorbed water
Hemicellulose		1729.82			1730.20	1731.23			1730.75	C=O stretching
2800-3000 (cm^{-1})										
Cellulose II						2851.46			2851.13	CH_2 asymmetrical and symmetrical stretching
Cellulose II	2896.80	2921.53	2918.04	2919.01	2920.50	2920.32	2919.23	2919.19	2918.90	CH_2 asymmetrical and symmetrical stretching
3000-3700 (cm^{-1})										
Crystalline cellulose	3333.93	3336.56	3319.15	3327.12	3335.02	3332.28	3304.01	3333.23	3329.90	Intra-molecular hydrogen bonding C(2)OH…O(6)

*The empty cell = no IR band/peak present at that region; Each IR value represent the mean value of the 4 scans

In contrast, Avicel exhibited the highest absorbance value at 3333 cm^{-1}, followed by NaOH and $NaClO_2$ pretreated SSB (Appendix Figure A3.2-B). Like Siroky et al. (2010) and Abidi et al. (2014), we assigned this band at approximately 3333 cm^{-1} to crystalline cellulose as it corresponds to the O(2)H…O(6) intramolecular hydrogen bond.

The untreated and lime pretreated SSB spectral profile was still similar to that seen on CC biomass. Poletto et al. (2012) claimed the peaks/bands are usually higher at this point if there are more OH groups in a sample. In addition, these OH groups could be associated with an increased number of hydrogen bonds formed in a sample. A mixture of intermolecular and intramolecular hydrogen bonds is considered to cause the broadening of the OH band in the IR spectra. This explanation by Poletto and co-workers could explain the difference of IR spectra of pretreated CC and SSB biomass 3333 cm^{-1}.

The IR spectral region ranging between 3000 and 2800 cm^{-1} is assigned to the CH_2 asymmetrical and symmetrical stretching vibrations (Abidi et al. 2014). The vibrations located at 2918 and 2850 cm^{-1} are suggested to originate from the non-cellulosic substances present in the primary cell-wall, for instance wax or pectin (Abidi et al. 2008, 2010; Siontorou et al. 2015). In the current study, the alkaline pretreated CC biomass and the untreated biomass exhibited one band at 2920 cm^{-1}, respectively, while Avicel exhibited a band at 2896 cm^{-1}. On the other hand, both untreated and $NaClO_2$ pretreated SSB biomass had bands at 2920 and 2850 cm^{-1}, while both lime and NaOH pretreated SSB had a band at 2919 cm^{-1}, respectively. Therefore, it was evident from the FTIR spectra that SSB samples were different to CC samples (Appendix Figures A-3.2 and Table 3.1). Interestingly, these bands were in the untreated SSB and CC biomass. The decrease or disappearance of these bands on the alkaline pretreated CC and SSB validate the removal of non-cellulosic substances (including lignin).

The IR spectral region ranging between 1800 and 1500 cm^{-1} has been reported to have four important peaks (Abidi et al. 2010; Siroky et al. 2010). The vibrations at 1738 cm^{-1} are attributed to C=O Stretching and the presence of this peak represented esterified uronic acids. The vibration at 1639 cm^{-1} is assigned to O-H bending of adsorbed water. The vibrations at 1620 and 1543 cm^{-1} are assigned to C=O (amide I) and amide II, respectively, which could originate from protein or pectic acid ester or amino acids, respectively. In the current study, we observed the water band in all SSB and CC samples at about 1640 cm^{-1} (Schwanninger et al. 2004; Abidi et al. 2010). A very small band was also seen at 1640 cm^{-1} in Avicel. In

addition, only untreated and $NaClO_2$ pretreated CC and SSB samples had the C=O valence vibration of uronic acid at about 1730 cm^{-1}. Poletto et al. (2012) stated that the bands at 1730 and 1600 cm^{-1} represent xylan from hemicelluloses and derived from molecular vibrations of the uronic acids. The untreated CC and SSB had an aromatic skeletal vibration of lignin band at about 1515 cm^{-1}. The 1515cm^{-1} band was also present in the $NaClO_2$ pretreated SSB, but absent in $NaClO_2$ pretreated CC. Our findings showed that the alkaline (NaOH and lime) pretreatment of CC and SSB was able to remove the acetyl and lignin functional groups. We suggest that alkaline pretreatments are better than oxidative pretreatments because they appear to remove both acetyl and lignin functional groups which oxidative pretreatment could not remove.

The peak region from 1500 to 1200 cm^{-1} is important for the identification of lignin, hemicellulose and cellulose (Adapa et al. 2011). Table 3.1 shows that untreated CC had bands at 1423, 1371 and 1243 cm^{-1} which are assigned to C-H bending, CH_2 scissoring and C-O stretching vibrations, respectively. The bands at 1423 and 1371 cm^{-1} represent cellulose, while the one at 1243 cm^{-1} represents lignin (Kim et al. 2003; Adapa et al. 2011; Abidi et al. 2014). Interestingly, lime and NaOH pretreated CC did not have this lignin band. The NaOH treated biomass had a similar profile to Avicel. These results suggest that lignin was completely removed or in very minimal amounts such that it could not be detected with FTIR. However, the lime pretreated samples had a broad band at 1416 cm-1 (Appendix Figure A-3.2-A). The FTIR results for lime pretreated biomass are in agreement with findings by Beukes & Pletschke, (2010) that lime solubilizes the lignin. We suggest that a broad peak could have formed due to the formation of the lignin-holocellulose complex. Lastly, $NaClO_2$ treated biomass had a similar profile to that of NaOH pretreated CC and Avicel in this region, except that the lignin at 1243 cm^{-1} was still present. Figure A3.2-B (Appendix) shows that lime, NaOH and $NaClO_2$ pretreated SSB had the same FTIR profiles as CC samples.

The last region of the FTIR spectra that was investigated was between 1200 and 650 cm^{-1}. Table 3.1 shows that this region is made up of holocellulose (Schwanninger et al. 2004). Pretreated CC and SSB biomass had similar FITR profiles in this region (Appendix Figure A-3.2).

3.4.3 Biomass crystallinity analysis

Crystallinity analysis of biomass can be quantified with FTIR. The crystallinity measure is based on the crystallinity FTIR indices. These indices are applied to cellulose-based polymers composed of either crystalline cellulose I or II or of mixtures of both components (Carrillo et al. 2004). The following spectral ratios: 1420/893 and 1375/2902 cm^{-1}, were used to determine the lateral order index (LOI), and total crystalline index (TCI) of lyocell and viscose-type fibres by Carrillo et al. (2004), while Poletto et al. (2012) used the following spectral ratios; 1430/898 and 1372/2900 cm^{-1}, for LOI and TCI in *Eucalyptus grandis* and *Pinus taeda*, respectively. The study by Poletto et al. (2012) showed that wood or agricultural based substrates can be used to calculate the LOI and TCI. Hence, the LOI and TCI of the Avicel (positive control) and pretreated CC and SSB were determined in this current study.

The TCI value (1.08) for Avicel was the highest compared to CC and SSB samples (Table 3.2). The TCI values of untreated, NaOH and $NaClO_2$ pretreated CC samples were 0.82, 0.82 and 0.91, respectively. The TCI values of untreated, NaOH and $NaClO_2$ pretreated SSB were 0.53, 0.84 and 0.74, respectively. The TCI values for both lime pretreated CC and SSB samples could not be determined due to the disappearance of the band at 1372 cm^{-1} (Appendix Figure A-3.2 and Table 3.2). In addition, LOI values of untreated, lime, NaOH, and $NaClO_2$ pretreated CC samples were 0.88, 2.39, 1.06 and 0.90, respectively. The LOI values of untreated, lime, NaOH, and $NaClO_2$ pretreated SSB were 0.9, 3.31, 0.91 and 0.79, respectively. It worth noting that the Avicel LOI value was the lower compared to that of both CC and SSB. Both CC and SSB samples had lower TCI values compared to LOI values. In fact, the untreated-CC biomass had a lower LOI value compared to the Avicel which shows that amorphous cellulose content in the CC samples increased. Our findings for CC are in agreement with other studies which reported that LOI increases with a decreasing degree of crystallinity (TCI) (Carrillo et al. 2004; Poletto et al. 2012). The results of the SSB samples were comparable to those of CC samples (Table 3.2).

The TCI values for lime pretreated biomass could not be determined due to the disappearance of the 1372 cm^{-1} band. However, the hydrogen bond intensity (HBI) of cellulose or biomass is closely related to the crystal system and the degree of intermolecular regularity. In fact, this parameter should also be used for indicating the crystallinity of the biomass (Oh et al. 2005; Poletto et al. 2012).

We used the HBI values to to support the TCI and LOI data, especially lime pretreated SSB and CC. The HBI values for lime pretreated CC and SSB were 1.61 and 2.2, respectively. The HBI value for lime pretreated CC and SSB were about 32.83% (or 0.78) and 33.33% (or 1.10) smaller than the LOI values, respectively. (Table 3.2). Furthermore, the untreated CC also had about 12.02% (0.22) higher HBI value compared to that of the lime pretreated CC. The $NaClO_2$ pretreated CC was an exception. The lime pretreated SSB was an exception because it had the highest HBI value. The untreated SSB had about 17.16% and 2.06% higher HBI values compared to the NaOH and $NaClO_2$ pretreated biomass, respectively. We conclude that chemical (alkaline and oxidative) pretreatment changed the structure of the SSB and CC biomass. Thus, lime and NaOH pretreated SSB or CC biomass were more amorphous compared to both untreated and $NaClO_2$ pretreated SSB or CC biomass. These data from Table 3.2 suggest that the amorphous nature of the alkaline pretreated SSB or CC biomass could be easly accessed by biomass degrading enzymes compared to those pretreated with $NaClO_2$, which are less amorphous.

Table 3.2: Substrate crystallinity ratios and hydrogen bond intensities of cellulose quantified by FTIR

Name of the sample	IR crystallinity ratio		Hydrogen bond intensity (HBI) of cellulose
	Total crystallinity index (TCI)	Latera order index (LOI)	
Avicel	1.08	0.10	1.77
CC-control	0.82	0.88	1.83
CC-Lime	*	2.39	1.61
CC-NaOH	0.82	1.06	1.96
CC-$NaClO_2$	0.91	0.90	1.80
SSB-control	0.53	0.91	1.94
SSB-Lime	*	3.31	2.20
SSB-NaOH	0.84	0.91	1.61
SSB-$NaClO_2$	0.74	0.79	1.90

*The crystallinity index could not be determined due to missing bands/peaks, see Appendices; Figure A3.2. The TCI, LOI and HBI ratios (values) represents IR mean values of 4 FTIR scans.

3.4.4 Biomass composition

Biomass compositional analysis of lime, NaOH and $NaClO_2$ pretreated SSB and CC was performed (according to NREL methods) to determine the amounts of lignin, cellulose and hemicellulose after pretreatments (Figure 3.2). The untreated/control SSB lignin, cellulose and hemicellulose content was about 29.33%, 17.75% and 16.3%, respectively. The untreated

CC lignin, cellulose and hemicellulose content was about 22.50%, 23.60% and 33.30%, respectively. As expected, alkaline pretreatments reduced the amount of the lignin and increased the amounts of the cellulose content (Figure 3.2).

The lime pretreated SSB lignin content was significately reduced (p-value < 0.011) compared to that of the control. In addition, the hemicellulose content was reduced to 11.50% and the cellulose content signicantly increased (p-value < 0.002) from 17.75% to 19.60% after SSB was pretreated with lime (Figure 3.2). In the case of lime pretreated CC, lignin content was significantly reduced by 6.00%, while cellulose content significantly increased by 4.00% compared to the control (both lignin and cellulose p-value < 0.01 for lime pretreated CC). The NaOH pretreated SSB had about an 18.00% reduction in the amount of lignin and a 5.00% increase in cellulose content compared to the control. The NaOH pretreated SSB hemicellulose content was 13.10%. The NaOH pretreated CC lignin content was reduced by about 11.00%. Interestingly, when compared to controls the cellulose and hemicellulose contents were increased by 10.00% and 2.00%, respectively. These results demonstrate that NaOH pretreatment was the most effective pretreatment compared to lime and $NaClO_2$ pretreatment because it significantly removed higher amounts of lignin (p-value <0.01) in both SSB and CC biomass samples. Subsequent to SSB and CC biomass pretreatment with NaOH, the cellulose content (p-value < 0.01) and hemicellulose content (p-value < 0.05) of the biomass significantly increased. $NaClO_2$ pretreatment reduced lignin content of SSB by 6.00%, while cellulose content increased by 4.00% and hemicellulose content did not change significantly compared to the control (Figure 3.2). The $NaClO_2$ pretreated CC only had a 2.00% and 6.00% reduction in lignin and hemicellulose content, respectively. The CC biomass cellulose content increased by about 5.00% after $NaClO_2$ pretreatment.

To validate these results, acid soluble lignin, acid insoluble lignin, arabinose, glucose and xylose were quantified (Table 3.3). Both the xylose and the arabinose represent the hemicellulose, while glucose represents the cellulose content of the biomass. After pretreating the SSB biomass with NaOH and $NaClO_2$, the amounts of acid insoluble lignin decreased significantly. In contrast, lime pretreated samples also had higher amounts of insoluble lignin (Table 3.3). Notably, $NaClO_2$ pretreated samples had high acid soluble amounts compared to untreated, lime and NaOH pretreated samples.

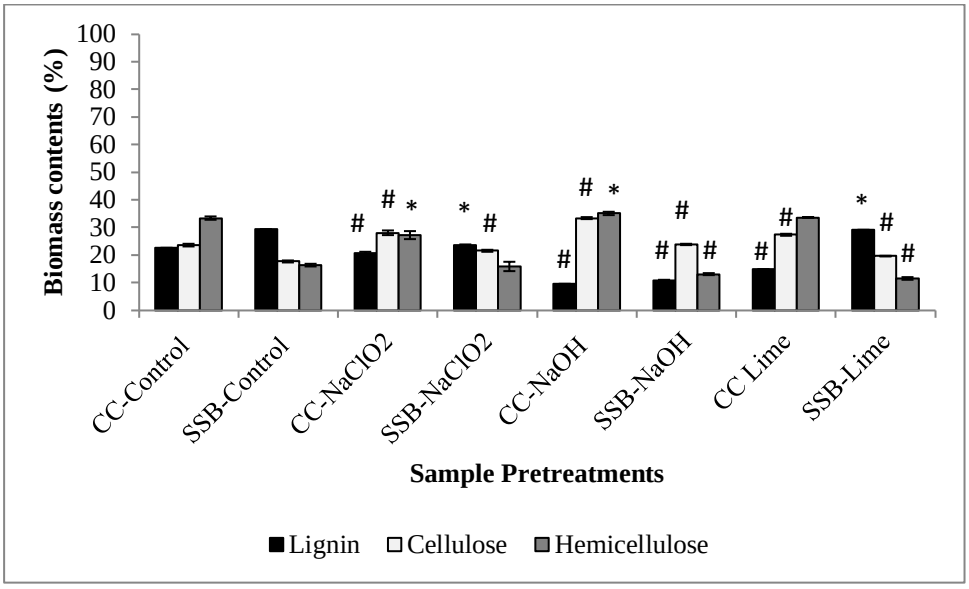


Figure 3.2. The analysis of lignin, cellulose and hemicellulose in untreated/control, lime, sodium chloride ($NaClO_2$) and sodium hydroxide (NaOH) pretreated corncob (CC) and sweet sorghum bagasse (SSB) samples. * represents p-value < 0.05 and # represents p-value < 0.01. The experiments were performed in triplicate and datapoints represent the means ± SD.

The glucose, xylose and arabinose contents increased significantly (p-vaalue < 0.01) after the CC and SSB biomass was pretreated with lime, NaOH and $NaClO_2$ (Table 3.3). The increasing amounts of glucose, xylose and arabinose confirmed that the percentage content of cellulose and hemicellulose were increasing in both SSB and CC biomass after pretreatment. The decreasing amounts of insoluble lignin confirm the removal of lignin after biomass was pretreated.

These results show that NaOH pretreatment was the best with regards to removal of lignin from the CC and SSB biomass. Beukes & Pletschke (2010) found that lime pre-treated sugar cane bagasse had about 10.00% less lignin and a higher content of sugar relative to that of the untreated samples. $NaClO_2$ pretreatment was also reported to remove lignin from sugar cane bagasse and increased the relative amount of the cellulose (Siqueira et al. 2013). In addition, NaOH pretreated corn stover and SSB had less lignin and increased amounts of cellulose and hemicellulose compared to lime, weak acid and ammonia pretreatments (Chen et al. 2009; Wu et al. 2011). These findings are in agreement with our findings that NaOH is a better

pretreatment method compared to lime and $NaClO_2$, because it removed more lignin and increased the amount of cellulose and hemicellulose.

Table 3.3: Simple sugars and lignin content of alkaline and oxidative pretreated corncob and sweet sorghum bagasse

Sample names	Acid insoluble lignin (%)	Soluble lignin (%)	Glucose (%)	Xylose (%)	Arabinose (%)
SSB-control	20.78 ± 0.07	8.3 ± 0.07	19.44 ± 0.42	14.96 ± 0.28	0 ± 0.0
SSB-Lime	24.30 ± 0.07#	5.13 ± 0.01#	21.62 ± 0.18#	10.76 ± 0.28#	0 ± 0.0
SSB-NaOH	11.56 ± 0.04#	5.09 ± 0.07#	39.64 ± 1.38#	18.62 ± 0.04#	3.62 ± 0.07#
SSB-NaClO2	13.57 ± 0.07#	9.95 ± 0.32*	23.95 ± 0.42#	13.70 ± 1.29	3.19 ± 0.24#
CC-Control	15.15 ± 0.04	7.55 ± 0.25	25.96 ± 0.54	32.61 ± 0.51	0 ± 0.0
CC-Lime	10.56 ± 0.13#	4.33 ± 0.06#	30.40 ± 0.40#	33.91 ± 0.14*	4.21 ± 0.06#
CC-NaOH	5.68 ± 0.06#	3.86 ± 0.06#	36.98 ± 0.46#	35.28 ± 0.65#	4.53 ± 0.07#
CC-NaClO2	8.97 ± 0.05#	11.77 ± 0.43#	30.52 ± 0.64#	26.23 ± 1.37#	4.14 ± 0.05#

The percentage value represents the mean ± SD, n=3

* represents p-value < 0.05 and # represents p-value < 0.01

3.5. Conclusion

This chapter described and discussed the changes that occurred on the CC and SSB biomass samples subsequent to alkaline (NaOH and lime) and oxidative ($NaClO_2$) pretreatment. Both methods are well studied; however, to the best of our knowledge, there are no studies that compared these methods with regard to lignin removal from the agricultural biomass substrates we intended to use. The results of our study demonstrated that both alkaline and oxidative pretreatments removed lignin from the biomass. Generally the removal of the lignin resulted in significantly increased amount of hemicellulose and cellulose in both CC and SSB biomass. The NaOH pretreatment removed significantly higher amounts of lignin compared to both the lime and $NaClO_2$ pretreatments. The FTIR results showed that $NaClO_2$ did not remove all the lignin from the biomass. SEM showed that the morphology of all the pretreated CC and SSB biomass samples had changed. The surfaces of pretreated CC and SSB samples had changed and pores appeared as lignin was removed. Thus, we conclude that

both treatments removed lignin. However, NaOH was selected as the best pretreatment method, because it removed most of lignin compared to lime and $NaClO_2$.

Chapter 4: Characterization of Glycosyl Hydrolases Derived from a Termite (Trinervitermes trinervoides) Metagenome

4.1. Introduction

The production of second-generation biofuels from agricultural biomass remains inefficient. A better understanding of lignocellulose digestion by termite-derived enzymes may assist to uncover and eventually overcome challenges in the conversion of lignified plant cell walls into soluble sugars (Brune 2014). The termite gut microflora act as an important reservoir for novel glycosyl hydrolases exhibiting efficient hydrolysis of lignocellulosic biomass (Ni and Tokuda 2013; Brune 2014). A metagenomic approach has been used previously to mine various glycosyl hydrolase (GH) and feruloyl esterase enzyme genes from the hindgut bacterial symbionts of *Trinervitermes trinervoides* termite species (Rashamuse et al. 2013, 2016).

As noted by Brune (2014), much of the research performed to date on termite derived glycoside hydrolase enzyme systems has focused on the heterologous expression and biochemical characterisation of endogenous (host) derived glucanases and glucosidases (Tokuda et al. 1997; Watanabe et al. 1998; Uchima et al. 2012; Feng et al. 2015). In contrast, other classes of glycoside hydrolases (exo-glucanases, xyloglucanases and xylanases), particularly those from microbial symbionts associated with the termite hindgut, have been poorly studied. Hence, the focus of our current study seeks to understand and comprehensively characterise selected GH enzymes (GH5D, GH5E, GH5H and GH45) from a termite hindgut metagenome produced by Rashamuse et al. (2016). Furthermore, Rashamuse et al. (2016) patialy characterised GH5D, GH5E, GH5H and GH45. We argue that partial charecterisation of these enzymes is not sufficient and could be misleading with regard to the biochemical classification (based on activities) and application of these enzymes. This necessitated comprehensive characterisation of these enzymes.

Multiple genes mined through metagenomic techniques have successfully provided initial tools to further our understanding with regards to termite enzyme research (Ni and Tokuda 2013). However, our knowledge is limited with respect to the potential for industrial applications of lignocellulolytic enzymes derived from a termite gut metagenome. Thus, a comprehensive characterisation of these glycosyl hydrolases could shed light on their mechanisms of action/hydrolysis and therefore help in unlocking their potential for application in the biofuel, biorefinery and food industries.

4.2. Aims and Objectives

4.2.1. Aims

To comprehensively characterize the purified GH5D, GH5E, GH5H and GH45 enzymes derived from a termite metagenome.

4.2.2. Objectives

1. To determine the biochemical properties of the purified GH enzymes.
2. To determine substrate specificities of these enzymes.
3. To determine the kinetic parameters of the GH enzymes.
4. To investigate and elucidate the hydrolysis patterns of the GH enzymes.
5. To investigate whether GH5H is a multifunctional enzyme with two catalytic centers.

4.3. Materials and Methods

4.3.1. Materials

The gene sequences of the four glycosyl hydrolases (GH5D, GH5E, GH5H, and GH45) from a termite gut metagenomic library used in the current study were previously reported with the following GenBank accession numbers *GH5D*: KT694321, *GH5E*: KT694322, *GH5H*: KT694325 and *GH45*: KT694340 (Rashamuse et al. 2016). The constructs were provided by CSIR Biosciences under the material transfer agreement (MTA).

The model polymeric substrates carboxymethyl-cellulose (CMC) and the *p*-nitrophenyl (*p*NP) based substrates *p*NP-α-L-arabinofuranoside, *p*NP-β-D-glucopyranoside, *p*NP-β-D-mannopyranoside, *p*NP-β-D-cellobioside, *p*NP-β-D-glucopyranoside and *p*NP-β-D-xylopyranoside were purchased from Sigma-Aldrich, South Africa. Beechwood xylan, soluble wheat flour xylan, insoluble wheat arabino xylan and xyloglucan, as well as the xylo-, cello- and xyloglucan oligomers were purchased from Megazyme, Ireland. All other chemicals used in this study were of analytical grade and were purchased from Sigma-Aldrich unless stated otherwise.

4.3.2. Determination of protein purity and concentration

The protein concentrations of the four glycosyl hydrolases were determined using the Bradford method (Bradford 1976) using bovine serum albumin as the standard (Appendix:

Figure A4.1A). The purity of the protein samples provided by Rashamuse et al. (2016) was analysed and confirmed by SDS-PAGE.

4.3.3. Standard assays for reducing sugars and *p*-nitrophenol

The cellulase and xylanase activities were determined using a modified DNS method (Miller 1959). The model substrates, including CMC, xyloglucan, wheat flour xylan and beechwood xylan were dissolved in 50 mM phosphate buffer (pH 7.0) and their final concentrations were maintained at 2% (w/v). The substrate [2%(w/v)] and enzyme (6 µg/mg) mixture was incubated at 50°C for 30 min. After the incubation, the product was mixed with DNS reagent at a ratio of 1: 2, then the mixture was heated on a digital dry heating bath (Labnet International, Inc) at 100°C for 5 min and allowed to cool on ice. The absorbance readings were taken at 540 nm with a Powerwave-X microplate reader (Bio-Tek Instruments) using KC Junior software. The total reducing sugars were quantified using glucose and xylose as suitable standards (Appendix: Figure A4.1 B and C).

The cellulase and xylanase activities were also tested using various *p*NP substrates (*p*NP-α-L-arabinofuranoside, *p*NP-β-D-glucopyranoside, *p*NP-β-D-mannopyranoside, *p*NP-β-D-cellobioside, *p*NP-β-D-glucopyranoside and *p*NP-β-D-xylopyranoside). The reaction mixture consisted of 5 mM final concentration of the substrate in a 50 mM phosphate buffer, pH 7.0, and the reaction was intiated by adding 50 µl of suitably diluted enzyme concentration. The mixture was incubated for 10 min at 50°C and the reaction was terminated with 1 M $NaCO_3$. The absorbance reading was taken at 405 nm with a Powerwave-X microplate reader (Bio-Tek Instruments) using KC Junior software. The enzyme activity was quantified using a standard curve that was generated with different concentrations of *p*-nitrophenol.

4.3.4. Biochemical properties of the recombinant enzymes

To determine the pH optima of the enzyme reactions, 6µg/mg enzyme samples were incubated in universal buffers between pH 3.0 and 9.0 (50 mM Tris, 50 mM boric acid, 33 mM citric acid, and 50 mM Na_2PO_4 adjusted with either HCl or NaOH to the required pH (Britton and Robinson 1931)), containing 1% (w/v) CMC for GH5E, GH5H and GH45 or xyloglucan for GH5H and GH45 or beechwood xylan for GH5H or *p*NP-C for GH5D, and assayed according to the standard DNS reducing sugar method. However, GH5D was assayed according to the *p*NP standard assay. Temperature optima were determined using the standard reducing sugar assay and the standard *p*NP assay for GH5D at set temperatures

ranging between 20 and 70°C using a dry heating block (Labnet International, Inc). The thermostability was inverstigated at 50°C or 37°C for a 72 h period.

4.3.5. Substrate specificity and kinetic parameters

The substrate specificities of the purified GH5D, GH5E, GH5H, and GH45 enzymes were determined by measuring the activities under standard assay conditions with beechwood xylan, oat-spelt xylan, wheat arabinoxylan, avicel, xyloglucan, locust been gum and CMC. Specific activities of these enzymes on *p*NP based substrate were determined with *p*NP-α-L-arabinofuranoside, *p*NP-β-D-glucopyranoside, *p*NP-β-D-mannopyranoside, *p*NP-β-D-cellobioside, *p*NP-β-D-glucopyranoside and *p*NP-β-D-xylopyranoside under standard conditions. For the determination of kinetic parameters, purified enzymes [6 μg/mg biomass] were incubated with either the polymer substrates (CMC, xyloglucan, and beechwood xylan) or *p*NP substrate (*p*NP-β-D-cellobiose) at various concentrations under optimal conditions. The experimental data (initial velocities at various substrate concentrations) were used to determine the kinetic constants (K_M and V_{max}) of the respective purified enzymes, using nonlinear regression analysis with Kaleidagraph version 4.5.

4.3.6. Substrate competition assay

The maximum velocities (V_{max}) and Michaelis constant (K_M) for xylanase, CMCase and xyloglucanase activities of GH5H were estimated using Kaleidagraph version 4.5. Beechwood xylan, CMC and xyloglucan were used at a concentration range of 1 to 16 mg/ml. The reactions were carried out using 50 mM citrate buffer at pH 5.5 and 50°C. The substrate competition assays were carried out by incubating GH5H [6 μg/mg biomass] and a mixture of two substrates (CMC and xylan or CMC and xyloglucan, and or xyloglucan and xylan) with different concentrations that added up to 20 mg/ml total substrate concentration (Table 4.3). The kinetic studies and calculations were carried out according to equations described by Chen et al. (2006). The substrate competition assays suggested that 1), if the observed velocity values are much closer to the theoretical velocity values for a single catalytic center then the enzyme has a single catalytic center or 2), if the observed velocity values are close to the theoretical velocity values for the double catalytic centers, then the enzyme has two catalytic centers.

4.3.7. Thin-layer chromatography

Chain length specificity was determined by enzyme mode of hydrolysis of CMC, beechwood xylan, xyloglucan, and xylo- or cello- or xyloglucan-oligosaccharides (1-6 degree of polymerization (DP)). The reaction mixtures were incubated in a dry heating bath at 50°C for 30 min. Samples were then centrifuged at 13 000 rpm for 5 minutes. The supernatants were analyzed using TLC (Silica Gel 60G F254 HPTLC plates from Merck, Darmstadt, Germany). Plates were developed twice with n-butanol: acetic acid: water (2:1:1, v/v/v). To detect carbohydrates, plates were briefly submerged in methanol containing 5% (v/v) sulfuric acid and 0.3% α-naphthol. Plates were then air dried and then heated at 120°C for 10 min.

4.3.8. Enzyme deactivation and inhibition by metal-ions

For deactivation assays, GH5E, GH5H, and GH45 enzymes [6 μg/mg biomass] were mixed with 2 mM metal-ions [Fe^{2+}, Cu^{2+}, Mn^{2+}, Ca^{2+}, Mg^{2+}, Zn^{2+}, Co^{2+} and EDTA (a chelating agent)] on ice. The mixtures were stored at 4°C for 14 hours or overnight. The enzyme-metal-ion mixture was then added into the substrate and the reaction was performed under standard assay conditions. For inhibition assays, the metal-ions (4 mM) were mixed with substrates [1% (w/v)], followed by the immediate addition of enzymes [6 μg/mg biomass] to commence the reaction under standard assay conditions.-

4.3.9. Enzyme inhibition by oligosaccharides

For oligosaccharide inhibition assays, GH5E, GH5H, and GH45 enzymes were mixed with xylo-, cello- or xyloglucan oligosaccharides and the substrates (Azo-CMC, Azo-beechwood and Azo-xyloglucan) to determine the effects of oligosaccharides on enzyme activity. The enzyme [6 μg/mg biomass] was added to the reaction mixture containing 2.5 mg/ml of oligosaccharides and 2% (w/v) of Azo substrate mixed in a citrate buffer (pH 5.5) and the reaction was performed at 50°C for 30 min. A ratio 2:1 of ethanol was added to the reaction mixture. The mixture was then vortexed vigorously and centrifuged at 5000 *g* for 10 min. The blue colour formed after enzyme hydrolysis was directly proportional to the enzyme activity. The absorbance was measured at 590 nm with a Powerwave-X microplate reader (Bio-Tek Instruments) using KC Junior software. Remazol Brilliant Blue R was used as a suitable standard.

4.3.9.1. Data Analysis

One way analysis of variance (ANOVA) was used to analyse and elucidate significant increases or decreases in GH5E, GH5H and GH45 activities due to metal ions. All pairwise comparison procedures were conducted using the Data analysis feature in Microsoft Excel.

4.4. Results and Discussion

All the termite glycosyl hydrolases (GH5D, GH5E, GH5H and GH45) used in the current study were derived from the hind gut microflora metagenome, heterologously expressed in *E. coli* and purified using histidine tag specific chromatography (Rashamuse et al. 2016). The purity of these enzymes was also confirmed by SDS-PAGE gel (Appendix: Figure A4.2). The initial concentration of GH5D, GH5E, GH5H, and GH45 was 1.24, 0.12, 0.10, and 0.18 mg/ml, respectively.

4.4.1. Substrate specificity assays of the recombinant enzymes

The enzyme substrate specificity was assayed using different polymeric and *p*NP-substrates (Table 4.1). GH5D was the only enzyme that exhibited activity on *p*NP-cellobioside; however, it also showed no activity on wheat arabinoxylan and CMC. GH5E displayed the highest activity on CMC followed by wheat flour xylan, beechwood xylan and xyloglucan (Table 4.1).

Rashamuse et al. (2016) reported that GH5D and GH5E are an exo-glucanase and endo-glucanase, respectively. Our results confirmed that GH5D is indeed an exo-acting enzyme, as it exhibited high activity on *p*NP-cellobioside, but displayed no activity on CMC, which is a substrate that is generally hydrolyzed by endo acting cellulases. GH5E demonstrated activity of 17.189 U/mg and 13.471 U/mg on CMC and wheat flour xylan, respectively. It also displayed a weak activity on both xyloglucan and beechwood xylan suggesting that it is an endoglucanase (Table 4.1). Valls et al. (2016) recently discovered a novel arabinoxylan-specific arabinofuranohydrolase from *Bacillus* sp. BP-7. This enzyme preferred Rye arabinoxylan with a specific activity of 27.70 U/mg, followed by wheat arabinoxylan with a specific activity of 14.70 U/mg. Interestingly, the arabinoxylan-specific arabinofuranohydrolase showed no activity on birchwood or beechwood xylan. Our results demonstrated that an GH5E (endo-glucanase) also showed activity on xylan substrates. Thus, we propose that the GH5E enzyme (endoglucanase deposited at Genbank with accession number KT694322) should be renamed an endo-glucanase/xylanase.

In addition, GH5H and GH45 showed the highest activity on xyloglucan. GH5H and GH45 displayed about 84.98 and 12.86 U/mg activities on xyloglucan, respectively. Our findings differ from the earlier report by Rashamuse et al. (2016) that GH5H and GH45 (Genebank accession numbers KT694325 and KT694340, respectively) are xylanase and endoglucanase/ xylanase enzymes, respectively. We suggest that these enzymes should rather be classified as

xyloglucanases. However, we will refer to them throughout this study as multifunctional enzymes (GH5H) and tri-functional enzymes (GH45), respectively. We report here that GH5H is a multifunctional enzyme that hydrolyses β-1,4-linkages of the substrates, CMC, beech wood xylan, wheat arabinoxylan and xyloglucan (Table 4.1). GH45 hydrolyses β-1,4-linkages of xyloglucan, CMC and beechwood xylan, thus GH45 was referred to as a trifunctional enzyme (kinetic studies supports these observations, see section 4.4.2). Such enzymes have been reported in the literature before. For instance, Morrison et al. (2016) reported a family 5 endo-glucanase that showed activity on xylan, in addition to weak β-glucosidase activity. This enzyme was shown to be a trifunctional enzyme like GH45.

Table 4.1. Substrate specificity assays on different model substrates. The reactions were conducted in triplicate, the values represent means ± SD.

Model Substrates (2%: w/v)		Enzymes activities (U/mg protein)		
	GH5D: Exo-glucanase	GH5E: Endo-glucanase	GH5H: multifunctional enzyme	GH45: Endo-glucanase/xylanase
CMC	N/A	17.19 ± 0.04	28.58 ± 0.02	4.75 ± 0.05
Beechwood xylan	N/A	7.48 ± 0.01	52.02 ± 0.07	4.55 ± 0.01
Wheat flour xylan	N/A	13.47 ± 0.16	67.37 ± 0.1	Not tested on this substrate
Xyloglucan	N/A	6.83 ± 0.01	84.97 ± 0.01	12.86 ± 0.19
Locus bean gum	N/A	N/A	N/A	N/A
pNP-G	N/A	N/A	N/A	N/A
pNP-X	N/A	N/A	N/A	N/A
pNP-C	6.89 ± 0.084	N/A	N/A	N/A
pNP-M	N/A	N/A	N/A	N/A

Values=means ± SD, n=3; N/A means there was no activity; U represents activity in µmol/min; *p*NP-G, -X, -C, M and Ac represent *p*-nitrophenyl-glucopyranoside, -cellobioside, -xylopyranoside, –mannopyranoside and acetate, respectively.

Our results support the argument that some of the GH5 enzymes are multifunctional enzymes, and have not been explored in great detail. As a result, they have not been characterized extensively - leading to ambiguity in their classification (Aspeborg et al. 2012). Finally, we suggest that the multifunctional enzyme originally classified by Rashamuse et al. (2016) as endo-1,4-D-xylanase (GH5H: accession number, KT694325), as well as the trifunctional enzyme (GH45; accession number, KT694340), be reclassified as xylo-glucanases with cellulase and xylanase activities.

4.4.2. Biochemical properties of the recombinant enzymes

The biochemical properties of GH5D, GH5E, GH5H and GH45 were assayed using all the substrates (except for wheat arabinoxylan) that were hydrolyzed by these enzymes. The temperature optima for GH5D and GH5E were 50 and 40°C, respectively (Table 4.2). For GH5H, the temperature optima on CMC, beech-wood xylan and xyloglucan were 40, 60 and 40 to 60°C, respectively. GH45 demonstrated temperature optima at 40°C for both CMC and xyloglucan.

The pH optima of GH5D and GH5E were pH 6.5 and 5.0, respectively. GH5H demonstrated broad range pH optima of pH 4.0, 5.0 to 6.5 and 6.0 to 8.0 on CMC, beech-wood xylan and xyloglucan, respectively. Feng et al. (2015) also found that the two termite β-glucosidase enzymes also had pH optima at 5.0 and temperature optima at 40°C. The endo-β-1,4-glucanase from the termite had a pH optimum of 5.8 and the enzyme retained more than 60% activity at pH 5.0 to 9.5 (Tokuda et al. 1997). These reports are also in line with our findings, except for GH5H, with a low pH optimum of 4.0 on CMC. Given that GH5D and GH5H were derived from a termite gut metagenome, we did not expect them to have temperature optima of 60°C. However, Tokuda et al. (1997) and Ni and Tokuda (2013) reported that recombinant endoglucanases from different termite species also had different temperature optima ranging from 37 to 70°C.

Rashamuse et al. (2016) and Ni and Tokuda (2013) showed that xylanases and endoglucanases had a pH optimal range of pH 5 to 7.5. In our study we also demonstrated that GH enzymes used displayed pH optima ranges between 4.0 and 8.5. We suggest that the different pH and temperature optima could be attributed to the different micro-habitats containing micro-organisms within the termite gut (Santana et al. 2015), given that our enzymes are derived from a termite gut metagenome. Tokuda et al. (1997) purified an endo-

glucanase that was only stable for 30 min at 55°C. All the enzymes used in the current study were stable for 30 min at 50°C, with the activity being lost completely after 5 h (Table 4.2). Interestingly, all the enzymes in the current study were stable at 37°C for over 72 h and displayed more that 80% residual activity.

The kinetic constants for GH5E, GH5H and GH45 are shown in Table 4.2. Kinetic constants for GH5D were not determined because its did not display the Mechelis-Menten profile. We suggested that GH5D did not display Mechelis-Menten profile due to product inhibition. The GH45 or GH5H k_{cat}/K_M and k_{cat} values from Table 4.2 showed that both enzymes displayed different affinities on CMC, beechwood xylan and xyloglucan. GH5H K_M values were 95.03, 4.98 and 2.10 mg/ml for CMC, beechwood xylan and xyloglucan, respectively. GH45 displayed a K_M of about 5.515 and 105.98 mg/ml on CMC and xyloglucan.

Table 4.2. Biochemical properties and kinetic constants of the novel termite enzymes

Biochemical properties of the enzymes					Kinetic constants				
Enzymes	Substrate	Temperature (°C)	pH	Thermostability*	K_M (mg/ml)	*Vmax* (U/mg protein)	R^2 value	k_{cat} (min^{-1})	k_{cat}/K_M (min^{-1}. mg/ml)
GH5D	pNP-C	50.00	6.50	50 and 37 °C for 72 h	N/A	N/A	N/A	N/A	N/A
GH5E	CMC	40.00	5.00	50 and 37 °C for 72h	11.13	19.84	0.99	862.69	77.48
GH5H	CMC	40.00	4.00	50 and 37 °C for 72	95.03	144.45	0.98	9028.12	94.99
	Beechwood xylan	60.00	5.00-6.50	50 and 37 °C for 72h	4.98	120.36		7522.50	1510.54
	Xyloglucan	40.00- 60.00	6.00-8.50	50 and 37 °C for 72h	2.10	180.32		11270.00	5366.67
GH45	CMC	40.00	6.50	50 and 37 °C for 72h	5.52	8.51	0.95	386.86	70.15
	Xyloglucan	40.00	6.50	50 and 37 °C for 72 h	105.98	44.62	0.95	2028.31	19.32

* Thermostability was performed at 50 and 37°C for 72 h. However, enzymes retained more that 80% relative activity at 37°C. U represents activity in µmol/min; N/A means activity was not tested.

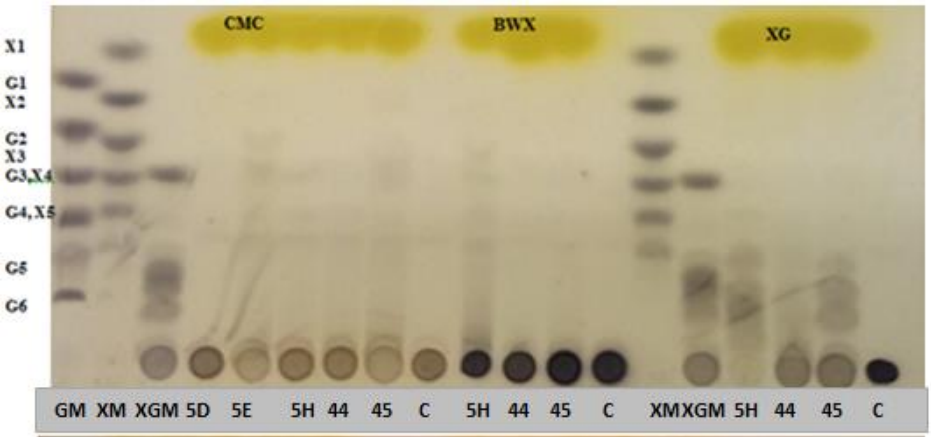
Even though some K_M values for GH5H and GH45 were high, the k_{cat} values were also high. These observations showed that GH5H and GH45 were able to hydrolyze CMC and xyloglucan, respectively. In addition, the k_{cat} and k_{cat}/K_M results showed that GH5H was also able to hydrolyze beechwood xylan and xyloglucan. Our results are in line with other studies that also show that some of enzymes' K_M values were high, but the high catalytic efficiency (k_{cat}) and k_{cat}/K_M values suggest that enzyme binds on to the substrate and forms product rapidly (Bai et al. 2013; Feng et al. 2015). Previously we reported that the termite, *Trinervitermes trinervoides*, harbours an enzyme (GH5H) that hydrolyse both xylan and xyloglucan (Bhattacharya et al. 2016). In the current study, we report for the first time that GH5H catalyses the hydrolysis of xyloglucan, CMC, xylan, indicating that these enzymes are true multifunctional enzymes.

4.4.3. Thin layer chromatography: chain-length specificity

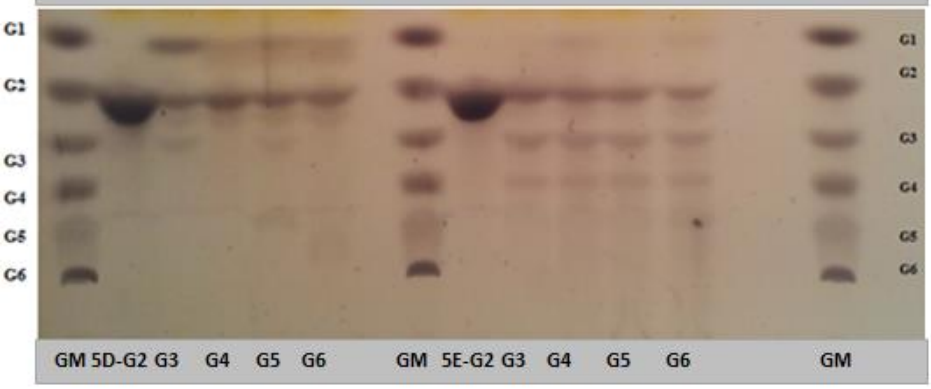
Thin layer chromatography (TLC) showed that GH5E, GH45, GH5H and GH5D preferred different chain-length specificities during biomass hydrolysis. The results showed that GH5D does not hydrolyse CMC (Figure 4.1A), but can catalyze cello-oligosaccharides (Figure 4.1B). GH5D hydrolysed cellohexaose and cellotetraose to cellobiose. It also hydrolysed cellotriose and cellopentaose to produce both cellobiose and glucose. The results show that the GH5D only acts on oligosaccharides with a DP longer than 2 and cleaves the oligomer into cellobiose only or both cellobiose and glucose (Figure 4.1B). Non-cellulosomal family 48 cellobiohydrolases from *Clostridium phytofermentans* ISDg (CpCel48) had high activity on Avicel and regenerated amorphous cellulose (RAC), but showed very low activity on CMC (Zhang et al. 2010).

The CpCel48 did not have any activity on *p*NPC, *p*NP-β-D-lactopyranoside (*p*NPL), and *p*-methylumbelliferyl-β-D-cellobioside (*p*MUC). Interestingly, CpCl48 showed no activity on cellotriose, but catalyzed cellotetraose to cellobiose, cellopentaose to cellobiose and triose, while it catalyzed cellohexaose to cellobiose, -triose and -tetraose. These results showed that this enzyme requires a polymer with a minimum DP of 4. In contrast, GH5D requires a cello-oligomer with a minimum DP of 3. Kern et al. (2013) reported a cellobiohydrolase (LqCel7B) from a marine wood borer, *Limnoria quadripunctata*, that exhibited very low activity on *p*-NPC and CMC, but which hydrolyzed both Avicel and RAC efficiently. LqCel7B and GH5D hydrolyzed cello-oligomers in a similar manner, except that GH5D did not hydrolyze cellobiose.

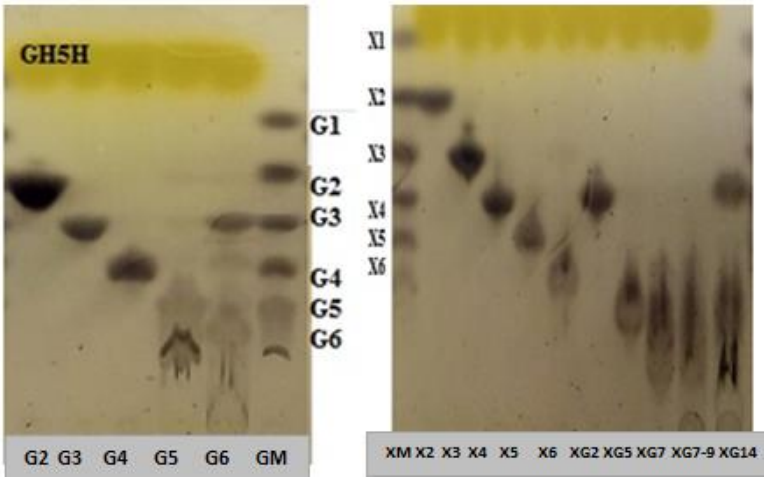
A


CMC
BWX
XG
X1
G1
X2
G2
X3
G3,X4
G4,X5
G5
G6
GM XM XGM 5D 5E 5H 44 45 C 5H 44 45 C XM XGM 5H 44 45 C

B


G1
G2
G3
G4
G5
G6
GM 5D-G2 G3 G4 G5 G6 GM 5E-G2 G3 G4 G5 G6 GM

C


GH5H
G1
G2
G3
G4
G5
G6
G2 G3 G4 G5 G6 GM
X1
X2
X3
X4
X5
X6
XM X2 X3 X4 X5 X6 XG2 XG5 XG7 XG7-9 XG14

D

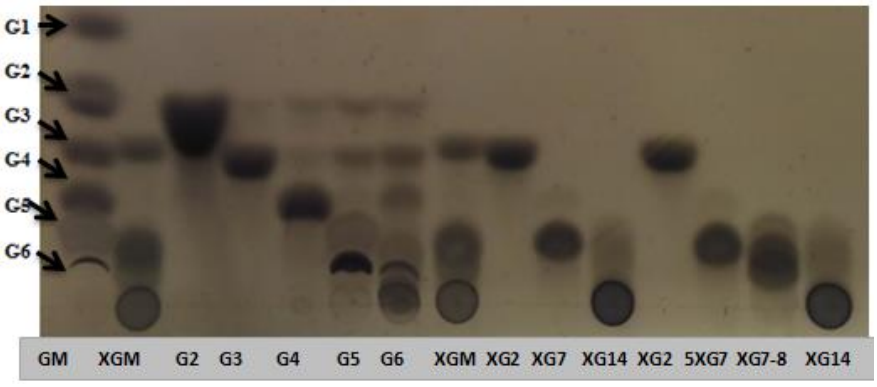

G1
G2
G3
G4
G5
G6
GM XGM G2 G3 G4 G5 G6 XGM XG2 XG7 XG14 XG2 5XG7 XG7-8 XG14

Figure 4.1. (A) Thin layer chromatography plate showing hydrolysis products of GH5D, GH5E, GH5H, GH44 and GH45 on carboxymethylcellulose (CMC), beechwood xylan (BWX) and xyloglucan (XG). The first three samples represent the standard mixtures; GM is the cello-oligosaccharides mixture (glucose to cellohexaose), XM is the xylo-oligosaccharides mixture (xylose to xylohexaose) and xyloglucan-oligosaccharides mixture (DP of 2,7,8,9 and 14). The controls for CMC, BWX and XG are indicated with the letter (C) on the plate. The cell-oligosaccharides chain length specificities of GH5D and GH5E during hydrolysis (B). (C) Indicates the chain length specificity of GH5H during cello-, xylo- and xyloglucan- oligosaccharides hydrolysis. XG represents the xyloglucan oligosaccharides with a degree of polymerization of 2, 7, a mixture of 7 to 9, and 14 and XGM represents a mixture of these xyloglucan oligosaccharides. The cello- and xyloglucan-oligosaccharides chain length specificity of GH45 during hydrolysis (D).

In contrast, GH5E hydrolyzed the CMC to produce cello-tetraose, cellotriose, and cellobiose (Figure 4.1A). GH5E also hydrolysed cellohexaose, pentaose, tetraose and triose to produce cellobiose, triose, and tetraose in all the reactions (Figure 4.1B). Our results showed the formation of cellotetraose when the enzyme hydrolysed cellotriose, respectively. These results were not expected.

In light of these results, we suggest that GH5E might also have transglycosylation activity on cellotriose and glucose, and as a result, further tests are required to elucidate the polymer cleaving mechanism of this enzyme.

Zhang et al. (2014) also reported on a processive endoglucanase enzyme that displayed transglycosylation activity. TLC results showed the endoglucanase could form cellotetraose, -triose, -biose and *p*-nitrophenylmaltotriose (*p*NPG3) when it hydrolysed *p*NPC. The authors suggested that the formation of larger molecules (rather than the one being hydrolysed) by the endoglucanase comfirmed that the enzyme has transglycosylation activity. This supports our hypothesis that GH5E displayed transglycosylation activity because it formed cellotetraose when it hydrolysed cellotriose (Figure 4.1B).

The two enzymes GH5H and GH45 are good polysaccharide cleaving enzymes. GH5H cleaved CMC and beechwood xylan mainly to products with a DP of 3-5, while cleaving xyloglucan to products with a DP of 5, 7-9 (Figure 4.1C). GH45 produces similar products to

GH5H when it acts on CMC and xyloglucan (Figure 4.1A and 4.1D). Beechwood xylan was an exception because it was not hydrolyzed into shorter oligosaccharide that could be resolved using TLC. GH5H catalyzed the cello and xylo-oligomers with a DP of 5 and 6 to produce tetraoses and trioses, respectively (Figure 4.1C). GH45 catalyzed cello-oligomers with a DP higher than 4, but could not catalyze cellobiose or triose (Figure 4.1D). From cello-oligomers with a DP of 4, GH45 produced cellobiose, while from cellopentaose it produced cello-oligomers with DP of 3, 2 and 4, and from hexaose, it produced cello-oligomers with DP of 3 and 2, respectively (Figure 4.1D). The products which formed when cellopentaose was catalyzed by GH45 were not expected because this enzyme requires a polymer with a DP of 4 for catalysis to occur. We expected that this enzyme would hydrolyze cellopentaose to cellotetraose and glucose. Furthermore, when GH45 hydrolysed cello-hexaose it appeared that it also required a minimum DP of 3 to bind to the polymer, which explains why this enzyme did not hydrolyse cellotriose.

GH5H also hydrolysed xyloglucan-oligosaccharides with a DP of 7-9 to products with a DP of 5, while the same enzyme hydrolysed the xyloglucan-oligomer with a DP of 14 to products that had a DP of 5 and 2 (Figure 4.1C). The enzyme did not hydrolyze any oligomer with a DP less than 7. In contrast, GH45 did not hydrolyze xyloglucan-oligosaccharides with a DP less than 14. The products formed when GH45 catalyzed a xyloglucan-oligosaccharide with DP of 14 were hepta-, octa- and nona-saccharides (Figure 4.1D). This is one of a few studies that used TLC to identify xyloglucan-oligosaccharides (Song et al. 2013). Some studies analyzed xyloglucan-oligosaccharides formed after hydrolysis of xyloglucan using MALDI-TOF MS or normal-phase HPLC. However, our results were comparable to the findings of these studies with regard to identification of xyloglucan-oligosaccharides (Ishida et al. 2007; Song et al. 2013; Vitcosque et al. 2016).

4.4.4. Glycosyl hydrolase enzymes mode of hydrolysis

TLC results showed very interesting observations with regard to enzyme-oligosaccharide chain length specificity. These results were used to propose a model for the mode of enzyme hydrolysis for GH5D, GH5E, GH5H and GH45. All enzymes evaluated in this study did not hydrolyse cellobiose. This was expected because these enzymes did not exhibited β-glucosidase activity. GH5D was active on cello-triose, -tetraose, -pentaose and -hexaose. GH5D enzyme hydrolyzed cellotriose to cellobiose and glucose; cello-pentaose to cellobiose and triose; however, GH5D could hydrolyse cello-triose further to cellobiose and glucose. Cellotetraose and cellohexaose were both hydrolyzed to cellobiose.

Rashamuse et al. (2016) classified GH5D as an exoglucanase and our finding shows that this enzyme is a processive-like cellobiohydrolase. Wilson and Kostylev 2012 argue that a processive enzyme is one that produces higher amounts of cellobiose than cellotriose. Our TLC results are in agreement with this argument, because only cellobiose was produced by GH5D (Figure 4.1B). In addition, Table 4.1 shows that GH5D had high activity on *p*NPC, a substrate that is hydrolyzed by cellobiohydrolase Cel7A from *Trichoderma reesei* (Kipper et al. 2005). In their study, Kipper and co-workers showed that a cellobiohydrolase, Cel7A, cleaved *p*NPC into *p*NP and cellobiose, indicating that it hydrolyses the polymer from the reducing end. This observation validates our findings that GH5D is not only an exo-glucanase but also a cellobiohydrolase-I.

GH5E showed a very interesting mode of hydrolysis when it hydrolysed cellotriose. Unlike GH5D, this enzyme produced cellobiose, -triose and -tetraose after one hour of hydrolysis of cellopentaose or cellohexaose. However, after 8 hours of hydrolysis, the enzyme only produced cellobiose. Our results suggests that: 1) GH5E hydrolysed cellotriose to cellobiose and then the glucose produced probably gets transglycosylated to cellotriose to form cellotetraose. The produced cellotetraose is then subsequently hydrolysed to cellobiose by GH5E. This enzyme hydrolysed cellotetraose and -hexaose to cellobiose. 2) Interestingly, on cellopentaose, GH5E produced cellobiose and cellotriose. The cellotriose produced could be hydrolysed to cellobiose through the process explained in 1). Thus, we hypothesised that GH5E has hydrolytic and trans-glycosylating properties.

Some of the large endoglucanase enzymes like GH5E (with trans-glycosylation properties) have been reported to be processive proteins(Kipper et al. 2005; Watson et al. 2009; Wilson and Kostylev 2012). This is in agreement with our findings, because, after 8 hours of hydrolysis, GH5E produced only cellobiose.

GH45 hydrolyzed cellotetraose to cellobiose, while it hydrolyzed cellopentaose to cellotriose and cellobiose. When GH45 hydrolyzed cellohexaose, three products were observed; cellobiose, -triose and -tetraose. This observation suggests that GH45 has a preference for long chain oligomers, because in the presence of cello-hexaose the enzyme did not hydrolyze cellotetraose further to cellobiose, as was the case where it hydrolyzed cellotetraose only.

GH5H hydrolyzed both cello- and xylo-oligosaccharides. This enzyme hydrolyzed cellopentaose and xylopentaose to bioses and trioses, respectively. In contrast, both cello-hexaose and xylo-hexaose were hydrolysed to trioses, respectively.

The results suggest that GH5H and GH45 (multifunctional enzymes) hydrolyze long chain oligomers in comparison to GH5E and GH5D (cellulases). The general observation is that the products of GH5H and GH45 were mainly trioses, while GH5D and GH5E produced cellobiose and glucose or cellobiose only. Thus, we suggest these enzymes can hydrolyze cellulose containing substrates in a synergistic manner. Therefore, we propose that these enzymes could be used in different industrial processes (see Chapters 6 and 7 for further details). GH5H and GH45 proved not to be processive enzymes. It must be noted that the mode of hydrolysis by GH5H and GH45 on xyloglucan was not elucidated due to the complex structure of xyloglucan.

4.4.5. Substrate competition

Substrate competition assays were performed to confirm whether GH5H has one or two different catalytic centers. The motivation for these assays was based on the fact that pH and temperature optima studies for the GH5H enzyme (section 4.4.2) showed different pH and temperature optima when the enzyme was catalyzing different substrates. As a result, substrate competition assays were conducted according to the method of Chen et al. (2006). The substrate competition assays suggested that 1), if the observed velocity values are much closer to the theoretical velocity values for a single catalytic center then the enzyme has a single catalytic center or 2), if the observed velocity values are close to the theoretical velocity values for the double catalytic centers, then the enzyme has two catalytic centers.

To test if GH5H has two or one catalytic center (s), we calculated the theoretical values from the absorbance using the equation that Chen et al. (2006) used in their studies. Our findings showed that GH5H behaved as though it has two different catalytic centers when catalyzing a xyloglucan and beechwood xylan mixture (Table 4.3). The enzyme showed this behavior when xyloglucan concentration was higher and beechwood xylan concentration lower in the mixture. When the xylan concentration was increased and the xyloglucan concentration decreased in the mixture, the enzyme behaved as if it had a single (same) catalytic center. This behavior of GH5H can be explained by what we observed in the kinetic studies (Table 4.2). GH5H catalyzed xyloglucan better at high concentrations compared to low concentrations. In addition, at low concentrations of xyloglucan, the activity of GH5H was

very low when compared to that observed on beechwood xylan. Based on this observation we concluded that GH5H has two different catalytic centers, because when GH5H catalyzed a mixture with xyloglucan at high concentration and beechwood xylan at low concentration, the observed values were similar/ comparable to those theoretical values conferring two different catalytic centers (Table 4.3).

Table 4.3. Determination of GH5H active centres by substrate competition assays using variying combinations of CMC, beechwood xylan and xyloglucan concentrations

No	Concentrations of the substrate in the reactions (mg/ml)		Total velocity of hydrolysis values (ΔA_{540nm}/ml/min)		
	Xyloglucan	Beechwood xylan	Observed values*	Theoretical values calculated for	
				Same active centre	Different active centres
1	16	4	0.31 ± 0.01	0.27	0.31
2	12	8	0.29 ± 0.00	0.26	0.32
3	8	12	0.25 ± 0.01	0.24	0.31
4	4	16	0.21 ± 0.01	0.23	0.28
	CMC	Beechwood xylan			
1	16	4	0.10 ± 0.01	0.05	0.09
2	12	8	0.14 ± 0.00	0.08	0.13
3	8	12	0.18 ± 0.00	0.11	0.17
4	4	16	0.19 ± 0.00	0.15	0.19
	Xyloglucan	CMC			
1	16	4	0.28 ± 0.02	0.22	0.27
2	12	8	0.21 ± 0.00	0.16	0.24
3	8	12	0.17 ± 0.00	0.12	0.19
4	4	16	0.16 ± 0.01	0.07	0.14

*Observed values represent the means ± SD

Another reason for the proposed two different catalytic centers (of GH5H) could be that, since xyloglucan consists of a glucan backbone that has branches of xylo-oligosaccharides (Pauly et al. 2001; Gilbert et al. 2008; Vitcosque et al. 2016), the catalytic centers that catalyse the hydrolysis of xylan (including beechwood xylan) may also catalyse the hydrolysis of xyloglucan branches. The hypothesis that both catalytic centers exhibited a preference for xyloglucan (due to its heterologous structure) was supported by the change in behavior when GH5H catalyzed the hydrolysis of a mixture of CMC and beechwood xylan.

Based on the observed values and the theoretical calculated values (Table 4.3), GH5H behaved like an enzyme with two catalytic centers because the observed values for the total velocity of hydrolysis were similar to the values of theoretical calculated total velocity of hydrolysis for an enzyme with two different catalytic centers (Chen et al. 2006).

The K_M value of GH5H for CMC was higher than the K_M value for beechwood xylan (Table 4.2), which suggests that GH5H tightly binds to beechwood xylan compared to CMC. This observation validates our hypothesis that GH5H behaves differently on these substrates. GH5H also behaved like an enzyme with two catalytic centers when it hydrolysed a CMC and xyloglucan mixture (Table 4.3). We expected GH5H to behave like an enzyme with single catalytic center when it catalyzed a mixture of CMC and xyloglucan. The GH5H still acted like an enzyme with two different catalytic centers showing that xyloglucan structure played a role.

4.4.6. Enzyme deactivation, activation and inhibition by metal-ions and oligosaccharides

The deactivation/activation of the enzyme is based on the principle that metal-ion interaction with the enzyme might reduce, stabilize or enhance the enzyme performance during hydrolysis (Henley and Sadana 1985; Polizzi et al. 2007). GH5H was stable across the range of the metal-ions used, retaining more than 80% residual activity when it hydrolysed the xyloglucan in the presence of Mn^{2+} (Figure 4.2A). GH5H residual activity following hydrolysis of CMC, beechwood xylan and xyloglucan did not change much, except that Co^{2+}, Fe^{2+} and EDTA activated the enzyme when it hydrolysed xyloglucan. GH45 showed two different profiles when it hydrolysed CMC and xyloglucan (Figure 4.2B). GH45 was deactivated in presence of Cu^{2+} and Zn^{2+}, activated by Mn^{2+} and Co^{2+} and stabilized by Fe^{2+}, Ca^{2+}, Mg^{2+} and EDTA (a metal-ion chelator), when CMC was used as the substrate. In contrast, GH45 was deactivated by all the metal-ions, except for Co^{2+} which stabilized GH45 when xyloglucan was used as the substrate. The GH5E enzyme was slightly activated by Fe^{2+}, Cu^{2+}, Ca^{2+} and Zu^{2+}. The Co^{2+} ions significantly activated GH5E, while Mn^{2+}, Mg^{2+} and EDTA stabilized it (Figure 4.2B).

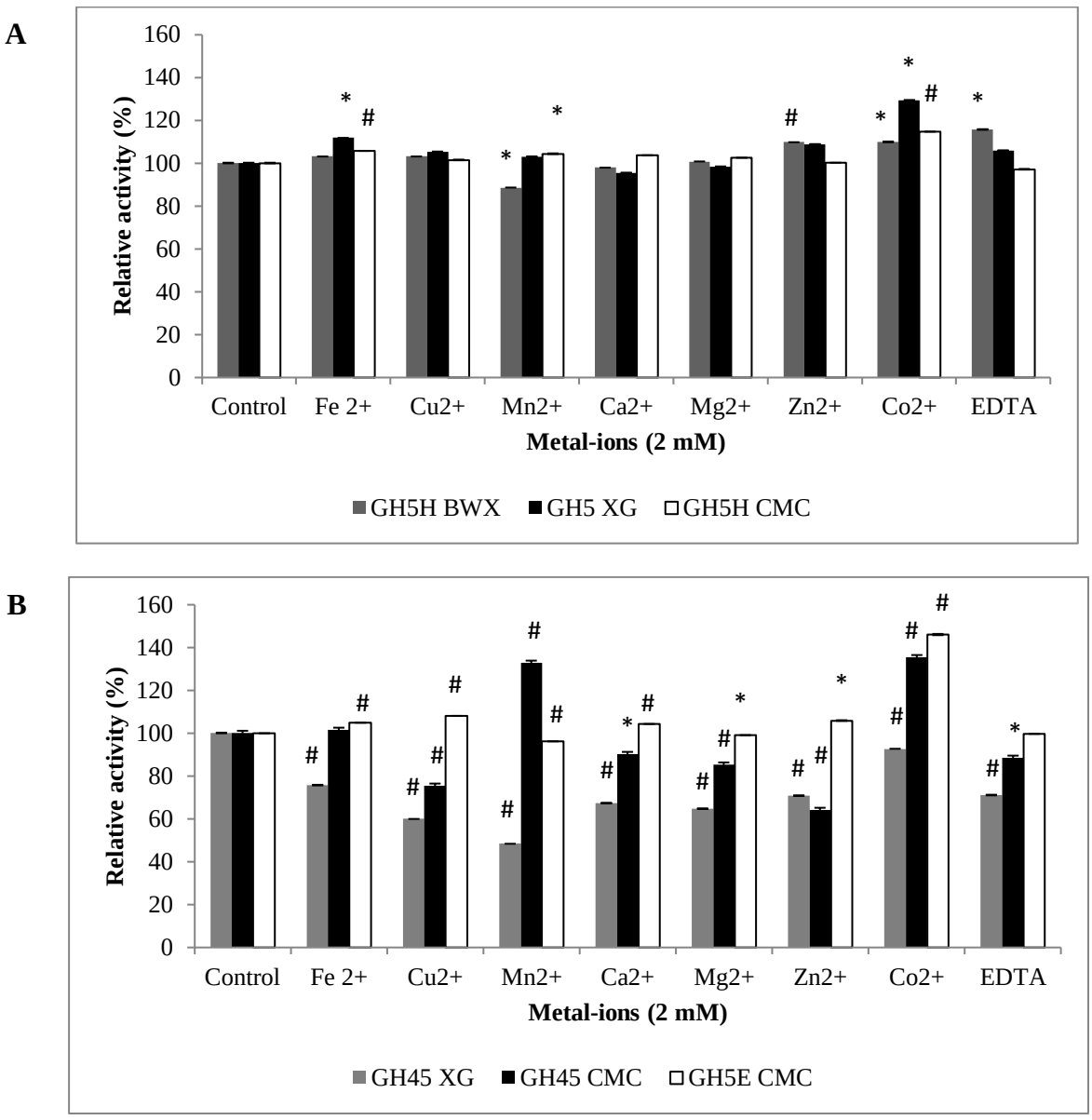


Figure 4.2. **Deactivation/activation effects of metal-ions and EDTA on GH5H (A), GH45 and GH5E (B) activities. The Fe^{2+}, Cu^{2+}, Mn^{2+}, Ca^{2+}, Mg^{2+}, Zn^{2+}, Co^{2+} and EDTA (chelating agent) were mixed with GH5H, GH45 and GH5E. The metal-ion enzyme mixture was added to the substrate (CMC, beechwood xylan and xyloglucan) and the effects of metals were measured relative to the activity of the control samples. The control did not have metal-ion treatment and its activity was considered to be 100%. *represents p-value < 0.05 and # represents p-value < 0.01. The experiments were performed in triplicate and the values represent means ± SD.**

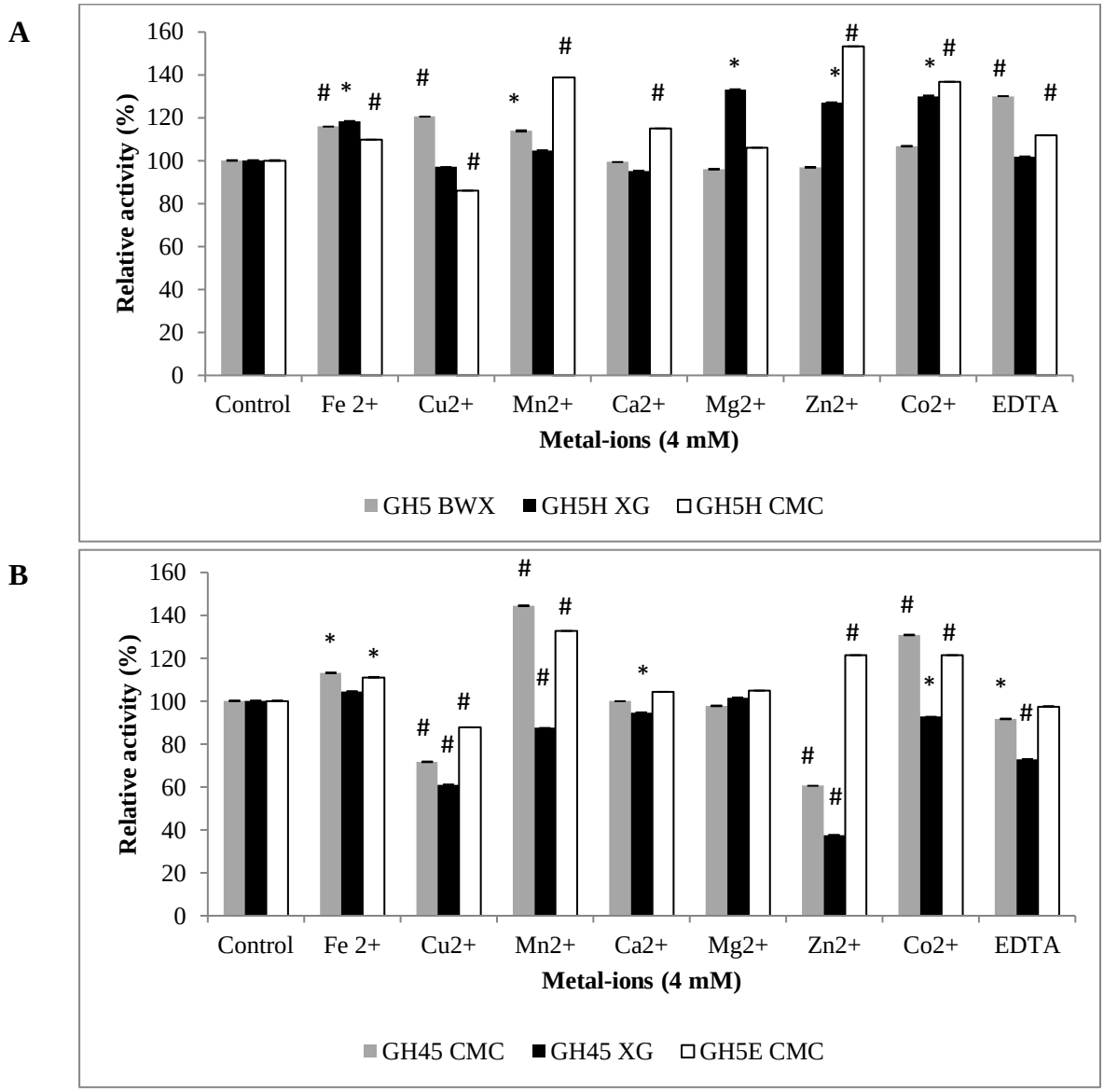


Figure 4.3. Inhibition/activation effects of metal-ions and EDTA on GH5H (A), GH45 or GH5E (B) activity. The Fe^{2+}, Cu^{2+}, Mn^{2+}, Ca^{2+}, Mg^{2+}, Zn^{2+}, Co^{2+} and EDTA (chelating agent) were mixed with CMC, beechwood xylan and xyloglucan. The GH5H, GH45 and GH5E were added to the metal-ion-substrate mixture, respectively. The activity of the enzymes was measured relative to the activity of the controls samples. The control did not have metal-ion treatment and its activity was considered to be 100%. *represents p-value < 0.05 and # represents p-value < 0.01. The experiments were performed in triplicate and the values represent means ± SD.

Pol et al. (2012) described a xyloglucanase that was stable in the presence of various metal-ions. This supports our findings because GH5H was stable over a range of metal-ions, except when it hydrolysed CMC in the presence of Cu^{2+} (Figure 4.3A). Unlike the xyloglucanase [used by Pol et al.(2012)], which was activated by Zn^{2+}, GH5H was stabilized by Zn^{2+} but activated by Co^{2+}. These observations support our claim that GH5H has some novel properties. The xyloglucanase (GH12) from thermophilic *Rhizomucor miehei* did not show an ability to interact with such a broad range of metal-ions (Song et al. 2013). Instead, xyloglucanase (GH12) was deactivated by Zn^{2+} and Co^{2+} with residual activities of about 33% and 80%, respectively.

The endoglucanase (GH5E) used in the current study was also stable over a range of metal-ions tested, compared to the endoglucanase from *Asperiguls niger* ANL301 which was activated by two metal-ions (Mn^{2+} and Fe^{2+}) and deactivated by other ions including EDTA (Nwodo et al. 2011). On the other hand, GH5E was not stabilized by EDTA and was significantly activated by Co^{2+}. *Asperigillus niger* and *Bacillus subtilis* are well known producers of GH-enzymes but their endoglucanases were only activated by one or two metal-ions (Nwodo et al. 2011; Wei et al. 2015). The fact that GH5E was stable over a range of metal-ions, support our point that GH5E also has novel properties. Thus, these termite metagenome derived enzymes can have vast industrial applications. In addition, the GH5E enzyme metal inhibition and deactivation studies have similar profiles (Figure 4.2B and 4.3B).

GH5E, GH5H, and GH45 were not inhibited by cello-oligosaccharides (see Figure 4.4A). Instead, GH5E was activated by glucose, cellobiose and cellotriose. GH5E had about 2.5 fold higher activity compared to the control (Figure 4.4A). Even though these results were not expected, they showed that GH5E is highly stable in the presence of shorter oligosaccharides that normally inhibit other GH enzymes (Andrić et al. 2010; El-Naggar et al. 2014). Furthermore, GH45 and GH5H were activated to about 1.2 fold, relative to their activities in the controls by cellotriose and cellotetraose, respectively (Figures 4.4B and D). The multifunctional GH5H was also found to be stable when xylo-oligosaccharides were used as inhibitors (Figure 4.4D). Both GH45 and GH5H were inhibited by xyloglucan oligosaccharides (Figure 4.4C and E). It's interesting to note that these enzymes were not inhibited by some of their products, suggesting that these enzymes might be good for industrial applications.

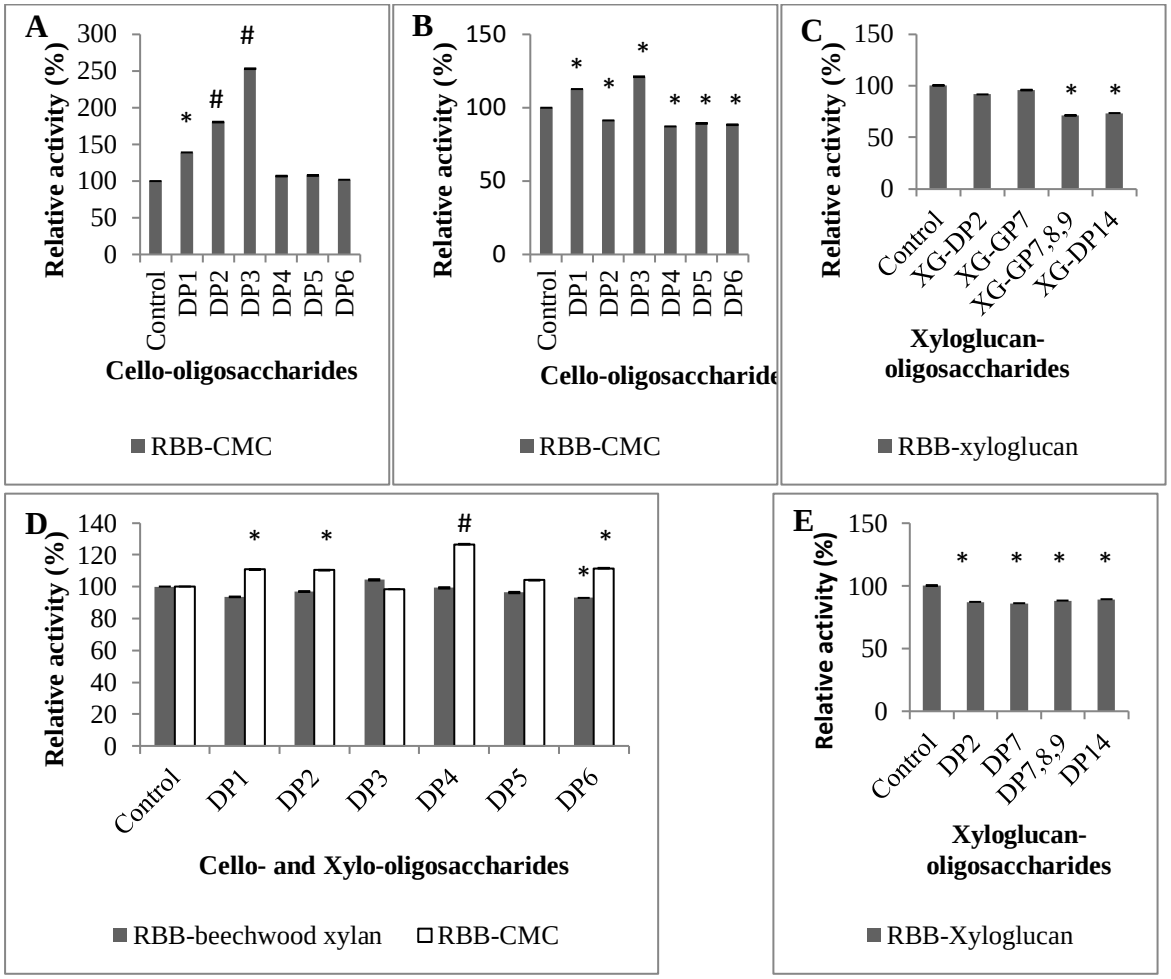


Figure 4.4. **A shows inhibition/activation effects of cello-oligosaccharides on GH5E activity. DP1 to DP6 represents glucose, cellobiose, cellotriose, cellotetraose, cellopentaose or cellohexaose which were used as inhibitors (graph A, B and E (gray bars)). Azo-CMC or remazol brilliant blue R (RBB)- CMC was used as substrates. Graphs B and C shows inhibition/activation of GH45 by cello- and xyloglucan-oligosaccharides in the reaction where RBB- CMC and RBB- xyloglucan were used as substrates, respectively. The XG-DP2 to XG-DP14 in graph C and E represents xylosyl-cellobiose, heptasaccharide, xyloglucan (hepta+octa+nona saccharides), or higher degree of polymerization xyloglucan oligosaccharides, respectively. Graph D and E show inhibition/activation of GH5H by xylo-, cello- or xyloglucan-oligosaccharides. The gray and white bars of graph D represents the reaction between GH5H and RBB-beechwood xylan or –CMC in the presence of the inhibitor, respectively. DP1-DP6 represents xylose, xylobiose, triose, tetraose, pentaose, and hexaose, respectively (blue bars). Graph E represents inhibition reaction between RBB- xyloglucan and GH5H in the presence of the inhibitor. All controls represent the reactions between the enzyme and the substrates without inhibitors. The relative activities were expressed relative to controls. *represents p-value < 0.05 and # represents p-value < 0.01 All experiments were in triplicate and the values represent means ± SD**

4.5. Conclusion

The aim of the research presented in this chapter was to comprehensively characterise pure GH5D, GH5E, GH5H, and GH45 enzymes derived from the termite (*Trinervitermes trinervoides*) hindgut metagenome. The temperature optima of these enzymes varied, however, at 37°C all enzymes were thermostable for 72 h. The enzymes displayed different pH optima, but at pH 5.5 all enzymes retained about 80% relative activity. We have concluded that GH5D is an exo-glucanase and GH5E is an endoglucanase with soluble wheat arabinoxylan hydrolysing activity. Both GH5H and GH45 proved to be xyloglucanase enzymes with cellulase and xylanase activities. GH5E and GH5H enzymes were stable in the presence of metal-ions and GH5E, GH5H and GH45 were stable in the presence of oligosaccharides. Substrate competition assays suggested that GH5H is a multifunctional enzyme with two catalytic centers. Thus, we suggest further studies regarding the catalytic centers of GH5H will be important in understanding its hydrolytic mechanisms.

The proposed mode of hydrolysis suggests that GH5D and GH5E could be processive enzymes. Based on the model and TLC results, we confirmed that GH5E also possessed transglycosylation activity. We conclude that based on their properties GH5E, GH5H, and GH45 can form core enzyme set that could hydrolyse polymeric substrate of complex substrate like agricultural biomass to simple sugars.

Chapter 5: Changes in Biomass Crystallinity are Necessary for GH5D Activity: A Comparative Study Between a Termite Metagenome Derived CBH-GH5D, Cel7 and Cel6

5.1. Introduction

Cellulose is the most abundant biopolymer on earth and is present in the majority of plant cell walls (Rose 2003). Cellulose consists of chains of glucose linked by β-1,4-glycosidic bonds. These chains are linked by strong hydrogen bonds which form parallel microfibrils (Rose 2003; Van Dyk and Pletschke 2012). The cellulose microfibrils are known to be very recalcitrant to enzymatic degradation, hence they are referred to as the crystalline regions of the biomass. Some parts of the cellulose structure are amorphous in nature and are easier to degrade through enzymatic means.

Lignocellulose is a very important feedstock for bioethanol production and understanding how the enzymes, in particular the cellobiohydrolases (CBHs), act on crystalline cellulose microfibrils, could assist in improving the production of glucose (Nidetzky et al. 1993; Ganner et al. 2012; Segato et al. 2012). Hence, in the current study, we elucidated how chemical and structural changes in cellulose affect the activities or performance of CBHs.
In addition to cellulose, β-D-glucans are also found in cereal plants such as barley. β-Glucan is a linear polymer made up of blocks of 1,4-linked β-glucose units that are separated by single 1,3-linkages (Johansson 2006; Synytsya and Novak 2014).

Most of the β-glucan (cellulose-like) segments exist as trioses and tetraoses, but longer 1,4-linked oligomers may exist in some plants. Based on the above mentioned features, we argue that crystalline cellulose is chemically and structurally different from β-glucan. Therefore, we anticipate that CBHs will function differently on these two polysaccharides. Interestingly, researchers have shown that both CBHI and CBHII, mostly from glycosyl hydrolase families (GH) 7 and 6, respectively, are capable of hydrolyzing β-glucan (Henrikssona et al. 1995; Takahashi et al. 2010). However, to the best of our knowledge, CBHs from other GH families (such as 5 or 74) have not been investigated for their activities against β-glucan polymers.

5.2. Aims and Objectives

5.2.1. Aims

To use untreated, phosphoric acid swollen (PASC) and sodium hydroxide (NaOH) pretreated Avicel, and β-glucan from barley, to comparatively evaluate the activity of a novel termite metagenome derived enzyme (GH5D) to a microbially derived Cel6 and a *Trichoderma longibranchiatum* Cel7. To elucidate GH5D's ability to hydrolyze β-1,4-linkages by using oligosaccharides and comparing its activity to that of CBHI and CBHII.

5.2.2. Objectives

1. To change the Avicel crystalline nature with phosphoric acid or NaOH and produce a less crystalline Avicel substrate.
2. To investigate if GH5D can hydrolyse the pretreated Avicel substrate better than the untreated Avicel and to compare GH5D activities to that of CBHI and CBHII.
3. To investigate if GH5D can hydrolyse the crystalline insoluble β-glucan from barley and compare its activities to that of CBHI and CBHII
4. To investigate if the GH5D hydrolysis products of oligo-saccharides are comparable to that of CBHI or CBHII using thin layer chromatography (TLC).
5. To inverstigate the binding of the GH5D and CBHs on untreated Avicel.
6. To perform synergy experiments with these enzymes and confirm their degree of synergism.

5.3. Materials and Methods

5.3.1. Phosphoric acid and NaOH pretreatments of Avicel

All cellulose-pretreated samples were prepared using a crystalline cellulose substrate (Avicel PH 101). A cellulose substrate solution of 5% (w/w) was prepared by mixing 0.5 g Avicel with 5 ml of 50% phosphoric acid - this was then incubated at 4°C for 16 h. The biomass was then stirred at room temperature for 1 and 3 h, respectively, depending on the severity of pretreatment required. These phosphoric acid pretreated samples were referred to as PASC-1 and PASC-3, respectively. A second pretreatment was performed with sodium hydroxide (NaOH). Avicel (0.5 g) was mixed with 20% (w/v) NaOH and the mixture was stirred at 60°C. Samples were taken after 0.5, 1.5, 3 and 16 h, respectively. After pretreatment, samples were washed with distilled water until the pH of the effluent was approximately 6.8. The samples were air dried and stored at room temperature in airtight containers for future use.

5.3.2. Scanning electron microscope (SEM) analysis

To determine the effectiveness of each pretreatment (lime, NaOH and $NaClO_2$) at modifying the surface area of the biomass, the dried untreated PASC-1, PASC-3, NaOH-05, NaOH-1.5, NaOH-3 and NaOH-16 h pretreated Avicel samples were mounted on scanning electron microscope (SEM) stubs using double sided graphite tape and sputter coated with gold using a Balzers Union sputtering device. The gold coated untreated, PASC and NaOH pretreated samples were viewed under a Tescan Vega scanning electron microscope at 20 kV. Digital images were captured using a Vega imaging system (Pinchuck 2009).

5.3.3. FTIR analysis

Functional group analysis of the untreated, PASC and NaOH pretreated Avicel was conducted via Universal Attenuated Total Reflectance Fourier Transform Infrared (UATR-FTIR) (PerkinElmer, USA). The FTIR was performed as described in Chapter 3 (section 3.3.3). The spectra of the untreated, PASC and NaOH pretreated samples were presented as absorbance values and each value represented the mean of four scans.

5.3.4. Enzyme activity assays

The PASC-1, PASC-3 and NaOH pretreated Avicel samples were mixed with GH5D, CBHI, and CBHII in sodium citrate buffer (pH 5.5). The reactions were initiated by incubating the substrate and enzyme mixture at 37°C for 48 h. Identical reactions were also set up in the same manner; however, 20 µg of Novozyme 188 β-glucosidase was added to 10 µg of GH5D, CBHI and/or CBHII, respectively. The β-glucosidase was added to the reaction to reduce enzyme product inhibition. GH5D, CBHI and CBHII activities were determined using a modified 3,5 dinitrosalicylic acid (DNS) method (Miller 1959). At the end of each reaction, samples were mixed with DNS reagent in a ratio of 1: 2, the mixture was then heated at 100°C for 5 min and allowed to cool on ice. The absorbance readings were taken at 540 nm with a Powerwave microplate reader (Bio-Tek Instruments) using KC Junior software. The total reducing sugars were quantified using a standard curve that was generated using glucose as a suitable standard. The GH5D activities were also quantified using para-nitrophenyl-β-D-cellobioside, according to the method described by Malgas et al. (2015).

5.3.5. Enzyme binding assays

CBHI, CBHII, and GH5D, at a concentration of 0.1 mg/ml were mixed with 1% (w/v) of untreated Avicel dissolved in 50 mM sodium citrate buffer. The mixture was incubated at 4°C for 2 hours and then centrifuged at 13 000 *g* for 15 min. The controls were similar to the experimental reactions, except that the substrate control did not contain the enzyme and the

enzyme control did not contain substrate. CBHI, CBHII and GH5D were allowed to bind to the substrate for 2 hours and the free enzyme concentrations in the supernatants were measured using Bradford assays (Bradford, 1976).

After the supernatants were removed, the pellets were treated in two different ways; the pellet in one tube was mixed with 100 µl of 2x Laemmli sample buffer (Bio-Rad), while the pellet in the other tube was washed twice with sodium citrate buffer pH 5.5 and the mixed with 100 µl of 2x Laemmli sample buffer and boiled 100°C for 5 min. Lastly, the supernatants from binding assays were mixed with Laemmli sample buffer at a ratio of 2:1 and boiled at 100°C for 5 min. To assess which enzyme was bound to untreated Avicel, the samples and controls were analysed on SDS-PAGE (Laemmli 1970).

5.3.6. Enzyme synergism studies

Individually, GH5D, CBHI and CBHII performed better on NaOH pretreated Avicel compared to PASC pretreated Avicel. Thus, NaOH pretreated Avicel (0.5 and 16 h pretreatments) were selected to perform synergy studies. Binary combinations of GH5D, CBHI and CBHII were formulated at ratios of 0, 25, 50, 75 and 100% with a total protein loading of 20 µg/mg in the reaction. β-glucosidase (10 µg/mg) was included in all reactions during the synergy studies. Synergy assays were initiated by adding the 25%:75%; 50%:50%; 75%:25% or 100% of GH5D and CBHI (or GH5D and CBHII) (or CBHI and CBHII) to the 1% (w/v) untreated, 0.5h and 16h NaOH pretreated biomass (rehydrated in sodium citrate buffer, pH 5.5). The total reducing sugars produced after 48 h were measured using the DNS method.

GH5D, CBHI and CBHII activities on insoluble Azo-β-glucan were investigated at 37°C for 16 h and the reaction was carried out in sodium citrate buffer at pH 5.5. The reaction was terminated by cooling on ice and then a 2:1 ratio of ethanol to hydrolysate was added to the reaction mixture. The mixture was then vortexed vigorously and centrifuged at 13 000 *x* g for 10 min. The blue colour observed after enzyme hydrolysis was directly proportional to the enzyme activity, with the absorbance measured at 590 nm with a Powerwave$_x$ microplate reader (Bio-Tek Instruments) using KC Junior software. Remazol Brilliant Blue R was used as a suitable standard.

5.3.7. Calculation of the degree of synergy (DS)

Van Dyk & Pletschke (2012) defined the degree of synergy (DS) as the total reducing sugars (TRS) produced by two enzymes working in synergy divided by the arithmetic sum of total reducing sugars produced by the individual enzymes. Thus, to determine the degree of

synergy (DS) the synergy assays were performed as mentioned above and the individual enzyme (GH5D, CBHI and CBHII) assays were performed using different enzyme loadings of 25%, 50% and 75%. The data was used to calculate the degrees of synergy (DS) according to the following equation:

DS = (TRS of A: B combination) / (TRS of enzyme A) + (TRS of enzyme B)

DS represents the degree of synergy; TRS represents the total reducing sugars produced, enzymes A and B represent GH5D, CBHI or CBHII.

5.3.8. Data analysis

One way analysis of variance (ANOVA) was used to analyse the activity of the enzymes. Evaluation was conducted to elucidate significant increases exhibited by the enzyme combinations with respect to GH5D, CBHI or CBHII and/or reducing sugar release compared to that released by 100% protein enzyme loading. All pairwise comparison procedures were conducted using the Data analysis feature in Microsoft Excel.

5.4. Results and Discussion

5.4.1. SEM and FTIR analysis of pretreated Avicel

Our main objective was to use PASC and NaOH pretreatment on Avicel to produce cellulose substrates with varying degrees of crystallinity, and then investigate if the extent of crystallinity affected the activities of CBHI, CBHII and GH5D. The untreated Avicel served as the control. We first pretreated the biomass with 50% (w/v) phosphoric acid to obtain swollen biomass. With regards to PASC treatment of cellulose, other studies have used varying strengths between 70 and 80% of phosphoric acid. However, as shown by Zhang et al. (2006), a highly concentrated phosphoric acid hydrolyzes cellulose during the treatment, resulting in biomass with small particle size, and in the case of RAC (regenerated amorphous cellulose), the biomass becomes highly soluble. After 1 and 3 h of PASC pretreatment (PASC-1 and PASC-3), the biomass was dried and the morphology of the biomass surface was studied using SEM. Appendix: Figures A5.1A-C show that the untreated Avicel biomass surface was more uneven (more irregular) with some parts being covered by layers (flakes) - compared to PASC-1 and PASC-3, which displayed a smooth surface with no/or very little flaky layers covering the biomass. There was no substantial difference between the untreated Avicel, PASC-1 and PASC-3 samples regarding the porosity of the biomass. Kumar et al.

(2001) and Zhang et al. (2006), using SEM analysis, showed the formation of the pores on the biomass due to acid hydrolysis when concentrated phosphoric acid (80%) was used to pretreat Avicel or cotton. These findings confirmed that we had produced a swollen biomass from Avicel after phosphoric acid pretreatment.

As early as 1983, Farid et al. (1983) showed that phosphoric acid and NaOH can decrease the crystallinity of cellulosic biomass. In our current study, we also pretreated the Avicel with 20% (w/v) NaOH. The biomass was pretreated with 0.5, 1.5, 3 and 16 h, with the resulting pretreated biomass being referred to as NaOH-0.5, NaOH-1.5, NaOH-3 and NaOH-16, respectively. Unlike phosphoric acid pretreatment, which appears to have a smoothening effect on the Avicel after pretreatment, the NaOH pretreatment appeared as if it was peeling off the Avicel layer by layer (Appendix: Figure A5.1 D-G). This peeling effect of the NaOH pretreatment on Avicel in our study has also been reported on natural lignocellulosic substrates such as Congo grass (Haque et al. 2016). After sodium hydroxide pretreatment, the biomass displayed more pores appearing on the surface of the biomass compared to untreated and phosphoric acid pretreated biomass. In addition, the NaOH-16 samples had a morphologically different surface compared to untreated Avicel. These findings were in agreement with a report by Selig et al. (2013) that illustrated that NaOH pretreatment changed Avicel from cellulose Iβ to produce cellulose II which is more amorphous. Thus, we confirmed that NaOH pretreated Avicel was indeed an amorphous substrate.

We further confirmed the SEM findings with FTIR analysis. FTIR was employed to investigate the decrease in crystallinity in the NaOH and PASC pretreated Avicel substrates. The total crystallinity index (TCI) and lateral order index (LOI) were used to demonstrate the effects of each pretreatment on the crystallinity of Avicel. The TCI values were calculated by taking the ratio of peak intensities at 1372 and 2900 cm^{-1}, while the LOI values were calculated from peak intensities at 1429 and 893 cm^{-1} (Selig et al. 2013). Table 5.1 shows that both the TCI and LOI values decreased in both the NaOH and PASC pretreated samples in comparison to untreated Avicel. NaOH-05 was the only exception since it displayed the same TCI and LOI values as the untreated biomass. Selig et al. (2013) and Kuo & Lee (2009) reported results that were similar to our study when they analyzed biomass pretreated with NaOH. Phosphoric acid pretreated Avicel also exhibited decreasing crystallinity values compared to the control (Jia et al. 2013). Based on SEM and FTIR results, it was clear that all

the pretreated samples had swollen structures that were less crystalline compared to untreated Avicel.

Table 5.1. Values of the crystallinity indexes (CI) for untreated, NaOH and PASC pretreated Avicel samples determined from UATR-IR spectra

Pretreated samples	Lateral order index (LOI)	Total crystallinity index (TCI)
Avicel (untreated/control)	2.166	1.64
NaOH-05	2.166	1.64
NaOH-1.5	0.54	0.974
NaOH-3	0.7	1.12
NaOH-16	0.57	0.988
PA-1	0.8	1.1
PA-3	0.916	1.04

5.4.2. Effect of crystallinity on enzyme activity

A pure CBHI from *Trichoderma longibrachiatum* (Megazyme), a CBHII from a microbial source supplied by Megazyme™ and a GH5D derived from a *Trinervitermes trinervoides* termite metagenome (Rashamuse et al. 2016), were used individually to hydrolyze untreated, PASC- and NaOH-pretreated Avicel. To eliminate the possibility of product inhibition, Novozyme 188 β-glucosidase was added to the second set of separate reactions that contained individual CBHI, CBHII and GH5D enzymes, respectively. Figure 5.1A shows the individual activities of CBHI, CBHII and GH5D on untreated and pretreated biomass, respectively.

Figure 5.1B shows the activities of the CBHI, CBHII and GH5D enzymes, which were applied with β-glucosidase to the untreated and pretreated Avicel substrates. The results from Figure 5.1A show that CBHI applied with β-glucosidase or without β-glucosidase exhibited higher activity on both the pretreated and untreated Avicel substrates compared to both CBHII and GH5D. The activity of CBHI was higher by about 2.5-3-fold compared to that of CBHII and GH5D when the reaction was performed without the addition of β-glucosidase. Interestingly, the activity of CBHII was improved by the addition of the β-glucosidase. The amount of the total reducing sugar produced by CBHI and CBHII in the presence of β-glucosidase increased, showing that β-glucosidase activity was very important for improving

the activity of the CBHs. The hydrolysis of NaOH-0.5 and NaOH-16 by CBHII, in particular, was greatly improved in the presence of the glucosidase (Figure 5.1B).

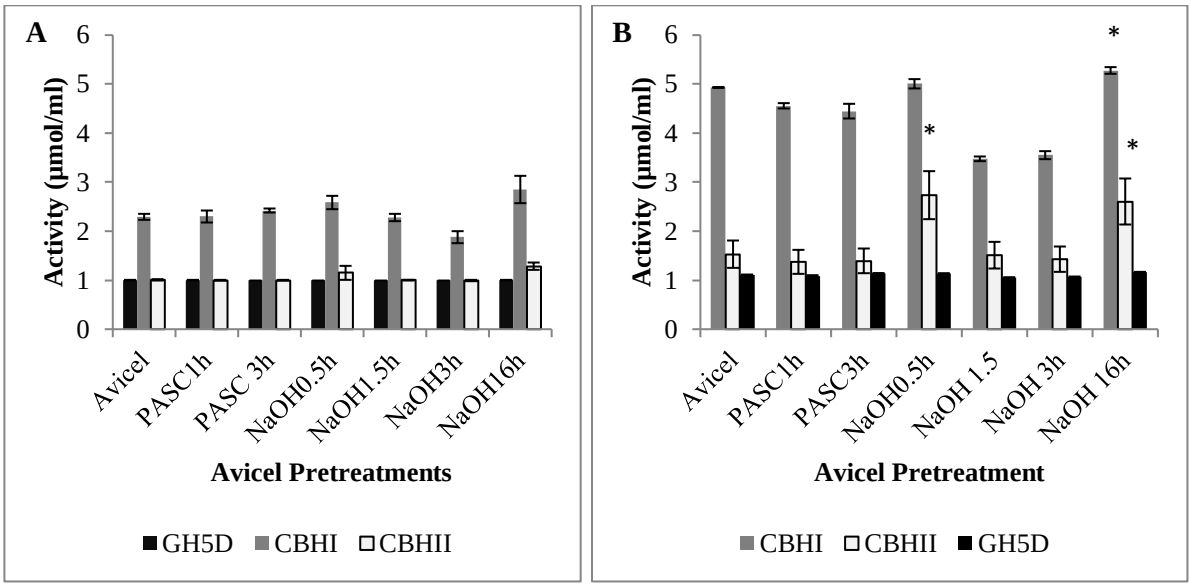


Figure 5.1. CBHI, CBHII and GH5D activity on the untreated, PASC and NaOH pretreated Avicel. A represents the individual enzyme activity without addition of β-glucosidase, Novozyme 188, and B represents the activity of individual enzyme activity with β-glucosidase, Novozyme 188. The enzyme loading was kept at a total of 20 µg and β-glucosidase was supplied at 10 µg in all the reactions. Avicel, PASC1h, PASC3h represent the control and phosphoric acid pretreated cellulose for 1 and 3 h, while NaOH 0.5, 1.5, 3 and 16 h represents the sodium hydroxide pretreated for 0.5, 1.5, 3 and 16 h, respectively. * represents p-values < 0.05. The experiments were performed in triplicate and data points represent the means ± SD.

It is also worth noting that after addition of β-glucosidase, CBHI activity on the phosphoric acid treated cellulose increased to such an extent that there was no significant difference between the reducing sugars produced from PASC-1 and PASC-3.

Our results are in agreement with other studies which showed that the addition of glucosidase in the reaction increased the enzyme activity of CBHs (Lantz et al. 2010; Van Dyk and Pletschke 2012). Nakamura et al. (2014) showed that cellobiohydrolase, Cel7A, from *T. reesei* or Cel7C and Cel7D from *Phanerochaete chrysosporium*, exhibited more activity on cellulose-II and cellulose-III compared to cellulose-I. Furthermore, it was reported

that an exo-cellulase was able to hydrolyze crystalline cellulose better than pretreated cellulose which was hydrolyzed by an endo-cellulase enzyme (Kanda et al. 1980). In the same light, we found that CBHII displayed more activity on NaOH pretreated Avicel compared to CBHI, which preferred a crystalline substrate. The low levels of activity of GH5D was unexpected (Figure 5.1), because it remained low on both the untreated and pretreated Avicel, and the addition of β-glucosidase did not make a significant difference with regards to its activity.

Interestingly, Ganner et al. (2012) used the same CBHI as we used in the current study to investigate its role in biomass hydrolysis. Their findings showed that CBHI was able to hydrolyze the crystalline regions of the untreated and phosphoric acid pretreated Avicel, but that it was not able to hydrolyze the amorphous regions of the biomass. They also observed that CBHII hydrolyzed the amorphous regions of the biomass, but was unable to hydrolyze the crystalline regions of the untreated Avicel – as was the case for CBHII in our study. From this data, we concluded that GH5D, an exo-glucanase/cellobiohydrolase derived from a termite metagenome, has a different substrate specificity compared to the CBHI from *Trichoderma longibrachiatum* (Megazyme) and CBHII from a microbial source (Megazyme).

5.4.3. GH5D and CBHs substrate binding analysis

The results showed that CBHI bound to the untreated Avicel. CBHII and GH5D did not bind to the untreated Avicel because the concentrations of both enzymes recovered from binding assays and the concentrations of their controls were similar (Figure 5.2A). The CBHI control had a protein concentration of about 0.28 mg/ml, but after the binding assay, no CBHI was detected. This confirms that CBHI was bound to the crystalline Avicel. These findings were validated by SDS-PAGE results (Figure 5.2B). The supernatant/hydrolysate for CBHII and GH5D enzymes recovered subsequent to binding assays displayed clear bands on the gel, while there was no display of a clear band for recovered hydrolysate of CBHI. In comparison, the CBHI control displayed intense bands on the gel and the bands intensity of CBHII and GH5D controls were compar to that of the CBHII and GH5D recovered from the hydrolysate subsequent to the binding assays. Figure 5.2B shows that only CBHI displayed clear bands of the protein that was recovered from the pellet before and after the wash. Thus, we concluded that CBHI's ability to bind to crystalline cellulose is important for the

hydrolysis of crystalline cellulose, because both GH5D and CBHII did not bind to this substrate and, as a result, failed to hydrolyze crystalline cellulose.

5.4.4. Cello-oligosaccharide hydrolysis

An interesting finding was that GH5D displayed very high activity on *p*-nitrophenyl-cellobioside (60.89µmol/min/mg). We then decided to investigate if GH5D could hydrolyze cello-oligosaccharides with a degree of polymerization (DP) of between 3 and 6, which is characteristic of cellobiohydrolases. The hydrolysis patterns of oligosaccharides produced by GH5D were compared to those produced by CBHI and CBHII.

After hydrolysis of cello-oligosaccharides by GH5D, CBHI and CBHII, the hydrolysates were spotted onto TLC plates. The TLC results showed that GH5D is a true cellobiohydrolase because it produced the same products as CBHI and CBHII (Figure 5.3A). GH5D produced glucose and cellobiose when hydrolyzing cellotriose, and mainly cellobiose when it hydrolyzed cellotetraose or hexaose. Cellotriose and cellobiose were produced when it hydrolyzed cellopentaose. These profiles were similar to those of CBHI, except that CBHI produced glucose when it hydrolyzed all the oligosaccharides (Figure 5.3A). CBHII was specific on its product because it produced mainly cellobiose and cellotriose with trace amounts of glucose (Figure 5.3A).

Other studies also found that cellobiohydrolases produce mainly cellobiose, then glucose and cellotriose from biomass (Vršanská and Biely 1992; Zhang et al. 2010; Kern et al. 2013). Given that glucose was produced by all the enzymes used in our study when they hydrolyzed the oligosaccharides, we quantified the glucose produced by these enzymes using the GOPOD kit from Megazyme. The amount of glucose measured by the kit (Figure 5.3B) correlated with the TLC results (Fig. 5.3A). From the TLC it was clear that CBHI produced more glucose compared to GH5D and CBHII. However, all three enzymes produced more glucose when hydrolyzing cellotriose and cellopentaose (odd number chains), but lower quantities when they hydrolyzed tetraose and hexaose (even numbered chains). From these results, we concluded that GH5D is an active cellobiohydrolase that hydrolyzes the β-1,4-linkage of *p*NPC and oligosaccharides. Thus, we decided to investigate whether the enzyme would work in a synergistic manner with commercial CBHI or CBHII because it was highly active on *p*NPC and cello-oligosaccharides.

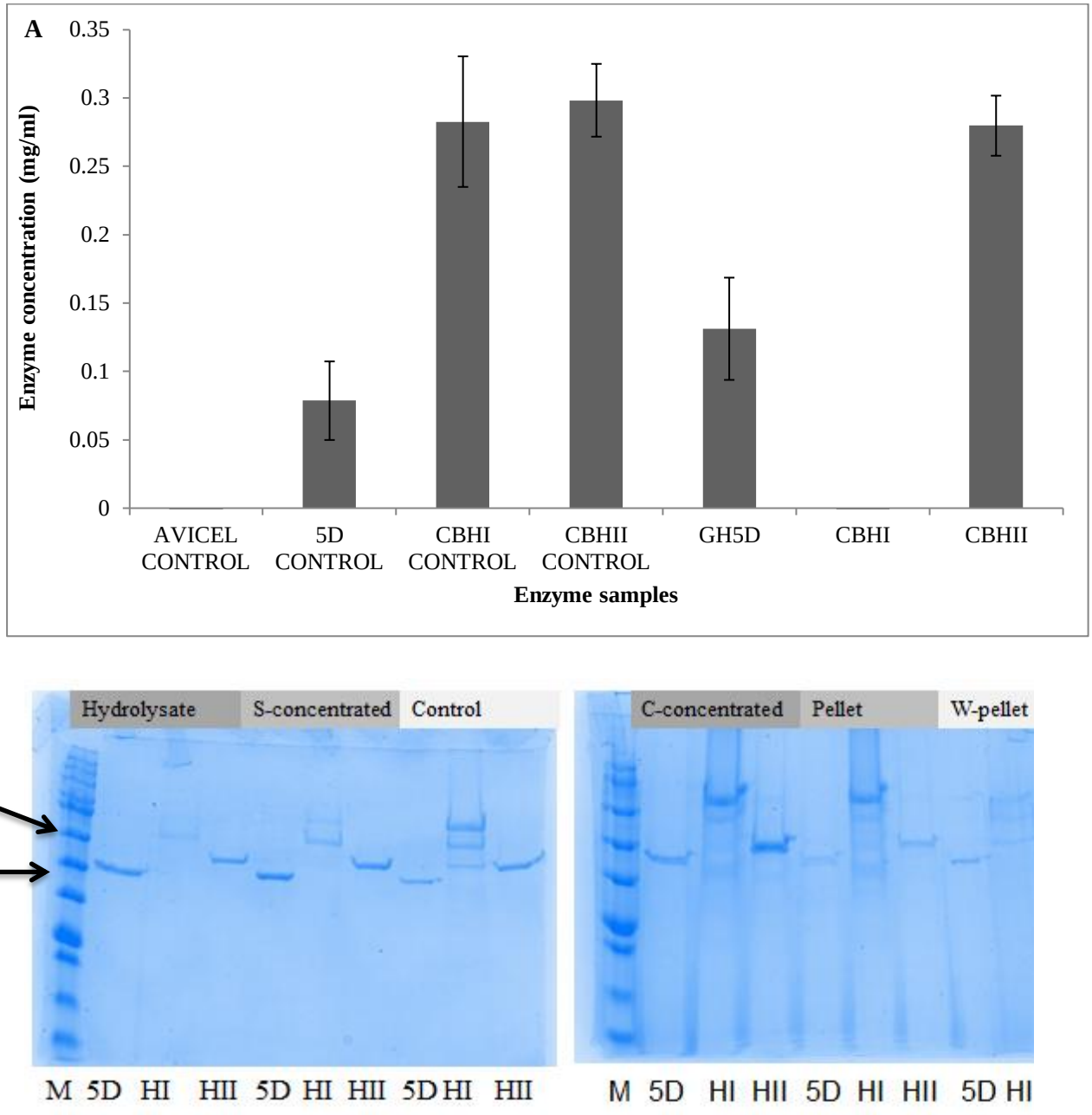


Figure 5.2. A. Concentrations of the cellobiohydrolase enzymes (CBHI, CBHII and GH5D) in the supernatant that was recovered from the binding assays. The experiments for determining enzyme concentrations were performed in triplicate and the error bars represent standard deviations. B. Analysis of the SDS-PAGE gel for the supernatant, pellet and control samples. The hydrolysate, S-concentrated, control, C-concentrated, pellet and W-pellet represent the supernatant, concentrated supernatant, enzyme control, concentrated enzyme control, enzyme recovered from the pellet and enzyme recovered from washed pellet, respectively. M, 5D, HI and HII represent the Bio-Rad protein molecular weight marker, GH5D, CBHI and CBHII enzyme, respectively.

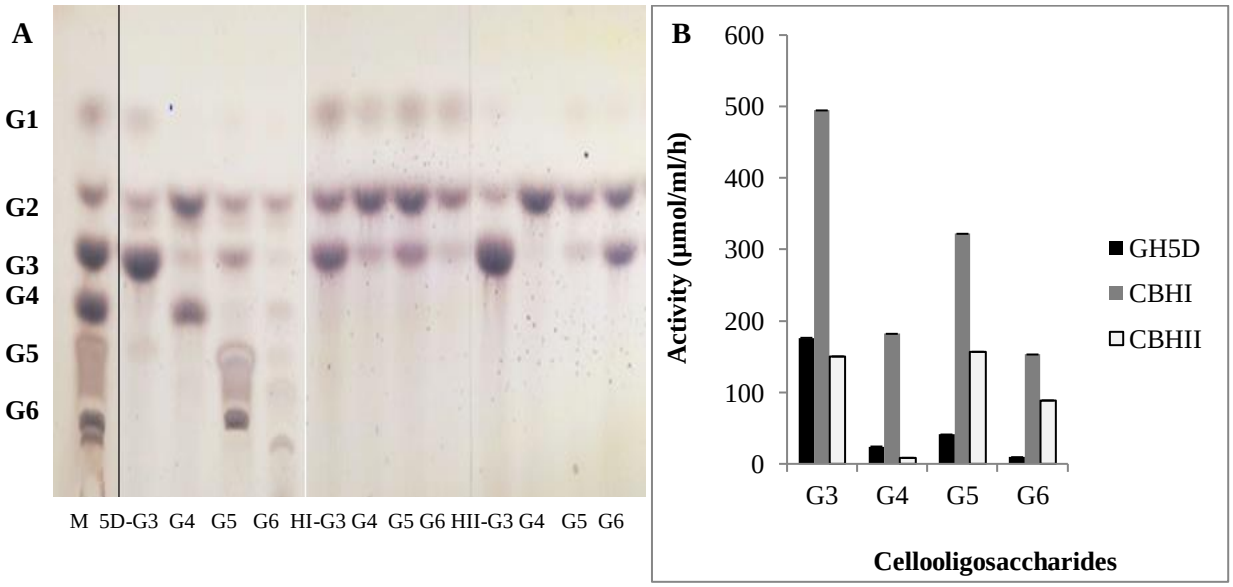


Figure 5.3. CBHI, CBHII and GH5D activities on cello-triose, -tetraose, -pentaose and -hexaose. The thin layer chromatography shows the products of CBHI, CBHII or GH5D hydrolysis of cello-oligosaccharides (A). G3 to G6 represent cello-triose, -tetraose, -pentaose and –hexaose, while 5D, HI and HII represent GH5D, CBHI and CBHII enzymes, respectively. The graph shows the quantified amounts of glucose produced from cello-oligosaccharide by each enzyme with beta-glucosidase (B). The glucose quantification experiments were performed in triplicate and datapoints represent the means ± SD.

5.4.5. CBHs and GH5D synergism on NaOH pretreated Avicel

Given that all enzymes used in this study exhibited high activity on the cello-oligosaccharides, we then investigated if there could be binary synergism between the two cellobiohydrolases and GH5D. The synergy and the degree of synergy (DS) between GH5D and either of the CBHs were compared to the synergy and DS between the CBHs. To perform these assays - untreated, 0.5 h and 16 h NaOH pretreated Avicel were used as substrates. Figure 5.1B shows that both the CBHs performed the best on these pretreated substrates, thus we decided to use NaOH pretreated Avicel as a substrate for subsequent studies and untreated Avicel was used as a control.

When synergy studies were carried out on Avicel as a substrate, there was no synergy displayed between CBHI and GH5D or CBHII and GH5D or CBHI and CBHII (Figure 5.4A). The DS as seen from Figure 5.4A did not reach or exceed 1.0, which means that there

was no synergy between all the enzymes used in the current study. In addition, there was no synergy between CBHI and GH5D or CBHII and GH5D when hydrolyzing Avicel that was pretreated with NaOH for 0.5 h. Figure 5.4B shows that the DS levels between CBHI and GH5D or CBHII and GH5D were below 1.0, meaning that there was no synergism between these enzymes. CBHI and CBHII formed synergy on Avicel that was pretreated with NaOH for 0.5 h when the enzyme combinations were CBHI75%: CBHII25%; CBHI50%: CBHII50% and CBHI25%: CBHII75%. Howeve, there was no statistical significance between CBHI100% and the CBHI and CBHII combinations because the p-value was less than 0.1. A DS level of above 1.0 also confirmed that there was synergism between CBHI and CBHII (Figure 5.4B). The highest enzyme activity of about 4.954 µmol/ml after 48 h was observed when CBHI and CBHII were combined at a ratio of 75% and 25%, respectively. However, the DS level was the highest (1.37) when CBHI and CBHII were combined in a ratio of 50% to 50% (1:1), respectively.

Similar results were also observed when these enzymes ydrolyzed Avicel that was pretreated with NaOH for 16 h. CBHI and GH5D or CBHII and GH5D did not show any synergism. Figure 5.4C shows that the DS levels of the former or the latter were below 1.0. Intrestingly, there was synergism between CBHI and CBHII (CBHI75%: CBHII25%; CBHI50%: CBHII50% and CBHI25%: CBHII75%) and the activities of the CBHI:CBHII combinations were statistically significant (p-value < 0.05) compaired to the activities of individual enzymes. These results were supported by DS levels which were greater than 1.0.

The activities for CBHI75%: CBH25% and CBHI50%: CBHII50% enzyme combinations were higher than the activities of any of the other enzyme combinations or those of individual enzymes. These results were expected because it has been reported that CBHI and CBHII purified from *Trichoderma reesei* displayed synergy on NaOH pretreated cellulose (Nidetzky et al. 1993). In addition, the CBHI core enzyme set [enzymes with both catalytic domains (CD) and cellulose binding domains (CBD)] did not exhibit synergy with either CBHI or CBHII enzymes that only possessed a CD. This observation could explain why there was no synergy found between GH5D and CBHI or GH5D and CBHII. We propose that GH5D failed to synergise with either CBHI or CBHI because it did not bind to the Avicel substrate (Figure 5.2). This results suggest that GH5D does not possess the CBD, which is important for facilitating the hydrolysis of the crystaline substrate such as Avicel.The activities of

GH5D and CBHI or CBHII combinations were always less than the sum of the individual activities.

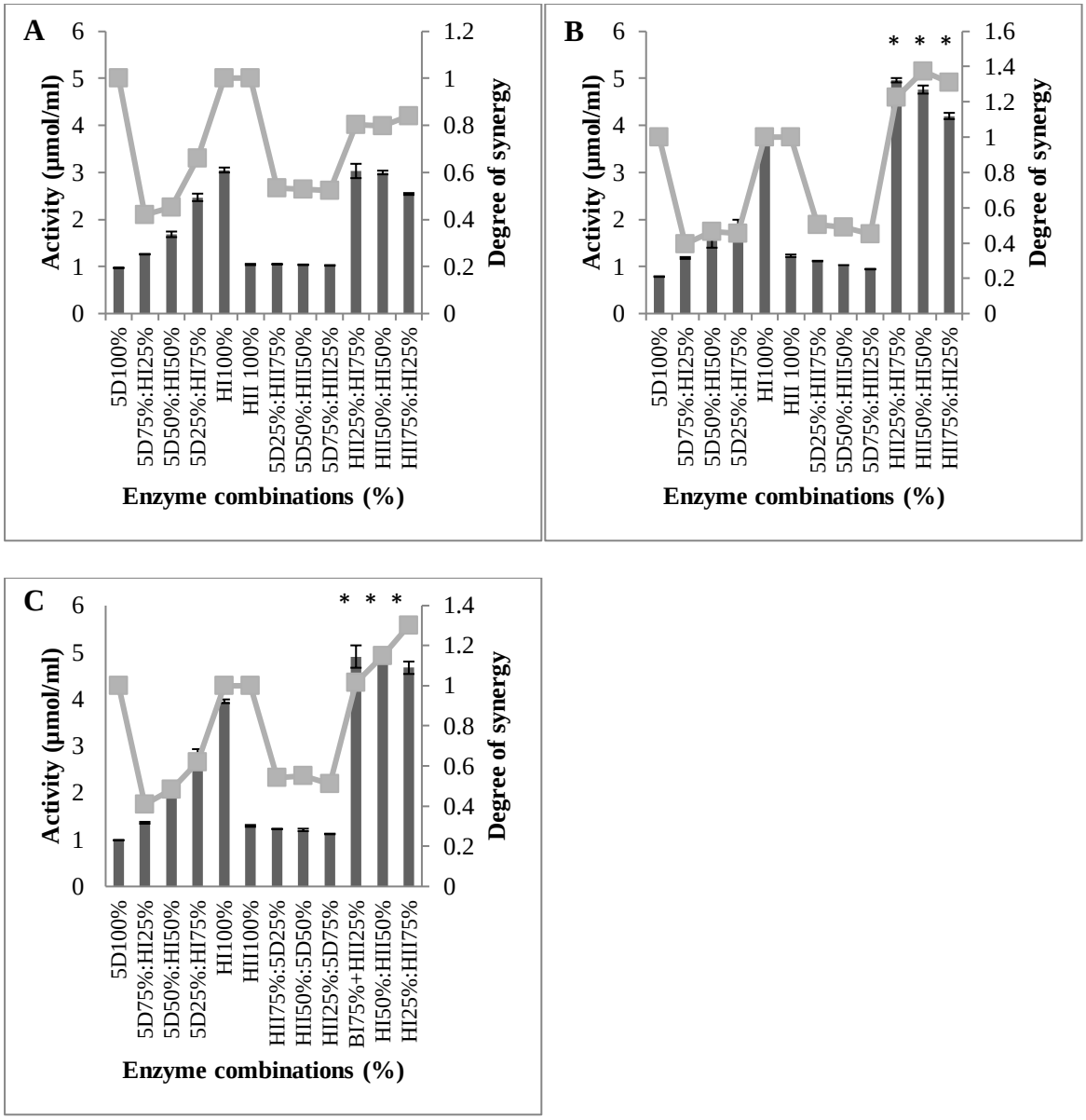


Figure 5.4. Analysis of the activity and the degree of synergy (SD) between CBHI, CBHII and GH5D. Activity (bars) and DS (line) were measured from total reducing sugars produced during the hydrolysis of untreated (A) and sodium hydroxide (NaOH) pretreated Avicel for 0.5 h (B) and 16 h (C). The 5D, HI and HII represent GH5D, CBHI and CBHII enzymes, respectively, while the percentage (%) represented the protein dosage of each enzyme in the enzyme combinations. * represents p-value < 0.05. The experiments were performed in triplicate and the datapoints represent the means ± SD.

Segato et al. (2012) showed that cellobiohydrolase I and cellobiohydrolase D (referred to as Cel7 and CelD) from *Aspergillus niveus* behaved differently during Avicel hydrolysis. Cel7 displayed the highest activity on Avicel and CelD had very low activity on the Avicel. The synergy assays conducted using the two enzymes showed that there was no synergy between Cel7 and CelD when the enzymes were added in a ratio of 1:1 or 50%:50%. The authors reported that the Cel7 concentration had to be four times higher, compared to the CelD concentration, for synergism to occur between Cel7 and CelD. However, an increased CBHI concentration relative to that of GH5D did not result in synergism in our study. Our results differed from that of Segato et al. (2012), because the CBHI used in our study was produced by a fungal source (Megazyme), while GH5D was derived from a termite gut metagenome (Rashamuse et al. 2016).

5.4.6. CBHs and GH5D synergism on beta-glucan

Beta glucan from barley is chemically different from cellulose in the sense that it consists of glucose residues that are linked by β-1,3- and β-1,4-linkages (Henrikssona et al. 1995; Takahashi et al. 2010; Synytsya and Novak 2014). Takahashi et al. (2010) showed that a *Magnaporthe oryzae* GH6 family cellobiohydrolase (referred to as MoCel6A) hydrolyzed β-glucan better than Avicel and PASC. CBHII activity on β-glucan has also been reported by Henrikssona et al. (1995). We, therefore, investigated whether GH5D derived from a termite metagenome, CBHI or CBHII could display activity on β-glucan.

Figure 5.5 shows that GH5D displayed higher activity compared to CBHI during the hydrolysis of insoluble β-glucan from barley. Interestingly, CBHII had the highest activity (7.35 μmol/ml) amongst the three cellobiohydrolases on insoluble β-glucan. CBHII activity was approximately 1.9-fold higher compared to GH5D activity and approximately 2.6-fold higher compared to CBHI activity. We concluded that GH5D displays activity on insoluble β-glucan. In addition, our results are in agreement with other studies which reported that CBHI and CBHII are active on β-glucan.

After confirming the activity of GH5D, CBHI and CBHII on insoluble β-glucan, we investigated if the enzymes would also work in synergy to break down insoluble β-glucan. The results displayed in Figure 5.5 show that GH5D and CBHI did not act in synergy. Instead, the activities of individual enzymes were higher compared to the activities of the two enzymes, GH5D and CBHI, used in combination. CBHII displayed synergism with GH5D at all enzyme combinations (CBHII75%:GH5D25% or CBHII50%:GH5D50% or CBHII25%:GH5D75%) considered. In addition, CBHII showed synergy with CBHI when the

enzymes were combined at CBHII50%:CBHI50% or CBHII25%:CBHI75% (not for CBHII75%:CBHI25%). The synergy displayed between CBHII75% and GH5D25% and CBHII100% resulted in significantly (p-value < 0.04) higher activities compared to other combinations that formed synergistic relationships. Therefore, we conclude that synergy between CBHII and GH5D was significantly higher than that between CBHI and CBHII. To the best of our knowledge, this is the first study that shows the synergistic activity of CBHI (GH6), CBHII (GH7) and GH5D on insoluble β-glucan from barley.

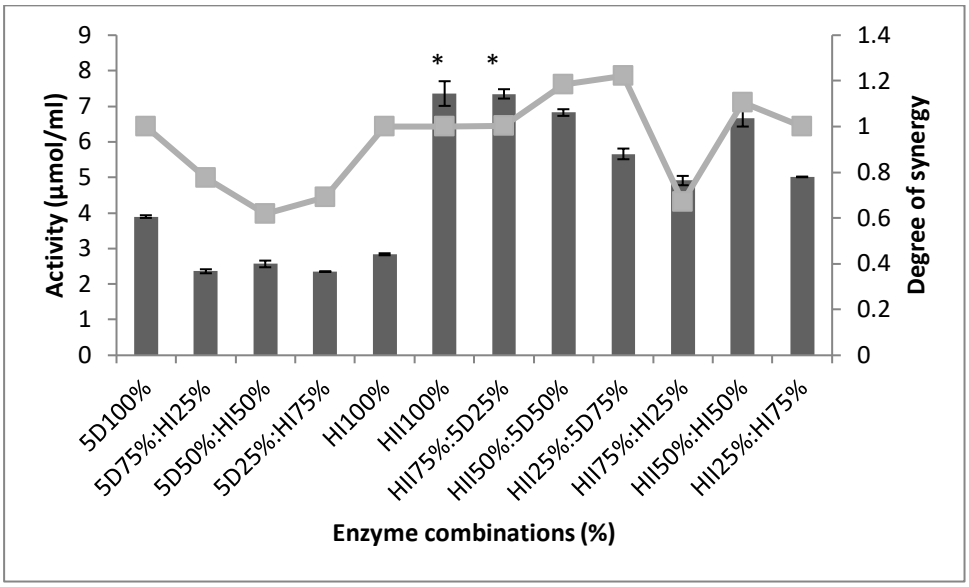


Figure 5.5. Analysis of the activity and the degree of synergy (DS) between CBHI, CBHII and GH5D. Activity and DS values were calculated from hydrolysates of insoluble Azo-beta-glucan from barley. The blue colour released was directly proportional to the enzyme activity. The 5D, HI and HII represent GH5D, CBHI and CBHII enzymes, respectively, while the % represent the protein dosage of each enzyme in the enzyme combinations. * represents p-value < 0.05. The experiments were performed in triplicate and datapoints represent the means ± SD.

5.4.7. The effects of GH5D on the activity of the CBHI and CBHII core enzyme sets

There was no synergy between GH5D and the CBHs during the hydrolysis of NaOH pretreated Avicel biomass (Figure 5.4). In addition, the same phenomenon was seen between GH5D and CBHI when they hydrolyzed insoluble β-glucan from barley (Figure 5.5). Due to

these findings, we investigated the effect of GH5D on the activity of CBHI and CBHII, which were fixed at 50%:50% ratio, while the concentration of GH5D was increased. The approach was motivated by several reports which used cellulases and hemicellulases at different concentrations and reported that an increase in enzyme concentration led to increased activity (Pallapolu et al. 2011; Van Dyk and Pletschke 2012; Inoue et al. 2014; Koppram et al. 2014).

Figures 5.6A and 5.6B show that the amounts of total reducing sugars and the glucose, respectively, decreased with increasing concentration of GH5D in the reaction. The addition of the highest concentration of GH5D (72 µg/g biomass) resulted in significantly decreased reducing sugars (p-value < 0.04) and glucose amounts (p-value < 0.02). About a 30% decrease in the activity of the CBHI50%: CBHII50% core enzyme set when it hydrolyzed 0.5 h NaOH pretreated Avicel in the presence of GH5D.

About a 20% decrease in the activity of CBHI50%: CBHII50% core enzyme set was observed when GH5D (72 µg/g biomass) was added to the reaction where untreated Avicel or 16 h NaOH pretreated Avicel was used as a substrate. The results confirmed that anti-synergy was observed between GH5D and the CBHs. In contrast, when the CBHI50%: CBHII50% core enzyme set was used to hydrolyze insoluble β-glucan from barley, the addition of GH5D in the reaction significantly (p-value < 0.05) increased the activity (Figure 5.6C). An increase in concentration of GH5D increased the reaction containing the CBHI50%: CBHII50% core enzyme set in a linear manner (possibly an additive effect), meaning that the addition of GH5D at the highest concentration (72 µg/g biomass) used in this study resulted in the highest total activity (Figure 5.6C). These results confirm synergism was observed between GH5D and CBHII or CBHI and CBHII but synergy was not formed between CBHI and GH5D. Furthermore, the results suggest that GH5D is a cellobiohydrolase that prefers to hydrolyze β-glucan over Avicel.

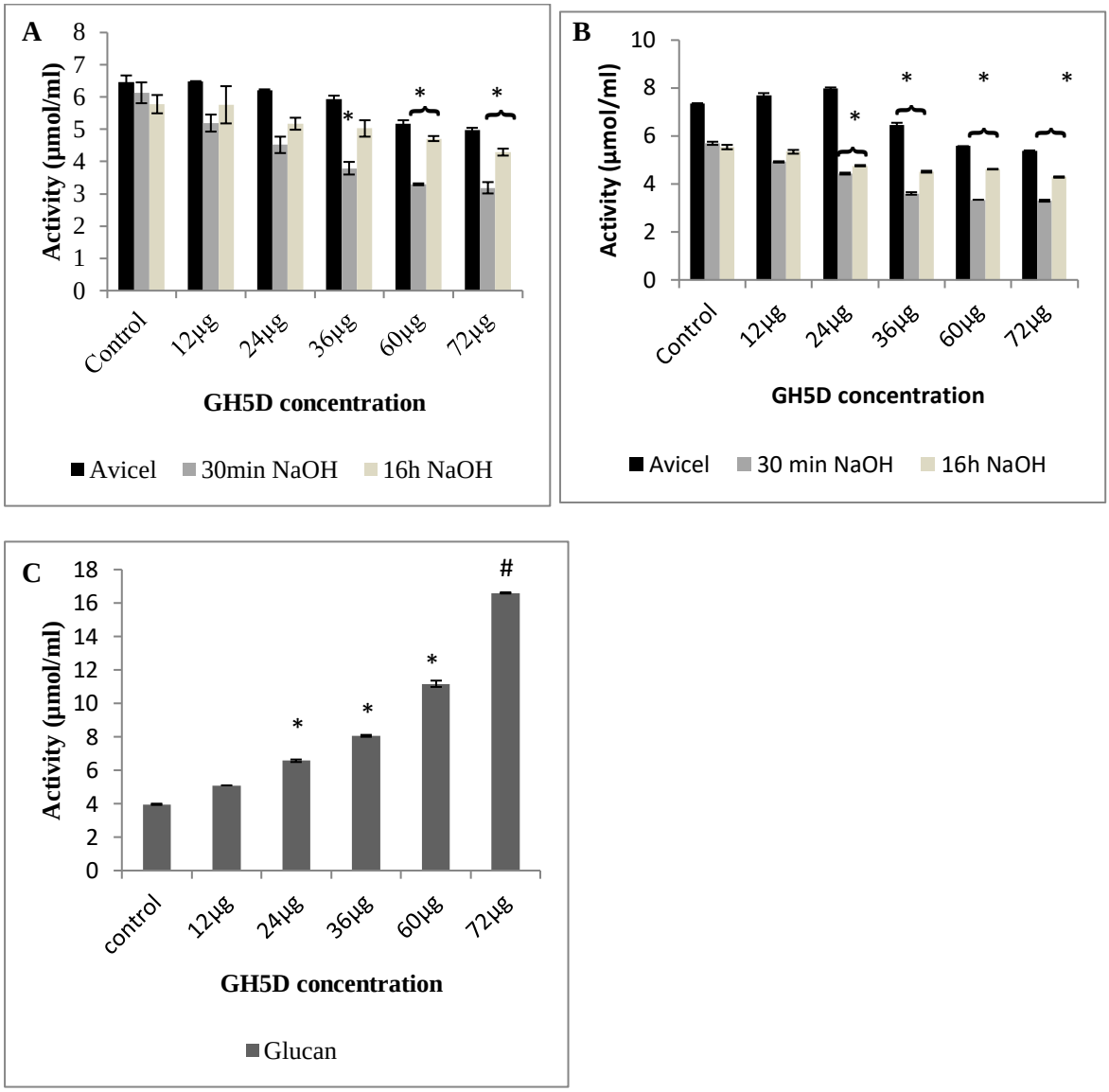


Figure 5.6. Analysis of the effects of GH5D on the fixed 50%CBHI: 50%CBHII core enzyme set protein loading, when hydrolyzing untreated and sodium hydroxide pretreated Avicel. The hydrolysis products, total reducing sugars (A) and glucose (B) were analyzed respectively. The effects of GH5D on the insoluble Azo-beta-glucan from barley was also analyzed (C). The concentration of the CBH core enzyme set was kept constant at 20 µl and that of GH5D was added in an increasing fashion from 12 µg to 72 µg. * represents p-value < 0.05 and # represents p-valuev < 0.01 The experiments were performed in triplicate and datapoints represent the means ± SD.

5.5. Conclusion

CBHI displayed higher activity on Avicel compared to CBHII and GH5D. Both CBHs displayed higher activity on the NaOH pretreated Avicel compared to PASC. CBHII displayed the highest activity on β-glucan, followed by GH5D. The CBHs only displayed synergy when they hydrolyzed NaOH pretreated Avicel. However, GH5D acted anti-synergistically with the CBHs during the hydrolysis of NaOH pretreated Avicel. In contrast, CBHII and GH5D displayed synergy on β-glucan as substrate. Increasing concentrations of GH5D decreased the activity of the CBH50%: CBHII50% core enzyme set during pretreated Avicel degradation, but enhanced the activity of the CBH core enzyme set during β-glucan hydrolysis. We therefore conclude that GH5D and CBHII are β-glucan specific cellobiohydrolases.

Chapter 6: Synergism of Glycosyl Hydrolases derived from termite Metagenome

6.1. Introduction

Cellulose is an abundant bio-polymer on earth and is composed of linked glucose residues (Saini et al. 2014; Lee et al. 2016). In agricultural feedstocks, hemicellulose is generally the second most abundant bio-polymer and is composed of heterologous (hexose and pentose) sugar residues (Saha, 2003). In most plants species, the major component of hemicellulose is xylan. The third most abundant polymer in the biomass is lignin (Hendriks and Zeeman 2009; Zhao et al. 2012).

The hydrolysis of a defined cellulose or hemicellulose substrate or a complex substrate such as lignocellulose requires a different combination of GH enzymes (Van Dyk and Pletschke 2012). The hydrolysis of a defined crystalline cellulose substrate requires the combination of an endo-glucanase and an exo-glucanase. β-glucanase is normally added to the endo and exo-glucanase mixture at constant concentrations to prevent product inhibition (Yang et al. 2014; Gupta et al. 2016). Synergism between cellulases and hemicellulases was also reported during the hydrolysis of lignocellulosic biomass (Olver et al. 2011; Yang et al. 2014). The hydrolysis of lignocellulosic biomass also requires auxiliary enzymes such as expansins, swollenins and AA9 lytic polysaccharide monooxygenases (LPMOs) which act in synergy with cellulases (Gupta et al. 2016).

Van Dyk & Pletschke (2012) argued that the synergy formed between cellulases or between cellulases and hemicellulases on a defined polymeric substrate is important for confirmation of synergism between enzymes, especially purified enzymes. However, they also highlight that enzyme synergy for lignocellulosic hydrolysis must be optimized on actual complex substrates such as agricultural feedstocks. Bai et al. (2013) used Avicel that was treated to produce regenerate amorphous cellulose (RAC), to study the synergism between CBH1 and CBH48 endoglucanase. CHB1 and CBH48 formed synergy and hydrolyzed biomass better in combination.

6.2. Aims and objectives

6.2.1. Aims

To construct a core enzyme set with GH5D, GH5E, GH5H and GH45 that could efficiently hydrolyze pretreated SSB and CC biomass. To apply this core enzyme set in a simultaneous, sequential or successive manner and elucidate which form of synergy (simultaneous, sequential or successive synergy) effectively hydrolyzes the biomass.

6.2.2. Objectives

1. To optimize synergy between the cellulolytic enzymes GH5E, GH5H and GH45 on CMC.
2. To optimize synergy between the xylanolytic enzymes GH5H and xylosidase for hydrolysis of insoluble wheat arabinoxylan.
3. To hydrolyze the alkaline and oxidative pretreated SSB and CC biomass with individual enzymes GH5D, GH5E, GH5H and GH45.
4. To establish a core enzyme set with GH5H, GH5E, GH5D and GH45 for hydrolysis of alkaline pretreated SSB and CC biomass.
5. To investigate which form of synergy- simultaneous, sequential or successive synergy- could hydrolyze the biomass efficiently.

6.3. Materials and Methods

6.3.1. Enzyme synergism on defined substrates

Enzyme synergism was performed with the polymeric substrates CMC and insoluble wheat arabinoxylan. The rationale for using defined substrates was to establish if the enzymes could form synergy during the hydrolysis of the substrates. The reactions were performed by mixing the cellulolytic and xylanolytic enzymes in different ratios (25%, 50%, 75% and 100%). The enzyme loading was kept at a constant concentration of 5 µg in the reaction mixture that contained 2% (w/v) of CMC, while insoluble xylan was kept at 1% (w/v). Both substrates were dissolved in citrate buffer at pH 5.5.

6.3.2. Synergy assays for the cellulolytic core enzyme set

The reactions performed with the cellulolytic core enzyme set (binary- or ternary-synergy and the enzyme load was kept at 30 µg/mg biomass) were allowed to run for 30 min at 37°C using CMC as substrate. To optimize binary-synergy combinations, the multifunctional GH5H and GH45 enzymes were mixed in different ratios (25%, 50%, 75% and 100%). The

optimized binary-synergy enzyme set was then used to form a ternary-synergy set in combination with GH5E. Again, different ratios of optimized binary-synergy mixture and GH5E were used to optimize the ternary-synergy cellulolytic core enzyme set. The binary- and ternary-synergy reactions were supplied with additional GH5D and β-glucosidase (Novozyme 188) at constant concentrations of 30 µg/mg biomass and 20 µg/mg biomass, respectively. GH5D was added at a constant concentration, because it hydrolyzed cello-oligomers, but failed to hydrolyze CMC (Chapter 3). We used GH5D as an ancillary enzyme together with β-glucosidase. To test the effects of GH5D as an ancillary enzyme, the optimized ternary-synergy combination of GH5E, GH5H and GH45 was supplemented with 12.13, 24.26, 36.4, 48.53 and 60.66 µg/mg biomass of GH5D, respectively. This reaction was referred to as quaternary-synergy. The total reducing sugars were measured by the standard DNS method and glucose was measured using a Megazyme kit (GOPOD kit).

6.3.3. Synergy assays with the xylanolytic core enzyme set

For xylanolytic synergy, GH5H and xylosidase from *Selenomonas ruminantium* were mixed in different ratios (25%, 50%, 75% and 100%) to hydrolyze insoluble wheat arabinoxylan. The total reducing sugars were measured according to DNS method and xylose was measured using a Megazyme kit (Xylose Kit).

6.3.4. Synergy studies on the complex substrate

To perform enzyme synergy assays, four enzymes [GH5D (exo-glucanase), GH5E (endoglucanase), GH5H and GH45 (both multifunctional enzymes)] that were characterized in Chapter 4 were used. The synergy assays were performed by dissolving 1% (w/v) of untreated, lime, NaOH and $NaClO_2$ pretreated biomass in 50 mM sodium citrate buffer at pH 5.5 and the reaction was carried out at 37°C for 48h. The core enzyme sets were formed using four enzymes, GH5D, GH5E, GH5H and GH45. The enzyme ratios used to form the core enzyme sets were 60%:20%:10% or10%. Out of 12 combinations, only the two best core enzyme sets [Set-E (GH5E 60%: GH5H 20%: GH44 10% and GH5D10%) and Set-H (GH5H 60%: GH5E 20%: GH44 10% and GH5D10%)] displayed higher activities. The best core enzyme sets were then used to perform simultaneous, successive and sequential synergy assays. The release of the total reducing sugars was measured by using DNS method and glucose was used as a suitable standard.

6.3.5. Data analysis

One way analysis of variance (ANOVA) was used to analyse the activity of the enzymes. Evaluation was conducted to elucidate significant increases exhibited by the core enzyme sets

[Set-E (GH5E 60%: GH5H 20%: GH44 10% and GH5D10%) and Set-H (GH5H 60%: GH5E 20%: GH44 10% and GH5D10%)]. All pairwise comparison procedures were conducted using the Data analysis feature in Microsoft Excel.

6.4. Results and discussion

Synergy studies were performed to establish an optimized cellulolytic core enzyme set optimized for CMC and a xylanolytic set optimised on insoluble wheat arabinoxylan. The purpose of these studies was to demonstrate the potential versatile industrial application of these purified glycosyl hydrolases from the termite gut metagenome. The ability of the enzyme core sets (GH5E, GH5H and GH45) to produce glucose was demonstrated on the amorphous polymeric substrate, CMC, while GH5H's ability to hydrolyze insoluble wheat arabinoxylan was demonstrated by the formation of synergy with β-xylosidase. In addition, a lignocellulolytic core enzyme set was then formed to hydrolyze the alkaline pretreated SSB and CC substrates.

6.4.1. Cellulolytic core enzyme set

The two multifunctional enzymes, GH5H and GH45, were mixed in the following ratios: 0:100, 25:75, 50:50, 75:25 and 100:0% protein loading to hydrolyze CMC. The best combination was 25% GH5H and 75% GH45. The mixture containing 25% GH5H and 75% GH45 (binary combination) was then mixed with GH5E to form a ternary combination. The best combination was found when GH5E was at 25% and the binary-set mixture concentration was at 75% (Figure 6.1).

The best combination released significantly higher amounts (p-value < 0.04) of total reducing sugar and glucose (1425.78 and 214.85 µg/ml/h) compared to 100% GH5E which only released about 1259.18 and 150.08 µg/ml/h, respectively. The total concentration of protein loaded in each of the binary- or ternary-synergy assays was 5 µg. GH5D and β-glucosidase (Novozyme 188) were added at a constant concentration of 30 and 20 µg respectively, to reduce product inhibition in both the binary- and ternary-synergy studies. The ternary-synergy study results showed that the total reducing sugars and glucose amounts increased significantly when the optimum ternary-synergy enzyme combination was used. Based on this finding, we investigated if an increase in GH5D concentration could increase glucose production.

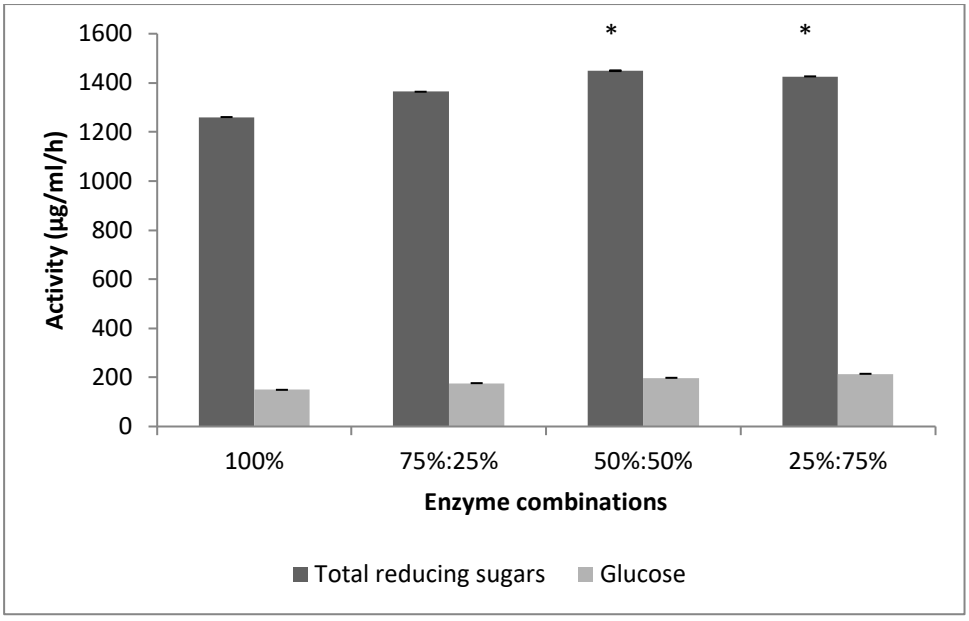


Figure 6.1. Ternary-synergy studies between GH5E and a mixture of multifunctional GH5H-25% and trifunctional GH45-75%. * represents p-value < 0.05. All experiments were performed in triplicate. The datapoints represent the means and the error bars represent the standard deviations.

Our previous results showed that GH5D could not hydrolyze the polymers, but was able to hydrolyze both cellotriose and cellopentaose to glucose and cellobiose (Chapter 4: section 4.4.1 and 4.4.3). We tested if an increasing concentration (12.13, 24.26, 36.4, 48.53 and 60.66 µg) of GH5D would improve its synergy with the optimized ternary-synergy set and subsequently improve glucose production from the polymeric substrate, CMC. This reaction was considered to be quaternary-synergy because of the different concentrations of GH5D added into the reaction. The assays were performed for 1 and 24 h, respectively. The amounts of total reducing sugar and glucose released after an hour significantly decreased (p-value < 0.03 for reducing sugar and p-value < 0.009 for glucose) as the concentration of the GH5D increased (Figure 6.2A). Interestingly, after 24h of CMC hydrolysis, the amounts of total reducing sugar and glucose were also decreasing (p-value < 0.03 for reducing sugars and p-value < 0.04 for glucose) as GH5D concentration added in a reaction was incresing (Figure 6.2B). The control (optimized ternary synergy core enzyme set without the addition of GH5D) amounts of glucose compared to amounts of total reducing sugars were not significantly different. The levels of glucose and total reducing sugars produced were reduced by more than 20% when 12.13 µg GH5D was added to the reaction and it continued to decrease significantly until more than 40% was reduced when GH5D (60.66 µg/mg biomass)

was added to the reaction. Therefore, we report for the first time that some enzymes (like GH5D) compete with ancillary enzymes, such as glucosidase, and reduce glucose and other sugar release.

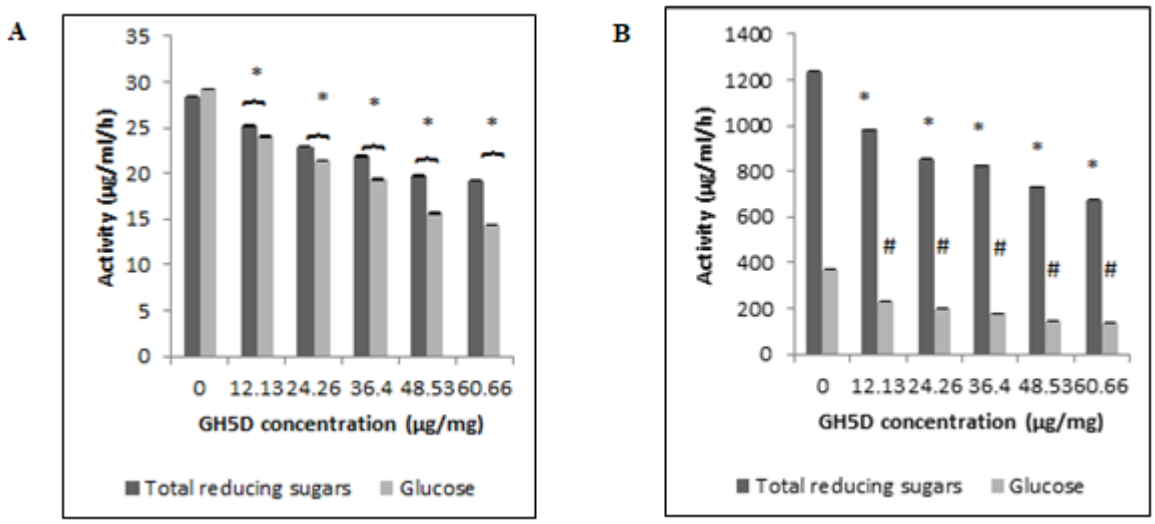


Figure 6.2. Quaternary-synergy studies and the effect of different concentrations of GH5D on the performance of the optimized core enzyme set (GH5E-25%, GH5H-25%, and GH45-50%). The 1 h assays are represented by A, while assays performed for 24h are represented by B. The dark gray bars in both graph A and B represent the total reducing sugars and light gray bars represents glucose amounts. The activity was expressed as the average activity of overtime period choosen, 1h and 24 h for A and B, respectively. * represents p-value < 0.05 and # represents p-value <0.01. All experiments were performed in triplicate. The datapoints represent the means and the error bars represent the standard deviations.

Van Dyk and Pletschke (2012) reported that the addition of a cellobiohydrolase increased the synergy between cellulosic enzymes. However, in our study, this was not the case. Instead, GH5D, which was classified as an exoglucanase, decreased glucose and total reducing sugar production. This was unexpected. However, our results showed that β-glucosidase could not access the cellobiose in the presence of GH5D (Figure 6.2). These observations suggest that GH5D inhibits the activities of both the core enzyme set and β-glucosidase (Novozyme 188). However, further studies are required to confirm whether or not the inhibition is due to GH5D competing for the substrate with the core enzyme set and glucosidase or if the total reducing sugars and glucose amounts decreased as a result of transglycosylation.

The absence of GH5D resulted in higher release of glucose and total reducing sugars from the hydrolysed biomass. This validated our hypothesis that the ternary-synergy enzymes (GH5E, GH5H and GH45) were effective hydrolases and could act synergistically on the biomass. Even though CMC is a polymeric model substrate, the optimized ternary-synergy catalyzed CMC and produced high amounts of total reducing sugars and glucose from this substrate. These glucose molecules could be fermented to produce biofuel or used in other industrial applications.

6.4.2. Xylanolytic synergy

The xylanolytic synergy study was performed using GH5H, a multifunctional enzyme with strong activity on both soluble and insoluble wheat arabinoxylan, and β- xylosidase. The aim of this experiment was to demonstrate that GH5H can produce xylo-oligosaccharides that can be converted to xylose in a synergy study. The enzyme ratios for the binary-synergy assays were as follows: 0:100, 25:75, 50:50, 75:25 and 0:100% of GH5H and xylosidase, respectively.

Huy et al. (2015) studied synergy between a xylosidase and a xylanase and showed that the synergy assays released higher amounts of total reducing sugars from barley compared to the individual enzymes. Our studies also show that synergy between GH5H and xylosidase significantly (p-value < 0.04) improved the release of total reducing sugar and xylose compared to the individual enzymes (Figure 6.3). Significantly high levels of total reducing sugar were released by the most effective enzyme synergy ratio of 75% GH5H and 25% xylosidase, while significantly high amounts of xylose were released by a combination of 50% GH5H and 50% xylosidase. Ravn et al. (2016) showed that xylanases were important for the removal of the "anti-nutritional effects of arabinoxylan". They demonstrated that the xylanase enzyme increased the amount of xylose and arabinose in wheat, rye and barley. We used insoluble wheat arabinoxylan as a substrate and found xylanase and xylosidase synergy effective at hydrolyzing the substrate. Thus, we suggest that GH5H could be well suited to the animal feed industry, especially if feedstocks like wheat, maize, barley or rye are used since they contain a high content of arabinoxylan. Furthermore, we propose that synergy assays are important for the removal of the "anti-nutritional effects" of arabinoxylan as we have demonstrated (Figure 6.3). Finally, our studies were focused on releasing xylose from the insoluble wheat arabinoxylan. Lei et al. (2016) studied the synergistic effects of xylanase, arabinofuranosidase and feruloyl esterase on arabinoxylan. The synergistic effects of xylanase and debranching enzymes were found to improve the release of xylo-

oligosaccharides. This observation is in agreement with our findings and hypothesis that synergy studies could improve the removal of "anti-nutritional effects" of arabinoxylan.

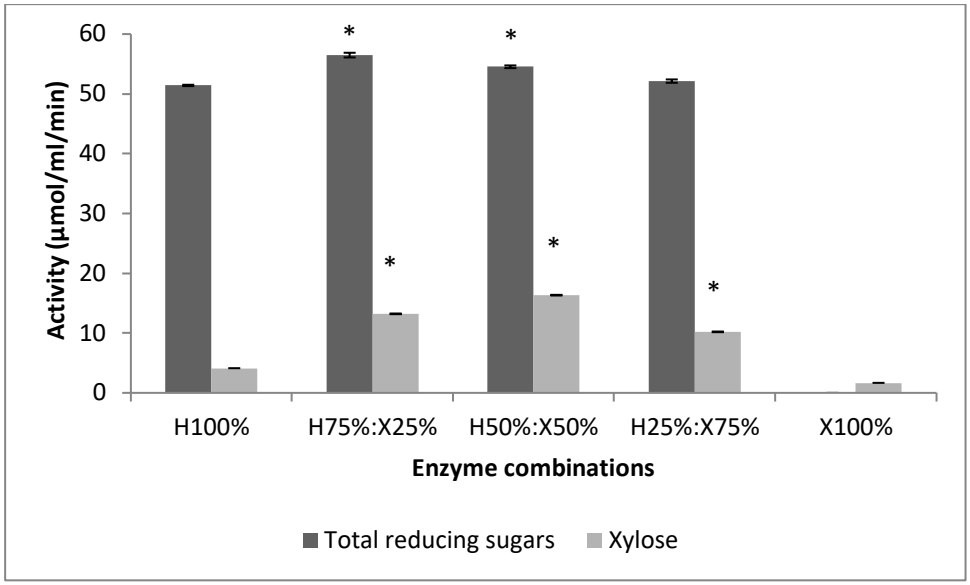


Figure 6.3. Synergism between multifunctional GH5H and xylosidase. The dark grey bars represent the total reducing sugars and the light grey bars represents xylose released from the hydrolysis of wheat arabinoxylan. The letters H and X represent GH5H and xylosidase, respectively. * represents p-value < 0.05. The error bars represent standard deviations and all experiments were performed in triplicate.

6.4.3. Lignocellulolytic synergy

The termite derived enzymes, GH5E, GH5H and GH45 displayed synergy on the polymeric substrate mentioned above. Based on these results, we investigated if these enzymes could form a lignocellulolytic core enzyme set that shows synergism on the alkaline or oxidative pretreated SSB and CC substrates. Application of the lignocellulolytic core enzyme set was tested in three different ways, with regards to simultaneous, successive and sequential synergy.

Simultaneous synergy reactions were defined as reactions wherein all the enzymes used to form a core enzyme set were added at the same time during the reaction. Sequential synergy assays were defined as the assays wherein one enzyme was added to the reaction at a time, allowed to hydrolyze CC or SSB for 12 h and then heat treated (denatured). Subsequent to heat treatment of the first enzyme, the second enzyme was added, and the same procedure that was applied to the first enzyme was employed to all the enzymes used in sequential assays. Successive synergy assays are when enzymes are added in a sequential manner, but

there is no heat treatment involved. Our focus was not to investigate which order of enzyme addition would work the best, but to investigate if there would be any changes to the total reducing sugar released, when we used the core enzyme sets simultaneously, successively or sequentially.

All the enzymes used in the current study were characterized previously (Chapter 4). All enzymes exhibited more that 80% relative activity at pH 5.5 and were all thermostable for more than 48 h at 37°C. Therefore, all the lignocellulolytic synergy studies were performed under these conditions.

The core enzyme sets were constructed by evaluating 12 different combinations that were arranged in the following ratios: 60%, 20%, 10% and 10%. This meant that each of the four enzymes (GH5D, GH5E, GH5H and GH45) was added once in a ratio of 60% or 20%, or 10% relative to others when the core enzyme set was constructed. Out of 12 combinations only two core enzyme sets were constructed in the following ratios, 60% GH5H, 20% GH5E, 10% GH44 and 10% GH5D (referred to as CES-H) and 60% GH5E, 20% GH5H, 10% GH44 and 10% GH5D (referred to as CES-E). Both the core enzyme sets were supplemented with 10% protein loading of *Aspergillus niger* β-glucosidase (Novozyme 188) and a *Selenomonas ruminantium* xylosidase, SXA. CES-E and CSE-H hydrolyzed the alkaline pretreated SSB and CC better than the oxidative pretreated SSB and CC (Figures 6.4 and 6.5). The CES-E displayed the highest activity (p-value < 0.009) on both NaOH and lime pretreated SSB. In addition the enzymes displayed higher activity when were used in a simultaneous manner, followed by successive and then sequential synergy (Figure 6.4A). These findings were expected because in the simultaneous synergy reactions, the enzymes worked together as a unit, but in both the successive and sequential synergy the assumption is that enzymes work independently (Sørensen et al. 2007; Han and Kim 2015). Figure 6.4 B shows that when enzymes used to form CES-E were added into the reaction in a sequential fashion (sequential synergy), they displayed higher activities on the CC samples compared to when they were added in a simultaneous and successive fashion. This was unexpected but it could be explained by the differences in the biomass samples used in the current study. Interestingly, CES-E displayed significantly higher activities (p-value < 0.02) on alkaline pretreated CC samples compared to the control/untreated samples.

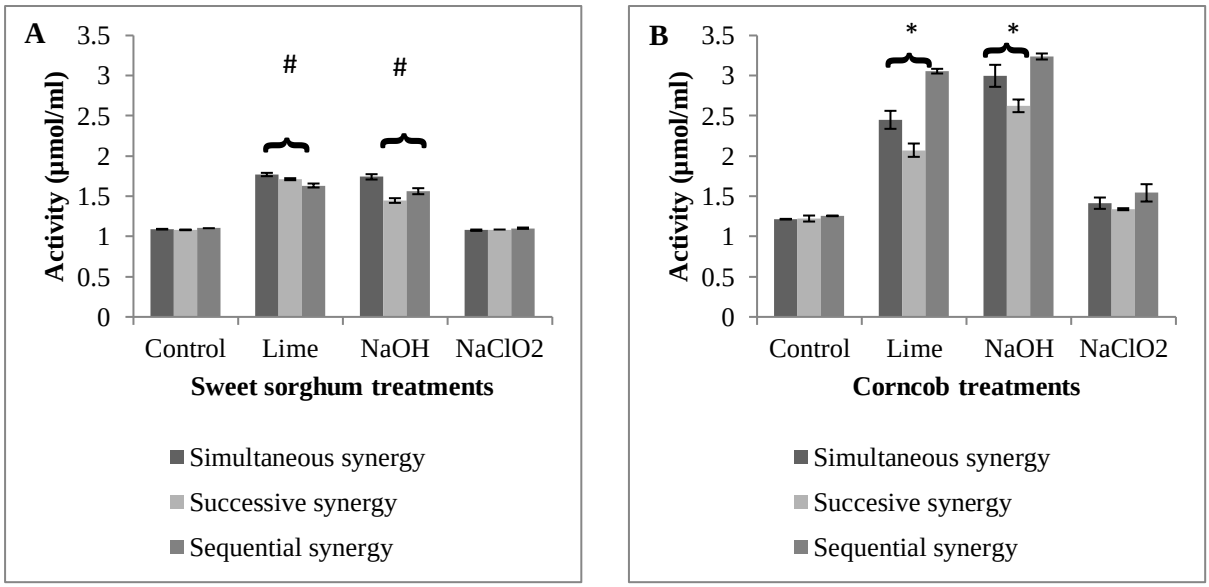


Figure 6.4. Synergism of termite derived enzymes on pretreated sweet sorghum bagasse and corncob. The amount of total reducing sugars (given as activity) produced after the sweet sorghum (A) and corncob(B) were hydrolyzed by the core enzyme set-E in a simultaneous, successive or sequential synergy manner. * represents p-value < 0.05 and # represents p-value <0.01. The error bars represent standard deviations and all experiments were performed in triplicate.

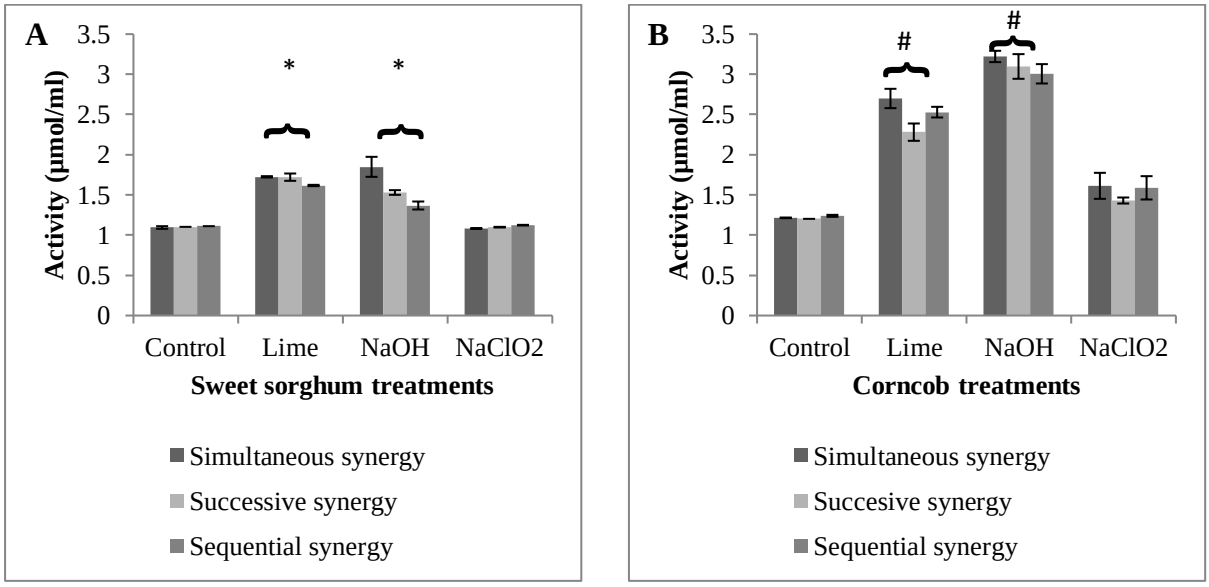


Figure 6.5. Synergism of termite derived enzymes on pretreated sweet sorghum bagasse and corncob. The amount of total reducing sugars produced after the sweet sorghum (A) and corncob (B) were hydrolyzed by the core enzyme set-H in a simultaneous, successive or sequential synergy manner. * represents p-value < 0.05 and # represents p-value <0.01. The error bars represent standard deviations and all experiments were performed in triplicate.

On alkaline pretreated SSB, the activity was higher with about 1.5-fold compared to control the (Figure 6.4A). There was no significant difference between the hydrolysis of control and the oxidative pretreated biomass by CES-E. These results show that CES-E behaves differently when it hydrolyzed the pretreated CC compared to pretreated SSB biomass. On the pretreated CC the sequential synergy out-performed simultaneous and successive synergy, while on pretreated SSB biomass, the simultaneous synergy out performed sequential and successive synergy.

In contrast, CSE-H did not behave differently when it hydrolyzed CC and SSB, as only simultaneous synergy exhibited the highest activity compared to both successive and sequential synergy (Figure 6.5). The similarity between the CES-H and CES-E sets was that

they hydrolyzed the CC better than the SSB and we suggest that this could be due to their structural differences (corncobis part of the fruit section of the plant, while sweet sorghum bagasse is mostly stem and leaves). CES-H displayed significantly higher activities on alkaline pretreated CC (p-value < 0.009) and SSB (p-value < 0.03) compared to the control. About a 1.8-fold increase was observed compared to control when the core enzyme set-H hydrolyseed the alkaline pretreated SSB biomass and dissplayed a 3-fold increase when it hydrolysed alkaline pretreated CC biomass . However, the activities with the $NaClO_2$ pretreated SSB and CC were not significantly different from those of the controls.

Álvarez et al. (2016) argued that physical and chemical pretreatments are essential to changes in the biomass composition. In addition, the pretreatments affect the hemicellulose and lignin structures. Thus, pretreatment directly impacts the enzyme loading used in the reaction. The current study is in line with this fundamental observation, because only alkaline pretreated CC and SSB samples were significantly hydrolyzed by the two constructed lignocellulolytic core enzymes sets. However, when hydrolyzing the $NaClO_2$ pretreated SSB and CC the core enzyme sets (CES-E and CES-H) did not show any significant differences relative to the controls. These findings are in agreement with those of Kumar et al. (2013) which stated that $NaClO_2$ makes the biomass more crystalline, which suggest that it will be more recalcitrant to enzyme hydrolysis.

In contrast, the use of alkali breaks down the ester and glycosidic side chains, resulting in the removal or structural alteration of lignin, biomass structural swelling and partial de-crystallization of the biomass (Brodeur et al. 2011). Our findings confirm this observation, Figure 6.4 and Figure 6.5 show that both alkali pretreated SSB and CC were degraded by the core enzyme set better compared than the control (untreated) and $NaClO_2$ pretreated SSB and CC samples.

6.5. Conclusion

The aim of the work presented in this chapter was to investigate the synergism between the purified the termite (*Trinervitermes trinervoides*) enzymes - GH5D, GH5E, GH5H, and GH45. The results showed that these enzymes displayed synergy when they hydrolysed the defined polymeric substrate CMC. In addition, GH5H displayed synergy with the commercial xylosidase during the hydrolysis of insoluble wheat arabinoxylan. Thus, we can conclude that these enzymes displayed a synergy on the defined polymeric substrate. Both synergy studies revealed that these enzymes, especially GH5H, may be attractive to the biofuel or feed

industry. In addition, the enzymes were used to construct a lignocellulolytic core enzyme set which was used to hydrolyse alkaline pretreated CC and SSB samples through three modes of synergy. We found that simultaneous synergy was the best for hydrolysis of alkaline pretreated SSB for both CES-E and CES-H. However, sequential synergy was the best when CES-E was used to hydrolyze alkaline pretreated CC, while CES-H applied simultaneously was the best for alkaline pretreated CC. All three modes of synergy were unable to improve the activities of the two core enzyme sets on $NaClO_2$ (oxidative) pretreated samples. We, therefore, conclude that lime and NaOH pretreatments are better than oxidative pretreatment for pretreating agricultural biomass samples.

Chapter 7: Synthesis of Alkyl Cellobiosides by GH5D Transglycosylation

7.1. Introduction

Alkyl glucosides and their derivatives are environmentally friendly compounds and industrially important non-ionic surfactants with high surface activity (Huy *et al*. 2013). Alkyl glucosides are also biodegradable and they have foaming control, wetting, detergent, and emulsifying properties. The chemical structures of alkyl-glucosides or –polyglucosides consist of aliphatic alcohols and glucose unit(s) obtained from feedstock biomass. Industry produces not pure alkyl monoglucosides but a complex mixture of alkyl mono-, di-, tri-, and oligo-glucosides and as a result, the industrial products are called alkyl polyglucosides (Rybinski and Hill 1998). The products are characterized by the length of the alkyl chain and the average number of glucose units linked to it, the degree of polymerization (DP).

The enzymatic synthesis of alkyl glucosides occurs through a process called transglycosylation. Activated glycosyl donors and an alcohol (as glycosyl acceptor) are used to generate a new glycosidic bond and water acts a competing nucleophile (Ito *et al*. 2007). A number of alkyl glucosides have been produced through enzymatic activity. These include, among others, methyl, ethyl, proponyl, butanyl, hexanyl or octanyl-D-glucosides (Ito *et al*. 2007; Kim *et al*. 2009; Dahiya *et al*. 2015). Bhatia *et al*. (2002) reported that β-glucosidase is one of the important enzymes responsible for synthesizing alkyl glucosides. Cyclodextrin glycosyltransferase is also known to synthesize the alkyl glucosides (Charoensapyanan *et al*. 2016).

Borodkin *et al*. (2016) reported that the GH1 beta glucosidase enzyme, with a $(\alpha/\beta)_8$ barrel structure, had transglycosylation activity. The authors also suggested that xylanase/cellulase (PDB 2HIS), a family 26 lichenase (PDB 2CIP), β-glucosidase (PDB 1UG6), Cel7A (PDB 4C4C), arabinofuranosidase (PDB) 2VRQ) and E-82 xylanase (PDB 2D24) are glycosyl hydrolase enzymes that also have transglycosylation properties. In this chapter, we investigated the role of the protein (enzyme) structure with regards to transglycosylation.

7.2. Aim and objectives

7.2.1. Aims

To use PRIMO for modeling the GH5D protein structure and to show that it has a tunnel-like active site. To synthesize alkyl cellobiose through GH5D transglycosylation reactions.

7.2.2. Objective

1. To model the protein structure of GH5D.
2. To investigate if GH5D can perform transglycosylation activity and produce alkyl cellobiosides.
3. To investigate if GH5D can produce variaty of the alkyl cellobiosides and identify them.
4. To investigate if the alkyl cellobiosides antibacterial activity of the.

7.3. Materials and Methods

7.3.1. Modeling GH5D enzyme structure

The GH5D peptide sequence was provided by Dr Rashamuse and the DNA sequence was uploaded to Genbank with the following number KT694321 (Rashamuse *et al*. 2016). The peptide sequence was then used as an input to search for the model template using PRotein Interactive MOdeling (PRIMO: https://primo.rubi.ru.ac.za/) software (Hatherley *et al*. 2016). The search for the template in PRIMO was set on default. The PRIMO automatically blasted proteins with a similar amino acid sequence to GH5D in the Protein Data Base (PDB). After selecting the preferred structure of PDB the PRIMO software then aligned the sequence using the T-COFFEE in 3D-COFFEE mode. The GH5D sequence was modeled using the PRIMO very slow refinement option which models the structure of the target protein using a selected PDB template. The quality of the structure was then tested by Protein Structure Analysis (ProSA) software (Wiederstein and Sippl 2017) and Verify3D software (Luthy *et al*. 1992). Finally, the conserved regions of the target and PDB template were investigated by aligning the sequences of the two proteins by multiple alignment sequence software (ClustalW).

7.3.2. Alkyl cellobiosides synthesis assays

The alkyl cellobiosides synthesis assays (through transglycosylation) were done using *p*-nitrophenyl cellobioside (*p*NPC) as a substrate. The standard reaction mixture was made up out of 5% alcohol (methanol, ethanol or propanol), 4 mM *p*NPC dissolved in 50 mM sodium citrate buffer (pH 5.5) and GH5D enzyme (10 µg/mg substrate). Thin layer chromatography (see Chapter 4: section 4.3.7 for detailed method) was used to confirm the production of

alkyl cellobiosides. After confirming the production of these compounds, the optimal conditions for producing the alkyl cellobiosides were investigated. Firstly, the effect of time on the production of alkyl cellobiosides was investigated between 1 and 6 hours. The effect of alcohol concentration on alkyl cellobioside production was measured between 1 and 30%, while the effect of the enzyme concentration was investigated between 6 and 24 µg. The optimal substrate concentration was determined between 0.5 and 4 mM. The alkyl glucosides produced during optimization was measured quantitatively by TLC and the *p*-nitrophenol released from the reaction was used to measure GH5D activity. The *p*-nitrophenol was measured as described in Chapter 4: section 4.3.3.

7.3.3. Antibacterial activity of alkyl cellobiosides

Filter paper discs (8 mm in diameter) were impregnated with a solution containing 2 mM of methyl, ethyl and propyl cellobiosides. The inoculum seed culture was grown until the OD reading at 600 nm reached 0.8. The seed culture of *Bacillus subtilis, Klebsiella spp* and *Staphylococcus aureus* (about 100 µl) were inoculated on LB agar plates and were allowed to adsorb for an hour at room temperature. An air-dried disc was placed on the surface LB agar containing *B. subtilis, Klebsiella* and *S. aureus*. *p*NPC and alcohol (methanol, ethanol and propanol) were used as appropriate controls. The plates were incubated at 37°C for 16 h. The clear zone of inhibition was calculated by measuring the diameter of the inhibition zone. The readings were taken in three different fixed directions of three different filter papers, and the average values of the inhibition zones were reported.

7.4. Results and discussion

7.4.1. GH5D structural modeling

The main aim of the work presented in this chapter was to demonstrate if the protein structure of termite metagenome derived exo-glucanase enzyme (GH5D) had a tunnel-like active site. Galactosidase is one of many GH enzymes that can perform transglycosylation reaction by using alcohols or sugars as acceptor molecules (Juers *et al*. 2012). Although galactosidase enzymes are tetramers, the third (central) domain structure proved to be a triose phosphate isomerase (TIM) or $(\alpha/\beta)_8$ barrel with the active site forming a deep tunnel at the C-terminal end of this barrel. These observations are the basis on which our hypothesis was derived: "The structure of GH5D has a tunnel-like active site, can perform transglycosylation by using alcohol as an acceptor molecule and form alkyl cellobiosides".

7.4.1.2. Peptide sequence analysis

PRIMO (https://primo.rubi.ru.ac.za/) software was used to search for the template that was used to model the protein structure of GH5D as described by Hatherley *et al*. (2016). GH5D peptide sequence blast in PRIMO resulted in 13 possible PDB templates (Figure 7.1). Two cellulases from family C (PDB ID: 1cec and 1 cen) had the highest sequence similarity of about 33% and 32% to GH5D, respectively. Both had coverage of 85% starting from the 41st to 360th amino acid and 1cec had a resolution of about 2.15 Å, while 1cen had a resolution of about 2.30 Å. The other 11 hits were divided into 5 thermophilic cellulases from GH5A (PDB ID: 3NCO, 1VJZ, 3MMU, 3AOF and 3AMG) and 6 bifunctional cellulases (PDB ID: 3RJX, 3RJY, 4U3A, 1VJZ, 4U3A, 4U5I, and 5BYW). 1Cec was selected as the template because of its particular features. Sequence comparison is a tool used to classify GH enzymes to different family and subfamilies (Domınguez *et al*. 1996). To identify and classify these enzymes, only a few strictly conserved amino acid positions within the entire family are used. For instance, Domınguez *et al.* (1996) reported that a GH5 endo-glucanase enzyme from *Clostridium thermocellum* has conserved amino acid at invariant positions which included the catalytic residues Glu140 and Glu280, and the polar planar groups of Arg46, His90, Asn139, His198 and Tyr200. In the current study, we performed sequence alignment of 1cec and GH5D. The results indicated that the peptide sequence of GH5D was longer than the peptide sequence of 1cec with about 45 amino acid residues (Figure 7.1).

GH5D had a similar amino acid sequence with 1cec at the conserved invariant positions, except that GH5D amino acids positions were affected in that GH5D's peptide sequence was longer than the sequence of 1cec. GH5D's Arg, His, Asn, Asn, Glu, and Glu amino acids were positioned at 81, 122, 183, 43, 196 and 225 on the peptide sequences, respectively. In comparison, the 1cec's Arg, His, Asn, Asn Glu and Glu amino acid were positioned at 46, 89, 140, 8, 153 and 281 on the peptide sequences, respectively. Thus, the peptide sequence comparison confirmed that GH5D enzyme is indeed a glycosyl hydrolase that belongs to family 5. Secondly, residue Arg46 is buried at the bottom of the active site cleft and has hydrogen bonding interactions with Asn9, Glu136, Asn139, His198 and Glu280 both in the complexed and unliganded structures (Dominguez *et al*. 1995; Domınguez *et al*. 1996). The authors argued that these five positions were either invariant or highly conserved in family 5 cellulases and they proposed that Arg46 plays an important structural and functional role in the 1cec enzyme. These findings suggest that 1cec is a good template for modelling the

structure of GH5D because important amino acids were conserved between them, including those that form an important part of the 3D protein structure.

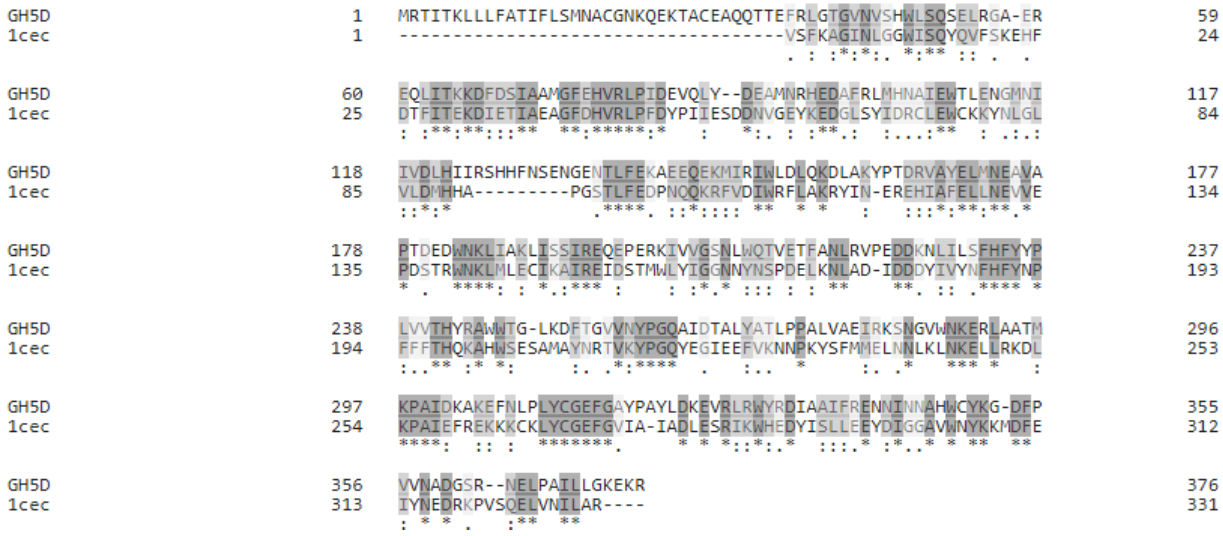


Figure 7.1. Amino acid sequence alignment of the termite metagenome derived GH5D enzyme and *Clostridium thermocellum* endo-glucanase (1cec). Different shades of gray shows similar amino acids in both the enzymes, peptide sequences; the numbers represent the position of the amino acids.

7.4.1.3. Structural analysis

The structure of the 1cec enzyme was determined by X-ray crystallography and the resolution was reported to be 2.15Å (Dominguez *et al.* 1995; Domınguez *et al.* 1996). The structure of 1cec was shown to be a cylindrical $(\alpha/\beta)_8$ barrel with two β-bulges at strands 3 and 7 (Figure 7.2 A). The enzyme's active site, which is situated on the carboxyl-terminal of the barrel, was located in the acidic tunnel-like structure. This tunnel-like structure was found to be a deep substrate-binding cleft (tunnel) and is formed by 54 amino acids adjacent to the active site cleft of the 1cec enzyme. These amino acid segments fold into a distinct sub-domain consisting of four α-helices and a short two stranded β-structure (Dominguez *et al.* 1995). The four α-helices and short two stranded β-structure are inserted between strand 6 and helix 6 of the α/β barrel and form a tunnel-like structure.

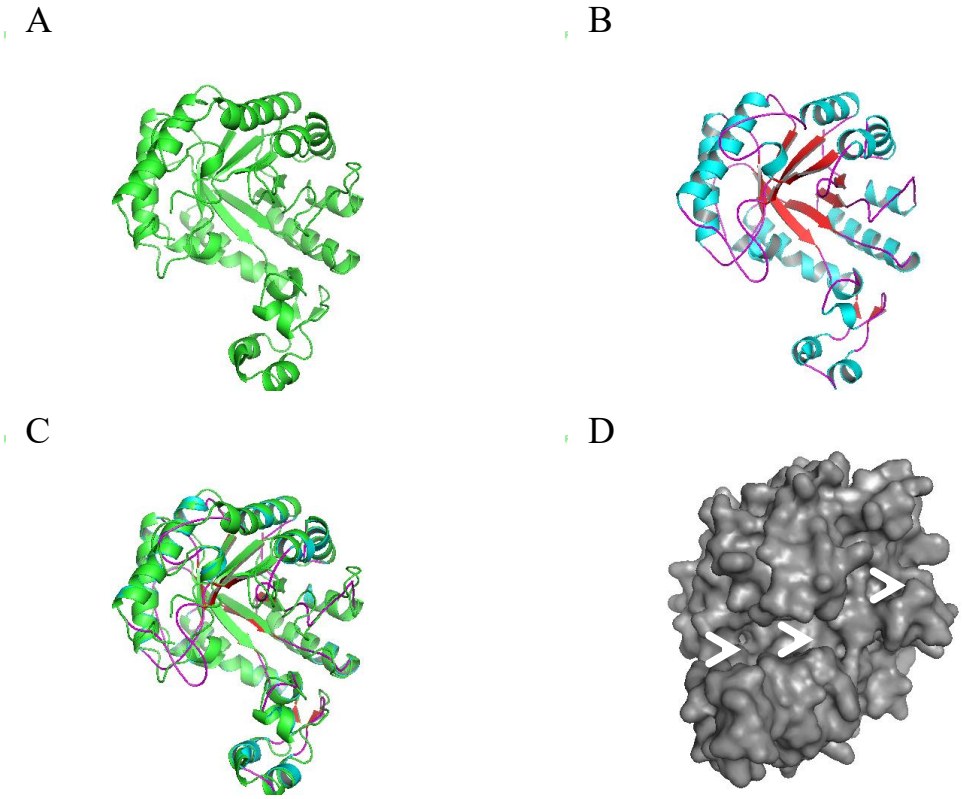


Figure 7.2. Protein structure of templete (A) endoglucanase (with potein database ID: 1CEC) used to model the structure of GH5D (B). The protein stractural alignment of 1cec and GH5D (C) and shows the surface topology of GH5D (D). The GH5D structure red, blue and purple colours represent α-sheets, β-helix and loops, respectively. The tunnel-like structure is indicated with white arrows.

The first enzyme to show the $(\alpha/\beta)_8$ barrel was triose phosphate isomerase. In addition, the $(\alpha/\beta)_8$ barrel was also reported in various glycosyl hydrolases such as β-galactosidase, neuraminidase, α and β-amylases, lichenases, xylanase and chitinases (Dominguez *et al.* 1995; Ducros *et al.* 1995; Domınguez *et al.* 1996; Juers *et al.* 2012). Some of these enzymes are notorious for performing transglycosylation when conditions are optimized for such a reaction. In particular, β-galactosidases and β-glucosidases are well known transglycosylases.

The GH5D protein structure was modelled with PRIMO software and PyMol software was used to show the cartoon form and the surface of the enzyme. Figure 7.2 B shows the cartoon form of the GH5D protein. The protein structure of the GH5D displayed a $(\alpha/\beta)_8$ barrel and

the alignment of GH5D and 1 cec showed that both enzymes displayed $(\alpha/\beta)_8$ barrel protein structure. The modelled protein structure of GH5D was then verified with ProSA and Verify3D software. Verify3D software showed that more than 80% of the amino acid residues had a score of less and equal to 0.2 in the 3D/1D profile. In addition, the ProSA Z-score for GH5D modeled protein structure was -8.2. Both the ProSA and Verify3D scores confirmed that the modelled GH5D protein structure is of good quality and it is possible to have such a protein in nature. The GH5D protein structure surface topology showed that GH5D active site is within a tunnel-like structure (Figure 7.2D). These findings confirm our hypothesis that GH5D is a protein which has a tunnel-like the active site.

7.4.2. Alkyl cellobioside synthesis

In the section above, we showed that the GH5D active site is situated within the tunnel-like structures. In addition, we hypothesized that enzymes with the active sites within tunnel-like structures, including GH5D, can perform transglycosylation by using the alcohol as an acceptor molecule to form alkyl cellobioside". In this section, we show that GH5D is able to perform transglycosylation reactions and that it produces different alkyl cellobiosides.

7.4.2.1. Alkyl cellobioside synthesis conditions optimization

The activity of the GH5D enzyme was tested using *p*NPC as a substrate. The results showed that the enzyme was active and that it displayed a specific activity of about 160 U/mg protein. After establishing that GH5D can hydrolyze *p*NPC we attempted to optimize the conditions which would allow the enzymes to perform trans-glycosylation. Firstly, the effect of alcohol concentration on GH5D activity was measured using different concentrations (1-30%) of methanol, ethanol and propanol, respectively. Low concentration of alcohol had less of an effect on the GH5D activity but higher concentrations of alcohol inhibited the enzyme (Figure 7.3.A). Compared to the control, 1% of methanol, ethanol or propanol reduced GH5D activity by more than 40%. In addition, 20% of methanol, ethanol or propanol reduced GH5D activity by more than 60%. Thus, 10% methanol and ethanol concentration were selected for further studies, while 5% propanol concentration was selected for further studies.

The effect of time on alkyl cellobioside production was investigated over 48 h and each reaction contained 10% methanol or ethanol and 5% propanol. Figure 7.3B shows that there was no significant difference in GH5D activity over time. The enzyme showed higher activity in the presence of 10% methanol followed by 5% propanol. In the presence of 10% ethanol, the enzyme showed lower activity compared to when the enzyme hydrolyzed *p*NPC

in the presence of the methanol and propanol. Based on these results it was concluded that the reactions should be conducted for less than 5 hours.

The third factor that was investigated was the concentration of GH5D enzyme. Different enzyme concentrations were used, but a lower concentration of about 6 µg exhibited the highest specific activity of about 200 U/ mg protein in the presence of 10% methanol while the addition of 12 µg protein concentration in the presence of ethanol or propanol resulted in the highest specific activity Figure 7.3. C. High enzyme loading did not improve the enzyme activity- instead, it reduced the activity of the enzyme.

Finally, the effect of the substrate concentration on GH5D activity was investigated. Figure 7.3D showed that GH5D activity increased with increasing substrate concentration. When 4 mM substrate was used the enzyme displayed higher activity (more than 100 U/mg protein) compared to the 0.5 mM where the enzyme displayed less than 12 U/mg activity. Interestingly, GH5D activity was more than 50 U/mg protein in all the reactions containing 2 mM in the presence of methanol, ethanol or propanol. Thus, 2 mM substrate concentration was chosen as an optimal substrate concentration.

Younis *et al.* (2012) also optimized the conditions for alkyl glucoside production. They argued that for optimizing the reaction for alkyl glucoside, is important to reduce the reaction time as well as the cost of the process. Hence, we have optimized the reactions for the production of methyl, ethyl and propyl cellobiosides, respectively. Figure 7.3 shows that our results were in agreement with Younis *et al.* (2012) regarding optimizing the reaction of alkyl cellobioside synthesis. For instance, GH5D was inhibited by alcohol concentrations which were higher than 10% (v/v) and its activity was comparable to that of a β-glucosidase which was inhibited by methanol, ethanol and propanol at concentrations above 15% (Christakopoulos *et al*. 1994).

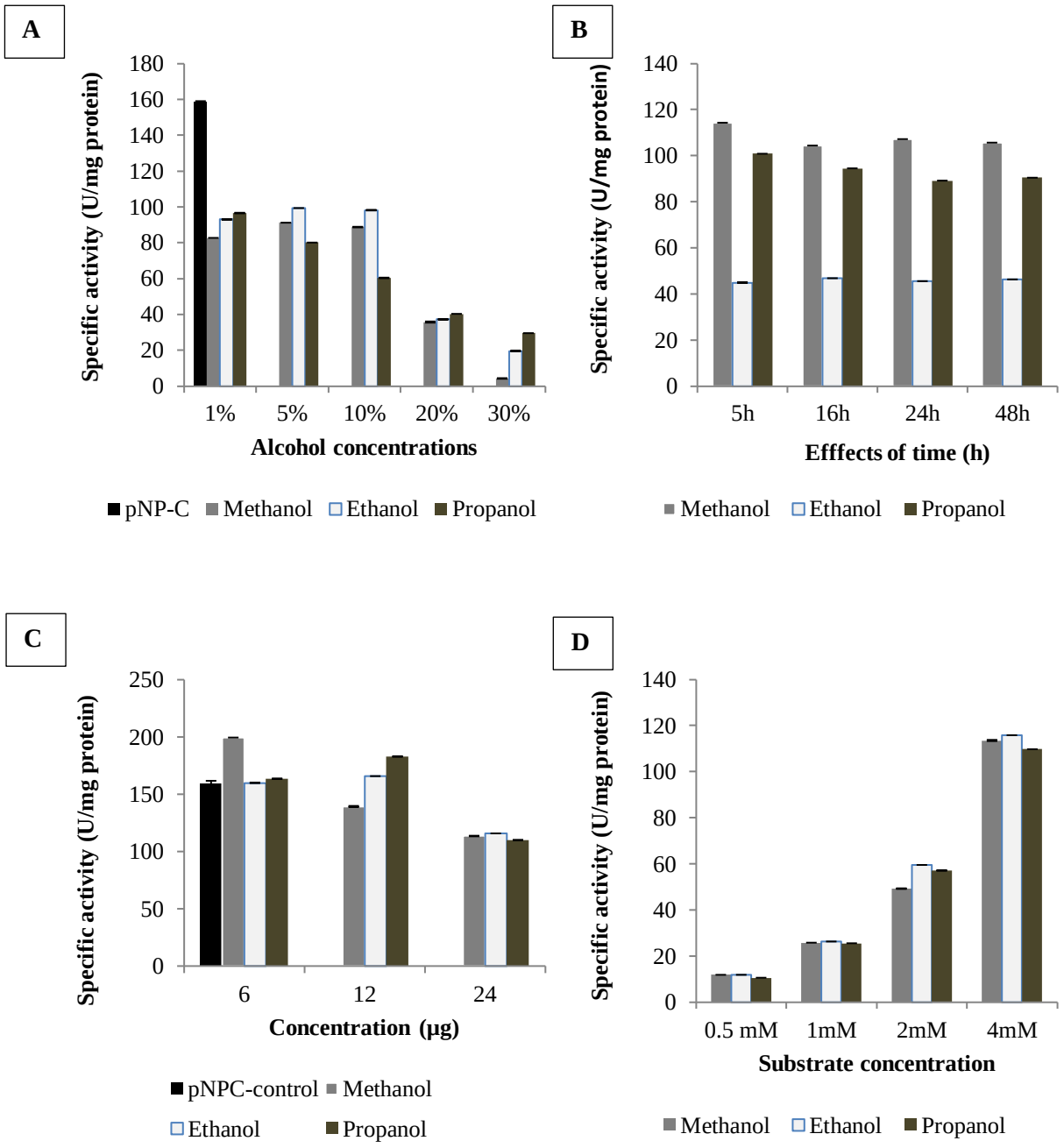


Figure 7.3. Optimization of alkyl cellobioside production. The effects of alcohol concentration, time, enzyme and substrate concentrations are shown in A, B, C and D, respectively.

7.4.2.2. Identification of the transglycosylation products

The reaction conditions for alkyl cellobioside production were optimized as described above. The products of each optimization reaction were identified with TLC. The results showed that

methanol and ethanol reduced the production of cellobioses and methyl or ethyl cellobiosides when they were added into the reactions at concentrations higher than 10% (v/v). Interestingly, when 1 -10% (v/v) alcohol was added to the reaction the cellobiose and methyl and ethyl cellobiosides were produced and the substrate (*p*NPC) was hydrolyzed completely (Figure 7.4A). The addition of propanol was inhibitory to GH5D compared to methanol and ethanol. GH5D completely hydrolyzed *p*NPC to cellobiose and propyl cellobioside when 1 and 5 % (v/v) propanol were added to the reaction, respectively. The TLC results were complementary to that of GH5D activity (Figure 7.3A).

In addition, TLC results showed that the GH5D hydrolyzed methyl, ethyl or propyl cellobioside to glucose and cellobiose after 12 h hours. For this reason, we decided to test the synthesis of alkyl cellobioside over a period of 6 h. Figure 7.4B shows that GH5D hydrolyzed *p*NPC completely and it produced methyl, ethyl and propyl cellobiosides, respectively. Dahiya *et al.* (2015) reported that production of the alkyl glycosides in their study was time dependent. They showed that alkyl glycoside production decreased after 12 h of the reaction. In addition, 6 µg of GH5D completely hydrolyzed *p*NPC and produced cellobiose, methyl, ethyl and propyl cellobiosides, respectively. However, when protein concentrations were increased to 12 and 24 µg, GH5D was not able to hydrolyze the substrate completely (Figure 6.4C). This phenomenon was also reported by Dahiya *et al.* (2015) - in their case the enzyme lost activity after 8 IU. Finally, Figure 7.4D shows that GH5D synthesized methyl, ethyl and propyl cellobioside optimally when the substrate concentration tested was the highest (4 mM) and when the substrate concentration was lower (0.5 mM) GH5D produced less alkyl cellobioside.

These results confirmed that our hypothesis that GH5D can perform transglycosylation. Interestingly, GH5D is able to synthesize methyl, ethyl and propyl cellobioside from *p*NPC and the mechanism of synthesis is described in the next section.

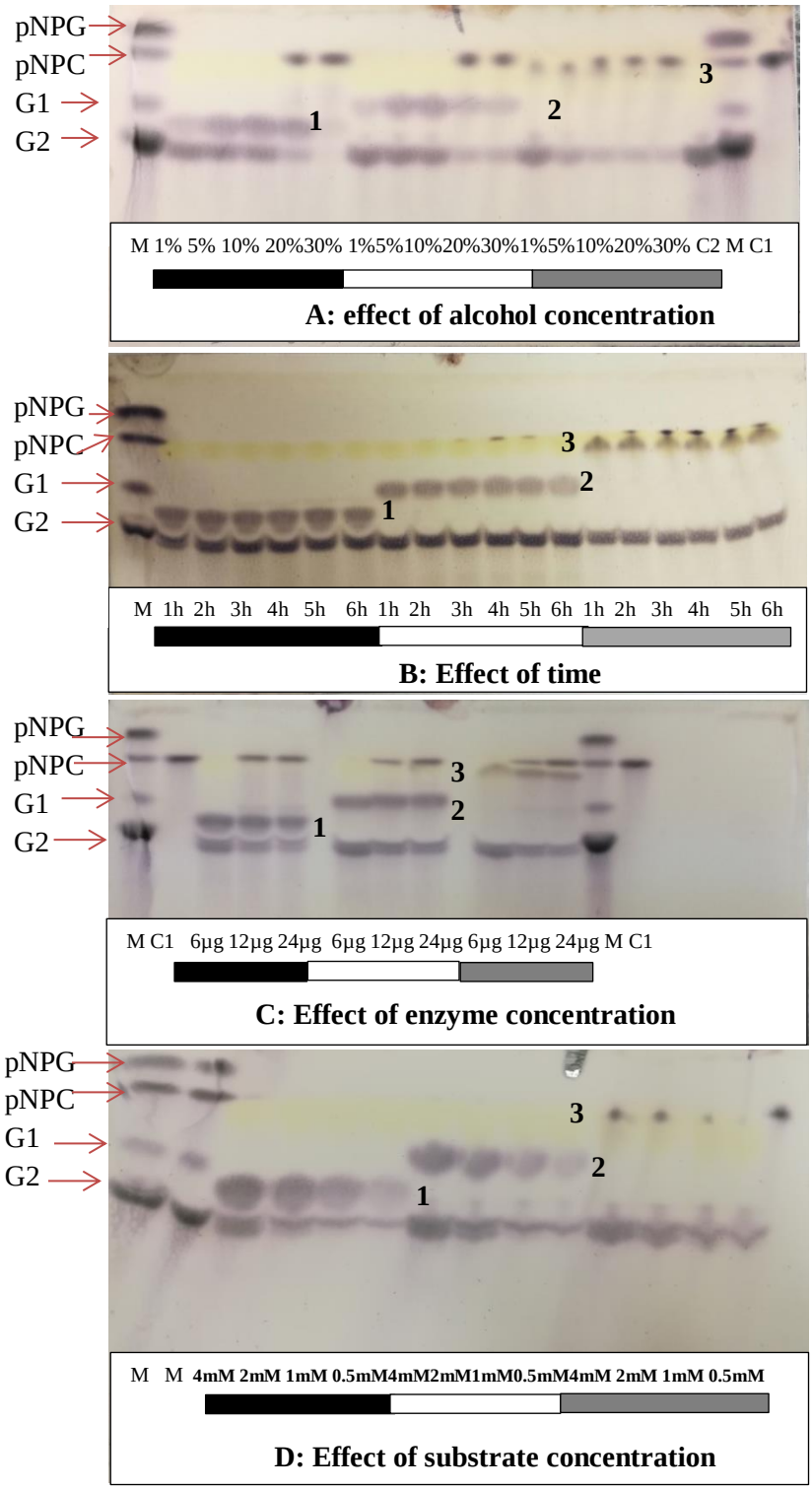


Figure 7.4. Analysis of alkyl cellobiosides produced through a GH5D trans-glycosylation reaction. The numbers 1, 2 and 3 represent methyl, ethyl and propyl cellobiosides, respectively. Para-nitrophenyl-glucopyranoside and cellobioside are represented by *p*NPG and *p*NPC. G1, G2, C1, C2 and M represent the glucose, cellobiose, substrate control, reaction control and marker, respectively. Black, white and gray bars represent the reactions that contained methanol, ethanol and propanol as acceptor molecules.

7.4.2.3. Mechanism of alkyl cellobioside synthesis

GH5D synthesized the methyl, ethyl and propyl cellobiosides through a two step reaction: Firstly, the enzyme hydrolyzed the *p*NPC and produced cellobiose and phenol. Secondly, GH5D used cellobiose produced from hydrolysis of *p*NPC as a glycosyl donor while the alcohol (methanol, ethanol or propanol) was used as an acceptor. The reaction whereby the cellobiose and alcohol molecules are joined to form alkyl cellobiosides is referred to as trans-glycosylation. The products of transglycosylation in our case were alkyl cellobiosides. Other studies have demonstrated that catalysts/enzymes such as a whole cell, β-glucosidase and cellobiose phosphorylase can produce alkyl glucosides (Kino *et al*. 2008; Dahiya *et al.* 2015; Boudabbous *et al*. 2017). To the best of our knowledge, CBH enzymes have not been reported to perform transglycosylation in the manner which GH5D does. Most of the enzymes hydrolyze polysaccharides or disaccharides to monosaccharide (especially glucose) first and then use monosaccharide as a glycosyl donor and alcohol as an acceptor (Rather and Mishra 2013).

7.4.2.4. Antimicrobial properties of alkyl cellobiosides

The antimicrobial properties of methyl, ethyl and propyl cellobiosides were tested on *Bacillus subtilis*, *Klebsiella spp* and *Staphylococcus aureus*. The antibacterial activity of the synthesized methyl, ethyl and propyl cellobiosides was determined by the disc diffusion method. The results were expressed as an average diameter of the inhibition zone. Figure 7.5 shows that no signicant inhibition of bacterial growth was found using alkyl cellobioside at the tested concentration. However, *Klebsiella spp* was inhibited by the ethyl and propyl cellobiosides (Figure 7.5A). The inhibition zone of the ethyl cellobiose was about 11 mm, while propyl cellobioside displayed inhibition zone of about 12.5 mm. In addition, methyl cellobioside did not inhibit the growth of *Klebsiella ssp*. Methyl and ethyl cellobiosides did not inhibit the growth of *B. subtilis*. The propyl cellobioside inhibited *B. subtilis*, displaying about an 11 mm inhibition zone on the agar plant (Figure 7.5B). The inhibition activity of methyl, ethyl or propyl cellobioside on the *S. aureus* was inconclusive because alcohol controls (methanol, ethanol and propanol) also inhibited *S. aureus* (results not shown).

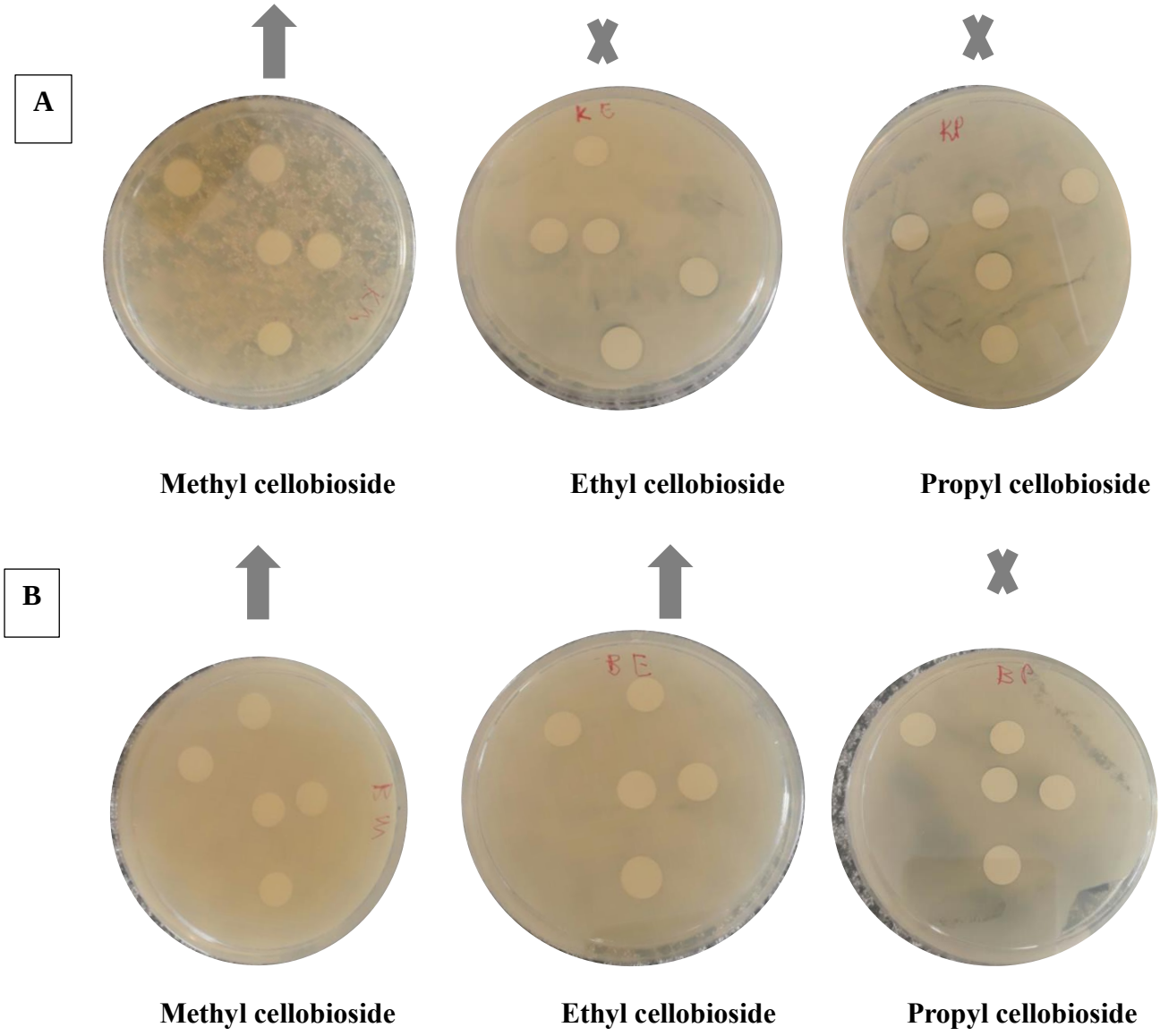


Figure 7.5. Antibacterial activity of methyl, ethyl and propyl cellobiosides tested against *Klebsiella* (A) and *Bacillis subtilis* (B). The symbols X indicate that the bacteria were inhibited and the arrow shows that the bacteria were not inhibited.

Generally, the alkyl glucoside or poly-glucosides have inhibitory properties and have been demonstrated to inhibit both Gram positive and Gram negative bacteria (Matsumura *et al*. 1990; Matin *et al*. 2014; Charoensapyanan *et al*. 2016). The inhibition of *E.coli* growth by propyl glycosides is attributed to their properties as non-ionic surfactants that can disrupt both the lipopolysaccharides and phospholipids in the Gram-negative bacterial cell wall. In addition, other alkyl glycosides, such as butyl glycoside and dodecyl mannoside, exhibited a broad spectrum of antimicrobial activity by inhibiting both Gram positive and Gram negative bacteria (Matsumura *et al*. 1990; Matin *et al*. 2014). These authors proposed that the antimicrobial activity resulted from the dominant role of alkyl glycosides in penetrating microbial cell membranes. We suggest that *B. subtilis* and *Klebsiella* growth was inhibited by the ethyl and propyl cellobiosides because of their properties, which include acting as non-ionic surfactants. These alkyl cellobiosides can penetrate the microbial cell membrane and denature this membrane.

7.5. Conclusion

The work presented in this chapter demonstrated that GH5D protein structure consists of a $(\alpha/\beta)_8$ barrel with the tunnel-like active site. It is reported in the literature that enzymes with this type of the protein structure is able to perform transglycosylation. In this chapter, we have confirmed that GH5D could perform transglycosylation, a process in which it produced methyl, ethyl and propyl cellobiosides. The alkyl cellobiosides are produced when the GH5D enzyme donates cellobiose glycosyl ends to alcohols (acceptors). The ethyl and propyl cellobiosides inhibited the growth of *Klebsiella* and *B. subtillis*, but methyl cellobioside did not inhibit any of the bacteria tested. Thus, we conclude that GH5D transglycosylation products (ethyl and propyl cellobiosides) have antimicrobial activity. In future we will investigate the surface activity, emulsifying and forming properties of these alkyl cellobiosides. We will use liquid chromatography mass spectrometry (LCMS) to quantify the amount of alkyl cellobiosides and try to scale up the reaction in a bioreactor for the production of these important compounds.

Chapter 8: General Discussion and Conclusion

The study presented in this thesis addresses some of the challenges facing glycosyl hydrolase research (such as the high pricing of GH enzymes, finding novel sources that possess highly active GH enzymes and using of these enzymes in synergy to release higher amount of simple sugars from the substrate). GH enzymes are a group of enzymes that are involved in carbohydrate conversion and metabolism. The proposed solutions to the challenges of high pricing of GH enzymes are to find novel sources of the GH enzymes and the most efficient ways of producing these enzymes cheaply. The current study used termite metagenome-derived enzymes endo-glucanase (GH5E), exo-glucanase (GH5D) and multifunctional enzymes (GH5H and GH45) to form the core enzyme sets CES-E and CES-H. Prior to the formation of the core enzyme sets, these enzymes were comprehensively characterized. The temperature optima of these enzymes varied; however, at 37°C all enzymes were thermostable for more than 48 h. The pH optima of all enzymes differed, but at pH 5.5 all the enzymes retained about 80% relative activity. We concluded that reactions using these enzymes should be conducted at a temperature of 37°C and a pH of 5.5 for 48 hours or not more than 72 hours.

Furthermore, we have concluded that GH5D is an exoglucanase and GH5E an endoglucanase with soluble wheat arabinoxylan activity. Both GH5H and GH45 proved to be xyloglucanase enzymes with cellulase and xylanase activities (i.e. multifunctional enzymes). These enzymes displayed novel features such as being stable in the presence of metal-ions and oligosaccharides. These features are important, because in industry the enzymes are used in the presence of some metal-ions. Enzymes that are stable in the presence of metal-ions are in high demand. We also propose that GH5H is a multifunctional enzyme with two catalytic centers. Thus, we suggest that further structural protein studies regarding the catalytic centers of GH5H will be important in understanding its hydrolytic mechanisms and whether each catalytic center interacts with the substrate of choice.

Given that GH5D is an exoglucanase enzyme, we assayed this enzyme together with commercial CBHI (Cel7) and CBHII (Cel6) enzymes and compared their abilities to hydrolyze crystalline cellulose. Because there are reports in literature where CBHs showed activity on β-glucan from barley, we conducted a comparative assay of the three enzymes on barley β-glucan. CBHI displayed higher activity on Avicel compared to β-glucan. Both CBHs displayed higher activity on NaOH pretreated Avicel compared to PASC. CBHII displayed

the highest activity on β-glucan, followed by GH5D. However, low activity of GH5D was observed on Avicel. The CBHs only displayed synergy when they hydrolyzed NaOH pretreated Avicel. This was not surprising, because commercial CBHs are kown to be effective on untreated or pretreated Avicel. In contrast, GH5D acted anti-synergistically with both CBHs on NaOH pretreated Avicel. Interestingly, CBHII and GH5D displayed synergy on β-glucan as substrate. To the best of our knowledge, this is the first report on synergy between a termite derived exoglucanase (GH5D) and a commercial CBHII on barley β-glucan. The degree of synergy confirmed that CBHII and GH5D displayed synergy compared to CBHI and GH5D on barley β-glucan. Furthermore, increasing concentrations of GH5D decreased the activity of the CBHI50%: CBHII50% core enzyme set during pretreated Avicel degradation but improved the activity of the CBH core enzyme set on β-glucan hydrolysis. We, therefore, concluded that GH5D and CBHII are β-glucan-specific cellobiohydrolases.

Prior to conducting synergy assays with GH5D, GH5E, GH5H and GH45, the SSB and CC biomass samples were pretreated with alkaline (NaOH and lime) and oxidative ($NaClO_2$). Both methods removed lignin from the SSB and CC biomass. Lignin removal exposes cellulose and hemicellulose; allowing enzymes degrading cellulose and hemicellulose to now access them. In addition, the removal of the lignin resulted in an increasing amount of hemicellulose and cellulose in both CC and SSB biomass. NaOH removed the highest amounts of lignin compared to both lime and $NaClO_2$. FTIR results showed that $NaClO_2$ did not remove all the lignin from the biomass. SEM showed that the morphology of all the pretreated CC and SSB biomass had changed. The surface of the pretreated CC and SSB had changed and the pores appeared as the lignin was removed. Thus, we concluded that both treatments removed lignin. However, NaOH was the best pretreatment method, because it removed most of the lignin compared to lime and $NaClO_2$.

The pretreated SSB and CC biomass samples were used to investigate synergism between the purified termite derived enzymes - GH5D, GH5E, GH5H, and GH45. Firstly, the results showed that these enzymes (GH5E, GH5H, and GH45) displyed synergy when they hydrolyzed the defined polymeric substrate CMC. In addition, GH5H displayed synergy with a commercial xylosidase during the hydrolysis of insoluble wheat arabinoxylan. Thus, we can conclude that these enzymes derived from a termite metagenome can display synergy on the defined polymeric substrate. Both synergy studies revealed that these enzymes, especially GH5H, may be attractive to the biofuel or feed industry.

In addition, GH5D, GH5E, GH5H, and GH45 were used to construct a lignocellulolytic core enzyme set which was used to hydrolyze alkaline or oxidatively pretreated CC and SSB through three modes of synergy. We found that simultaneous synergy was the best for hydrolysis of alkaline pretreated SSB for both CES-E and CES-H. However, sequential synergy was the best option when CES-E was used to hydrolyze alkaline pretreated CC, while CES-H applied simultaneously was the best for alkaline pretreated CC. All three modes of synergy were unable to improve the activities of the two core enzyme sets on $NaClO_2$ (oxidative) pretreated biomass. We therefore suggest that lime and NaOH pretreatments are the best for pretreating agricultural biomass over oxidative pretreatment.

Finally, the TLC results for GH5D indicated that this enzyme was capable of performing transglycosylation. Following on these observations, we comprehensively investigated the ability of GH5D to perform transglycosylation. Given that literature has shown that enzymes that are capable of performing transglycosylation reactions have a tunnel-like active site, we firstly modeled the protein structure of the GH5D. Its modeled protein structure was a $(\alpha/\beta)8$ barrel with a tunnel-like active site. Secondly, we confirmed that GH5D can perform transglycosylation, a process in which it produced methyl, ethyl and propyl cellobiosides. The alkyl cellobiosides were produced when GH5D donated cellobiose glycosyl ends to alcohols (acceptors). After confirming that GH5D could perform transglycosylation, we investigated whether the produced methyl, ethyl and propyl cellobiosides displayed anti-microbial properties. The ethyl and propyl cellobiosides inhibited the growth of *Klebsiella* and *B. subtillis*, but the methyl cellobioside did not inhibit any of the bacteria. Thus, we concluded that the GH5D transglycosylation products (ethyl and propyl cellobiosides) have antimicrobial activity.

We conclude that the *Trinervitermes trinervoides* metagenome-derived GH enzymes implemented in our study are novel enzymes with promising features for feed and biofuel industries. GH5H, a multifunctional enzyme that was highly active with a range of substrates (CMC, xylan and xyloglucan), and GH5D, an exoglucanase enzyme that could perform transglycosylation (producing methyl, ethyl and propyl cellobiosides), may be commercialized because of their rare properties. In addition, CES-E and CES-H enzyme core sets should be employed in a simultaneous fashion during hydrolysis of SSB biomass. During the hydrolysis of CC biomass the CES-E and CES-H enzyme core sets should rather be employed in a simultaneous and sequential fashion, respectively. Finally, methyl, ethyl and propyl cellobioside anti-microbial properties should, in future, be investigated using a

broader range of Gram-negative and Gram-positive bacteria. However, the preliminary studies showed that the ethyl and propyl cellobiosides were inhibitory to *Klebsiella* and *B. subtillis*.

Future studies should investigate the effects of a combination of feruloyl esterases (FAE's) and the CES-E enzymes core set, or FAE's and the CES-H enzyme core set, on SSB and CC biomass. These FAE enzymes should also be derived from the termite metagenome. The rationale for an addition of FAE to the CES-E or CES-H enzyme core set is based on the fact that CC and SSB contain ferulic acids which form part of the lignin structure. We, therefore, propose that FAE enzyme addition to both optimized enzyme core sets could improve the amounts of sugar released.

In addition, the surface activity, emulsifying and forming properties of methyl, ethyl and propyl cellobiosides produced by GH5D through transglycosylation should be investigated. The antimicrobial properties of these alkyl cellobiosides should be quantified (e.g. the IC_{50} values should be determined), because in the current study we could only qualitatively measure the zone of clearance. Future studies should determine the statistical significance of the effects of alkyl cellobiosides antimicrobial properties. Liquid chromatography-mass spectrometry (LCMS) should be employed to quantify the amounts of the alkyl cellobiosides and the reaction should be scaled-up in a bioreactor for the production of methyl, ethyl and propyl cellobiosides.

In summary, and to the best of our knowledge, this is the first study that demonstrated that GH5D, an exoglucanase derived from a termite metagenome, could not hydrolyse the established substrate for exoglucanases (CBHs) - Avicel. In addition, the polymeric substrates (CMC, xyloglucan and beechwood xylan) could not be hydrolysed by GH5D. However, GH5D displayed high activity on barley β-glucan. This observation suggests that GH5D is a β-glucan-specific enzyme. The Biochemical properties of GH5D, GH5E, GH5H and GH45 confirmed that the pH related properties of the enzymes were similar to those used in bioprocessing industries. However, the enzymes displayed a lower thermostability (37ºC). This low temperature preference could help conserve energy in the bioprocessing industry. In addition, this is the first study that constructed the termite metagenome-derived core enzyme sets (CES-E and CES-H) that could effectively break down SSB and CC biomass sample to simple sugars. The CES-E and CES-H core enzyme sets performed better on alkaline pretreated SSB and CC samples, compared to when they hydrolysed $NaClO_2$

pretreated SSB and CC samples. Lastly, GH5D produced methyl, ethyl and propyl cellobiosides through a transglycosylation reaction. These findings not only confirmed that termites are novel reservoirs of carbohydrate degrading enzymes, but also highlighted the contribution of termite derived enzymes to the ongoing development of new, improved enzymes for a diverse array of biotechnological applications.

References

Abidi N, Cabrales L, Haigler CH. Changes in the cell wall and cellulose content of developing cotton fibers investigated by FTIR spectroscopy. Carbohydr Polym.2014;100:9–16.

Abidi N, Cabrales L, Hequet E. Fourier transform infrared spectroscopic approach to the study of the secondary cell wall development in cotton fiber. Cellulose. 2010;17(2):309–20.

Abidi N, Hequet E, Cabrales L, Gannaway J, Wilkins T, Wells LW. Evaluating cell wall structure and composition of developing cotton fibers using fourier transform infrared spectroscopy and thermogravimetric analysis noureddine. J Appl Polym Sci. 2008;107:476–86.

Abot A, Arnal G, Auer L, Lazuka A, Labourdette D, Lamarre S, Trouilh L, Laville E, Lombard V, Henrissat B, O'Donohue M, Hernandez-Raquet G, Dumon C, Leberre VA. CAZyChip : dynamic assessment of exploration of glycoside hydrolases in microbial ecosystems. BMC Genomics; 2016;17(671):1–12.

Adapa PK, Schonenau LG, Canam T, Dumonceaux T. Quantitative analysis of lignocellulosic components of non-treated and steam exploded barley, canola, oat and wheat straw using fourier transform infrared spectroscopy. J Agric Sci Technol. 2011;1:177–88.

Ali SS, Wu J, Xie R, Zhou F, Sun J. Screening and characterizing of xylanolytic and xylose-fermenting yeasts isolated from the wood-feeding termite, *Reticulitermes chinensis*. PLoS Biol. 2017;12(7):1–21.

Álvarez C, Reyes-Sosa FM, Díez B. Enzymatic hydrolysis of biomass from wood. Microb Biotechnol. 2016;9(2):149–56.

Andrić P, Meyer AS, Jensen PA, Dam-Johansen K. Reactor design for minimizing product inhibition during enzymatic lignocellulose hydrolysis: I. Significance and mechanism of cellobiose and glucose inhibition on cellulolytic enzymes. Biotechnol Adv. 2010;28(3):308–24.

Arakawa G, Watanabe H, Yamasaki H, Maekawa H, Tokuda G. Purification and molecular cloning of xylanases from the wood-feeding termite, *Coptotermes formosanus* Shiraki. Biosci Biotechnol Biochem. 2009;73(3):710–8.

References

Aspeborg H, Coutinho PM, Wang Y, Brumer H, Henrissat B. Evolution, substrate specificity and subfamily classification of glycoside hydrolase family 5 (GH5). BMC Evol Biol. 2012;12(1):186.

Azuma J, Nishimoto K, Koshijima T. Studies on digestive system of termites. Wood Res. 1984;70.

Bai A, Zhao X, Jin Y, Yang G, Feng Y. A novel thermophilic β-glucosidase from *Caldicellulosiruptor bescii*: Characterization and its synergistic catalysis with other cellulases. J Mol Catal B Enzym. 2013;85–86:248–56.

Bastien G, Arnal G, Bozonnet S, Laguerre S, Ferreira F, Fauré R, Henrissat B, Lefèvre F, Robe P, Bouchez O, Noirot C, Dumon C, O'Donohue M. Mining for hemicellulases in the fungus-growing termite *Pseudacanthotermes militaris* using functional metagenomics. Biotechnol Biofuels. 2013;6(78):1–15.

Batista-Garcia AR, Sanchez-Carbente M del R, Talia P, Jackso SA, O'Leary ND, Dobson ADW, Folch-Mallol JL. From lignocellulosic metagenomes to lignocellulolytic genes : trends, challenges and future prospects. Biofuels, Bioproducts Biorefining. 2016;10:864–82.

Bensah EC, Mensah M. Chemical pretreatment methods for the production of cellulosic ethanol: technologies and innovations. Int J Chem Eng. 2013:21.

Beukes N, Pletschke BI. Effect of lime pre-treatment on the synergistic hydrolysis of sugarcane bagasse by hemicellulases. Bioresour Technol. 2010;101(12):4472–8.

Bhatia Y, Mishra S, Bisaria VS. Microbial β-Glucosidases : Cloning, Properties, and Applications. Crit Rev Biotechnol. 2002;22(4):375–407.

Bhattacharya A, Mafa M, Rashamuse K, Pletschke BI. A bi-functional xylanase/xyloglucanase, GH5H, from termite metagenome improves the hydrolysis of pretreated biomass. 2016;(1):213–6.

Bignell DE, Jones DT. A taxonomic index, with names of descriptive authorities of termite genera and species : An accompaniment to biology of termites : A modern synthesis. J Insect Sci. 2011;14(81):1–33.

References

Borodkin V, Fang W, Castro-lo J, Rovira C. A trapped covalent intermediate of a glycoside hydrolase on the pathway to transglycosylation. Insights from experiments and quantum mechanics / molecular mechanics simulations. J Chem Soc.2016;132:3325-3332.

Botha CEJ. Developmental plant anatomy. 2008. The Virtual Plant. Blackwell Publishing.

Boudabbous M, Ben I, Walid H, Mariem S. Trans-glycosylation capacity of a highly glycosylated multi- specific β-glucosidase from *Fusarium solani*. Bioprocess Biosyst Eng. 2017;40(4):559–71.

Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem. 1976;72:248–54.

Britton HTS, Robinson AR. CXCVIII, Universal buffer solutions and the dissociation constant of veronal. J Chem Soc. 1931;(1456):1456–62.

Brodeur G, Yau E, Badal K, Collier J, Ramachandran KB, Ramakrishnan S. Chemical and physicochemical pretreatment of lignocellulosic biomass: A Review. Enzyme Res. 2011:e787532.

Brune A. Symbiotic digestion of lignocellulose in termite guts. Nat Rev Microbiol. 2014;12:168–80.

Cairo JPLF, Leonardo FC, Alvarez TM, Ribeiro DA, Buchli F, Costa-Leonardo AM, Carazzolle MF, Costa FF, Leme AFP, Pereira GAG, Squina FM. Functional characterization and target discovery of glycoside hydrolases from the digestome of the lower termite *Coptotermes gestroi*. Biotechnol Biofuels. 2011;4(50).

Carrillo F, Colom X, Suñol JJ, Saurina J. Structural FTIR analysis and thermal characterisation of lyocell and viscose-type fibres. Eur Polym J. 2004;40(9):2229–34.

Charoensapyanan R, Ito K, Rudeekulthamrong P, Kaulpiboon J. Enzymatic synthesis of propyl-α-glycosides and their application as emulsifying and antibacterial agents. Biotechnol Bioprocess Eng. 2016;21:389–401.

Chen HG, Yan X, Liu XY, Wang MD, Huang HM, Jia XC, Wang J. Purification and characterization of novel bifunctional xylanase, XynIII, isolated from *Aspergillus niger* A-25. J Microbiol Biotechnol. 2006;16(7):1132–8.

References

Christakopoulos P, Bhat M k., Kekos D, Macris BJ. Enzymatic synthesis of trisaccharides and alkyl β-D-glucosides by the transglycosylation reaction of β-glucosidase from *Fusarium oxysporurn*. J Biol Macromol. 1994;16(6):331–4.

Dahiya S, Ojha S, Mishra S. Biotransformation of sucrose into hexyl-α-D-glucopyranoside and -polyglucosides by whole cells of *Microbacterium paraoxydans*. Biotechnol Lett. 2015;37:1431–7.

Dheeran P, Sachin NN, Jaiswal YK, Adhikari DK. A novel thermostable xylanase of *Paenibacillus macerans* IIPSP3 isolated from the termite gut. J Ind Microbiol Biotechnol. 2012;39:851–60.

Do TH, Nguyen TT, Nguyen TN, Le QG, Nguyen C, Kimura K, Truong NH. Mining biomass-degrading genes through Illumina-based de novo sequencing and metagenomic analysis of free-living bacteria in the gut of the lower termite *Coptotermes gestroi* harvested in Vietnam. J Biosci Bioeng. 2014;118(6):665–71

Domınguez R, Souchon H, Lascombe M, Alzari PM. The crystal structure of a family 5 endoglucanase mutant in complexed and uncomplexed forms reveals an induced fit activation mechanism. J Mol Biol. 1996;257:1042–51.

Dominguez R, Souchon H, Spinelli S, Dauter Z, Wilson K., Chauvaux S, et al. A common protein fold and similar active site in two distinct families of β-glycanases. Nat Struct Biol. 1995;2(7):569–79.

Donohoe BS, Decker SR, Tucker MP, Himmel ME, Vinzant TB. Visualizing lignin coalescence and migration through maize cell walls following thermochemical pretreatment. Biotechnol Appl Biochem. 2008;101(5):913–25.

Ducros V, Czjzek M, Belaich A, Gaudin C, Fierobe H, Belaich J, Davies GJ, Haser R. Crystal structure of the catalytic domain of a bacterial cellulase belonging to family 5. Structure. 1995;3(9).

Eggleton P. Termites and trees : a review of recent advances in termite phylogenetics. 2001;48:187–93.

El-Naggar NEA, Deraz S, Khalil A. Bioethanol production from lignocellulosic feedstocks based on enzymatic hydrolysis: Current status and recent developments. Biotechnology. 2014;13(1): 1–21.

Farid MA, Shaker HM, Ei-diwany AI. Effect of peracetic acid, sodium hydroxide and phosphoric acid on cellulosic materials as a pretreatment for enzymatic hydrolysis. Enzym Microb Technol. 1983;5:421–4.

Feng T, Liu H, Xu Q, Sun J, Shi H. Identification and characterization of two endogenous β-glucosidases from the termite *Coptotermes formosanus*. Appl Biochem Biotechnol. 2015;176(7):2039–52.

Gabhane J, William S, Vaidya A, Anand D, Wate S. Pretreatment of garden biomass by alkali-assisted ultrasonication: effects on enzymatic hydrolysis and ultrastructural changes. J Environ Heal Sci Eng. 2014;12(1):76.

Ganner T, Bubner P, Eibinger M, Mayrhofers C, Planks H, Nidetzky B. Dissecting and reconstructing synergism: *In situ* visualization of cooperativity among cellulases. J Biol Chem. 2012;287(52):43215–22.

Gilbert HJ, Stålbrand H, Brumer H. How the walls come crumbling down: recent structural biochemistry of plant polysaccharide degradation. Curr Opin Plant Biol. 2008;11(3):338–48.

Gupta VK, Kubicek CP, Berrin J, Wilson DW, Couturier M, Berlin A, Filho EXF, Ezeji T. Fungal enzymes for bio-products from sustainable and waste biomass. Trends Biochem Sci. 2016;41(7):633–45

Han NS, Kim T. Synergistic action modes of arabinan degradation by exo- and endo-arabinosyl hydrolases. J Microbiol Biotechnol. 2015;25(2):227–33.

Han Q, Liu N, Robinson H, Cao L, Qian C, Wang Q, Xie L, Ding H, Wang Q, Huang Y, Li J, Zhou Z. Biochemical characterization and crystal structure of a GH10 xylanase from termite gut bacteria reveal a novel structural feature and significance of its bacterial ig-like domain. Biotechnol Bioeng. 2013;110(12):3093–103.

Handelsmanl J, Rondon MR, Goodman RM, Brady SF, Clardy J. Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. Chem Biol. 1998;5(10).

Haque MA, Barman DN, Kim MK, Yun HD, Cho KM. Cogon grass (*Imperata cylindrica*), a potential biomass candidate for bioethanol: Cell wall structural changes enhancing hydrolysis in a mild alkali pretreatment regime. J Sci Food Agric. 2016;96(5):1790–7.

Hatherley R, Brown DK, Glenister M, Bishop T. PRIMO : an interactive homology modeling pipeline. 2016;11(11):1–20.

Hendriks ATWM, Zeeman G. Pretreatments to enhance the digestibility of lignocellulosic biomass. Bioresour Technol. 2009;100:10–8.

Henley JP, Sadana A. Categorization of enzyme deactivations using a series-type mechanism. Enzyme Microb Technol. 1985;7(2):50–60.

Henrikssona K, Telemad A, Suorttia T, Reinikainen T, Jaskarib J, Teleman O, Poutanen K. Hydrolysis of barley (1,3), (1,4)-beta-D-glucan by a cellobiohydrolase II preparation from *Trichoderma reesei* hydrolysis. Carbohydr Polym. 1995;26(94).

Henrissat B, Bairoch A. Updating the sequence-based classification of glycosyl hydrolases. Biochem J Lett. 1996;316:695–6.

Hoon T, Hyun T. Overview of technical barriers and implementation of cellulosic ethanol in the U.S. Energy. 2014;66:13–9.

Huy N, Jae T, Sohng K. Recent biotechnological progress in enzymatic synthesis of glycosides. 2013;1329–56.

Huy ND, Nguyen CL, Seo J, Kim D, Park S. Putative endoglucanase PcGH5 from *Phanerochaete chrysosporium* is a β-xylosidase that cleaves xylans in synergistic action with endo-xylanase. Journal of Bioscience and Bioengineering. 2015; 119(4): 416-420

Inoue H, Decker SR, Taylor LE, Yano S, Sawayama S. Identification and characterization of core cellulolytic enzymes from *Talaromyces cellulolyticus* (formerly *Acremonium cellulolyticus*) critical for hydrolysis of lignocellulosic biomass. Biotechnol Biofuels 2014;7(1):151.

Ishida T, Yaoi K, Hiyoshi A, Igarashi K, Samejima M. Substrate recognition by glycoside hydrolase family 74 xyloglucanase from the basidiomycete *Phanerochaete chrysosporium*. FEBS J. 2007;274(21):5727–36.

References

Ito J, Ebe T, Shibasaki S, Fukuda H, Kondo A. Production of alkyl glucoside from cellooligosaccharides using yeast strains displaying *Aspergillus aculeatus* β-glucosidase 1. J Mol Catal B Enzym. 2007;49:92–7.

Jia X, Chen Y, Shi C, Ye Y, Wang P, Zeng X, Wu T. Preparation and characterization of cellulose regenerated from phosphoric acid. J Agric Food Chem. 2013;61(50):12405–14.

Johansson L. Structural analyses of (1→3),(1→4)-β-D-glucan of oats and barley. 2006; Thesis: University of Helsinki.

Juers DH, Matthews BW, Huber RE. LacZ β-galactosidase : Structure and function of an enzyme of historical and molecular biological importance. Protein Sci. 2012;21:1792-1807

Kamarullah SH, Mydin MM,Omar WSAW,Harith SS, Noor BHM, Alias NZA, Manap S, Mohamad R. Surface morphology and chemical composition of napier grass fibers. Malaysian J Anal Sci. 2015;19(4):889–95.

Kanda T, Wakabayashi K, Nisizawa K. Modes of action of exo- and endo-cellulases in the degradation of celluloses I and II. J Biochem. 1980;87(6):1635–9.

Karl ZJ, Scharf ME. Effects of five diverse lignocellulosic diets on digestive enzyme *Reticulitermes flavipes*. Arch Biochem Biophys. 2015;90(2):89–103.

Kern M, McGeehan JE, Streeter SD, Martin RN a, Besser K, Elias L, Eboralla W, Malyonb GP, Paynec CM, Himmeld ME, Schnorre K, Beckhamf GT, Cragg SM, Brucea NC, McQueen-Masona SJ. Structural characterization of a unique marine animal family 7 cellobiohydrolase suggests a mechanism of cellulase salt tolerance. Proc Natl Acad Sci USA. 2013;110(25):10189–94.

Kim I, Han J. Optimization of alkaline pretreatment conditions for enhancing glucose yield of rice straw by response surface methodology. Biomass and Bioenergy. 2012;46:210–7.

Kim JS, Lee YY, Kim TH. A review on alkaline pretreatment technology for bioconversion of lignocellulosic biomass. Bioresour Technol 2016;199:42–8.

Kim TH, Kim JS, Sunwoo C, Lee YY. Pretreatment of corn stover by aqueous ammonia. Bioresour Technol. 2003;90(1):39–47.

Kim Y-M, Kim B-H, Ahn J-S, Kim G-E, Jin S-D, Nguyen T-H, Kim D. Enzymatic synthesis of alkyl glucosides using *Leuconostoc mesenteroides* dextransucrase. Biotechnol Lett. 2009;31:1433–8.

Kino K, Satake R, Morimatsu T, Kuratsu S, Sato M, Kirimura K. A new method of synthesis of Alkyl β-glycosides using sucrose as sugar Donor. Biosci Biotechnol Biochem. 2008;72(9):2415–7.

Kipper K, Väljamäe P, Johansson G. Processive action of cellobiohydrolase Cel7A from *Trichoderma reesei* is revealed as "burst" kinetics on fluorescent polymeric model substrates. Biochem J. 2005;385(2):527–35.

Koppram R, Tomás-Pejó E, Xiros C, Olsson L. Lignocellulosic ethanol production at high-gravity: Challenges and perspectives. Trends Biotechnol. 2014;32(1):46–53.

Kumar R, Hu F, Hubbell CA, Ragauskas AJ, Wyman CE. Comparison of laboratory delignification methods, their selectivity, and impacts on physiochemical characteristics of cellulosic biomass. Bioresour Technol. 2013;130:372–81.

Kumar V, Kothari S, Banker GS. Effect of the agitation rate on the generation of low-crystallinity cellulose from phosphoric acid. J Appl Polym Sci. 2001;82(11):2624–8.

Kuo CH, Lee CK. Enhancement of enzymatic saccharification of cellulose by cellulose dissolution pretreatments. Carbohydr Polym. 2009;77(1):41–6.

Lantz SE, Goedegebuur F, Hommes R, Kaper T, Kelemen BR, Mitchinson C, Wallace L, Ståhlberg J, Larenas EA. *Hypocrea jecorina* CEL6A protein engineering. Biotechnol Biofuels. 2010;3(1):20.

Lee J, Saddler J, Parameswaran B. Pretreatment of biomass. Bioresour Technol. 2016;199:1.

Lei Z, Shao Y, Yin X, Yin D, Guo Y, Yuan J. Combination of xylanase and debranching enzymes specificto wheat arabinoxylan improve the growth performance and gut health of broilers. J. Agric. Food Chem. 2016;64(24): 4932-4942

Lima MA, Lavorente GB, da Silva HK, Bragatto J, Rezende C, Bernardinelli OD, deAzevedo ER, Gomez LD, McQueen-Mason SJ, Labate CA, Polikarpov I. Effects of pretreatment on morphology, chemical composition and enzymatic digestibility of eucalyptus bark: a

potentially valuable source of fermentable sugars for biofuel production - part 1. Biotechnol Biofuels. 2013;6(1):75.

Lima T de A, Pontual EV, Dornelles LP, Amorim PK, Sa RA, Coelho LCBB, Napoleão TH, Paiva PG. Digestive enzymes from workers and soldiers of termite *Nasutitermes corniger*. Comp Biochem Physiol Part B. 2014;176:1–8.

Liu X, Che M, Xie L, Zhan S, Zhou Z, Huang Y, Wang Q. Metatranscriptome of the protistan community in *Reticulitermes flaviceps*. Insect Sci. 2016;23:543–7.

Liu X, Xie L, Liu N, Zhan S, Zhou X, Wang Q. RNA interference unveils the importance of *Pseudotrichonympha grassii* cellobiohydrolase, a protozoan exoglucanase, in termite cellulose degradation. Insect Mol Biol. 2017;26(2):233–42.

Luthy R, Bowie J, Eisenberg D. Assessment of protein models with three-dimensional profiles. Nature. 1992;356:83–5.

Malgas S, Dyk JS, Pletschke BI. A review of the enzymatic hydrolysis of mannans and synergistic interactions between β-mannanase , β-mannosidase and α-galactosidase. World J Microbiol Biotechnol. 2015a.

Malgas S, van Dyk SJ, Pletschke BI. β-Mannanase (Man26A) and α-galactosidase (Aga27A) synergism – A key factor for the hydrolysis of galactomannan substrates. Enzyme Microb Technol. 2015b;70:1–8.

Mathew L, Sharad V, Divya A, Rishi S, Mandhan P. Cost-effective screening and isolation of xylano-cellulolytic positive microbes from termite gut and termitarium. 3 Biotech. 2017;7(2):1–7.

Matin MM, Bhuiyan MMH, Azad AKMS. Synthesis and antimicrobial evaluation of some n-butyl α - and β-d-glucopyranoside derivatives synthesis and antimicrobial evaluation of some n-butyl α- and β-d-glucopyranoside derivatives. RGUHS J Pharm Sci. 2014;3(1).

Matsumura S, Irnai K, Yoshikawa S, Kawada K, Uchibori T. Surface activities , biodegradability and antimicrobial properties of n-alkyl glucosides, mannosides and galactosides. J Am Oil Chem Soc. 1990;67(12):996–1001.

Matt C, Haubruge E, Thonart P, Pauw E De, Portetelle D, Vandenbol M. Characterization of a new beta glucosidase/beta xylosidase from the gut microbiota of the termite (*Reticulitermes santonensis*). FEMS Microbiol Lett. 2017;314:147–57.

Mattéotti C, Bauwens J, Brasseur C, Tarayre C, Thonart P, Destain J, Francis F, Haubruge E, De Pauw E, Portetelle D, Vandenbol M. Identification and characterization of a new xylanase from Gram-positive bacteria isolated from termite gut (*Reticulitermes santonensis*). 2012;83:117–27.

Miller LG. Use of dinitrosaiicyiic acid reagent for determination of reducing sugar. Anal Chem. 1959;31(3):426–8.

Modenbach A. Sodium hydroxide pretreatment of corn stover and subsequent enzymatic hydrolysis : An investigation of yields, kinetic modeling and glucose recovery. 2013; Thesis: University of Kentucky.

Morrison JM, Elshahed MS, Youssef NH. Defined enzyme cocktail from the anaerobic fungus *Orpinomyces* sp. strain C1A effectively releases sugars from pretreated corn stover and switchgrass. Sci Rep. 2016;6:29217.

Moxley G, Zhang Y-HP. More accurate determination of acid-labile carbohydrates in lignocellulose by modified quantitative saccharification. Energy and Fuels. 2007;21(6):3684–8.

Nakamura A, Watanabe H, Ishida T, Uchihashi T, Wada M, Ando T, Igarashi K, Samejima M. Trade-off between processivity and hydrolytic velocity of cellobiohydrolases at the surface of crystalline cellulose. J Am Chem Soc. 2014;136(12):4584–92.

Naumoff DG. Hierarchical Classification of Glycoside Hydrolases. Biochem. 2011;76(6):622–35.

Ni J, Tokuda G. Lignocellulose-degrading enzymes from termites and their symbiotic microbiota. Biotechnol Adv. 2013;31(6):838–50.

Nidetzky B, Hayn M, Macarron R, Steiner W. Synergism of *Trichoderma reesei* cellulases while degrading different celluloses. Biotechnol Lett. 1993;15(1):71–6.

Nwodo S, Obinna C, Uzoma CN, Veronica AO. Kinetic study and characterization of 1,4-β-endoglucanase of *Aspergillus niger* ANL301. Dyn Biochem Process Biotechnol Mol Biol. 2011;5(2):2–7.

O'Connor RT, DuPre EF, Mitcham D. Applications of infrared absorption spectroscopy to investigations of cotton and modified cottons: Part I: Physical and Crystalline Modifications and Oxidation. Text Res J. 1958;28(5):382–92.

Olver B, Dyk JS Van, Beukes N, Pletschke BI. Synergy between EngE , XynA and ManA from *Clostridium cellulovorans* on corn stalk , grass and pineapple pulp substrates. 2011;187–92.

Pallapolu VR, Lee YY, Garlock RJ, Balan V, Dale BE, Kim Y, Mosier NS, Ladisch MR, Falls M, Holtzapple MT, Sierra-Ramirez R, Shi J, Ebrik MA, Redmondg T, Yangg B, Wyman CE, Donohoeh BS, Vinzant TB Elander RT, Hamesi B, Thomasi B, Warner RE. Effects of enzyme loading and β-glucosidase supplementation on enzymatic hydrolysis of switchgrass processed by leading pretreatment technologies. Bioresour Technol. 2011;102(24):11115–20.

Panagiotopoulos IA, Bakker RR, De Vrije T, Koukios EG, Claassen PAM. Pretreatment of sweet sorghum bagasse for hydrogen production by *Caldicellulosiruptor saccharolyticus*. Int J Hydrogen Energy. 2010;35(15):7738–47.

Pauly M, Qin Q, Greene H, Albersheim P, Darvill A, York WS. Changes in the structure of xyloglucan during cell elongation. Planta. 2001;212(5–6):842–50.

Pinchuck SC. The ultrastructure and histology of the defensive epidermal glands of some marine pulmonates. 2009; Thesis: Rhodes University.

Pol D, Menon V, Rao M. Biochemical characterization of a novel thermostable xyloglucanase from an alkalothermophilic *Thermomonospora* sp. Extremophiles. 2012;16(1):135–46.

Poletto M, Pistor VV, Santana RMC, Zattera AJ. Materials produced from plant biomass: part II: evaluation of crystallinity and degradation kinetics of cellulose. Mater Res. 2012;15(3):421–7.

Polizzi KM, Bommarius AS, Broering JM, Chaparro-Riggers JF. Stability of biocatalysts. Curr Opin Chem Biol. 2007;11(2):220–5.

Prins RA, Kreulen DA.. Comparative aspects of plant cell wall digestion in insects. Anim Feed Sci Technol. 1991;32:101–18.

Rabelo SC, Filho RM, Costa AC. Lime pretreatment and fermentation of enzymatically hydrolyzed sugarcane bagasse. Appl Biochem Biotechnol. 2013;169(5):1696–712.

Rajarapu SP, Scharf ME. Saccharification of agricultural lignocellulose feedstocks and protein-level responses by a termite gut-microbe bioreactor. Front Energy Res. 2017;5:1–10.

Rashamuse K, Ronneburg T, Sanyika W, Mathiba K, Mmutlane E, Brady D. Metagenomic mining of feruloyl esterases from termite enteric flora. Appl Microbiol Biotechnol. 2013;98(2):727–37.

Rashamuse K, Sanyika TW, Mathiba K, Ngcobo T, Mtimka S, Brady D. Metagenomic mining of glycoside hydrolases from the hindgut bacterial symbionts of a termite, *Trinervitermise trinervoides* and the characterisation of a multimodular β-1, 4-Xylanase (GH11). Biotechnol Appl Biochem. 2016;1–13.

Rather MY, Mishra S. β-Glycosidases : An alternative enzyme based method for synthesis of alkyl-glycosides. Sustain Chem Process. 2013;1(7):1–15.

Ravn JL, Martens HJ, Petterssona D, Pedersen NR. A commercial GH 11 xylanase mediates xylan solubilisation and degradation in wheat, rye and barley as demonstrated by microscopy techniques and wet chemistry methods. Animal Feed Science and Technology. 2016;219:216–225.

Rose JKC. The Plant Cell Wall. Annu. Plant Rev. 2003; 8.

Rybinski W Von, Hill K. Alkyl Polyglycosides-Properties and Applications of a new class of surfactants. Angew Chemie Int Ed. 1998;37:1328–45.

Saadeddin A. The complexities of hydrolytic enzymes from the termite digestive system. Crit Rev Biotechnol. 2014;8551(2):115–22.

Saha BC. Hemicellulose bioconversion. J Ind Microbiol Biotechnol. 2003;279–91.

Saini JK, Saini R, Tewari L. Lignocellulosic agriculture wastes as biomass feedstocks for second-generation bioethanol production : concepts and recent developments. 3 Biotech Process. 2014;5(4):337–53.

Santana RH, Catao ECP, Lopes FAC, Constantino R, Barreto CC, Kruger RH. The gut microbiota of workers of the litter-feeding termite *Syntermes wheeleri* (Termitidae: Syntermitinae): Archaeal, bacterial, and fungal communities. Microb Ecol. 2015;545–56.

Scharf ME, Karl ZJ, Sethi A, Boucias DG. Multiple levels of synergistic collaboration in termite lignocellulose digestion. PLoS One. 2011;6(7):1–7.

Scharf ME, Kovaleva ES, Jadhao S, Campbell JH, Buchman GW, Boucias DG. Functional and translational analyses of a beta-glucosidase gene (glycosyl hydrolase family 1) isolated from the gut of the lower termite *Reticulitermes flavipes*. Insect Biochem Mol Biol. 2010;40(8):611–20.

Scharf ME. Termites as targets and models for biotechnology. Annu Rev Entomol. 2015;60:77–102.

Schwanninger M, Rodrigues JC, Pereira H, Hinterstoisser B. Effects of short-time vibratory ball milling on the shape of FT-IR spectra of wood and cellulose. Vib Spectrosc. 2004;36(1):23–40.

Segato F, Damasio ARL, Gonçalves T, Murakami MT, Squina FM, Polizeli M, Mort AJ, Prade RA. Two structurally discrete GH7-cellobiohydrolases compete for the same cellulosic substrate fiber. Biotechnol Biofuels. 2012;5(1):21.

Selig MJ, Thygesen LG, Johnson DK, Himmel ME, Felby C, Mittal A. Hydration and saccharification of cellulose I-β, II and IIII at increasing dry solids loadings. Biotechnol Lett. 2013;35(10):1599–607.

Sethi A, Kovaleva ES, Slack JM, Brown S, Buchman GW, Scharf ME. A GHF7 cellulase from the protist symbiont community of *Reticulitermes flavipes* enables more efficient lignocellulose processing. Arch Insect Biochem Physiol. 2013;84(4):175–93.

Siontorou CG, Bidikoudi M, Chandrinou C, Boukos N, Falaras P, Fardis M, Apostolopoulos G, Batzias F, Sidiras D. Spectroscopic assessment of biomass derived adsorbents for oil spill cleaning. Int'l J Res Chem Metall Civ Engg. 2015;2(1):49–54.

Siqueira G, Várnai A, Ferraz A, Milagres AMF. Enhancement of cellulose hydrolysis in sugarcane bagasse by the selective removal of lignin with sodium chlorite. Appl Energy. 2013;102:399–402.

Siroky J, Blackburn RS, Bechtold T, Taylor J, White P. Attenuated total reflectance Fourier-transform infrared spectroscopy analysis of crystallinity changes in lyocell following continuous treatment with sodium hydroxide. Cellulose. 2010;17(1):103–15.

Sleator RD, Shortall C, Hill C. Metagenomics. Lett Appl Microbiol. 2008;47(5):361–6.

Sluiter JB, Ruiz RO, Scarlata CJ, Sluiter AD, Templeton DW. Compositional analysis of lignocellulosic feedstocks. Review and description of methods. J Agric Food Chem. 2010;58(16):9043–53.

Song S, Tang Y, Yang S, Yan Q, Zhou P, Jiang Z. Characterization of two novel family 12 xyloglucanases from the thermophilic *Rhizomucor miehei*. Appl Microbiol Biotechnol. 2013;97(23):10013–24.

Sørensen HR, Pedersen S, Meyer AS. Synergistic enzyme mechanisms and effects of sequential enzyme additions on degradation of water insoluble wheat arabinoxylan. Enzyme Microb Technol. 2007;40(4):908–18.

Synytsya A, Novak M. Structural analysis of glucans. Ann Transl Med. 2014;2(2):17.

Takahashi M, Takahashi H, Nakano Y, Konishi T, Terauchi R, Takeda T. Characterization of a cellobiohydrolase (MoCel6A) produced by *Magnaporthe oryzae*. Appl Environ Microbiol. 2010;76(19):6583–90.

Tao S, Khanizadeh S, Zhang H, Zhang S. Anatomy, ultrastructure and lignin distribution of stone cells in two *Pyrus* species. Plant Sci. 2009;176(3):413–9.

Todaka N, Lopez CM, Inoue T, Saita K. Heterologous expression and characterization of an endoglucanase from a symbiotic protist of the lower termite , *Reticulitermes speratus*. Appl Biochem Biotechnol. 2010;160:1168–78.

Togashi H, Kato A, Shimizu K. Enzymatically derived aldouronic acids from *Eucalyptus globulus* glucuronoxylan. Carbohydr Polym. 2009;78(2):247–52.

References

Tokuda G, LO N, Watanabe H. Marked variations in patterns of cellulase activity against crystalline- vs . carboxymethyl-cellulose in the digestive systems of diverse , wood-feeding termites. Physiol Mol Plant Pathol. 2005;30:372–80.

Tokuda G, Watanabe H, Hojo M, Fujita A, Makiya H, Miyagi M. Cellulolytic environment in the midgut of the wood-feeding higher termite *Nasutitermes takasagoensis*. J Insect Physiol. 2012;58(1):147–54.

Tokuda G, Watanabe H, Matsumoto T, Noda H. Cellulose digestion in the wood-eating higher termite, *Nasutitermes takasagoensis* (Shiraki): distribution of cellulases and properties of endo-beta-1,4-glucanase. Zoolog. Sci. 1997. p. 83–93.

Tokuda G, Watanabe H. Hidden cellulases in termites : revision of an old hypothesis. Biology. 2007;3:336–9.

Tsakagoshi H, Nakamura A, Ishida T, Touhara KK, Otagiri M, Moriya S, Samejima M, Igarashi K, Fushinobu S, Kitamoto K, Arioka M. Structural and biochemical analyses of glycoside hydrolase family 26-mannanase from a symbiotic protist of the termite *Reticulitermes speratus*. J Biol Chem. 2014;289(15):10843–52.

Tsukagoshi H, Nakamura A, Ishida T, Otagiri M, Moriya S, Samejima M, Igarashi K, Fushinobu S, Kitamoto K, Arioka M. The GH26 β -mannanase RsMan26H from a symbiotic protist of the termite *Reticulitermes speratus* is an endo-processive mannobiohydrolase : Heterologous expression and characterization. Biochem Biophys Res Commun. 2014;452(3):520–5.

Turkmen N, Sari F, Velioglu YS. Effects of extraction solvents on concentration and antioxidant activity of black and black mate tea polyphenols determined by ferrous tartrate and Folin-Ciocalteu methods. Food Chem. 2006;99(4):835–41.

Uchima CA, Tokuda G, Watanabe H, Kitamoto K, Arioka M. Heterologous expression in *Pichia pastoris* and characterization of an endogenous thermostable and high-glucose-tolerant β-glucosidase from the termite *Nasutitermes takasagoensis*. Appl Environ Microbiol. 2012;78(12):4288–93.

Valls A, Diaz P, Pastor FIJ, Valenzuela S V. A newly discovered arabinoxylan-specific arabinofuranohydrolase. Synergistic action with xylanases from different glycosyl hydrolase families. Appl Microbiol Biotechnol. 2016;100(4):1743–51.

Van Dyk JS, Pletschke BI. A review of lignocellulose bioconversion using enzymatic hydrolysis and synergistic cooperation between enzymes-factors affecting enzymes, conversion and synergy. Biotechnol Adv. 2012;30(6):1458–80.

Vitcosque GL, Ribeiro LFC, de Lucas RC, da Silva TM, Ribeiro LF, de Lima Damasio AR, et al. The functional properties of a xyloglucanase (GH12) of *Aspergillus terreus* expressed in *Aspergillus nidulans* may increase performance of biomass degradation. Appl Microbiol Biotechnol. Applied Microbiology and Biotechnology; 2016;100(21):9133–44.

Walia A, Guleria S, Mehta P, Chauhan A. Microbial xylanases and their industrial application in pulp and paper biobleaching : a review. 3 Biotech. 2017;7(1):1–12.

Wang Z, Li R, Xu J, Marita JM, Hatfield RD, Qu R, Cheng JJ. Sodium hydroxide pretreatment of genetically modified switchgrass for improved enzymatic release of sugars. Bioresour Technol. 2012;110:364–70.

Warnecke F, Luginbühl P, Ivanova N, Ghassemian M, Richardson TH, Stege JT, et al. Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. Nature. 2007;450(7169):560–5.

Watanabe H, Noda H, Tokuda G. A cellulase gene of termite origin. Nature. 1998;394:1–2.

Watanabe H, Tokuda G. Cellulolytic systems in insects. Ann Rev Entomol. 2010;55:609–32.

Watson BJ, Zhang H, Longmire AG, Young HM, Hutcheson SW. Processive endoglucanases mediate degradation of cellulose by *Saccharophagus degradans*. J Bacteriol. 2009;191(18):5697–705.

Wei KSC, Teoh TC, Koshy P, Salmah I, Zainudin A. Cloning, expression and characterization of the endoglucanase gene from *Bacillus subtilis* UMC7 isolated from the gut of the indigenous termite *Macrotermes malaccensis* in *Escherichia coli*. Electron J Biotechnol. 2015;18(2):103–9.

Wiederstein M, Sippl MJ. ProSA-web : interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res. 2017;35:407–10.

Wilson DB, Kostylev M. Cellulase Processivity. Biomass Convers. 2012;908(7):73–82.

Wu L, Arakane M, Ike M, Wada M, Takai T, Gau M, Tokuyasu K. Low temperature alkali pretreatment for improving enzymatic digestibility of sweet sorghum bagasse for ethanol production. Bioresour Technol. 2011;102(7):4793–9.

Yang B, Dai Z, Ding S-Y, Wyman CE. Enzymatic hydrolysis of cellulosic biomass. Biofuels. 2014;2:421–49.

Younis M, Mishra S, Verma V, Chand S. Biotransformation of methyl-β-D-glucopyranoside to higher chain alkyl glucosides by cell bound β-glucosidase of *Pichia etchellsii*. Bioresour Technol. 2012;107:287–94.

Yuki M, Kuwahara H, Shintani M, Izawa K, Sato T, Starns D, Hongoh Y, Ohkum M. Dominant ectosymbiotic bacteria of cellulolytic protists in the termite gut also have the potential to digest lignocellulose. Environ Microbiol. 2015;17(12):4942–53.

Yuki M, Oshima K, Suda W, Sakamoto M, Lida T, Hattori M, et al. Draft Genome Sequence of *Bacteroides reticulotermitis* Strain JCM 10512 T , isolated from the gut of a termite. Genome Announc. 2014;2(1):2.

Zhang C, Wang Y, Li Z, Zhou X, Zhang W, Zhao Y, Ohkuma M. Characterization of a multi-function processive endoglucanase CHU-2103 from *Cytophaga hutchinsonii*. Appl Microbiol Biotechnol. 2014;98(15):6679–87.

Zhang XZ, Zhang Z, Zhu Z, Sathitsuksanoh N, Yang Y, Zhang YHP. The noncellulosomal family 48 cellobiohydrolase from *Clostridium phytofermentans* ISDg: Heterologous expression, characterization, and processivity. Appl Microbiol Biotechnol. 2010;86(2):525–33.

Zhang Y-HP, Cui J, Lynd LR, Kuang LR. A Transition from cellulose swelling to cellulose dissolution by o-phosphoric acid: Evidence from enzymatic hydrolysis and supramolecular Structure. Biomacromolecules. 2006;7:644–8.

Zhao X, Zhang L, Liu D. Biomass recalcitrance. Part I : The chemical compositions and physical structures affecting the enzymatic hydrolysis of lignocellulose. Biofuels, Bioprod. Bioref. 2012;465–82.

Appendices

Appendix A: Chapter 3

Results

Figure A 3.1 illustrates the violet colour formed due to phloroglucinol in the control CC and SSB samples. After lime pretreatment both the CC and SSB biomass displayed less intense violet colour. The colour intensity decreased further when CC and SSB biomass samples were pretreated by NaOH and $NaClO_2$. The colour intensity was directly proportional to amounts of lignin present in the biomass (Botha 2008; Tao et al. 2009). Thus, we concluded that lime, NaOH and $NaClO_2$ removed lignin from both CC and SSB biomass. These results were also consistent with the SEM results (Figure A3.1), which showed that the CC and SSB biomass surfaces became more porous after pretreatment when compared to the controls. SEM showed that the CC and SSB control samples were covered by a smooth layer and that there were no pores visible on the biomass. These results support our conclusion that the CC and SSB biomasses were covered by lignin.

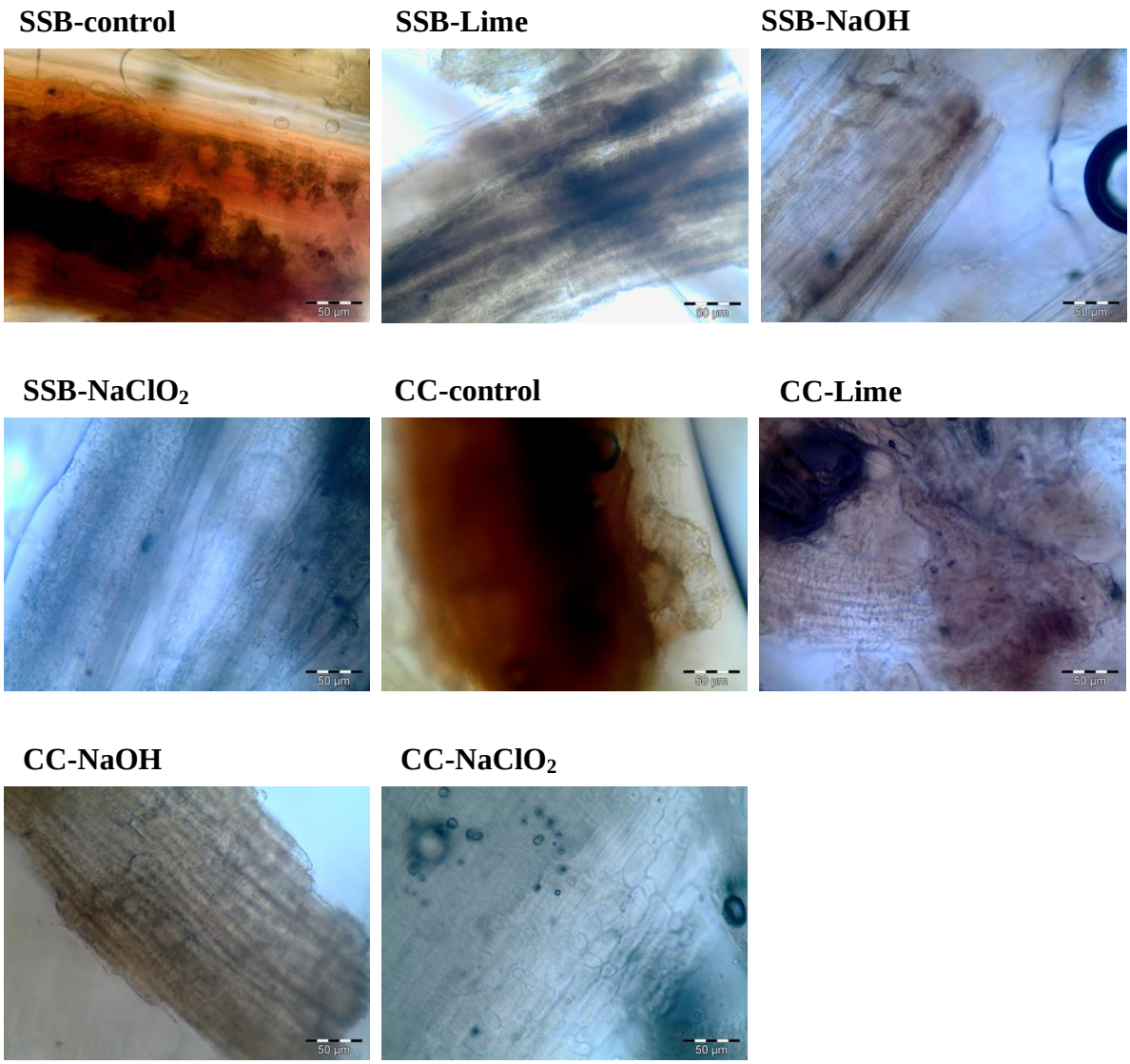


Figure A3.1. Analysis of lignin removal from the sweet sorghum biomass after alkaline or oxidative pretreatment. SSB represents sweet sorghum bagasse that was pretreated with lime, sodium hydroxide (NaOH) and sodium chlorite ($NaClO_2$). CC represents corncobthat was pretreated with lime, sodium hydroxide (NaOH) and sodium chlorite ($NaClO_2$). Controls were not treated with chemicals. The disappearance of violet colour from the biomass is a direct indication of the lignin removal.

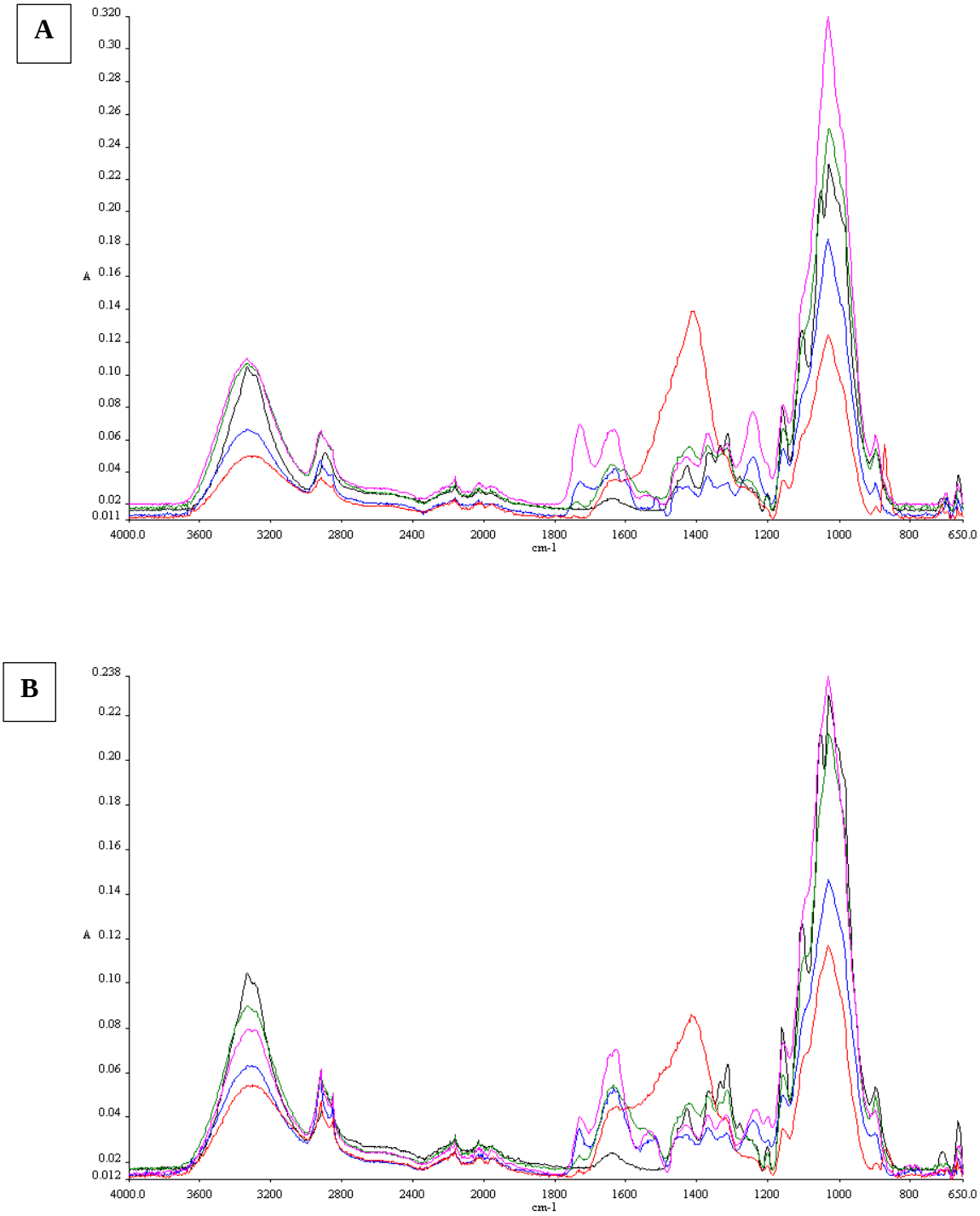


Figure A3.2. The FTIR profiles of the alkaline and oxidative pretreated corncoband sweet sorghum biomass samples, represented by A and B, respectively. The black, blue, red, green and pink colored lines of the spectra represent Avicel/control, untreated, lime, NaOH and $NaClO_2$ pretreated corn cob, respectively.

Appendix B: Chapter 4

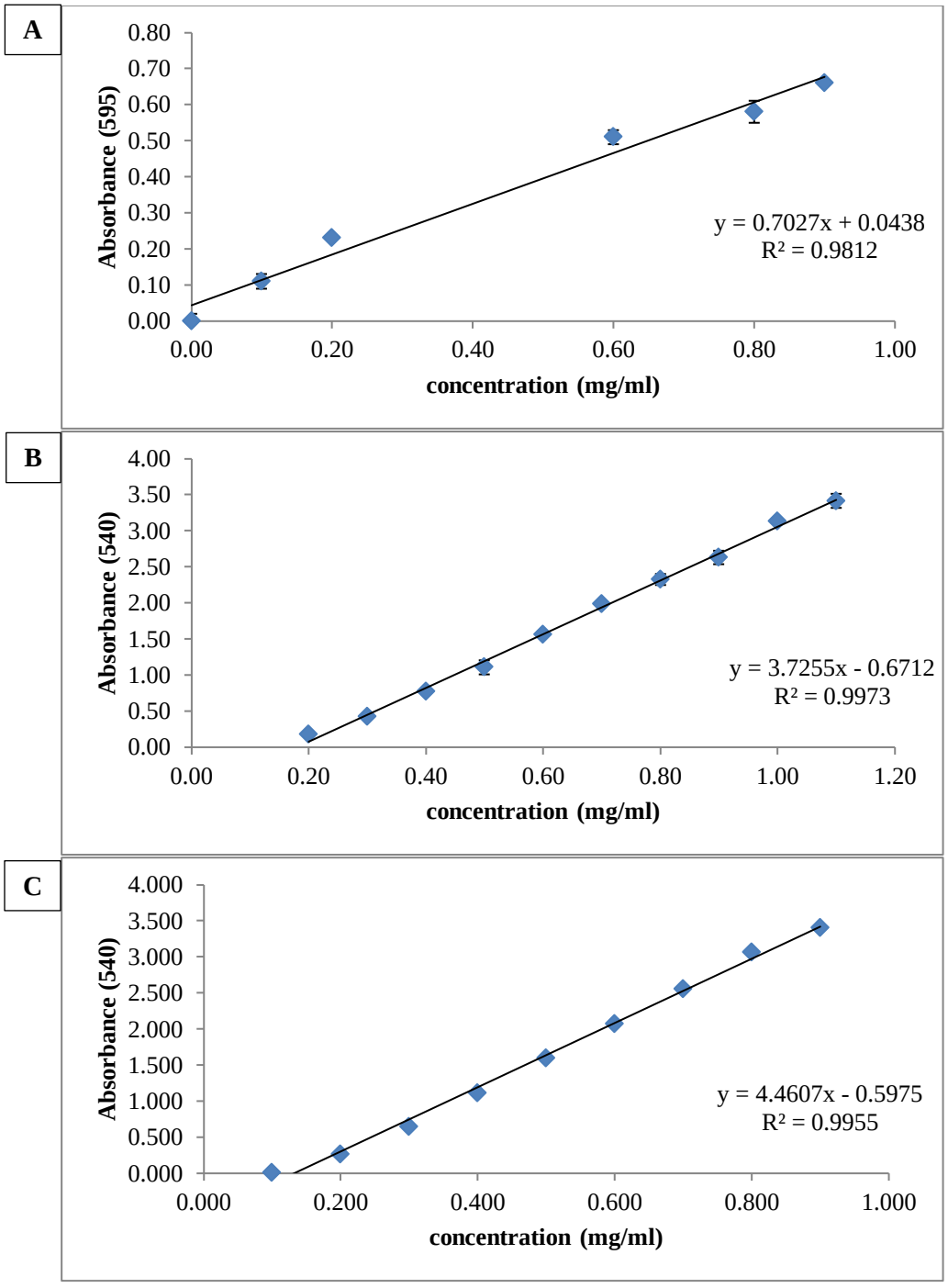


Figure A4.1. A represents the protein standard curve developed with the Bradford method. Bovine serum albumin was used as a suitable standard. B and C represent the sugar standard curves which were developed with the DNS method. Glucose and xylose were used as suitable standards. The datapoints represent means ± SD, n=3.

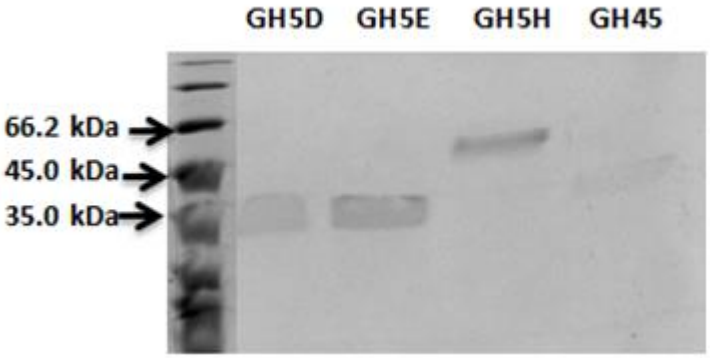
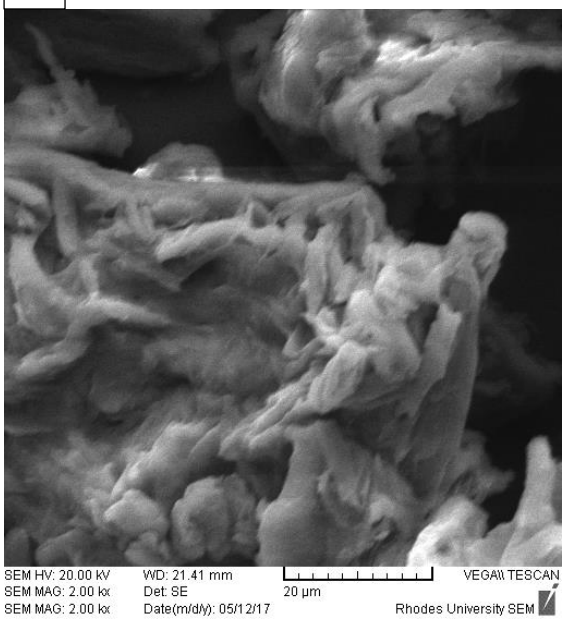


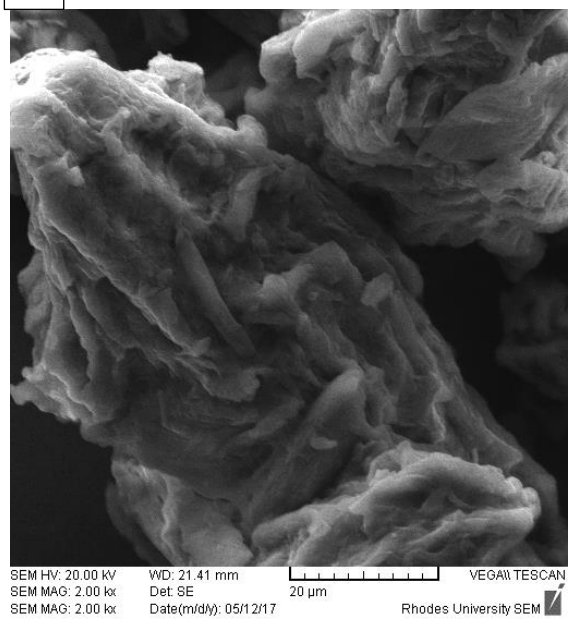
Figure A4.2. SDS-PAGE for purified exo-glucanase (GH5D), endo-glucanase (GH5E), xylanase (GH5H) and endoglucanase/xylanase (GH45)

Appendix C: Chapter 5

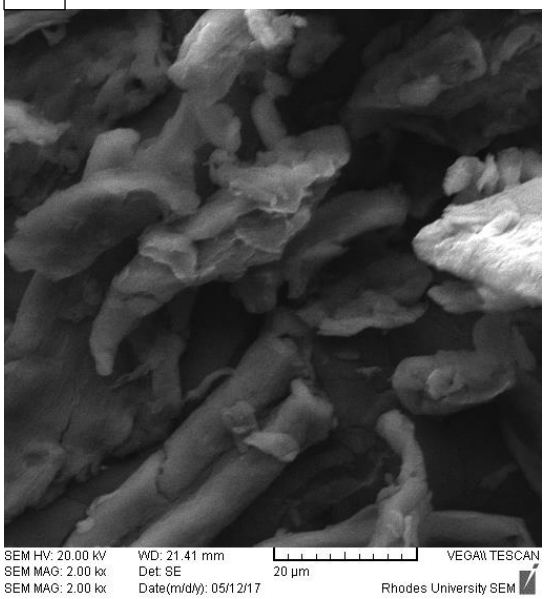
A



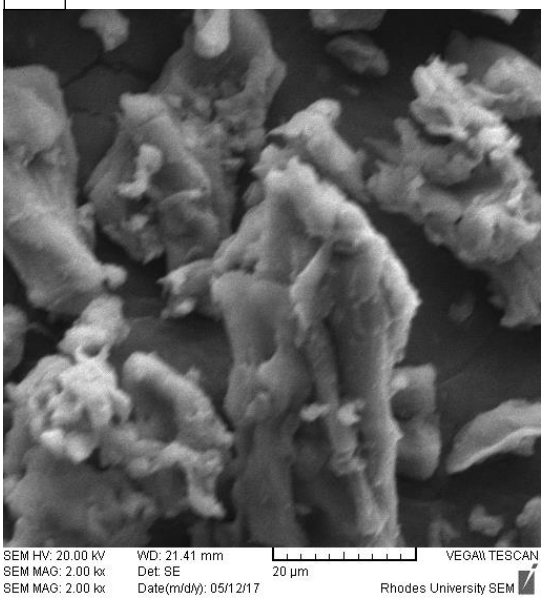
B



C



D



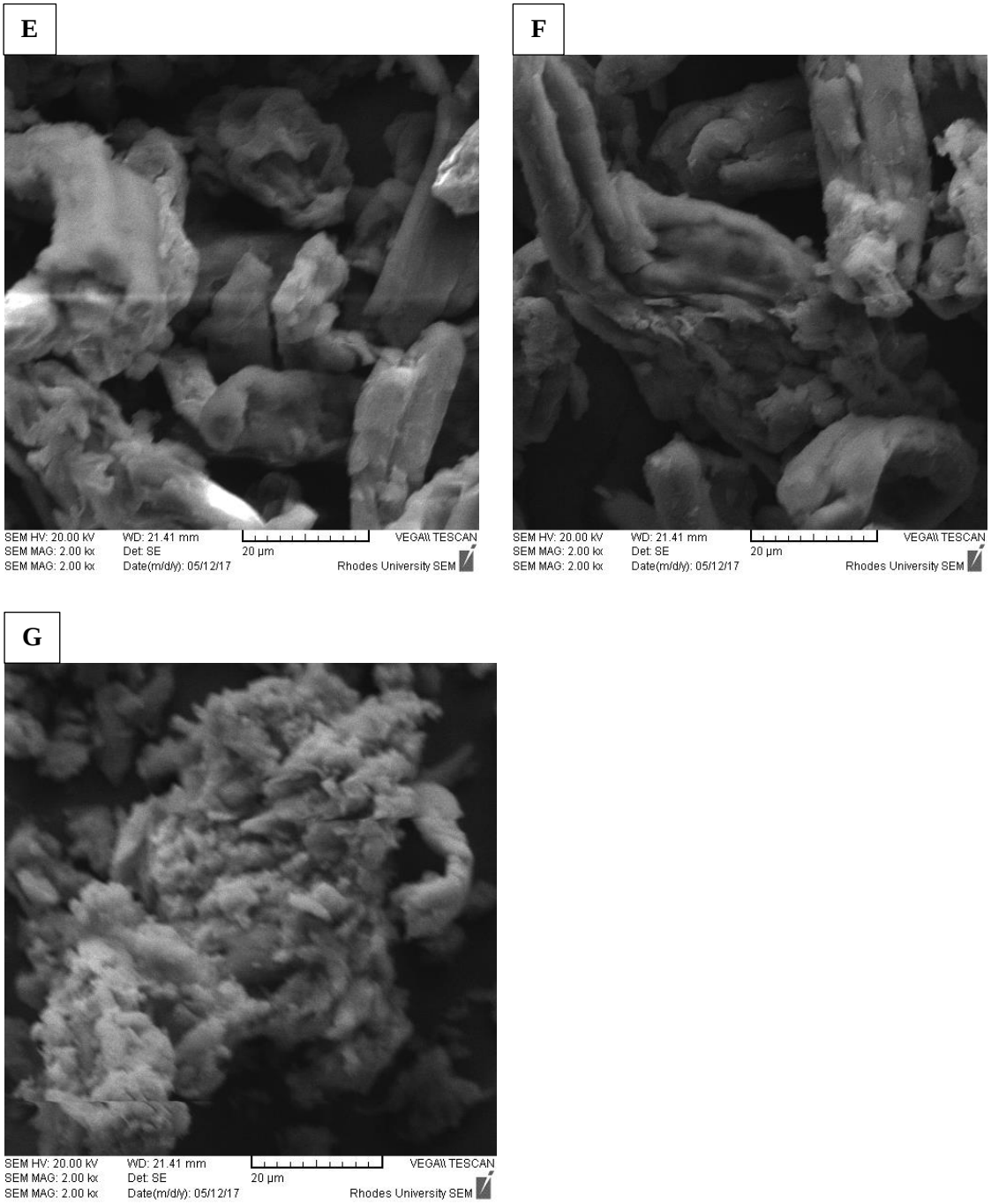


Figure A5.1. Scanning electron microscopy analysis of untreated, phosphoric acid and NaOH pretreated Avicel. A, B and C represent untreated Avicel, 50% (w/v) phosphoric acid pretreated Avicel for 1 and 3 h, respectively. D, E, F and G represent 20% (w/v) sodium hydroxide pretreated Avicel for 0.5, 1.5, 3 and 16 h.