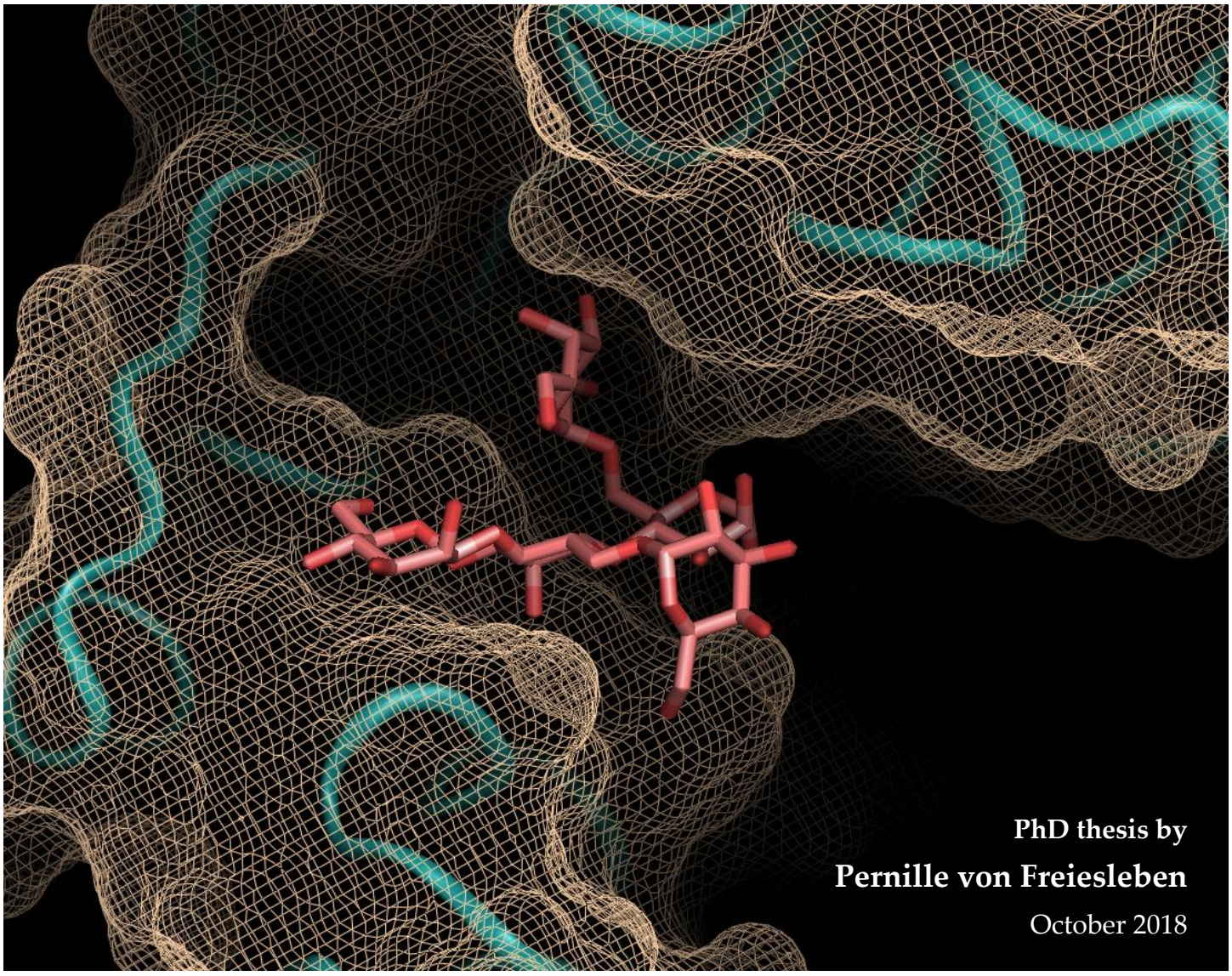
Characterization of fungal endo-β(1→4)-mannanases for biomass conversion

von Freiesleben, Pernille

Publication date:
2018

Document Version
Publisher's PDF, also known as Version of record



PhD thesis by

Pernille von Freiesleben

October 2018

Characterization of fungal endo-β(1→4)-mannanases for biomass conversion

DTU Bioengineering
Department of Biotechnology and Biomedicine

Cover: The crystal structure of a GH26 endomannanase from *Yunnania penicillata* in complex with α-6^2-6^1-di-galactosyl-mannotriose, spanning the -4 to -2 subsites of the active site cleft (6HPF). The picture was drawn using *CCP*4*mg*.

Characterization of fungal endo-β(1→4)-mannanases for biomass conversion

PhD thesis by

Pernille von Freiesleben

Supervisors:

Anne S. Meyer, Professor, Head
Protein Chemistry and Enzyme Technology Division
Department of Biotechnology and Biomedicine, DTU Bioengineering
DTU Building 221, DK-2800 Kgs Lyngby, Denmark

Kristian B. R. M. Krogh, Science Manager
Department of Protein Biochemistry and Stability, Novozymes
Krogshøjvej 36, DK-2880 Bagsværd, Denmark

Henrik Stålbrand, Professor
Department of Biochemistry and Structural Biology, Center for Molecular Protein Science
Lund University, PO Box 124, SE-221 00 Lund, Sweden

Preface

The work associated with this thesis has been conducted at Department of Biotechnology and Biomedicine, at the Technical University of Denmark and at the two departments: Department of Protein Biochemistry and Stability and Department of Biomass Technology, at Novozymes during the period June 2014 – October 2018. The work has been funded by Novozymes and by the BioValue SPIR, Strategic Platform for Innovation and Research on value added products from biomass, which is co-funded by The Innovation Fund Denmark, case no: 0603 – 00522B.

The PhD project has been a learning experience and interesting period for me. For wise suggestions and advice, I would like to thank my supervisors Anne S. Meyer, Henrik Stålbrand and Kristian B.R.M. Krogh. I have been encouraged by your enthusiastic and visionary view on the field of mannanases and enzyme application in general. Also, a special thanks to Anne Stenbæk, Henning Jørgensen, Thomas Holberg Blicher and Lars Anderson for supervision about softwood saccharification trials, structure analysis and ligand docking and general knowledge on mannanases and enzyme applications. I am grateful to Nikolaj Spodsberg, Dorthe Hasling Christiansen and Melahat Balcioglu for the work associated with expressing an extensive amount of interesting fungal endomannanases and assisting me in expressing a few of them on my own. For assisting me in the use and optimization of the DASH method I will like to thank Katja Salomon Johansen, Rune Halvorson and Paul Dupree. I would also like to express my appreciation to Trine Holst Sørensen, Silke Flindt Badino and the other members of Peter Westh's group (RESAB and TEMPEN projects) for good advices, enjoyable times and for assisting me in some of the experimental work. Acknowledgements to all my great colleagues at Novozymes for assisting and for being great colleagues. Great thanks to current and former colleagues at DTU, particularly Jesper Holck, Jane Wittrup Agger, Caroline Mosbech and Birgitte Zeuner, who have all helped with MS and Dionex analyses. In general, I would like to thank all my co-authors and external collaborators for all the inspiring inputs, scientific discussion, constructive collaboration, for helping with experiments and in writing and publishing manuscripts. A special thank goes to Olga V. Moroz, Elena Blagova, Gideon J. Davies and Keith S. Wilson at the University of York. Finally, I am very grateful to my family and friends, without your help and support I could not have made this Thesis.

_____

Pernille von Freiesleben

October 8th 2018

Table of contents

Summary

Biomass from wood is an abundant renewable resource and is rapidly gaining interest as a raw material that can be converted into fuels and other value-added chemicals. Historically, research has mainly focused on cellulose, but interest in utilizing all components of the wood, including the lignin and the hemicellulose, has increased in recent years. In softwoods the dominant hemicellulose is *O*-acetylated galactoglucomannans, which constitutes up to 25 % by weight of the wood dry matter. Endo-β(1→4)-mannanases (endomannanases) catalyze the hydrolysis of β-mannans including galactoglucomannans. However, galactose substitutions on the mannan backbone have been shown to compromise enzymatic degradation of mannans. In addition, the recalcitrant nature of the wood biomass also seems to challenge the enzyme catalyzed degradation. To overcome these challenges and enhance the industrial softwood saccharification, it is important to understand the molecular background of the natural specificity and substrate interactions of endomannanases. This thesis aimed to improve knowledge on the natural specificity of fungal GH5 and GH26 endomannanases. Another aim was to evaluate performance of fungal endomannanases on softwood saccharification and to examine whether any differences in performance might correlate with specific enzyme characteristics.

The influence of galactose substitutions in hydrolysis of mannans catalyzed by enzymes was examined using five fungal endomannanases from GH family 5 and 26. The initial hydrolysis rate and the degree of conversion of two galactomannans, locust bean gum and the more heavily substituted guar gum, were determined for all five enzymes. Product profiles were analyzed using the DNA sequencer-Assisted Saccharide analysis in High throughput (DASH) method. In addition, the accommodation of galactopyranosyl residues in the enzymes active site clefts was analyzed by docking analyses using previously determined crystal structures and new homology models. Based on these investigations, it was shown that the fungal GH26 endomannanases, including the novel *Anid*Man26 from *Aspergillus nidulans*, had a very open active site cleft and accommodated galactopyranosyl residues in at least the -2, -1 and +1 subsites. This novel structural feature enabled full conversion of guar gum galactomannan and resulted in the production of multiple galactomanno-oligosaccharides from guar gum hydrolysis. In contrast, the fungal GH5 endomannanases were more restricted by the galactose substitutions, as indicated by a lower conversion of guar gum and in general from a more narrow active site cleft.

To investigate if the capability to accommodate multiple galactopyranosyl moieties in the active site cleft is beneficial in softwood saccharification, ten fungal endomannanases of GH5 and GH26 were assessed on a variety of pure mannans as well as in enzymatic cellulose saccharification of pretreated softwood. The results obtained emphasize that the saccharification performance of

endomannanases varies significantly, but cannot be predicted by initial rates on pure mannans or the ability to accommodate galactose substitutions. Rather, the enzymes ability to act on the mannan in complex with the lignocellulosic matrix that determined its performance. The best performing endomannanase in softwood saccharification was *Tres*Man5A from *Trichoderma reesei*. The catalytic efficiency of the core module and the presence of the CBM1 both played important roles in the superior performance.

Substrate binding amino acids in the fungal GH26 endomannanses were identified by solving and analyzing the crystal structure of a novel GH26 endomannanase from *Yunnania penicillata* in complex with α-6^2-6^1-di-galactosyl-mannotriose (MGG). The capability of accommodating multiple galactopyranosyl side-groups in the binding cleft was found to be conserved among the eight fungal GH26 endomannanases examined, as seen from the identified substrate binding amino acids which were highly conserved among these enzymes. Furthermore, all eight GH26 endomannanases reached full conversion of guar gum. *Ypen*Man26A mutated variants were designed based on the few variations found in the substrate binding amino acids among the studied GH26 endomannanases. A novel mass spectrometry-based method was used to determine k_{cat}/K_M of the *Ypen*Man26A wild type and the D37T mutant on α-6^4-6^3-di-galactosyl-mannopentaose (MGGMM). The results showed that this single amino acid substitution in the -2 subsite of *Ypen*Man26A compromised interactions with the galactose side group in this subsite. A more pronounced structural difference between the GH26 endomannanases was found in the area of the core module approaching the CBM35. The enzymes carrying a CBM35 all seem to have an α-helix, which allows ordered interactions with the binding domain, whereas the enzymes without a CBM had a flexible surface loop.

The results emphasize that great diversity exists in the specificity of fungal endomannanases and that this specificity affects endomannanase performance in mannan biomass conversion. The data also supported the assumption that structural features play significant roles in substrate recognition, i.e. galactose-accommodation, of endomannanases. The results have potential to affect future selection of endomannanases for industrial applications as well as the design of future screening campaigns.

Dansk sammenfatning

Træ er en udbredt bioresurse, og interessen for at bruge denne ressource som biobrændsel og til andre biobaserede produkter vokser. Historisk har forskningen på dette felt hovedsageligt fokuseret på cellulose, men i de senere år er interessen for at udnytte alle træets bestanddele steget herunder ligning og hemicellulose. I nåletræ er den dominerende hemicellulose acetyleret galaktoglukomannan, som udgør op til 25 vægtprocent af træets tørvægt. Endo-$\beta(1 \rightarrow 4)$-mannanaser (endomannanaser) katalyserer hydrolysen af β-mannaner, herunder galaktoglukomannan. Imidlertid er det vist, at galaktose sidegrupperne på mannan hovedkæden delvist forhindrer den enzymatiske nedbrydelse af substratet. Derudover reducerer træmassens genstridige natur også den enzymkatalyserede nedbrydning. For at overkomme disse forhindringer og øge industriel forsukring af nåletræ, er det vigtigt at forstå den molekylære baggrund for endomannanasernes naturlige specificitet og substrat genkendelse. Denne afhandling sigter mod at øge vores viden om den naturlige specificitet af fungale GH5 og GH26 endomannanaser. Et yderligere formål er at undersøge de fungale endomannanaser i nåletræs forsukring samt at undersøge, om forskelle i effektivitet korrelerer med bestemte enzym karakteristika.

I arbejdet der førte til denne afhandling blev indflydelsen af galaktosesidegrupper på enzymkatalyseret hydrolyse af mannaner undersøgt for fem fungale endomannanaser fra GH familie 5 og 26. Den initiale hydrolysehastighed og omdannelsen af to galaktomannaner fra johannesbrød mannan og den mere substituerede mannan fra guar blev bestemt for alle fem enzymer. Nedbrydningsprodukter blev analyseret med metoden DNA-sekventerings Assisteret Sukker analyse med Høj gennemstrømning (DASH). Ud fra tidligere krystalstrukturer og nye homologimodeller blev det analyseret, hvordan mannaner med sidekæder kan passe i bindingskløfter. Ud fra disse undersøgelser blev det vist, at fungale GH26 endomannanaser, inklusiv det nye *Anid*Man26A enzym fra *Aspergillus nidulans*, har en meget åben bindingskløft og kan rumme galaktosesidekæder i -2, -1 og +1 bindings subsite. Bindinger i denne GH26 endomannanase muliggjorde fuld hydrolyse af guar galaktomannan med tilhørende produktion af mange galaktomannooligosakkarider. De undersøgte fungale GH5 endomannanaser blev hæmmede af galaktosesidegrupperne, hvilket afspejles i en lavere omdannelse af guar mannan grundet en mere snæver bindingskløft.

Ti fungale endomannanaser fra både GH5 og GH26 blev evalueret på forskellige rene mannan substrater men også i enzymatisk forsukring af nåletræ, for at undersøge om evnen til at rumme mange galaktosesidegrupper i bindingskløften medfører en bedre forsukring af nåletræ. Resultaterne understregede, at endomannanasers effektivitet i forsukring af nåletræ varierer

signifikant, og at effektiviteten kan ikke forudsiges fra initiale hastigheder på rene mannan substrater eller fra evnen til at rumme galaktosesidekæder i bindingskløften. I stedet så det ud til, at effektiviteten blev afgjort af hvor gode enzymerne var til at nedbryde mannanen når den befandt sig i lignocellulose matricen. Den endomannanase der havde den bedste ydeevne i forsukring af nåletræ var *Tres*Man5A fra *Trichoderma reesei*. Både effektiviteten af det katalytiske domæne og CBM1 spillede en vigtig rolle i den overlegne ydeevne.

De substratbindende aminosyrer i de fungale GH26 endomannanaser blev identificeret ved at løse og analysere krystalstrukturen af en ny fungal GH26 endomannanase fra *Yunnania penicillata* i kompleks med α-6^2-6^1-di-galaktosyl-mannotriose (MGG). Evnen til at rumme mange galaktosesidegrupper i substratbindingskløften viste sig at være konserveret blandt de otte undersøgte fungale GH26 endomannanaser. Dette viste sig dels ved, at de identificerede substratbindende aminosyrer i *Ypen*Man26A var konserveret i de andre GH26 endomannanaser, og ved de alle opnåede fuld omdannelse af guar gummi. *Ypen*Man26A varianter blev designet ud fra de få variationer der blev fundet i de substratbindende aminosyrer mellem de undersøgte GH26 endomannanaser. Enzymernes Micahelis-Menten kinetiske parametre blev undersøgt på johannesbrød mannan og guar mannan samt på α-6^4-6^3-di-galaktosyl-mannopentaose (MGGMM), v.h.a. en ny massespektrometribaseret metode. Med denne metode blev det vist, at en enkelt aminosyre ændring i -2 subsitet i *Ypen*Man26A reducerede interaktionen med galaktosesidekæden i dette subsite. En mere udtalt strukturel forskel mellem GH26 endomannanaserne blev fundet i det område af det katalytiske domæne, der støder op til deres CBM35. Enzymer, der bærer et CBM35, ser alle ud til at have en α-helix, der tillader ordnede interaktioner med bindingsdomænet, hvorimod enzymerne uden et CBM har et fleksibelt overfladeloop.

Resultaterne understreger, at der findes stor specificitets diversitet for fungale endomannanaser, og at specificiteten påvirker enzymernes ydeevne i omdannelsen af mannanholdig biomasse. Data verificerede også antagelsen om, at strukturen af bindingskløften i endomannanaser spiller en signifikant rolle i substratgenkendelsen, fx. ift. at kunne akkommodere galaktose sidekæder. Resultaterne har potentiale til at påvirke fremtidig udvælgelse af endomannanaser til industriel anvendelse og til at påvirke, hvordan fremtidige screeningskampagner bliver designet.

List of publications

The thesis is based on the following articles and manuscripts, which are appended at the end of the thesis.

I. An *Aspergillus nidulans* GH26 endo-*β*-mannanase with a novel degradation pattern on highly substituted galactomannans
Pernille von Freiesleben, Nikolaj Spodsberg, Thomas H. Blicher, Lars Anderson, Henning Jørgensen, Henrik Stålbrand, Anne S. Meyer and Kristian B.R.M. Krogh. *Enzyme and Microbial Technology*. 83: 68-77; 2016

II. Boosting of enzymatic softwood saccharification by fungal GH5 and GH26 endomannanases
Pernille von Freiesleben, Nikolaj Spodsberg, Anne Stenbæk, Henrik Stålbrand, Kristian B.R.M. Krogh and Anne S. Meyer. *Biotechnology for Biofuels*. 11: 194; 2018

III. Crystal structure and substrate interactions of an unusual fungal non-CBM carrying GH26 endo-*β*-mannanase from *Yunnania penicillata*
Pernille von Freiesleben, Olga V. Moroz, Elena Blagova, Mathias Wiemann, Nikolaj Spodsberg, Jane W. Agger, Gideon J. Davies, Keith S. Wilson, Henrik Stålbrand, Anne S. Meyer, and Kristian B. R. M. Krogh; *Submitted manuscript*

Other publications

β-Mannanase-catalyzed synthesis of alkyl mannosides
Johan Morrill, Anna Månberger, Anna Rosengren, Polina Naidjonoka, Pernille von Freiesleben, Kristian B.R.M. Krogh, Karl-Erik Bergquist, Tommy Nylander, Eva N. Karlsson, Patrick Adlercreutz and Henrik Stålbrand. *Applied Microbiology and Biotechnology*. 102: 5149-5163; 2018

Polypeptides having mannanase activity and polynucleotides encoding same. Patent WO2015040159
Kristian B.R.M. Krogh, Nikolaj Spodsberg, Klaus Gori and Pernille von Freiesleben

List of abbreviations

α9	α-helix 9 in *Podospora anserina* GH26 endomannanase A
ABI 3730xl	96 capillary array DNA sequencer
APTS	9-aminopyrene -1, 4, 6-trisulfonate
BSA	Bovine serum albumin
CAZy	Carbohydrate active enzymes database
CBM	Carbohydrate-binding module
CE-LIFE	Capillary electrophoresis with laser-induced fluorescence
DASH	DNA sequencer-assisted saccharide analysis in high throughput
DE	Dextran unit
DM	Dry matter
DP	Degree of polymerization
DSC	Differential scanning calorimetry
Endoglucanase	Endo-β(1→4)-glucanase
Endomannanase	Endo-β(1→4)-mannanase
Glucosidase	Exo-β-glucosidase
GH	*O*-glycoside hydrolase
HPLC	High-performance liquid chromatography
LC	Liquid chromatography
Mannan	β-1,4-mannan
Mannosidase	Exo-β-mannosidase
G	α-galactosyl-mannose
MG	α-6^1-galactosyl-mannobiose
MGG	α-6^2-6^1-di-galactosyl-mannotriose
MGGM	α-6^3-6^2-di-galactosyl-mannotetraose
MGGMM	α-6^4-6^3-di-galactosyl-mannopentaose
MGMM	α-6^3-galactosyl-mannotetraose
MMG	α-6^1-galactosyl-mannotriose
MRM	Multiple reaction monitoring
MS	Mass spectrometry
PACE	Polysaccharide analysis by carbohydrate gel electrophoresis
PKC	Palm kernel press cake

Enzymes

*Anid*Man26A	*Aspergillus nidulans* GH26 endomannanase A
*Anid*Man5A	*Aspergillus nidulans* GH5 endomannanase A
*Anid*Man5B	*Aspergillus nidulans* GH5 endomannanase B
*Anid*Man5C	*Aspergillus nidulans* GH5 endomannanase C
*Anig*Man5A	*Aspergillus niger* GH5 endomannanase A
*Asti*Man26A	*Ascobolus stictoideus* GH26 endomannanase A
BM2	*Aspergillus niger* GH2 mannosidase clade A
*Bova*Man26A	*Bacteroides ovatus* GH26 endomannanase A
*Bova*Man26B	*Bacteroides ovatus* GH26 endomannanase B
*Cjap*Man26C	*Cellvibrio japonicus* GH26 endomannanase C
*Cjap*Man5A	*Cellvibrio japonicus* GH5 endomannanase A
*Cvir*Man26A	*Collariella virescens* GH26 endomannanase A
*Mthe*Man26A	*Mycothermus thermophiles* GH26 endomannanase A
*Mt*Man26A	*Myceliophthora thermophila* GH26 endomannanase A
*Ndes*Man26A	*Neoascochyta desmazieri* GH26 endomannanase A
*Pans*Man26A	*Podospora anserina* GH26 endomannanase A
*Pans*Man26A core	The catalytic domain of *Podospora anserina* GH26 endomannanase A
*Pans*Man5A	*Podospora anserine* GH5 endomannanase A
*Tres*Man5A	*Trichoderma reesei* GH5 endomannanase A
*Tres*Man5A core	The catalytic domain of *Trichoderma reesei* GH5 endomannanase A
*Wsp.*Man26A	*Westerdykella sp.* GH26 endomannanase A
*Ypen*Man26A	*Yunnania penicillate* GH26 endomannanase A

1 Hypotheses and objectives

The aim of this thesis was to contribute to the understanding of fungal endomannanase diversity in relation to specificity, with focus on differences between fungal GH5 and GH26 enzymes. A further aim was to evaluate performance of fungal endomannanases in softwood saccharification and to assess if performance differences correlated with specific enzyme characteristics. The reason for targeting fungal enzymes was the industrial preference for expressing enzymes for biomass saccharification in a fungal expression strain, in case a promising candidate for softwood saccharification could be identified.

Galactose substitutions on mannan backbones have long been known to negatively affect enzymatic hydrolysis of mannans. The significance of galactose substitutions on mannan substrates has been studied for fungal GH5 endomannanases but not in detail for fungal GH26 endomannanases. The work of this thesis began by formulating hypothesis H1 to address the compromising effect of galactose substitutions, which potentially limits the utilization of endomannanases in applications using galactomannans or softwood galactoglucomannans. Hypothesis H2 and H3 were formulated based on the findings from H1.

H1 Fungal GH26 endomannanases may include enzymes that are less hindered by galactose substitutions on the mannan backbone than fungal GH5 endomannanases. If this hypothesis is valid, GH5 and GH26 fungal endomannanases will exhibit different degradation patterns and kinetics on galactomannans. Since the effect of the galactose substitutions must relate to how the substrate is fitted in the active site cleft of the enzymes, differences in enzyme kinetics and substrate degradation must be due to structural differences of the enzymes.

H2 Fungal endomannanases of GH5 and GH26 differ in their capacity to catalyze removal of galactoglucomannans from cellulose microfibrils, and thus in turn may have different effects on enzymatic cellulose saccharification.

H3 The ability to accommodate multiple galactopyranosyl moieties in the active site cleft is a common feature among fungal GH26 endomannanases and semi-conserved residues associated with this feature can be identified.

To examine the proposed hypotheses, the following objectives were set:

Obj. 1. To analyze the enzyme kinetics and degradation patterns on galactomannans and the active site cleft architecture of fungal GH5 and GH26 endomannanases (using the

sequence collection at Novozymes as a starting point for selecting the enzymes to be expressed and examined).

Obj. 2. To measure the performance of selected fungal GH5 and GH26 endomannanases in softwood saccharification and evaluate which enzyme characteristics seem to be important for the measured performance.

Obj. 3. To experimentally identify amino acids involved in binding galactomannans in the active site cleft of fungal GH26 endomannanases, but also to evaluate if the substrate binding amino acids are conserved and to analyze how non-conserved amino acids affect the catalytic rate and substrate affinity.

2 Introduction to lignocellulosic biomass

In the next sections, general features of the three main types of polymers in lignocellulose – cellulose, lignin and hemicellulose - are briefly introduced. Focus will be on hemicelluloses and particularly on mannans.

2.1 Softwood composition and structure

Softwood contains three main polymers in the plant cell walls, cellulose, hemicellulose and lignin, that are linked together in complex structures to provide the tree with mechanical strength and a transport system for water and nutrients (reviewed in Timell 1967; Gibson 2012). Plant cell walls consist of two major parts: the primary cell wall, formed during growth and cell division, and the secondary cell wall, formed inside the primary cell wall when the plant cell reaches its final size and shape. The primary cell wall is a thin, flexible and highly hydrated layer, while the secondary wall is thicker and adds rigidity and strength. Secondary cell walls make up the major part of wood, but are not present in all plant cells (reviewed in Timell 1967). Both layers contain a complex aggregate of cellulose microfibrils and hemicelluloses that interact by hydrogen bonds. The secondary cell wall can be divided into three layers (S_1, S_2 and S_3), each with a different organization of the microfibrils (reviewed in Timell 1967). In wood, significant amounts of lignin are formed at the end of the cell wall thickening (formation of the secondary cell wall) (reviewed in Timell 1967), after which the plant cell may die while the cell walls remain to support the plant, with the lumen being used for water and nutrient transport (reviewed in Gibson 2012). Besides lignocelluloses, softwood cell walls contain structural glycoproteins and other polysaccharides, e.g. pectins (reviewed in Timell 1967; Gibson 2012).

2.2 Cellulose

Cellulose, the most abundant polysaccharide in nature, consists of β-1,4-linked D-glucopyranosyl units in a linear polymeric chain. The Degree of Polymerization (DP) varies with origin and treatment of the biomass, and typically ranges from 300-1700 in some types of wood pulp to 800-10000 or more in some plant fibers such as cotton (reviewed in Klemm et al. 2005). The cellulose chains are held together by intermolecular hydrogen bonds forming insoluble cellulose microfibrils. The microfibrils are remarkably stable and highly resistant to degradation (reviewed in Wolfenden & Snider 2001). The microfibrils are arranged in parallel orientation and the order of this arrangement has traditionally been used to divide cellulose into crystalline (highly ordered) or amorphous (less ordered) regions, although several regions fall into an intermediate range between the two extremes (Park et al. 2010). Cellulose has been used for a

long time in several major industrial applications such as paper production and cotton textile weaving (reviewed in Klemm et al. 2005). More recently, cellulose has attracted attention as a renewable raw material in biorefineries for production of bio-based products (reviewed in Cherubini 2010), the aim being to replace fossil-based products in areas such as fuels, chemicals and polymers (reviewed in Oh et al. 2015).

2.3 Lignin

After cellulose, lignin is the most abundant polymeric organic substance in the plant world (reviewed in Crestini et al. 2010). Lignin's are aromatic polymers formed by oxidative coupling of mainly three phenylpropanoid precursors. When incorporated into the lignin polymers, the monomeric alcohol precursors are known as *p*-hydroxyphenyl, guaiacyl, and syringyl units (reviewed in Vanholme et al. 2010). These units are connected by ether bonds and carbon-carbon bonds to form lignin polymers (reviewed in Crestini et al. 2010). Softwood contains about 30 % lignin which is mainly composed of guaiacyl units with minor amounts of *p*-hydroxyphenyl units. The lignin content in hardwood is somewhat lower and consists mainly of guaiacyl and syringyl units in approximately equal ratios (Espiñeira et al. 2011). Historically, lignin is considered unsuitable for use for biomass processing and is primarily used for combustion to generate energy. However, increased research focuses on generating high-value products from this resource (reviewed in Ragauskas et al. 2014).

2.4 Hemicelluloses

Hemicelluloses are estimated to account for one third of all components available in plants (reveiwed in Moreira & Filho 2008). Hemicelluloses generally have a lower DP, reported to be between 70 – 200, compared to cellulose (reviewed in Gibson 2012). Hemicelluloses are a heterogenous group of polysaccharides with a backbone composed of mainly β-1,4-linked D-pyranosyl residues. The major hemicelluloses are divided into four groups of structurally different polysaccharides classified according to their backbone monosaccharides: xylans, mannans, xyloglucans, and mixed-linkage β-glucans. The backbone residues of hemicelluloses can be modified by various side groups such as (methyl)glucuronic acid, arabinose, galactose, ferulic acid, and acetyl residues, which makes some hemicelluloses branched and with varying structure and water solubility (reviewed in Ebringerová 2006; Scheller & Ulvskov 2010).

Mixed-linkage β-glucans, here referring to glucose based plant polysaccharides other than cellulose, are found in the cell walls of grains such as oat and barley. They have mixed β-1,3- and β-1,4-linked D-glucopyranosyl units in the backbones (reviewed in Rahar et al. 2011). In Xyloglucans, the glucose backbones can be substituted with α-1,6-linked D-xylopyranoside groups, are the major hemicellulose in the primary cell walls of most land plants. The xylose

residues can carry additional sugars, especially galactose or fucose (reviewed in Fry et al. 2008). Potential applications of glucans include prebiotics (reveiwed in Charalampopoulos et al. 2002), but they may also have other health beneficial properties (reviewed in Rahar et al. 2011).

Xylans have backbones consisting of *β*-1,4-linked D-xylopyranoside units and varying degrees of side chains of arabinose, glucuronic acid and other sugars. *O*-Acetylation can occur on C2 and C3 (reviewed in Teleman et al. 2000; Thorsheim et al. 2015). They are the dominant hemicelluloses in grasses and hardwoods such as birch and oak, where they constitute 25-35 % of the wood dry matter (reviewed in Ebringerová & Heinze 2000) but are also found to a lesser extent in softwoods (Várnai et al. 2011). Grain may also contain 30-50 % xylan (reviewed in Izydorczyk & Biliaderis 1995). An even substitution pattern on the xylan backbone in vascular plants has recently been found crucial for the interactions with the cellulose and essential for development of normal secondary plant cell walls (Grantham et al. 2017). Xylans, and xylan derived oligosaccharides, have recently been found to be valuable as dietary fibers due to their potential prebiotic and other health promoting properties (reviewed in Broekaert et al. 2011).

2.5 *β* -Mannans

β-1,4-Mannans (mannans) can be classified in four subfamilies based on their chemical characteristics: linear mannans, galactomannans, glucomannans and galactoglucomannans (Figure 1). These classes of mannans are found in a variety of biomasses (Table 1) (reviewed in Dhawan & Kaur 2007).

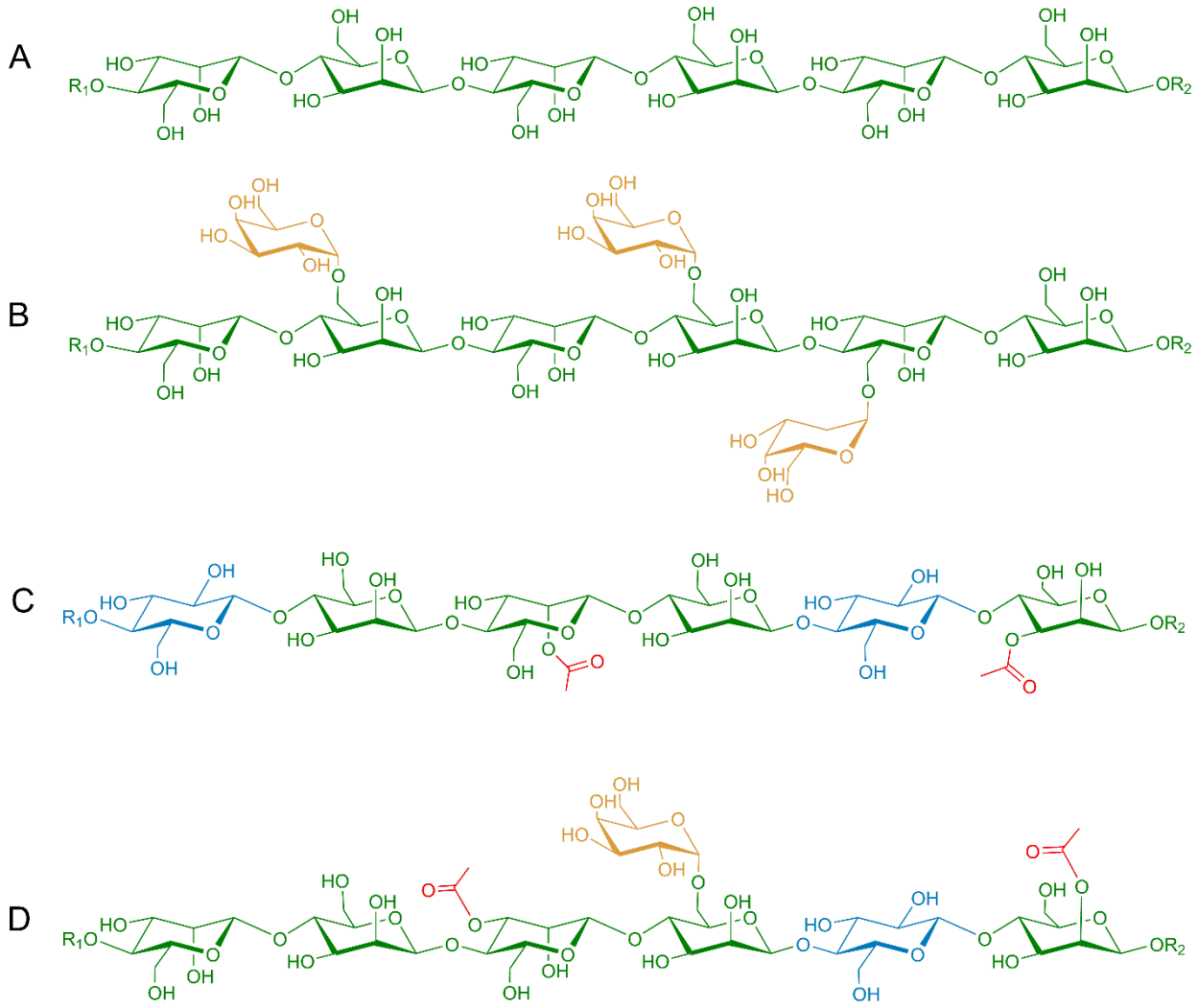


Figure 1: The four *β*-mannan classes: linear mannan (**A**), galactomannan (**B**), *O*-acetylated glucomannan (**C**) and *O*-acetylated galactoglucomannan (**D**). R_1 and R_2 indicate extensions of the polysaccharide chain towards the non-reducing and reducing end, respectively. According to the consortium for functional glycomics notation (Raman et al. 2006), mannose units are colored green, glucose units blue, and galactose units yellow. *O*-acetyl groups are colored red.

Table 1: Selected natural β-mannan characteristics and sources

β-mannans	**Source**	**Substitution pattern**	**Reference**
Linear mannans	Ivory nut	(< 2 % Gal)	(Aspinall et al. 1953) (Timell 1957)
	Green coffee bean	Man:Gal, 50:1	(Wolfrom et al. 1961)
	Palm kernel	Man:Gal, 20:1	(Jørgensen et al. 2010)
Galactomannans	Locust bean gum *	Man:Gal, 4:1	(McCleary 1985)
	Guar gum	Man:Gal, 2:1	(McCleary 1985)
Glucomannans	Konjac gum	Man:Glc, 1.6:1, Ac	(Kishida et al. 1978) (Katsuraya et al. 2003) (Ratcliffe et al. 2005)
	Hardwood	Man:Glc, 2:1, Ac	(Moreira & Filho 2008)
Galactoglucomannans	Softwood	Man:Glc:Gal, 3.5-4.5:1:0.5-1.1, Ac	(Lundqvist et al. 2003) (Willför et al. 2003) (Xu et al. 2010)

* Also known as carob galactomannan

Linear mannans (Figure 1A) have a main chain of β-1,4-linked D-mannopyranosyl residues and contain less than 5 % of α-1,6-linked D-galactopyranosyl residues (reviewed in Moreira & Filho 2008). Because of the low substitution, linear mannans resemble cellulose and are water insoluble. Linear mannans are found in ivory nuts, algae, green coffee beans and palm kernels (reviewed in Moreira & Filho 2008). Linear mannans with two different DPs (DP 10-13 and DP 39-40) have been extracted from ivory nut (Aspinall et al. 1953; Timell 1957). In palm kernel press cake (PKC), which is a residue from palm oil extraction, around 42 % of the dry matter is linear mannan (Jørgensen et al. 2010).

Galactomannans (Figure 1B) consist of the same β-1,4-linked mannan backbone as linear mannans, but have a higher degree of single α-1,6-linked D-galactopyranosyl residues attached along the chain and are therefore mostly water soluble (reviewed in Moreira & Filho 2008). Differences in the distribution of galactosyl units along the backbone vary between different sources. The side chains prevent intermolecular interactions between backbone molecules in different chains. Guar gum, produced from the seeds of the guar plant (*Cyamopsis tetragonolobus*), and locust bean gum, produced from the seeds of the carob tree (*Ceretonia siliqua*), are significant sources of galactomannans. Guar gum contains more galactopyranosyl groups (Man:Gal, 2:1)

than locust bean gum (Man:Gal, 4:1) (reviewed in Moreira & Filho 2008). In locust bean gum, the distribution of galactopyranosyl residues is irregular with a high proportion of unsubstituted blocks. In guar gum, in contrast, the galactopyranosyl residues are more ordered and found mainly in pairs and triplets with few non-substituted regions (McCleary 1985). The DP of locust bean gum and of guar gum has been reported to be 1500 and 900, respectively (McCleary 1985). Galactomannans have a thickening effect, which is a property that has been employed in a variety of industrial food applications (discussed in Kök 2010).

In glucomannans (Figure 1C), the backbone is constituted of both β-1,4-linked D-mannopyranosyl and D-glucopyranosyl residues. Mannose is a C-2 epimer of glucose. Hardwoods contain 3-5 % glucomannan which in general does not have galactopyranosyl side groups, but may be acetylated (reviewed in Moreira & Filho 2008; Schröder et al. 2009; Chauhan et al. 2012). Hardwood hemicelluloses have a DP around 150 to 200 (reviewed in Moreira & Filho 2008). Glucomannans are also found in the primary cell walls of some dicots (flowering plants whose seeds have two embryonic leaves) (reviewed in Carpita & Gibeaut 1993; Moreira & Filho 2008). One such source is konjac gum from tubers of *Amorphophallus konjac*, which has a Man:Glc ratio of approximately 1.6:1 (Katsuraya et al. 2003) and an observed DP around 6000 (Kishida et al. 1978). Konjac is acetylated and is therefore quite water soluble (Ratcliffe et al. 2005). Konjac glucomannan is used as a food thickener and gelling agent (reviewed in van Zyl et al. 2010).

The most abundant type of mannan is *O*-acetylated galactoglucomannan (Figure 1D) (reviewed in Lundqvist et al. 2003). This is the dominant hemicellulose in softwoods such as spruce (*Picea abies*) where it comprises 15-25 % of the wood Dry Matter (DM) (reviewed in Timell 1967; Lundqvist et al. 2003). galactoglucomannan consist of a β-1,4 linked backbone of D-mannopyranosyl and D-glucopyranosyl units. The mannose residues can be decorated with single α-1,6-linked D-galactopyranosyl side groups and up to 32 % of them can be acetylated on C-2 and/or C-3 (Xu et al. 2010). The DP of galactoglucomannans from spruce has been reported to be around 150-200 (reviewed in Timell 1967; Lundqvist et al. 2003), and the typical Man:Glc:Gal ratio in Norway spruce is 3.5-4.5:1:0.5-1.1 (Lundqvist et al. 2002; Willför et al. 2003; Bååth et al. 2018). Variations in the ratios depend on the raw material, but extraction method and pretreatment can also reduce the amount of backbone decorations. Galactoglucomannan can be extracted from spruce wood by heat fractionation (Lundqvist et al. 2002; Lundqvist et al. 2003) but it can also be recovered from industrial biomass side streams by filtration, chromatography and/or enzymatic treatment (Andersson et al. 2007). Softwood has significant potential as feedstock for renewable energy production and biorefining due to its abundancy, low cost, and lack of competition with the food and feed industry **(Paper II)**.

3 Introduction to mannan degrading enzymes

A variety of enzymes are involved in catalyzing the biosynthesis, degradation and modification of plant polysaccharides. One major class is the enzymes catalyzing the hydrolysis of *O*-glycosidic linkages, called *O*-Glycoside Hydrolases (EC 3.2.1-) (GHs) (reviewed in Sinnott 1990). Many organisms that feed on plant material have GHs such as cellulases (reviewed in Bayer et al. 1998) and hemicellulases, including mannanases; the former enzymes act on cellulose and the latter on hemicellulose. Wood degrading fungi and soil inhabiting filamentous fungi and bacteria are significant producers of mannan degrading enzymes (reviewed in van Zyl et al. 2010), most likely due to the high mannan content in softwood (Lundqvist et al. 2003) and other mannan containing plant materials. More recently, it has been shown that the genomes of some human gut bacteria also harbor mannan degrading enzymes, possibly to utilize mannans in our diet (Kulcinskaja et al. 2013; Reddy et al. 2016). Furthermore, a marine mollusk (the blue mussel *Mytilus edulis*) (Larsson et al. 2006), an earthworm (*Eisenia fetida*) (Ueda et al. 2018), beetles (*Gastrophysa viridula* and *Callosobruchus maculatus*) (Busch et al. 2017) and a protist of the termite *Recticulitermes speratus* (Tsukagoshi et al. 2014) have also been found to express endomannanases.

GHs have both endo- and exo- modes of action. Endo-acting GHs catalyze the hydrolysis of internal glycosidic bonds in the polysaccharide chain (reviewed in Davies & Henrissat 1995). Conversely, exo-acting GHs predominantly catalyze the cleavage of one or two monosaccharide units at a time often from the the non-reducing end of the saccharide chain (reviewed in Davies & Henrissat 1995; Reddy et al. 2013). Another enzyme classification scheme is the enzyme commission (EC) number, which is based on the chemical reaction being catalyzed. Enzymes from different organisms and with different folds are assigned the same EC number as long as they catalyze the same chemical reaction. Enzymes in EC class 3 are hydrolases, with GHs in subclass 3.2 whereas enzymes in subclass 3.1 catalyze the hydrolysis of ester bonds. Mannan degrading enzymes comprise both endo-and exo-acting GHs as well as an esterase: endo-β(1→4)-mannanases (endomannanases, EC 3.2.1.78), exo-β-mannosidases (mannosidases, EC 3.2.1.25), exo-β-glucosidases (glucosidases, EC 3.2.1.21), α-galactosidases (EC 3.2.1.22), and acetyl mannan esterases (EC 3.1.1.6) (reviewed in Moreira & Filho 2008; Malgas et al. 2015a). Certain endo-β(1→4)-glucanases (endoglucanases, EC 3.2.1.4) primarily attacking celluloses have also been shown to be active on glucomannans (Mikkelson et al. 2013). The mannan degrading enzymes attack mannan substrates at specific structures outlined in Figure 2. Endomannanases catalyze the random hydrolysis of mannosidic bonds within the mannan backbone. Their cleavage is affected by the extent and pattern of galactose and /or acetyl substitutions and the extent and pattern of glucose residues in the mannan backbone (McCleary & Matheson 1983; Dilokpimol et al. 2011; Bååth et al. 2018) (**Paper I, II** and **III**). Endomannanases are important enzymes for

facilitating the solubilization of the mannan polymer and for release of mannan from a substrate matrix, such as the lignocellulose in softwood (Tenkanen et al. 1997; Várnai et al. 2011). Mannosidases catalyze the hydrolysis of mannooligosaccharides and mannobiose, released by endomannanases. Several mannosidases are capable of degrading longer mannooligosaccharides with DP 2-6 (Ademark et al. 1999), and a few mannosidases may also aid in breaking backbone glyosidic linkages resulting in release of mannose from the non-reducing end of mannan based polymers (Reddy et al. 2013). Most often, a mannosidase can cleave up to a galactose substitution, but not beyond (Ademark et al. 1999). Glucosidases catalyze the release of glucose residues from the non-reducing end of glucomannooligosaccharides, which are released by hydrolysis of glucomannans and galactoglucomannans by endomannanases (reviewed in Moreira & Filho 2008). *α*-galactosidases catalyze the release of the galactose substitutions from galacto(gluco)mannans and galactosylated oligosaccharides (Reddy et al. 2016), while acetyl mannan esterases catalyze the release of the acetyl side groups (Tenkanen et al. 1993). These enzymes work in a coordinated interplay to degrade mannan substrates and may have synergistic effects (Malgas et al. 2015b).

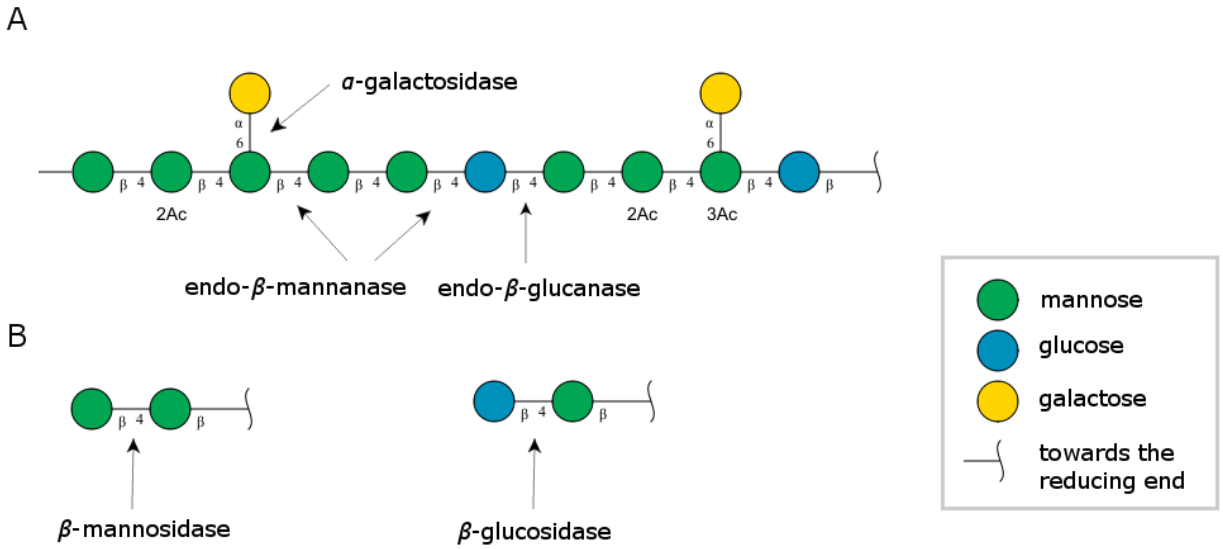


Figure 2: Schematic illustration of acetylated galactoglucomannan and glycoside hydrolases involved in hydrolysis of the backbone (**A**) and released oligosaccharides (**B**). Monosaccharides are shown using the consortium for functional glycomics notation (Raman et al. 2006). The figure shows general glycoside linkage specificity of each type of enzyme. A given enzyme may for instance be restricted by neighboring backbone monomers and/or substitutions, exemplified by the varying influence of galactopyranosyl substituents and potentially by backbone glucopyranosyl units on endomannanase activity (McCleary & Matheson 1983; Tenkanen et al. 1997). Illustration adapted from **Paper II**.

3.1 Endomannanases (EC 3.2.1.78)

3.1.1 Family classification and nomenclature

GHs are classified into GH families in the Carbohydrate-Active enZYmes database (CAZy) (www.cazy.org) (Lombard et al. 2014) according to amino acid sequence similarities, which gives a structural and mechanistic classification so that GHs with structurally related catalytic modules belong to the same family (Henrissat 1991; Henrissat & Bairoch 1996). There are currently 153 GH families in CAZy (www.cazy.org). To date, most endomannanases, including the ones studied in this thesis, have been classified into families GH5 and GH26, although some have recently also been found in GH113 (Zhang et al. 2008; Xia et al. 2016; You, Qin, Yan et al. 2018) and GH134 (Shimizu et al. 2015; Jin et al. 2016; Sakai et al. 2017; Sakai et al. 2018; You, Qin, Li et al. 2018). The endomannanases in family GH113 seem to be intracellular and so far are mostly from bacterial origin (You, Qin, Yan et al. 2018), whereas the endomannanases in family GH134 are found predominantly in fungi (Shimizu et al. 2015; Sakai et al. 2017; You, Qin, Li et al. 2018), but also in bacteria as seen in CAZy (www.cazy.org) (Lombard et al. 2014; Jin et al. 2016; Sakai et al. 2018).

GHs are also classified into clans based on their structural fold and conserved catalytic amino acids (Henrissat & Davies 1997). Until now, 17 different folds have been found among the GHs. Most known endomannanases, including those analyzed in the work of this thesis, belong to the largest GH clan, the clan GH-A (www.cazy.org) (Lombard et al. 2014). The GHs in clan GH-A share the $(\beta/\alpha)_8$-TIM barrel fold (Henrissat et al. 1995). The newly identified GH134 endomannanases have a lysozyme-like fold and catalyze the hydrolysis of the mannan backbone via an inverting mechanism, which is in contrast to GH5, 26 and 113 endomannanases (Jin et al. 2016).

In 1998, a scheme for naming GHs that enclosed the activity and family relation in the enzyme name was proposed and accepted. A capital letter was added to the end of the name to distinguish GHs from the same origin, activity and family (e.g. Man5A, Man5B etc. for GH5 endomannanases from the same organism) (Henrissat et al. 1998). To distinguish between GHs from different organisms, four letters from the organism's name are added to the beginning of the enzyme name throughout this thesis.

3.1.2 Catalytic machinery

GHs are classified into two major mechanistic classes. One class hydrolyzes the *O*-glycosidic bond with retention of the anomeric carbon at the reducing end of the polysaccharide (i.e. $\beta\rightarrow\beta$) while the second class does so with inversion (i.e. $\beta\rightarrow\alpha$) (reviewed in Sinnott 1990; Zechel & Withers 2000; Withers 2001; Davies & Henrissat 1995). Endomannanases from GH family 5, 26 and 113 hydrolyze the *O*-glycosidic bond with retention of the anomeric configuration (Figure 3). The retaining mechanism usually requires two catalytic carboxylates, a nucleophile and an

acid/base (reviewed in Zechel & Withers 2000; Withers 2001). In clan GH-A, the catalytic carboxylates are reported to be (mainly) glutamates and in general the transition states believed to be oxocarbenium ion-like (reviewed in Sinnott 1990; Withers 2001).

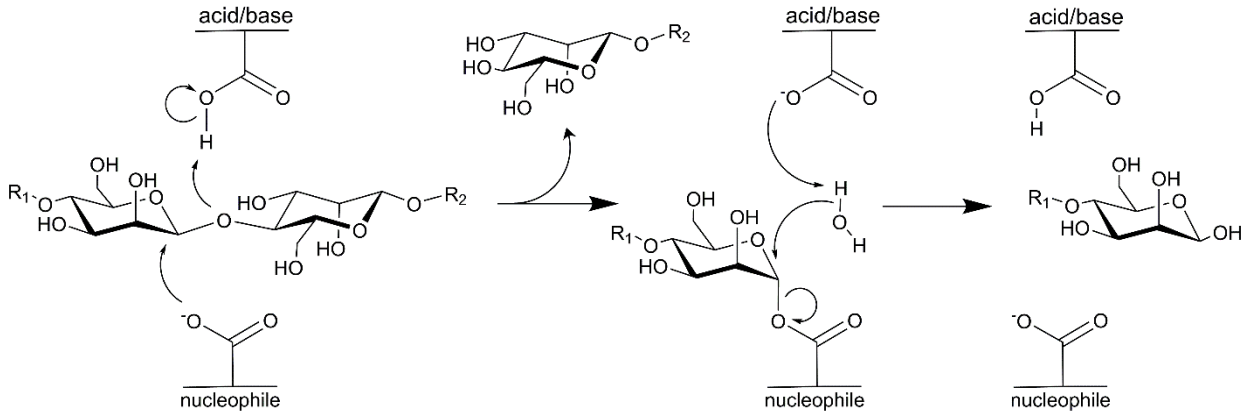


Figure 3: The retaining catalytic mechanism. In the first step, the nucleophile attacks the anomeric carbon. A covalent intermediate with inverted anomeric configuration is formed and the aglycone product is released. In the second step, water performs another nucleophilic attack on the anomeric carbon, which results in inversion of the configuration and release of the glycone product with overall anomeric retention. R_1 and R_2 represent extensions of the polysaccharide chain towards the non-reducing and reducing end, respectively.

In the retaining mechanism, the first step is a nucleophilic attack on the anomeric carbon by the nucleophile carboxylate residue, resulting in a covalent enzyme-glycoside intermediate and the release of the aglycone product (reviewed in Sinnott 1990; Withers 2001). For clan A endomannanases and mannosidases it has been suggested that the transition state has a boat-($B_{2,5}$)-like conformation (Ducros et al. 2002; Tailford et al. 2008). The step inverts the anomeric configuration of the glycosyl unit bound to the nucleophile. The acid/base carboxylate residue serves as proton donor in breaking the glycosidic bond (reviewed in Zechel & Withers 2000). The second step is a deglycosylation step where the anomeric carbon in the covalent intermediate is attacked by a nucleophilic water molecule and the glycone product is released from the enzyme (reviewed in Sinnott 1990; Withers 2001). In this step, the acid/base residue deprotonates the glycosyl acceptor, facilitating its nucleophilic attack. In this step, a second anomeric inversion takes place, which results in overall retention of the configuration of the anomeric carbon at the newly formed product (reviewed in Sinnott 1990). Some endomannanases can also perform transglycosylation, where a saccharide participates in the deglycosylation step instead of the water molecule. The transglycosylation results in the formation of a new *O*-glycosidic bond between two saccharides (reviewed in Sinnott 1990; Withers 2001).

3.1.3 Active site structure

Catalytic activity is affected by substrate binding at sites distant from the bond undergoing hydrolysis. Amino acids that bind one sugar monomer are grouped into a subsite (reviewed in Davies et al. 1997). In particular, aromatic residues are known to form hydrophobic platforms which interact with the sugar units (reviewed in Asensio et al. 2013). The active site of GHs contains several subsites where the sugar units of the polysaccharide substrate can bind. According to established nomenclature, these subsites are numbered by positive and negative integers (Figure 4) (Davies et al. 1997). The negative numbered subsites are positioned in the non-reducing end of the polysaccharide substrate (glycone binding subsites) and the positive subsites are located in the reducing end (aglycone binding subsites). The hydrolysis occurs between sugar units bound in the -1 and +1 subsite (Davies et al. 1997). In endomannanases, efficient hydrolysis often requires substrate binding to at least four subsites, which is why mannotrioses are mostly degraded slowly, (discussed by Dilokpimol et al. 2011 and references there in). The favored conformation of the galactomannan backbone is a flat ribbon with a two-fold axis, which places neighboring galactopyranosyl units on opposite side of the backbone. This means that every second galactopyranosyl moiety (if present) will point away from the enzyme, whereas galactopyranosyl residues in between may sterically clash with enzyme and thus prevent binding and subsequent hydrolysis (McCleary & Matheson 1983).

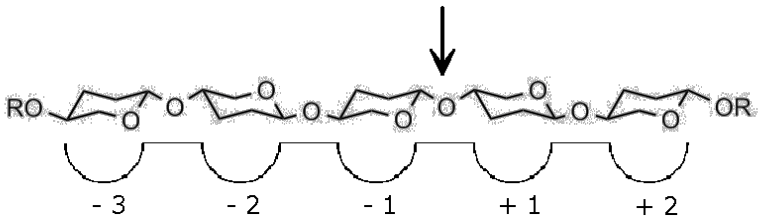


Figure 4: Schematic illustration of substrate binding in the active site subsites of endomannanases. The illustrated sugar binding subsites are labeled from -3 to +2. Hydrolysis occurs between subsite -1 and +1 as indicated by the arrow.

The enzymatic mode of action differs depending on the structure of the active site. Endo-enzymes, including endomannanases, often use an open active site cleft, in contrast to exo-enzymes the majority of which use a pocket shaped active site (reviewed in Davies & Henrissat 1995). Exo-acting GHs often have one or a few glycone subsites, corresponding to the hydrolysis of one mannose unit, while endo-acting GHs may have several glycone subsites. In endo-acting GHs, the substrate may bind to these subsites in different ways, depending on the architecture and substrate affinity in a specific enzyme, which results in different hydrolysis product profiles (Hekmat et al. 2010). Changes in subsite affinities, e.g. by mutagenesis, can affect product profiles, as seen in a study by Rosengren et al where a mutation, R171K, in the -2 subsite of a GH5 endomannanase from *Trichoderma reesei* decreased the affinity for mannooligosaccharides (by increase in K_M) and altered the binding mode and its transglycosylation capacity (Rosengren et

al. 2012). In **Paper III** the role of the -4 and -2 subsite was studied in GH26 endomannanase *Ypen*Man26A from *Yunnania penicillata*. **T**he binding mode may also be affected by the positioning and flexibility of loops in the active site. These loops may e.g. contribute to activity at lower temperatures (Kim et al. 2014) or they can change the mode of action from endo to exo by blocking subsites in distal regions of the active site cleft (Cartmell et al. 2008; Gonçalves et al. 2012) (discussed in **Paper I**).

3.1.4 Structural modularity

Each endomannanase has a catalytic module which harbors the active site. Some endomannanases also carry one or more additional modules, and Carbohydrate-Binding Modules (CBMs) is a common and also the most well-characterized type of non-catalytic module. CBMs are involved in the recognition and binding of carbohydrates and often seem to potentiate the activity of their attached catalytic module (reviewed in Gilbert et al. 2013). CBMs have been proposed to have three main roles in polysaccharide hydrolysis by GHs: (a) bringing the catalytic module to the proximity of the substrate, (b) maintaining proximity of the catalytic module and the target substrate, and (c) some CBMs maybe also mediate non-catalytic disruption of polysaccharide structures (reviewed in Boraston et al. 2004; Sunna 2010; Gilbert et al. 2013). Like the core modules, the CBMs are also classified into families based on sequence identity in CAZy (www.cazy.org) (Lombard et al. 2014). Today, there are 84 CBM families with different specificities listed.

CBMs are most often connected to the core module with a linker peptide in the N- and/or C-terminal. The linker between the catalytic module and a CBM is usually flexible, as seen in the few crystal structures of bacterial modular GH10 xylanases from *Cellvibrio japonicus* and *Streptomyces olivaceoviridis* (Fujimoto et al. 2002; Pell et al. 2004). In particular, linkers in fungal enzymes carrying a CBM1, including some endomannanases, are often long, flexible and highly glycosylated, which also makes them difficult to crystalize (Receveur et al. 2002). In contrast, the CBM35 in the GH26 endomannanase, *Pans*Man26A from *Podospora anserine,* and the CBM3c in some GH9 endoglucanases, including Cel9I from *Clostridium thermocellum*, are attached through a linker wrapped around the core domain (Couturier et al. 2013; Petkun et al. 2015).

3.2 Fungal endomannanases

To date, most known fungal endomannanases are found in family GH5 and to a smaller extent in family GH26 and GH134 (www.cazy.org) (Lombard et al. 2014). The work of this thesis has focused on the fungal GH26 endomannanases (**Paper I, II** and **III**), but has also compared their characteristics to known fungal GH5 endomannanases (**Paper I** and **II**). All endomannanases characterized in this work are from the fungal phylum Ascomycota. This selection was made on

the basis mainly of industrial experience with the utilization of enzymes from ascomycetes (including *Trichoderma reesei*) and partly also due to the desire to express in a fungal ascomycete cellulase expression strain in case of success. To date, fungal endomannanases from ascomycetes are by far the most studied (Jørgensen et al. 2010; Couturier et al. 2011; Várnai et al. 2011; Rosengren et al. 2014; Katsimpouras et al. 2016). But a few other fungal enzymes, such as the endomannanases ManA, ManB and ManC from the anaerobic fungus *Piromyces* (phylum: Chytridiomycota), have also been investigated (Fanutti et al. 1995; Millward-Sadler et al. 1996).

Limited knowledge is available that relates the sequence similarity classification of GH5 and GH26 to structural and functional differences among the two types of enzymes. Genome analysis has revealed that some organisms have endomannanases from both GH5 and GH26. The potentially different biological roles (substrate preferences) have been addressed for the bacterium *Cellvibrio japonicus* (Tailford et al. 2009; Zhang et al. 2014) and the fungus *Podospora anserina* (Couturier et al. 2013; Marchetti et al. 2016). Based on studies of the bacterial endomannanases from *C. japonicas,* it has been proposed that GH26 enzymes primarily attack soluble mannans, whilst the GH5 counterparts primarily attack insoluble mannans, including plant cell wall mannans (Tailford et al. 2009; Zhang et al. 2014). However, it is unclear whether this perception is valid for the fungal GH26 endomannanases (Couturier et al. 2013).

3.2.1 GH5 endomannanases

GH family 5 is a diverse family with many known enzyme activities. Some of the annotated proteins are cellulases, xylanases and endomannanases. Currently, the family is divided into 56 subfamilies based on sequence similarities as displayed in CAZy (www.cazy.org)(Lombard et al. 2014). Endomannanases in the GH5 family are classified into subfamily 7, 8, 10 and 17 (Aspeborg et al. 2012). GH5_7 and GH5_10 contain mainly eukaryotic endomannanases; GH5_7 predominantly contains endomannanases from fungi (Dilokpimol et al. 2011; Rosengren et al. 2012; Couturier et al. 2013) and plants (Wang et al. 2014; Wang et al. 2015), whereas GH5_10 contains several endomannanases from marine mollusks (Larsson et al. 2006; Mizutani et al. 2012). In contrast, GH5_8 and GH5_17 currently contain predominantly bacterial endomannanases (Hogg et al. 2003; Tailford et al. 2009; Aspeborg et al. 2012). The GH5 endomannanases studied in this thesis are all fungal enzymes and belong to GH5_7 (**Paper I** and **II**).

Several structures of GH5 endomannanases from various subfamilies have been solved (among these see Hilge et al. 1998; Sabini et al. 2000; Akita et al. 2004; Dias et al. 2004; Larsson et al. 2006; Tailford et al. 2009; Gonçalves et al. 2012; Mizutani et al. 2012; Couturier et al. 2013), and the *Tres*Man5A from *Trichoderma reesei* (Sabini et al. 2000) (Figure 5) is one of the most studied fungal GH5 endomannanases (**Paper I** and **II**) (Stålbrand et al. 1993; Stålbrand et al. 1995; Hägglund et

al. 2003; Rosengren et al. 2012), including in application studies that show degradation of galactoglucomannan from a lignocellulosic matrix by *Tres*Man5A (Rättö et al. 1993; Tenkanen et al. 1997; Várnai et al. 2011) (**Paper II**). Another fungal GH5 endomannanase used in application studies is *Anig*Man5A from *Aspergillus niger*, which was shown to degrade the linear mannan in PKC (Jørgensen et al. 2010) and which was also tested as a candidate for boosting of softwood saccharification in **Paper II**.

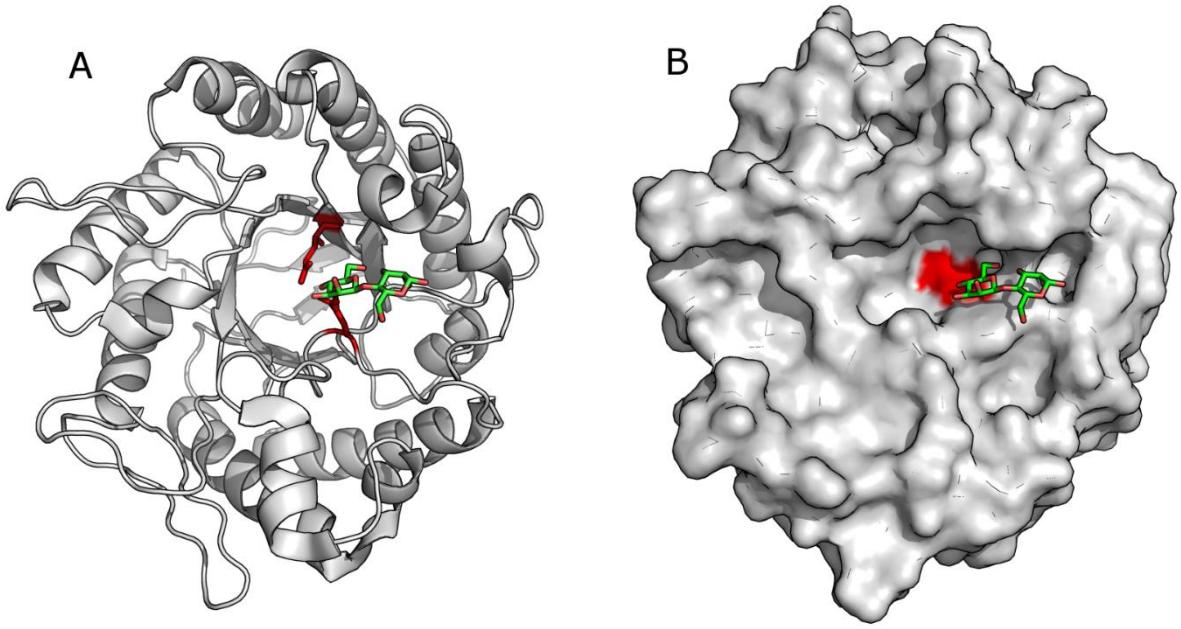


Figure 5: Crystal structure of *Tres*Man5A, a GH5 endomannanase from *Trichoderma reesei*. Cartoon representation showing the $(\beta/\alpha)_8$ barrel fold characteristic of clan A glycoside hydrolases (**A**). Surface representation of *Tres*Man5A (**B**). Both illustrations show the location of the two catalytic glutamates (red) and mannobiose bound in subsite +1 and +2 (green). The crystal structure of *Tres*Man5A (1QNR) was solved by Sabini et al (2000).

Many fungal GH5 endomannanases seems to have five subsites (often -3 to +2), and have mannobiose and mannotriose as their dominant hydrolysis products (McCleary & Matheson 1983; Dilokpimol et al. 2011; Chauhan et al. 2012; Couturier et al. 2013). Characterized endomannanases often show higher initial hydrolysis rates on locust bean gum compared to guar gum, and it has been proposed that the lower activity on guar gum is caused by the larger amount of galactopyranosyl residues (Dilokpimol et al. 2011; Malgas et al. 2015b). Studies investigating the accommodation of galactopyranosyl residues in the active site clefts of endomannanases, including a fungal GH5 endomannanase from *A. niger* and the *Tres*Man5A, found that they could be accommodated in subsite -1, but not in subsite -2 and +1 (McCleary & Matheson 1983; Tenkanen et al. 1997). In fact it has been highlighted that galactopyranosyl residues were absent from the non-reducing, terminal mannopyranosyl group in earlier identified hydrolysis products (McCleary 1979; McCleary et al. 1983; McCleary & Matheson 1983). Some GH5 endomannanases have been shown to be capable of transglycosylation (Larsson et al. 2006; Dilokpimol et al. 2011;

Couturier et al. 2013; Rosengren et al. 2014). The accommodation of galactopyranosyl residues in the active site cleft of fungal GH5 and GH26 endomannanases is examined in **Paper I**.

Some organisms have multiple GH5 endomannanases. This phenomenon has been studied in *Aspergillus nidulans*, which has a variety of different GH5 endomannanases, as indicated by those so far studied; *Anid*Man5A, *Anid*Man5B and *Anid*Man5C (Coutinho et al. 2009; Dilokpimol et al. 2011; Rosengren et al. 2014). These three GH5 endomannanases were found to have different characteristics. Of the three enzymes, *Anid*Man5C had the highest k_{cat} (220 s^{-1}) on locust bean gum, but was sensitive to galactose substitutions and glucose residues in the backbone (Dilokpimol et al. 2011; Rosengren et al. 2014). This sensitivity was seen as a drop in activity when tested on guar gum galactomannan and konjac glucomannan compared to locust bean gum (Dilokpimol et al. 2011). *Anid*Man5A showed lower k_{cat} (90 s^{-1}) on locust bean gum but appeared to be less sensitive to changes in the backbone composition and to the degree of substitution in the mannan substrates (Dilokpimol et al. 2011; Rosengren et al. 2014). *Anid*Man5B was found to have a high transglycosylation capacity and significantly lower k_{cat} (15 s^{-1}) on locust bean gum compared with *Anid*Man5A and *Anid*Man5C (Rosengren et al. 2014). Differences between *Anid*Man5A and *Anid*Man5C are further examined in **Paper I**.

Many fungal GH5 endomannanases are modular, and typically carry a CBM1. CBM1 is known to confer cellulose binding and increase the mannan hydrolysis of complex substrates, such as softwood and ivory nut extractions containing both mannan and cellulose, by endomannanases carrying this module (Hägglund et al. 2003; Pham et al. 2010).

3.2.2 GH26 endomannanases

The family GH26, currently containing predominantly bacterial endomannanases, appears less diverse than the GH5 family and has so far not been divided into subfamilies. Compared to GH5 endomannanases, fewer crystal structures of GH26 endomannanases have been solved (among these see Hogg et al. 2001; Ducros et al. 2002; Le Nours et al. 2005; Cartmell et al. 2008; Tailford et al. 2009; Couturier et al. 2013; Tsukagoshi et al. 2014; Bågenholm et al. 2017), and there is only one solved structure of an eukaryotic GH26 endomannanase, namely that of the *Pans*Man26A which carries a N-terminal CBM35 (Couturier et al. 2013). Apart from *Pans*Man26A, only a few other fungal GH26 endomannanases have been characterized, including *Mt*Man26A from *Myceliophthora thermophila* (Katsimpouras et al. 2016) and ManA, ManB and ManC from the anaerobic fungus *Piromyces* (Millward-Sadler et al. 1996). The solution and analysis of a new unique crystal structure of a fungal GH26 endomannanase is presented in **Paper III** and both **Paper I** and **II** also shed light on the functional characteristics of fungal GH26 endomannanases and their potential use in softwood saccharification.

*Pans*Man26A has been shown to have mannotetraose as the main product from ivory nut hydrolysis, with smaller amounts of mannotriose and mannose. Studies also revealed that *Pans*Man26A needs binding in 5 subsites for efficient hydrolysis, and predominantly binds mannopentaose from the -4 to the -1 subsite (Couturier et al. 2013). Tailford et al. state that GH26 endomannanases have strict selectivity for mannose in the -2 subsite, in contrast to GH5 endomannanases which can accommodate either glucose or mannose in subsite -2. It is suggested that GH26 endomannanases may therefore not be able to hydrolyze glucomannans as efficiently (Tailford et al. 2009). Furthermore, no native transglycosylation capacity has been detected among GH26 endomannanases (Couturier et al. 2013). The accommodation of galactopyranosyl residues had not been analyzed for fungal GH26 endomannanases before the work described in this thesis (**Paper I** and **III**). However, among the bacterial GH26 endomannanases there seems to be variation in their ability to accommodate these substitutions in the active site cleft, as exemplified by *Bova*Man26A and *Bova*Man26B from *Bacteroides ovatus* (Bågenholm et al. 2017).

Most known fungal GH26 endomannanases have a CBM35, which is an approximately 15 kDa binding domian (Couturier et al. 2013; Katsimpouras et al. 2016). The CBM35 family is known to include members that bind β-mannans, uronic acids, β-1,3-galactan or α-1,6-galactopyranosyl residues on carbohydrate polymers (Montanier et al. 2009; Correia et al. 2010). The binding site of CBM35s has been reported to be located between the loops connecting the β-strands and not on the concave surface presented by the β-strands as observed in CBM1 (Montanier et al. 2009; Correia et al. 2010). During the work described in this thesis, *Pans*Man26A was characterized both with and without its CBM35 (**Paper I** and **II**), and the differences between natural GH26 catalytic modules carrying and not carrying a CBM35 were investigated in **Paper III**.

3.3 Application of endomannanases

The presence of mannans in different biomasses suggests a variety of applications where endomannanases can be employed. This section focuses on the potential use of endomannanases in softwood hydrolysis. First, a list of selected application studies is presented which explores the utilization of endomannanases in biomass processing (Table 2).

Table 2: Selected application studies utilizing endomannanases in biomass processing

Application studies utilizing endomannanases in biomass processing	**Reference**
Bleaching of softwood pulps (paper production)	(Buchert et al. 1993)
Clarification of coffee extracts	(Sachslehner et al. 2000)
Clarification of fruit juices	(Nadaroglu et al. 2015)
Production of bioethanol and feed by hydrolysis and fermentation of PKC *	(Jørgensen et al. 2010)
Production of mannooligosaccharides as a prebiotic	(Al-Ghazzewi et al. 2007)
Production of partially hydrolyzed guar gum, a dietary fiber	(Li et al. 2017)
Synthesis of alkyl glycosides, biodegradable surfactants	(Morrill et al. 2018)
Bioethanol production by softwood saccharification and fermentation	(Várnai et al. 2011), Paper II

* Palm kernel press cake (PKC)

Traditionally endomannanases are used in the pulp and paper industry. Since Galactoglucomannans are the major hemicellulose in softwood, endomannanases are applied for aiding in bleaching of softwood pulps (Buchert et al. 1993). In the food industry, endomannanases can be used to clarify coffee extracts (Sachslehner et al. 2000) and fruit juices (Nadaroglu et al. 2015) containing mannans, and alkaline endomannanases have found application in laundry products as stain removal boosters, because mannans are present in our diet (reviewed in Chauhan et al. 2012). In the feed industry, mannans are found in ingredients such as soybean meal and palm kernel meal. These meals have some common properties, e.g. high fiber content and high viscosity, which limits their utilization in the intestine. Adding endomannanases to the meal helps the breakdown of mannans and thus improves the release of the encapsulated nutrients (reviewed in Chauhan et al. 2012). A promising example is the use of PKC, a residue from palm oil extraction, for production of both bioethanol and feed. While efficient hydrolysis of the mannan rich biomass into fermentable monomer sugars is good for bioethanol production, the same enzymatic processing of PKC also results in a residual product with increased protein content, which is beneficial for the feeding value of the biomass (Jørgensen et al. 2010). Due to the prebiotic potential of some mannooligosaccharides, endomannanases could possibly also be used to produce these by hydrolysis of mannans (Al-Ghazzewi et al. 2007). Additionally, endomannanases with high transglycosylation capacity could possibly be used to enzymatically synthesize novel mannoconjugates such as hexyl mannosides with potential surfactant properties (Morrill et al. 2018).

Due to the abundance of softwood and galactoglucomannans, endomannanases have potential application in biorefinery processes of this polysaccharide (reviewed in Cherubini 2010; Gilbert et al. 2008). One example is the use of endomannanases in combination with cellulases (and xylanases) for softwood saccharification and second generation bioethanol production (reviewed in Yamabhai et al. 2016), Mannose is an attractive monosaccharide in bioethanol production because it can be fermented into ethanol using conventional yeasts. Enzymatic degradation of softwood to fermentable monomeric sugars is still challenging however, due to the complex composition and inhomogeneous architecture of softwood (Khatri et al. 2018). Enzymatic hydrolysis of cellulose is prevented not only by lignin but also by hemicellulose, mannans and xylans in the absence of relevant accessory enzymes (Berlin et al. 2005; Várnai et al. 2011). The hemicelluloses are closely associated with the cellulose fibrils and with lignin (Eriksson et al. 1980; Várnai et al. 2011; Ono et al. 2017; Khatri et al. 2018). Traditionally, severe pretreatment methods, which leave only small amounts of mannan in the pretreated substrate, are used on the recalcitrant softwood substrate to allow enzymatic saccharification. However, new pretreatment methods that maximize hemicellulose, including mannan, recovery have emerged with the aim of making bioethanol production from wood substrates more sustainable and efficient (Chandra et al. 2015; Chandra et al. 2016). Endomannanases when applied together with cellulases and xylanases have increased the total hydrolysis of pretreated softwood substrates (Várnai et al. 2013). However, the enzymatic capacity has been demonstrated and studied mostly with selected fungal GH5 endomannanases. The available literature in particular includes several studies with *Tres*Man5A (Rättö et al. 1993; Tenkanen et al. 1997; Várnai et al. 2011; Inoue et al. 2015). A few studies have also shown increased glucose release from wood substrates when cellulase (and xylanase) cocktails were supplemented with fungal GH26 endomannanases *Pans*Man26A and *Mt*Man26A (Couturier et al. 2011; Katsimpouras et al. 2016). In **Paper II**, softwood saccharification with both fungal GH5 and GH26 endomannanases is analyzed and compared.

As outlined above, the utilization of endomannanases is being studied in several industrial processes. and should preferably result in significant savings in resources for the benefit of both the industry and the environment. The results presented and discussed in the following chapters can lead to better selection and optimization of fungal endomannanases in softwood saccharification as well as in other established or new industrial applications.

4 Experimental design and selected methods

The scientific work conducted in this PhD thesis commenced with hypothesis H1 (**Paper I**):

H1 Fungal GH26 endomannanases may include enzymes that are less hindered by galactose substitutions on the mannan backbone than fungal GH5 endomannanases. If this hypothesis is valid, GH5 and GH26 fungal endomannanases will exhibit different degradation patterns and kinetics on galactomannans. Since the effect of the galactose substitutions must relate to how the substrate is fitted in the active site cleft of the enzymes, differences in enzyme kinetics and substrate degradation must be due to structural differences of the enzymes.

To investigate this hypothesis, a study was carried out with an uncharacterized fungal GH26 endomannanase from *Aspergillus nidulans*, *Anid*Man26A, and a few other well-known fungal endomannanases: *Pans*Man26A, *Tres*Man5A, *Anid*Man5A and *Anid*Man5C. The enzymes were carefully selected to include members from both the GH5 and the GH26 family and to present GH5 endomannanases, namely *Anid*Man5A and *Anid*Man5C, with known differences in their hydrolytic activity on galactomannans. *Anid*Man5B was not included in this study because of its low k_{cat}, and because we were not assessing transglycosylation. To enable the study to focus on enzyme specificity, and not stability, all analyses were performed at 37 °C. This was particularly important because the studied GH26 endomannanases were found to be less stable, based on a lower melting temperature (*Tm*), than the GH5 endomannanases. The pH was set to 5 throughout the work in this thesis because the tested endomannanases had full or close to full residual activity at this pH and because pH 5 is often used in industrial biomass degradation, including softwood saccharification. Initial rate measurements and degree of conversion on soluble mannan substrates were performed as end-point measurements after 15 minutes hydrolysis. The period of the initial phase was verified by having multiple enzyme doses that attained different degrees of conversions (dose-response studies). In addition to assessing the rate and the degree of conversion, hydrolysis product profiles from galactomannan hydrolysis was evaluated using the DNA sequencer-assisted saccharide analysis in high throughput (DASH) method. The architecture of the enzyme active sites was assessed using available crystal structures and new homology models.

Based on the findings in **Paper I**, two new hypotheses were formulated and work with these was conducted in parallel. The second hypothesis (H2, **Paper II**) was as follows:

H2 Fungal endomannanases of GH5 and GH26 differ in their capacity to catalyze removal of galactoglucomannans from cellulose microfibrils, and thus in turn may have different effects on enzymatic cellulose saccharification.

In examining H2 we decided to focus on the less characterized fungal GH26 endomannanases by analyzing eight such enzymes, including *Anid*Man26A and *Pans*Man26A (also studied in **Paper I**). Their performance in softwood saccharification was compared to that of two GH5 endomannanases, namely *Tres*Man5A and *Anig*Man5A. Use of these GH5 endomannanases in potential industrial applications had previously been studied and reported. During the work with H2 it became clear that 37 °C was too high a temperature for the prolonged reaction times used in softwood saccharification if we were to draw conclusions on the enzymes specificity and not their stability. Consequently, the temperature was lowered to 30 °C during the prolonged softwood saccharification reactions. This study was designed to evaluate which type of endomannanase specificity was most suitable for softwood saccharification. With this knowledge, it would probably be possible to produce a more thermostable variant at a later stage.

Studies related to H3:

H3 The ability to accommodate multiple galactopyranosyl moieties in the active site cleft is a common feature among fungal GH26 endomannanases and semi-conserved residues associated with this feature can be identified.

In York, crystallization experiments were set up with the only two naturally occurring non-CBM carrying fungal GH26 endomannanases that were available, namely *Anid*Man26A (**Paper I**) and *Ypen*Man26A from *Yunnania penicillata,* which were identified as the fastest of the analyzed endomannanases on soluble mannans in **Paper II**. Since both *Ypen*Man26A and *Anid*Man26A seemed to hydrolyze soluble mannans faster than *Pans*Man26A (3ZM8), carrying a CBM35, it was our intention to solve a structure for a non-CBM carrying fungal GH26 endomannanase. For unknown reasons, *Anid*Man26A did not crystalize. The solved crystal structure of *Ypen*Man26A was analyzed with a focus on substrate interactions and compared with other GH26 endomannanases. Two variants were designed for *Ypen*Man26A to assess the molecular details of some of the substrate binding amino acids that were not conserved among the investigated fungal GH26 endomannanases in relation to kinetic rates and affinity on highly galactosyl-substituted mannan substrates. A new mass spectrometry (MS) based method was developed to allow real-time measurements of MGGMM degradation. The advantage of this method over existing HPAEC methods was mainly the ability to automatically sample and analyze the samples continuously during the degradation. Measuring of reducing ends was never an

alternative for these experiments because of the high background from the reducing ends in MGGMM itself.

Three of the methods used are of particular novelty or importance for the thesis work, namely the DASH method, protein crystallization and structure solution as well as the developed MS method.

4.1 DASH

DNA sequencer-assisted saccharide analysis in high throughput (DASH) is a method for characterizing GHs hydrolysis product profiles or analyzing plant cell wall polysaccharides, using a 96 capillary array DNA sequencer (ABI 3730xl) (Li et al. 2013).

The ABI 3730xl is normally employed to analyze the sequence of bases in a DNA fragment using capillary electrophoresis with laser-induced fluorescence (CE-LIFE). In CE, analytes migrate through electrolyte solutions in the thin capillaries under the influence of an electric field. The analytes are separated by their m/z ratio, but also the "bulkiness" of analytes affects the migration because of changes in viscosity. CE-LIF is a powerful technique that offers high detection sensitivity and good resolving capacity for the analysis of DNA bases and also for the analysis of 9-aminopyrene -1, 4, 6-trisulfonate (APTS) labelled carbohydrates (Evangelista et al. 1995; Li et al. 2013). One advantage of using CE-LIF in the ABI 3730xl is that samples can be analyzed in parallel. In theory, 96 samples, one microtiter plate, can be analyzed in one go. Moreover, the ABI 3730xl can detect up to 5 specific fluorescent emission wavelengths, which potentially allows spectral separation of standards and sample oligosaccharides (Li et al. 2013). However, the mobilities of APTS-labelled saccharides show small but significant variations, both when electrophoresed simultaneously in separate capillaries but also from run to run. To overcome this problem, mobility markers are introduced and used to normalize (align) the electropherogram traces which are the output from the ABI 3730xl. The eight mobility markers used in the work of this PhD thesis (**Paper I** and **III**) are short peptides with a large range of mobilities allowing accurate alignment over a wide range of saccharides (DP 1 to 35, if the right settings are used). The mobility markers studied were labelled with another fluorophore than the sample oligosaccharides, here the DY-481XL, which fluoresces at a longer wavelength (650 nm) than APTS (512 nm). The fluorescence of mobility markers and oligosaccharides can easily be resolved spectrally by the DNA sequencer, which ensures that there is effectively no detection of either fluorophore in the detection channel of the other fluorophore. The mobility markers enable the standardization of mobility between different equipment and experiments and they make the identification of peaks from aligned DASH traces more accurate, which is essential for the identification of peaks migrating close together (Li et al. 2013).

Data analysis in DASH resembles that of known separation techniques, such as high-performance liquid chromatography (HPLC). In the case of DASH, the peak area from the electropherogram traces can be determined and used to estimate amounts of observed saccharides. The method provides a possibility for absolute quantification of carbohydrates by adding internal standards. However, this work is labor intensive, and requires the incorporation of internal standards in every sample. Furthermore, oligosaccharides with different reducing end sugar (e.g. xylose vs glucose), or different lengths, have different labeling efficiencies, which should be taken into account (Li et al. 2013). The method is most often used as a semi-quantitative method giving the correct ratio between identified sugars in a fingerprint (Mortimer et al. 2015; Grantham et al. 2017) (**Paper I** and **II**).

DASH is an alternative to polysaccharide analysis by carbohydrate gel electrophoresis (PACE). PACE is a method that also uses reductive amination of the reducing end of oligosaccharides with a fluorophore, followed by separation by polyacrylamide gel electrophoresis (Goubet et al. 2002). PACE is simple, requiring little specialist equipment, but is a relatively low through-put method requiring considerable user time. Liquid chromatography (LC) or MS, which are labor intensive and unsuitable for screening a large number of samples, can also be used for characterization of hydrolysis product profiles from GHs (Vinzant et al. 2001). The work carried out in this thesis (**Paper I** and **III**) represents the first use of DASH for analyzing galactomannooligosaccharides.

4.2 Protein crystallization and structure solution

X-ray crystallography is the most common method for determination of protein structures (**Paper III**). The wavelengths of X-rays are in the order of ångströms (Å), which make them ideal for studying protein structures because their bond lengths are in the same order, around 1.5 Å to be precise. X-ray crystallography requires protein crystals, which are protein molecules, arranged regularly in a three-dimensional lattice. The smallest repeating unit in a crystal, which may contain one or more protein molecules and is called the unit cell (reviewed in Williamson 2012). Depending on the symmetry of the crystal lattice, the unit cell may contain several units related to each other by symmetry operations. The full symmetry of the lattice – lattice translations and operations relating the units within the unit cell – constitute the space group. There are 230 space groups described in the international tables for crystallography, but only 65 are possible for chiral biological molecules because some of the symmetry operations, e.g. mirror planes and inversions, are excluded. The asymmetric unit of a space group is the part of the unit cell which can be used to generate the entire unit cell by the symmetry of the space group. This is often a single protein molecule or even a protein domain in cases with e.g. homodimers (reviewed in Rupp 2009).

4.2.1 Protein crystallization

The first step in protein crystallization is to force the protein gently out of solution. The solubility of a protein can be reduced by adding precipitants to the solution and by removing solvent. Reducing the solubility moves the protein into the metastable region of the phase diagram, which is a necessary condition for phase separation and crystal formation. Once the solubility limit of the protein is exceeded, the solution becomes supersaturated and, given a nucleation event, the excess protein molecules separate from solution into a protein rich phase in equilibrium with the saturated protein solution. In some cases, the protein self-assembles into crystals, but solid precipitates or liquid phases (protein oils) are also common protein-rich phases. Low supersaturation favors controlled crystal growth, while high supersaturation favors spontaneous nucleation resulting in less ordered protein structures (reviewed in Asherie 2004; Rupp 2009).

It is usually not possible to predict conditions favoring protein crystallization, which is why several hundred crystallization trials are set up. One of the widely used procedures for achieving supersaturation is the vapor diffusion technique which is performed in sitting-drop or hanging-drop format. In vapor diffusion setups, a small droplet of concentrated protein is usually mixed 1:1 by volume with the so called "mother liquor" (containing precipitants, buffers and possible additives) and either hung upside down on a glass slide or placed on a raised platform above the well to be equilibrated against a much higher volume of the mother liquor in the reservoir. The system is then sealed and vapor diffusion occurs to equilibrate the water content in the droplet with the reservoir solution. During vapor diffusion, precipitant and protein concentration increase in the crystallization drop and supersaturation is achieved (reviewed in Rupp 2009). Crystal growth can be controlled by separating the phases of nucleation and growth. This can be achieved by inserting smaller crystals, or crystal seeds, directly into the crystallization drop (Chayen 2005). This method, called seeding, induces heterogenous nucleation at low supersaturation, which is more favorable for controlled crystal growth. Furthermore, it allows nucleation to happen under one set of conditions and crystal growth under another set of conditions, which can giver better crystals in less time (Shaw Stewart et al. 2011).

4.2.2 Structure solution and refinement

Today, X-ray data from crystal structures are predominantly obtained at a synchrotron. A synchrotron is a large storage ring (up to 1 km in diameter) with high-energy particles under high vacuum. The particles generate X-rays when they are accelerated (reviewed in Williamson 2012).

When X-rays pass through matter, some of them will interact with electrons in the matter and be scattered by them. This is called X-ray diffraction. The scattering is in all directions but in most directions, the scattered X-rays from different atoms will interfere and cancel each other so no

signal result. When the X-ray beams scattered from adjacent crystal planes are in phase, constructive interference will occur, which creates a diffraction spot. This constructive interference occurs in precise directions and gives rise to a diffraction pattern that contains distinct spots arranged in a regular pattern. The diffraction pattern changes as the incoming angle of the X-ray beam changes. To observe all possible spots, it is necessary to collect diffraction patterns from all possible orientations of the crystal. However, crystal symmetry often means that only a specific range of angles need to be covered. The output after data collection is a set of images with diffraction patterns.

Each spot in a diffraction pattern has an amplitude and a phase. To obtain the distribution of electrons in the asymmetric part of the unit cell, it is necessary to calculate the Fourier transformation of the so-called structure factors, or F values, which represent the reflection amplitudes and phases. However, only the amplitudes can be measured from the reflection intensities and there is no direct way to measure the phases. Several methods are used in protein crystallography to determine the phases. Molecular replacement is one of them. Molecular replacement is often faster and easier than experimental methods, but requires a solved structure of a similar protein (sequence identity of 30 % or higher). The structure of the chosen homolog is rotated and translated in the unit cell or asymmetric unit until the solution with the best fit between calculated diffraction data from the homolog and the observed data from the unknown structure is obtained. With the increasing number of structure models available in the protein data bank, molecular replacement is a method of increasing popularity (reviewed in Rupp 2009).

The primary result of an X-ray diffraction experiment is an electron density map. The protein structure can be derived by fitting atoms into the electron density. Fitting is done iteratively, with checks being made against the electron density at each stage. The accuracy with which this can be done depends on the resolution of the data. In a structure with a resolution of 2.5 Å or worse, it is difficult to resolve individual atoms. At 1.2 Å it can be possible to see protons. The CCP4 (Collaborative Computational Project, number 4) software suite is a collection of programs which can be used for macromolecular structure determination by X-ray crystallography (Winn et al. 2011).

The aim of protein crystallography is to produce a model of the protein of interest which explains the electron density map as accurately and completely as possible. Refinement is the process of optimizing the model parameters. In CCP4 v. 7.0 refinement is performed by the program REFMAC5 (Murshudov et al. 2011; Winn et al. 2011). Refinement goes together with rounds of model building which add or subtract parts of the model and apply large structural changes that are beyond the reach of refinement. Model building can be done with programs like Buccaneer and RAPPER or more manually using e.g. Coot (Emsley et al. 2010; Winn et al. 2011). Validation is ensuring that all aspects of the model are supported by the diffraction data, as well as following

known features of protein chemistry. Validation is an integral part of the process of structure solution, and should be carried out continuously (reviewed in Winn et al. 2011).

There is a one-to-one correspondence between a protein structure and its diffraction pattern. If the structure is known, the diffraction pattern can be calculated from it. This provides a check on the correctness of a solved structure, because the back-calculated diffraction pattern can be compared with the experimentally observed diffraction pattern. The fractional difference between the two is known as the R factor. For a good structure, the R factor is generally lower than 20 %, depending on the resolution of the data. The R factor is expected to be around 1/10 of the resolution. However, the R factor alone does not provide a reliable quality check. Crystal structures are always refined, which means that atomic positions are moved to produce the best fit with the electron density map. Therefore it is not possible from the R factor alone to tell the difference between a genuinely good structure and one produced by over-refining a bad structure. The solution is to randomly select a small proportion of the data, which is not used in the refinement. Calculation of a second R factor is then done using these data, and the number is called Rfree. If the refinement is a real improvement, then Rfree should also improve, and be similar to the R factor. A good structure will typically have an R factor of less than 20 % and an Rfree within 5 % of R (reviewed in Wlodawer et al. 2008; Rupp 2009; Williamson 2012).

4.3 Endomannanase kinetics in real-time by MS

Mass spectrometry (MS) makes it possible to directly and quantitatively monitor product formation or substrate depletion during a chemical reaction and is thus a highly relevant method for assessment of enzyme kinetics (Perna et al. 2018).

In MS, chemical species are ionized and then sorted and detected based on their mass to charge ratio. Electrospray ionization is an often used ionization technique for the analyses of proteins and other biopolymers, but also for smaller polar molecules. The ions are produced by applying high voltage to a liquid nebulizer stream to creating aerosols with the ionized molecules (reviewed in de Hoffmann & Stroobant 2007). Many quantitative MS assays use the selected reaction monitoring methodology. Here, electrospray ionization is followed by two stages of mass selection (tandem MS): a first stage (MS1) selects the mass of the intact analyte (parent ion), and after fragmentation of the parent ion by collisions with gas atoms, a second stage (MS2) selects a specific fragment of the parent. Multiple reaction monitoring (MRM) is based on identification and quantification of multiple specific fragment ions from the predetermined precursor ion (or ions). The advantage of this targeted approach is that the two mass filters produce a very specific and sensitive response. Moreover, when selecting and following a specific analyte, the ion count can be measured and used for quantification by peak integration

in a simple one-dimension chromatographic separation of the sample (reviewed in Anderson & Hunter 2006).

One advantage of using MS is the direct monitoring of the reaction in real time. MS is receiving increasing attention as a quantitative tool, and within the field of proteomics the quantitative use of MRM is particularly widespread (reviewed in Cox & Mann 2011). The objective for the work in this thesis (**Paper III**) was to develop a real time, sensitive and accurate methodology to assess endomannanase activity on MGGMM and specifically to assess the catalytic efficiency of two variants. In this scenario MRM and quantification was feasible because the reactions were well-defined and had a known starting component. Kinetics were determined as an on-line measurement of substrate depletion.

5 Results and discussion

The overall aim of this thesis is to contribute to understanding of fungal endomannanase diversity in relation to specificity, with focus on differences between fungal GH5 and GH26 enzymes. A further aim is to evaluate performance of fungal endomannanases in softwood saccharification and to assess if performance differences correlated with specific enzyme characteristics. In this chapter, the results obtained are outlined and discussed in relation to hypothesis H1, H2 and H3 of the PhD thesis.

5.1 The influence of galactose substitutions in enzymatic mannan hydrolysis (Paper I)

Galactose substitutions on mannan backbones have long been known to negatively affect enzymatic hydrolysis of mannans. The significance of galactose substitutions on mannan substrates has been studied for fungal GH5 endomannanases (McCleary 1979) but not in detail for fungal GH26 endomannanases. The thesis work began by addressing the significance of galactose substitutions on guar gum and locust bean gum mannans on the kinetics and substrate degradation patterns of GH26 endomannanases compared to known GH5 endomannanases. Limitations caused by galactose substitutions can potentially limit the utilization of endomannanases in applications using galactomannans or galactoglucomannans, including the abundant acetylated galactoglucomannan. The scientific work in this section addresses the following hypothesis and corresponding objective, but also includes more general characterization of the enzyme stability and pH profiles:

H1 Fungal GH26 endomannanases may include enzymes that are less hindered by galactose substitutions on the mannan backbone than fungal GH5 endomannanases. If this hypothesis is valid, GH5 and GH26 fungal endomannanases will exhibit different degradation patterns and kinetics on galactomannans. Since the effect of the galactose substitutions must relate to how the substrate is fitted in the active site cleft of the enzymes, differences in enzyme kinetics and substrate degradation must be due to structural differences of the enzymes.

Obj. 1. To analyze the enzyme kinetics and degradation patterns on galactomannans and the active site cleft architecture of fungal GH5 and GH26 endomannanases (using the sequence collection at Novozymes as a starting point for selecting the enzymes to be expressed and examined).

The hydrolysis of galactomannans was analyzed using five fungal endomannanases (Table 3) of which two were from family GH26, namely the previously uncharacterized *Anid*Man26A from *Aspergillus nidulans* and *Pans*Man26A, as well as three well known GH5 endomannanases. *Pans*Man26A and *Tres*Man5A were characterized both in their full length, i.e. with the CBM35 and CBM1 attached, respectively, and as truncated variants without the CBM. The fungal endomannanases were recombinantly expressed in *A. oryzae* (a few by the PhD student, the rest by Novozymes) and purified to electrophoretic purity.

Table 3: Fungal GH5 and GH26 endomannanases assessed in galactomannan degradation

Origin	Enzyme name	GH family	CBMs	Mw [a] (kDa)	pHopt [b]	Relative activity ($\frac{pH_5}{pH_{opt}}$)	Tm [c] (°C)	Sequence ID
A. nidulans	*Anid*Man26A	26	No	35.2	6 (5–7)	0.93	53	Q5AWB7
P. anserina	*Pans*Man26A	26	CBM35	49.8	6 (5-7)	0.97	57	B2AEP0
P. anserina	*Pans*Man26A core	26	No	34.4	5 (5-7)	1	58	(B2AEP0)
T. reesei	*Tres*Man5A	5	CBM1	45.2	4 (3-7)	0.93	81	Q99036
T. reesei	*Tres*Man5A core	5	No	38.8	4 (3-6)	0.88	78	(Q99036)
A. nidulans	*Anid*Man5A	5	No	40.7	4 (4-7)	0.99	70	Q5B7X2
A. nidulans	*Anid*Man5C	5	No	43.5	4 (3-6)	0.86	70	Q5AZ53

[a] Theoretical (non-glycosylated protein). [b] pH optimum at 37 °C and pH interval with 80 % relative activity in brackets.
[c] The thermal midpoint (*Tm*) at pH 5.0.

The melting temperature (the thermal midpoint, *Tm*) of the analyzed GH26 endomannanases was found to be lower (between 53 – 58 °C) than the *Tm* of the analyzed GH5 endomannanases (between 70 – 81 °C), which indicated a possible difference in thermal stability between the two families (Table 3). *Tm* was determined as the top of the denaturation peak obtained by differential scanning calorimetry (DSC) with a constant heating rate of approximately three degrees every minute. In total, the *Tm* of 44 fungal GH26 endomannanases and 4 fungal GH5 endomannanases were assessed during this PhD study. All assessed GH26 endomannanases had a *Tm* between 50 and 68 °C while the *Tm* of the GH5 endomannanases ranged from 70-81 °C (more data are given in Table 4, but not all data are shown). The thermostability did not seem to be influenced by the presence of a CBM irrespective of whether the CBM belonged to CBM1 (*Tres*Man5A) or CBM35 (*Pans*Man26).

5.1.1 Galactomannan hydrolysis by fungal endomannanases

To assess catalytic rates of the enzymes and their ability to attack different parts of the heterogenous galactomannan substrates, the initial rate of hydrolysis and the degree of conversion were estimated for the selected endomannanases on locust bean gum and the more densely substituted guar gum at 37 °C and pH 5 with 2.5 mg/ml substrate (Figure 6). These studies were performed as dose-response studies where end-point endomannanase activity was measured by reducing end assay after 15 min of hydrolysis. The initial rate was calculated from the first linear dose-response relation at low doses, and the maximal degree of conversion was defined as having been reached when the conversion degree did not increase with increasing enzyme dose (i.e. a plateau was reached).

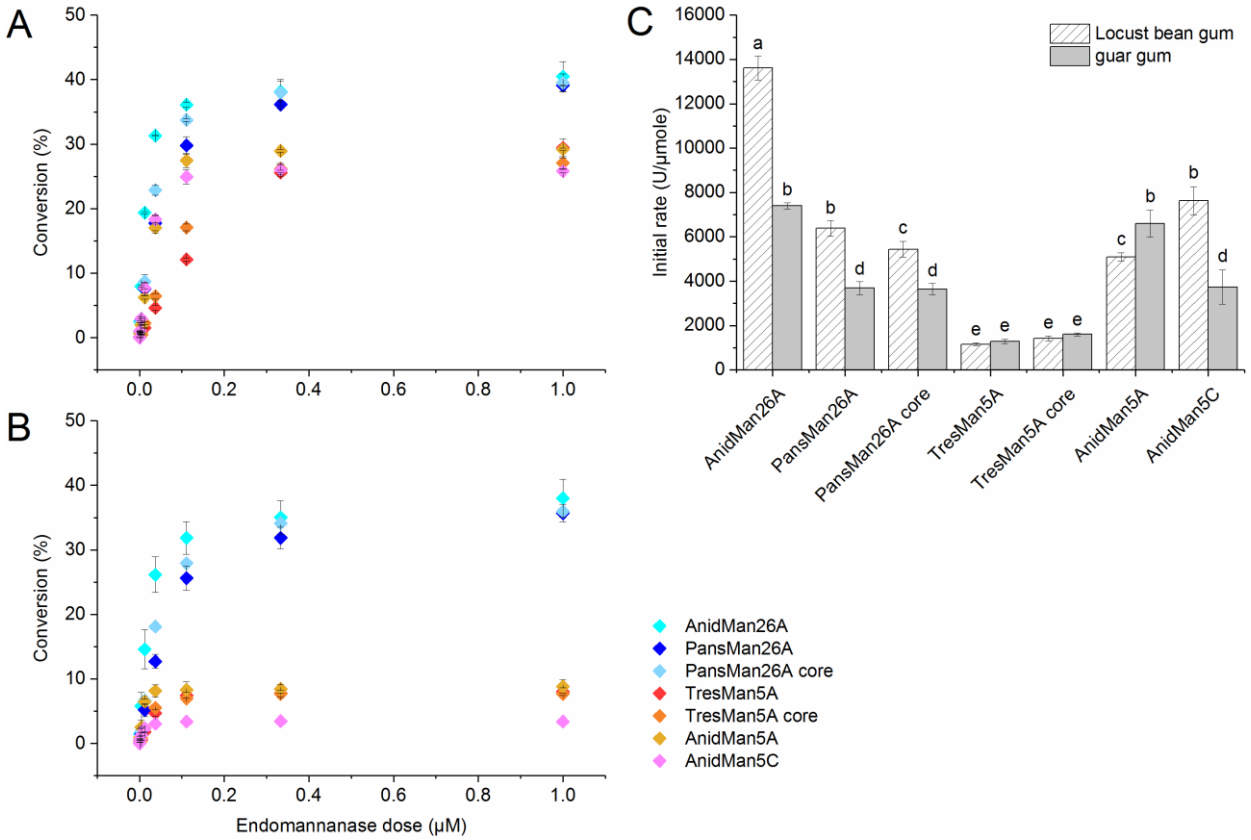


Figure 6: Conversion (%) of locust bean gum (**A**) and guar gum (**B**) and initial hydrolysis rates (U/μmole) on both substrates (**C**) by fungal endomannanases (37 °C and pH 5). Conversion was defined as released reducing ends relative to the theoretical monomeric yield. A one-way ANOVA analysis indicates that the maximal degree of conversion measured for 1 μM endomannanase *Anid*Man26A and *Pans*Man26A was significantly different from the maximal degree of conversion by *Tres*Man5A, *Anid*Man5A and *Anid*Man5C ($p < 0.01$) for both substrates. In C, each of the letters (a-e) represents a group of initial rates which are significantly different from initial rates belonging to other groups (ratings are assigned based on a 95 % confidence interval for means based on a pooled SD of 411). All values are given as mean values ± SD ($n = 2$).

The tested fungal GH26 endomannanases reached a high degree of galactomannan conversion compared to the GH5 endomannanases (Figure 6). The difference was particularly pronounced on the highly substituted guar gum on which *Anid*Man26A and *Pans*Man26A (with and without CBM35) reached a 3-4 times higher conversion (35-40 %) than *Tres*Man5A (with and without CBM1) and *Anid*Man5A (8-10 %), whereas *Anid*Man5C reached only 3 % conversion. To the best of our knowledge, 35 – 40 % conversion of guar gum has not been obtained with endomannanase catalyzed hydrolysis before the work of this thesis, whereas 10 % conversion is in agreement with earlier studies (McCleary 1979). Because the degree of conversion was defined as released reducing ends relative to the theoretical monomeric yield of the substrates, 30-40 % conversion can be considered full conversion for an endo-enzyme which does not hydrolyze the polymers to monomers. For the less substituted locust bean gum, the difference was less pronounced but still significant, because the tested GH26 endomannanases reached 38-40 % conversion and the GH5 endomannanases reached 26-29 % conversion (Figure 6). The same discrimination between GH5 and GH26 endomannanases was not observed when assessing their initial rate of hydrolysis (Figure 6). *Anid*Man26A had the highest observed initial rate on both locust bean gum and guar gum of all tested endomannanases (13600 and 7400 U/μmole, respectively) whereas *Tres*Man5A was found to have the lowest initial rates (1000-2000 U/μmole) and also the lowest discrimination between the two substrates. *Pans*Man26A, *Anid*Man5A and *Anid*Man5C had initial rates at the same level (4000-8000 U/μmole). *Anid*Man5C were found to be more restricted by the extra galactose substitutions in guar gum compared to *Anid*Man5A, as have also been described by Dilkopimol et al. (2011). The presence of the CBM (either CBM1 or CBM35) did not seem to affect the maximal degree of conversion on locust bean gum or on guar gum. However, on locust bean gum a significantly higher initial rate was observed for *Pans*Man26A carrying a CBM35 than for the truncated *Pans*Man26A without a CBM35. Since no significant difference was observed between these two enzymes on guar gum (Figure 6), it is possible that *Pans*CBM35 interacts with the β-mannan backbone, which is more accessible in locust bean gum, and not with the galactopyranosyl units.

To evaluate which parts of the substrates were catalytically degraded by the enzymes, the hydrolysis product profiles at enzyme individual maximal degree of conversion (Figure 6) was analyzed by DASH (Figure 7). The nomenclature used for oligosaccharides in this thesis is similar to that suggested by Fry et al (Fry et al. 1993). Oligosaccharides are listed from the nonreducing end to the reducing end and backbone moieties bearing a substitution are only represented by the substitution. Using this nomenclature, α-6^4-6^3-di-galactosyl-mannopentaose is named MGGMM. Due to the nature of the DASH analysis, the size of peaks can only be compared relatively within each sample but not across samples.

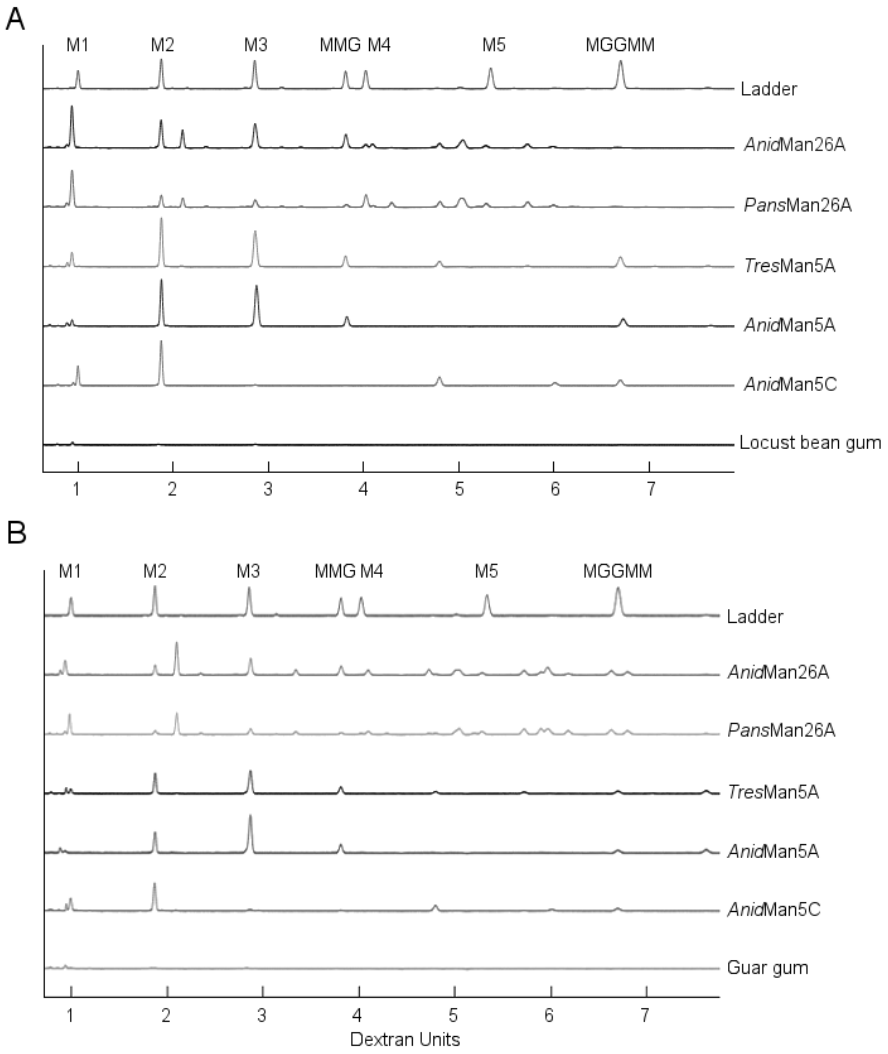


Figure 7: Product profiles from locust bean gum (**A**) and guar gum (**B**) hydrolysis by fungal endomannanases. Aligned electropherograms of product profiles at the maximal degree of conversion (cf. Figure 6). Migration of oligosaccharides is given in dextran units (DE). A ladder was run containing mannose (M1, 0.9 DE), mannobiose (M2, 1.87 DE), mannotriose (M3, 2.85 DE), α-6^1-galactosyl-mannotriose (MMG, 3.81 DE), mannotetraose (M4, 4.02 DE), mannopentaose (M5, 5.33 DE), and α-6^4-6^3-di-galactosyl-mannopentaose (MGGMM, 6.7 DE).

The DASH analysis showed that analyzed endomannanases had different hydrolysis product profiles (Figure 7). The two GH26 endomannanases, *Anid*Man26A and *Pans*Man26A, were found to produce multiple galactomanno-oligosaccharides based on the assumption that oligosaccharides migrating in between known mannooligosaccharides observed in the ladder (Figure 7) must be galactomanno-oligosaccharides. In the product profiles from guar gum hydrolysis in particular, many unknown galactomanno-oligosaccharides were observed in the profiles from the GH26 enzymes. The main product migrated to 2.1 dextran units (DE). This product was not observed among any of the tested GH5 endomannanases. The three GH5 endomannanase, *Tres*Man5A, *Anid*Man5A and *Anid*Man5C, produced mostly mannooligosaccharides without substitutions, and in agreement with published literature

(Dilokpimol et al. 2011; Rosengren et al. 2012), no unsubstituted mannooligosaccharides larger than mannotriose was observed. This result indicated that these three GH5 enzymes cleave mannotetraose or longer mannooligosaccharides. Surprisingly, *Anid*Man5C was found to produce mainly mannobiose, with low amounts of few other products. Such profile is unexpected for an endo-enzyme.

The assessment of galactomannan degradation demonstrated that the two tested GH26 endomannanases attacked areas of the galactomannans which were heavily substituted and inaccessible to the tested GH5 endomannanases. Phrased differently, *Anid*Man26A and *Pans*Man26A seemed to accommodate more galactopyranosyl residues in their active site cleft, based on the various galactose-substituted hydrolysis products (Figure 7). This ability again explains the high degree of conversion obtained with guar gum (Figure 6).

5.1.2 Accommodation of galactopyranosyl moieties in the active site cleft

To investigate which subsites in the active site cleft of *Anid*Man26A and *Pans*Man26A accommodate galactopyranosyl substitutions, the unknown dominant product from guar gum hydrolysis, which migrates to 2.1 DE, was identified. This was done by degrading α-6^1-galactosyl-mannotriose (MMG) with different doses of an *A. nidulans* GH2 clade A mannosidase (BM2, A2QWU9) followed by analysis of the released products using DASH (Figure 8).

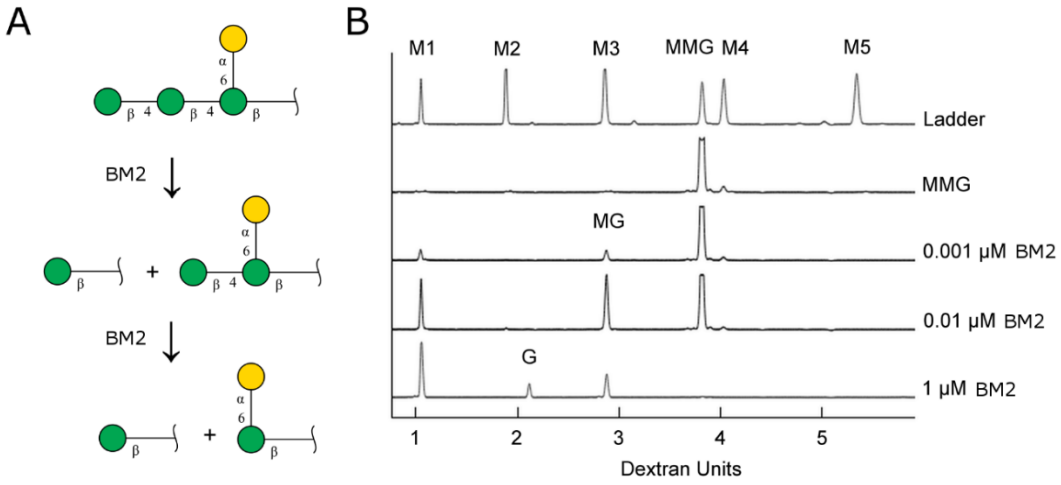


Figure 8: Degradation scheme (**A**) and aligned electropherograms (**B**) of the hydrolysis of α-6^1-galactosyl-mannotriose (MMG) by an *A. niger* GH2 clade A mannosidase (BM2). Sugars are shown using consortium for functional glycomics notation (Raman et al. 2006). Migration of oligosaccharides is given in dextran units (DE). A ladder was run containing: mannose (M1, 0.9 DE), mannobiose (M2, 1.87 DE), mannotriose (M3, 2.85 DE), MMG (3.81 DE), mannotetraose (M4, 4.02 DE), and mannopentaose (M5, 5.33 DE). Other abbreviations: α-6^1-galactosyl-mannobiose (MG) and α-galactosyl-mannose (G).

The degradation of MMG by *Anig*BM2 resulted in release of mannose from the non-reducing end of the oligosaccharide, and the products formed were those expected from a clade A mannosidase (Ademark et al. 1999; Reddy et al. 2013) (Figure 8). At low enzyme concentration α-6^1-galactosyl-mannobiose (MG) together with mannose were observed as a product, which migrated to 2.85 DE, the same migration as mannotetraose M3. At high BM2 concentration, some α-galactosyl-mannose (G) appeared in the product profile, and migrated to 2.1 DE. Because of identical migration, this oligosaccharide was assumed to be the same as the unknown dominant product from guar gum hydrolysis with *Anid*Man26A and *Pans*Man26A. Before the work described in this PhD study, the shortest characterized galactomannooligosaccharide produced by endomannanase-catalyzed hydrolysis of galactomannans or galactoglucomannans were MG (McCleary & Matheson 1983; McCleary et al. 1983; Tenkanen et al. 1997). To produce G from guar gum or locust bean gum, *Anid*Man26A and *Pans*Man26A must be able to accept galactopyranosyl side-groups both in the -1 and +1 subsites.

To further asses the accommodation of galactopyranosyl residues in the active site cleft of *Anid*Man26A, M5 and MGGMM were degraded and the product profiles analyzed by DASH (Figure 9).

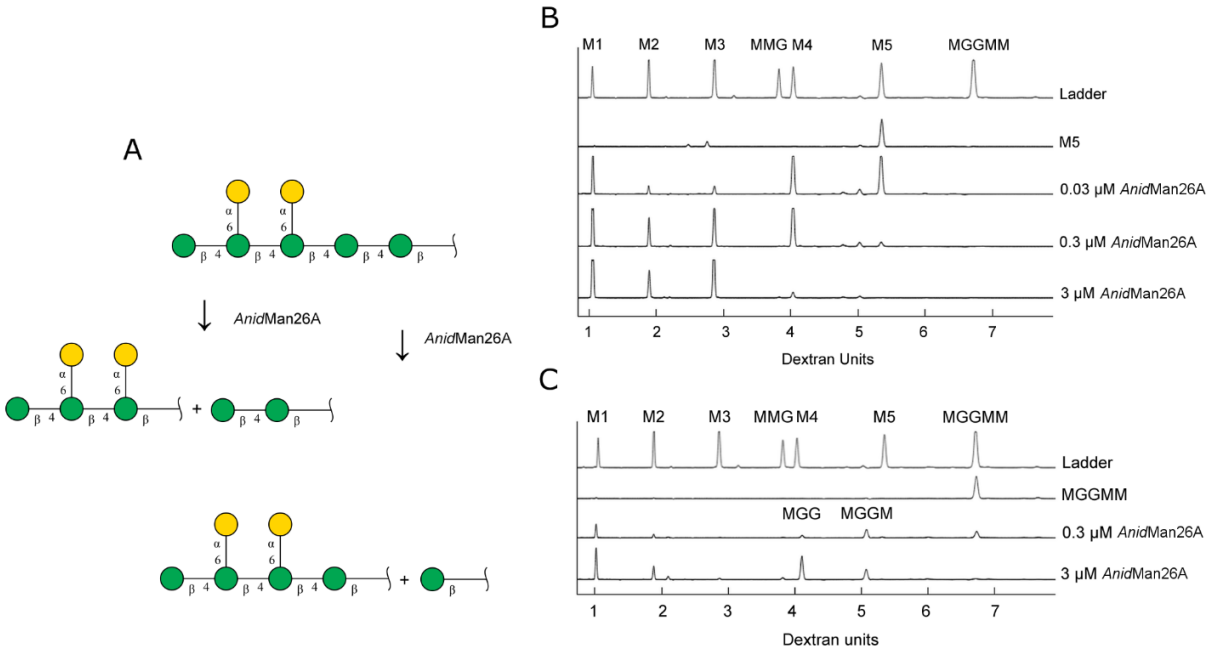


Figure 9: Proposed binding modes and subsequent hydrolysis of α-6^4-6^3-di-galactosyl-mannopentaose (MGGMM) by *Anid*Man26A (**A**) and aligned electropherograms of the hydrolysis of mannopentaose (**B**) and MGGMM (**C**) by *Anid*Man26A. Sugars are shown using the consortium for functional glycomics notation (Raman et al. 2006). Migration of oligosaccharides is given in dextran units (DE). A ladder was run containing: mannose (M1, 0.9 DE), mannobiose (M2, 1.87 DE), mannotriose (M3, 2.85 DE), α-6^1-galactosyl-mannotriose (MMG, 3.81 DE), mannotetraose (M4, 4.02 DE), mannopentaose (M5, 5.33 DE), and MGGMM, (6.7 DE). Other abbreviations: α-6^3-6^2-di-galactosyl-mannotetraose (MGGM) and α-6^2-6^1-di-galactosyl-mannotriose (MGG).

The degradation of M5 at a low concentration (0.03 μM) of *Anid*Man26A resulted in the production predominantly of mannose and mannotetraose and to a smaller extent mannobiose and mannotriose. Based on these results, *Anid*Man26A seems to have the same dominant binding mode as *Pans*Man26A (47 % sequence identity) (Couturier et al. 2013) which has a similar product profile on Mannopentaose. *Pans*Man26A has been shown to bind mainly from the -4 to the +1 subsites and to a smaller extent from the -3 to the +2 subsites. When compared together, the active site cleft of *Anid*Man26A (homology model) and *Pans*Man26A (3ZM8) are very similar, and it seems likely that *Anid*Man26A will also bind its substrate from the -4 to the + 1 subsites (**Paper I**). Degradation of MGGMM by *Anid*Man26A resulted in the production of mannose and to a smaller extent mannobiose together with two unknown galactomannooligosaccharides which migrated to 4.10 and 5.06 DE, respectively. A degradation scheme was proposed using the assumed binding from the -4 to the +1 subsites (Figure 9A). It is strongly believed that *Anid*Man26A binds MGGMM from the -4 to the +1 subsites, and cleaves mannose from the reducing end to produce α-6^3-6^2-di-galactosyl-mannotetraose (MGGM, 5.06 DE). It also appears that the enzyme to a lesser extent may bind MGGMM from the -3 to the +2 subsites to produce mannobiose and α-6^2-6^1-di-galactosyl-mannotriose (MGG, 4.10 DE). However, *Anid*Man26A was also shown to degrade MGGM to MGG and mannose (Figure 9C), presumably by binding from the -3 to the +1 subsites. In the proposed degradation scheme (Figure 9A), galactopyranosyl units must be accommodated in the – 3, -2 and -1 subsites of *Anid*Man26A.

The experimental results are supported by a comparison of the architecture of the active site cleft of the tested endomannanases using available crystal structures and homology models with docked galactomannooligosaccharides (Figure 10). Compared to the analyzed fungal GH5 endomannanases, *Anid*Man26A and *Pans*Man26A had a more open active site cleft (Figure 10). The docking analyses suggest that galactopyranosyl residues could be accommodated in the – 2, the -1 and the +1 subsites (Figure 10D and E). The active site structure even suggests that the galactose side groups can be accommodated in subsites beyond the -2, -1 and +1 subsites, however, this was not analyzed by docking experiments. For *Tres*Man5A and *Anid*Man5A the structural analysis strongly indicated that they could accommodate a galactopyranosyl unit in the – 1 subsite, as reported by Tenkanen et al for *Tres*Man5A (Tenkanen et al. 1997), but not in the -2 or +1 subsites (Figure 10A and B). The model of *Anid*Man5C exhibited a narrower active site cleft with no room for galactopyranosyl moieties (Figure 10C). This result might explain the low maximal degree of conversion obtained on guar gum (3 %) (Figure 6) and the pronounced difference observed in the initial rates on locust bean gum versus guar gum (Figure 6). Moreover, the glycone region of *Anid*Man5C appeared to be closed by a loop preventing binding beyond the -2 subsite. This structural feature fits with the product profiles determined in this PhD study where *Anid*Man5C produced mainly mannobiose from galactomannans (Figure 7).

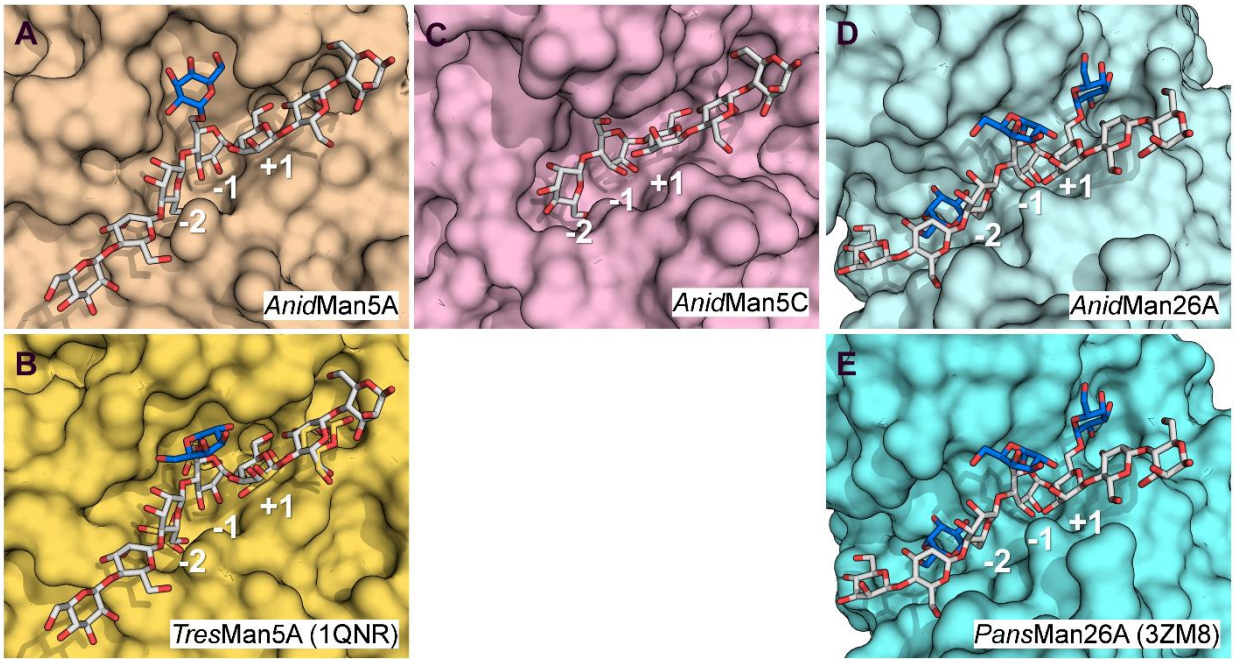


Figure 10: Surface views of the active site cleft of fungal endomannanases. In the binding cleft, mannoheptaose (white) is depicted with galactose side groups (blue) in the -2, -1 and +1 subsites, if accommodated by the enzyme. The mannopyranosyl unit in the -1 subsite was arranged in the same skewed boat conformation, as observed in the ligand in *Cjap*Man26C from *C. japonicus* (2VX6) (Cartmell et al. 2008), which is believed to be a prerequisite for catalysis.

5.1.3 Summary

In addressing the hypothesis, H1, relevant to this part of the PhD study, it was found that *Anid*Man26A and *Pans*Man26A had a novel degradation pattern on highly substituted galactomannan, thus accepting H1. These two fungal GH26 endomannanases reached around 40 % conversion of guar gum in contrast to the tested GH5 endomannanases which reached only 3-10 % conversion on the same substrate. *Anid*Man26A and *Pans*Man26A were also found to produce a variety of galactomannooligosaccharides from galactomannans, which indicated that they can attack highly substituted parts of the substrate. These results were supported by structural analyses that showed that *Anid*Man26A and *Pans*Man26A had a much more open active site than *Tres*Man5A, *Anid*Man5A and *Anid*Man5C. Experiments to degrade small, well-defined galactomannooligosaccharides and docking analyses showed that *Anid*Man26A and *Pans*Man26A accommodated galactopyranosyl units at least in the -2, -1 and +1 subsites, but most likely also beyond these subsites. The DASH method was also found to be very useful for the assessment of endomannanase product profiles on galactomannans because of high resolution and separation of similar (galacto)manno-oligosaccharides with identical Mw.

5.2 Endomannanase performance in softwood saccharification (Paper II)

The finding that *Anid*Man26A and *Pans*Man26A had a novel degradation pattern on galactomannans led to two new scientific hypotheses. One of these hypotheses concerned the interactions between the fungal GH26 endomannanases and the substrate and whether the open active site structure of *Anid*Man26A and *Pans*Man26 is unique or a more general trend for fungal GH26 endomannanases (Section 5.3). However, this paragraph will focus on the other hypothesis that refers to the application of endomannanases in softwood saccharification and whether the ability to accommodate multiple galactopyranosyl moieties, or another characteristic, is particularly valuable for endomannanase performance in this application. The following hypothesis will be addressed:

H2 Fungal endomannanases of GH5 and GH26 differ in their capacity to catalyze removal of galactoglucomannans from cellulose microfibrils, and thus in turn may have different effects on enzymatic cellulose saccharification.

Obj. 2. To measure the performance of selected fungal GH5 and GH26 endomannanases in softwood saccharification and evaluate which enzyme characteristics seem to be important for the measured performance.

Because of the novelty of the GH26 functionality, it was decided to expand the number of fungal GH26 endomannanases in the PhD study and to compare their performance with two GH5 endomannanases which had previously been proved useful in application studies, namely *Tres*Man5A (Rättö et al. 1993; Tenkanen et al. 1997; Várnai et al. 2011) and *Anig*Man5A (Jørgensen et al. 2010). Eight wild type enzymes, including six which were previously uncharacterized, were selected for investigation and subsequent recombinant expression assessment (Table 4) based on a phylogenetic sequence comparison of more than 50 fungal GH26 endomannanases. The new fungal endomannanases were recombinantly expressed in *A. oryzae* (by Novozymes) and purified to electrophoretic purity.

Table 4: Fungal GH5 and GH26 endomannanases assessed in softwood saccharification

Origin	Domains	Mw [a] (kDa)	pH_{opt} [b]	Relative activity pH_5/pH_{opt}	*Tm* [c] (°C)	*t* ½ [d] (h)	Sequence ID	Identity [e] (%)
GH26								
Collariella virescens (*Cvir*Man26A)	CBM35-GH26- CBM1	57.9	6 (5-7)	0.97	62	Stable	BBW45415	76
Mycothermus thermophiles (*Mthe*Man26A)	CBM35- GH26	52.1	5 (5-8)	1.00	68	91 ± 0.3	MH208368	76
Podospora anserina (*Pans*Man26A)	CBM35- GH26	49.8	6 (5-7)	0.97	57	90 ± 5.5	B2AEP0	100
Podospora anserina (*Pans*Man26A core)	GH26	34.4	5 (5-7)	1.00	58	103 ± 2.2	(B2AEP0)	100
Neoascochyta desmazieri (*Ndes*Man26A)	CBM35- GH26	48.7	5 (4-7)	1.00	65	267 ± 24.1	MH208367	60
Westerdykella sp. (*Wsp.*Man26A)	CBM35- GH26	50.4	6 (6-7)	0.79	58	59 ± 4.3	MH208369	55
Ascobolus stictoideus (*Asti*Man26A)	CBM35-GH26-CBM1	59.4	7 (5-7)	0.80	61	81 ± 7.7	BBW45412	55
Aspergillus nidulans (*Anid*Man26A)	GH26	35.2	6 (5-7)	0.93	53	10 ± 0.1	Q5AWB7	52
Yunnania penicillata (*Ypen*Man26A)	GH26	34.5	6 (5-8)	0.87	50	21 ± 0.1	MH899111	48
GH5								
Trichoderma reesei (*Tres*Man5A)	GH5-CBM1	45.2	4 (4-5)	0.93	81	Stable	Q99036	36
Trichoderma reesei (*Tres*Man5A core)	GH5	38.8	4 (3-6)	0.88	78	Stable	(Q99036)	36
Aspergillus niger (*Anig*Man5A)	GH5	39.8	4 (3-5)	0.85	87	137 ± 10.0	BCK48306	30

[a] Theoretical (non-glycosylated protein). [b] pH optimum at 37 °C and pH interval with 80 % relative activity in brackets. [c] The thermal midpoint (*Tm*) at pH 5. [d] Half-lives (*t* ½) at 30 °C. No decay was observed for *Cvir*Man26A and *Tres*Man5A during the 48 h incubation period. [e] Homology to *Pans*Man26A for which the structure is known (3ZM8, Couturier et al. 2013).

The GH26 endomannanases had pH optima in the range of pH 5-7 and *Tm* between 50 to 68 °C. The only two wild type core GH26 enzymes, *Anid*Man26A and *Ypen*Man26A (the latter from *Yunnania penicillata*), had the lowest *Tm* at 50 °C and 53 °C, respectively. The 30 °C stability data do not show the same difference in robustness between the GH5 and the GH26 endomannanases as indicated by the *Tm* values, but the two wild type core GH26 endomannanases, *Anid*Man26A and *Ypen*Man26A, were still the least stable of the analyzed enzymes.

5.2.1 Endomannanase activity on pure mannans

To assess activity level of the investigated endomannanases and compare enzyme performance in softwood saccharification with performance on pure, well defined mannans, initial rates of hydrolysis on locust bean gum, guar gum, konjac glucomannan as well as on acetylated and deacetylated spruce galactoglucomannan were determined at pH 5, 37 °C using 2.5 mg/ml substrate (Figure 11). The initial rate study was designed with a dose-response set-up as described in section 5.1.1. No endoglucanase cross-activity was found for the endomannanases, *Cvir*Man26A, *Pans*Man26A, *Anig*Man5A, *Tres*Man5A, when assessed on barley *β*-glucan and carboxylmethyl cellulose. It is therefore expected that the tested endomannanases do not degrade cellulose during softwood saccharification.

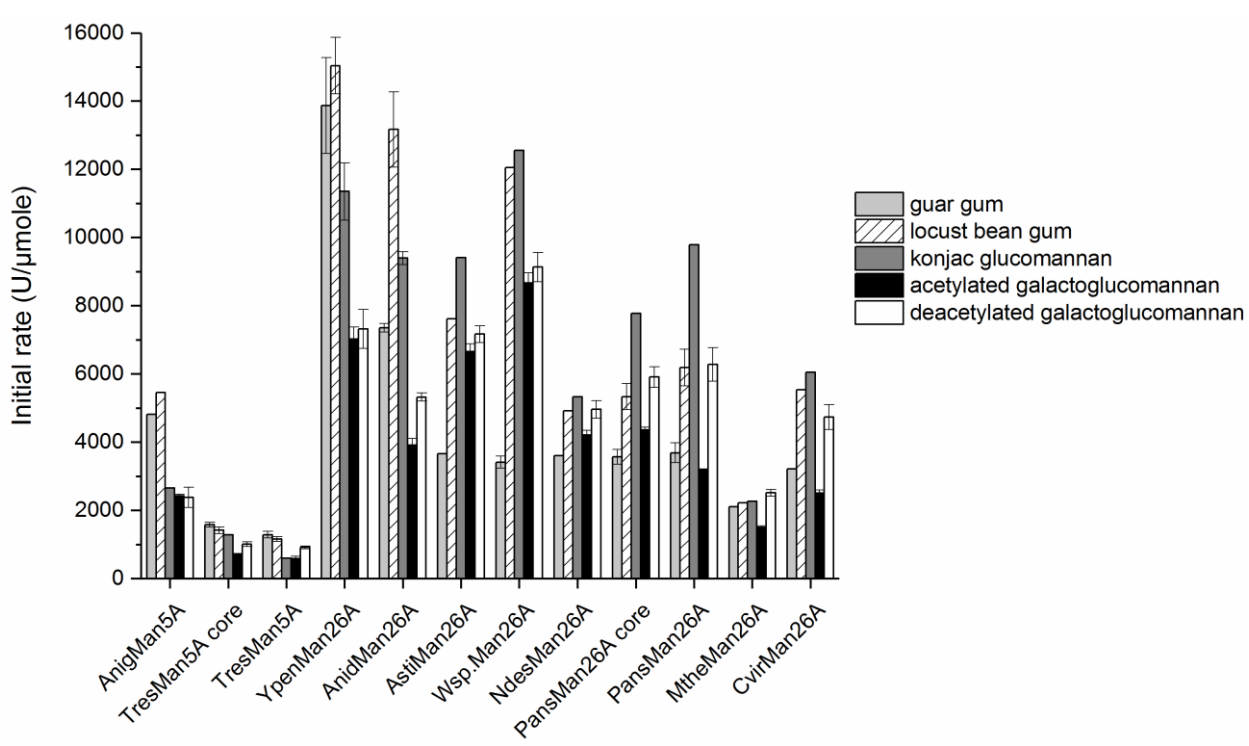


Figure 11: Initial reaction rates (U/µmole) by fungal endomannanases on mannans (37 °C and pH 5). Values are given as mean values ± SD (*n*=1-7).

The endomannanases exhibited different activity levels on the pure mannan substrates, but no consistent discrimination between substrate preferences of the enzymes was evident (Figure 11). *Ypen*Man26A had a significantly higher initial rate than all the other enzymes on the two galactomannans, locust bean gum (15050 U/μmole) and guar gum (13850 U/μmole). In contrast, *Wsp*Man26A had the highest initial rates on the glucomannans, konjac glucomannan (12550 U/μmole), acetylated galactoglucomannan (8650 U/μmole) and deacetylated galactoglucomannan (9150 U/μmole) (Figure 11). Deacetylation of galactoglucomannan doubled the rate for a few GH26 endomannanases (*Pans*Man26A and *Cvir*Man26A) but did not generally affect rates. The lowest initial rates were observed for *Tres*Man5A on acetylated galactoglucomannan and on konjac glucomannan (600 U/μmole) but *Tres*Man5A had generally low initial rates on all the tested mannans.

5.2.2 Endomannanase assisted softwood saccharification

The efficiency of the investigated endomannanases for saccharification of softwood was assessed by adding equal molar amounts of each endomannanase to Cellic ® CTec3 (a cellulase and xylanase mix). To this mix was also added BM2 (Table 5) to allow released mannooligosaccharides to be degraded to mannose. Mannose together with glucose and xylose released by the enzymes in Cellic ® CTec3 were measured by HPLC. The applied substrate was pretreated lodgepole pine (*Pinus contorta*), with 52 % cellulose, 12 % mannan, 6 % xylan, and 27 % lignin remaining after pretreatment (**Paper II**). The last 3 % came from galactose and arabinose side-groups on mannans and xylans, respectively, in approximately equal amounts.

Table 5: Applied doses of Cellic ® CTec3, accessory enzymes and BSA in softwood saccharification

Set-up [a]	**Cellic ® CTec3** [b] **mg EP/g DM**	**BM2 mg EP/g DM**	**Endomannanase or BSA mol/g DM**
(1)	10	1	$1.26 \cdot 10^{-8}$
(2)	50	1	$1.26 \cdot 10^{-7}$

[a] Two set-ups were used. Comparing endomannanases at 30 °C and 50 °C (**1**). Increased enzyme doses to evaluate saccharification at a higher degree of conversion (**2**). [b] Cellic ® CTec3 and the *A. niger* GH2 clade A mannosidase (BM2) doses are given as mg enzyme protein (EP)/g dry matter (DM) and not as mg product.

The saccharification was performed with 2 % DM at 30 °C to allow assessment of endomannanase performance based on specificity and not only stability. When assessed on locust bean gum, Cellic ® CTec3 itself exerted weak mannan degrading activity. The endomannanase addition levels were 10 times higher than this background activity. The release of glucose, mannose and xylose, respectively, was quantified at 24, 48 and 144h. Data are visualized for all investigated enzymes (Table 4) after 24h hydrolysis (Figure 12) and for a few selected endomannanases at all

time points (Figure 13). The best performing candidates was also assessed at 50 °C to evaluate their performance under industrially relevant conditions (Figure 13). BSA was added as a protein control to ensure that any differences in release of monosaccharides were not due to increased levels of protein as is sometimes observed in lignocellulose hydrolysis (- BSA binds non-productively) (Eriksson et al. 2002).

In a direct comparison of Cellic ® CTec3 with Cellic ® CTec3 plus BM2 plus BSA after 24 h hydrolysis, the mannosidase itself profoundly increased the release of mannose from 0.07 to 0.5 g/l (0.43 g/l increase) and the release of glucose from 3.06 to 3.18 g/l (0.12 g/l increase) (Figure 12). The increased mannose and glucose release is most likely due to BM2 activity on soluble galactoglucomanno-oligosaccharides in the mixtures. By removing mannopyranosyl units from the nonreducing end of these oligosaccharides, the BM2 will expose glucopyranosyl residues in the nonreducing end, which can be released by β-glucosidase activity from the Cellic ® CTec3. The released amount of mannose and glucose upon addition of BM2 corresponds to a Man:Glc ratio of 3.6:1. This ratio is in agreement with reported Man:Glc ratios in softwood (spruce) galactoglucomannans (Lundqvist et al. 2002; Lundqvist et al. 2003; Willför et al. 2003; Bååth et al. 2018) (Table 1), and suggests that no additional cellulose was degraded. The galactoglucomannano-oligosaccharides were most likely released by low endomannanase activity present in the Cellic ® CTec3 preparation and/or by weak glucomannan degrading capacity by some endoglucanases in this enzyme cocktail (Mikkelson et al. 2013).

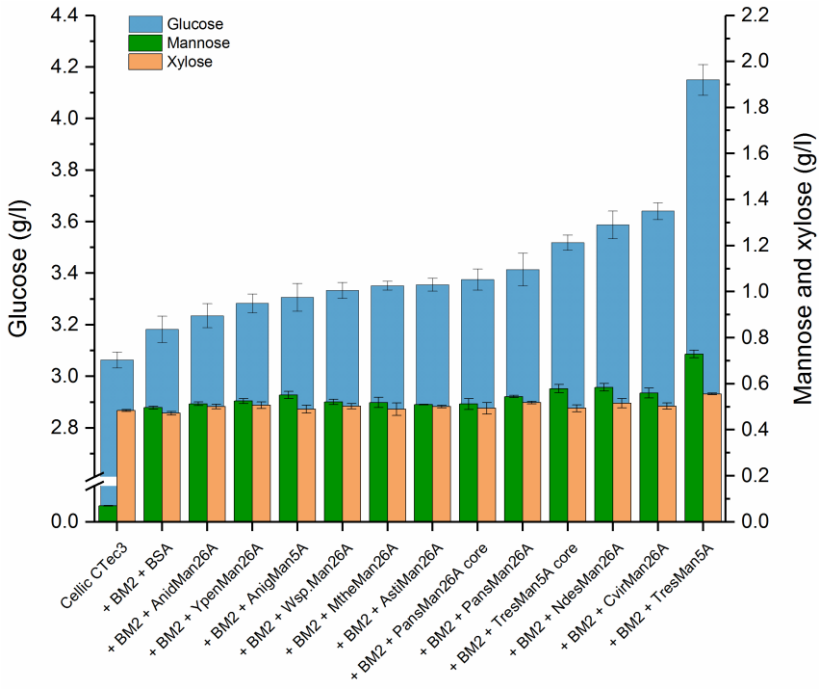


Figure 12: Softwood saccharification yields after 24 h at 30 °C. Endomannanases or BSA were added in equal molar amounts to Cellic ® Ctec3 plus an *A. niger* GH2 clade A mannosidase (BM2). Glucose, mannose and xylose yields (g/l) are given as mean values ± SD (n=3).

Supplementation of Cellic ® CTec3 with endomannanase significantly boosted the release of glucose for all tested enzymes, with *Tres*Man5A as the best performing candidate (Figure 12). After 24h enzyme treatment, the release of glucose and mannose obtained with *Tres*Man5A addition was 30 % (1 g/l increase) and 15 % (0.23 g/l increase) higher, respectively, than that of the control (Cellic ® CTec3 + BM2 + BSA, Figure 12), and much higher than those obtained with the other endomannanases. The relative amounts of released glucose and mannose (Man:Glc, 0.2:1) infer that the released glucose did not derive solely from hydrolyzed galactoglucomannan, but also from the cellulose fraction. The second-best performing enzyme was *Cvir*Man26A, which contains both a CBM35 and a CBM1, with a glucose yield of 88 % of that obtained by *Tres*Man5A (Figure 12). No obvious trends in the effect of GH5 versus GH26 endomannanases could be discerned. *Tres*Man5A was the superior enzyme, but glucose yields obtained with the other GH5 endomannanase, *Anig*Man5A, were in the low-to-middle range. For *Tres*Man5A, presence of CBM1 improved the release of both mannose and glucose. However, both *Cvir*Man26A and *Asti*Man26 with a CBM1 caused release of medium levels of glucose but there were no evident differences in their mannose release compared to the other endomannanases. No significant effect of the presence of CBM35 was observed for *Pans*Man26A.

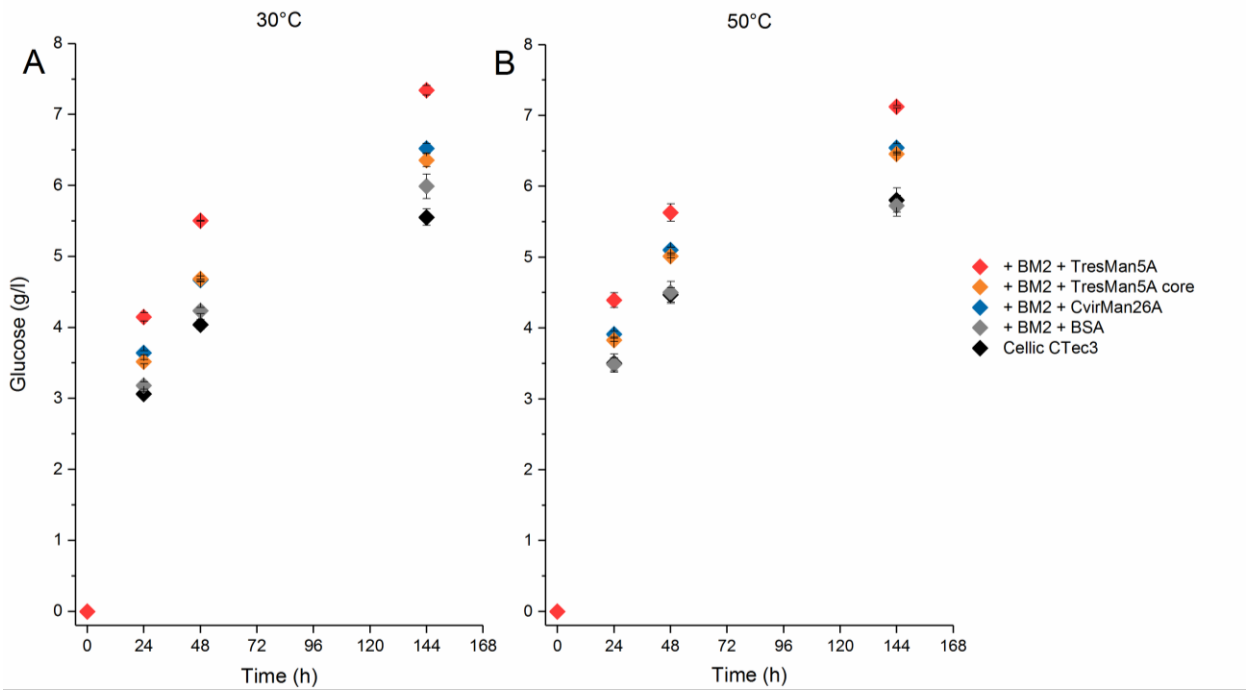


Figure 13: Softwood saccharification with selected endomannanases at 30 °C (**A**) or 50 °C (**B**). Endomannanases or BSA were added in equal molar amounts to Cellic ® Ctec3 plus an *A. niger* GH2 clade A mannosidase (BM2). Glucose yields (g/l) are given as mean values ± SD (n = 3).

When the two endomannanases, *Tres*Man5A and *Cvir*Man26A, were assessed at 50 °C, the time curves of the enzymatic glucose release at 50 and 30 °C were in complete agreement (Figure 13), and the ranking of enzyme performance was similar at the two reaction temperatures. The lack of increase in hydrolytic rate by cellulases in Cellic ® CTec3 with temperature (Q10 close to 1 between 30 and 50 °C) was not expected because the cellulases in Cellic ® CTec3 are stable at 30 °C as well as at 50 °C during prolonged reactions. However, because of the low substrate concentration used in the present study (2 % DM ~ 1 % cellulose), the data is in agreement with data published by Westh et al. (Sørensen et al. 2015; Westh et al. 2018). These authors have shown that at low Avicel concentrations, reduction in substrate affinity caused by heating (increase in K_M) cancels thermoactivation (increase in k_{cat}). The effective accessible substrate concentration in the present study may have been even lower because not all cellulose is equally accessible.

Since industrial lignocellulose conversion is usually performed at 50 °C, the data strongly indicate that addition of *Tres*Man5A to commercial cellulase preparations can efficiently boost glucose yields in industrial softwood saccharification reactions.

5.2.3 Correlation between the released monosaccharides

With the Cellic ® CTec3 and endomannanase doses used in the previous softwood saccharification trials (Table 5, set-up 1), the maximal degree of conversion was approximately 60 % of glucose (7.3 g/l of the available 11.5 g/l of glucose were released) after 144 h (Figure 13). To assess the softwood saccharification at a higher degree of conversion, the enzyme loadings were increased, i.e. the addition levels of *Tres*Man5A and *Cvir*Man26A and the Cellic ® CTec3 dose were increased (Table 5, set-up 2). At the higher enzyme doses, 85 % cellulose, 60 % mannan, and 81 % xylan conversion was obtained at 30 °C after 144 h.

The data obtained from all softwood saccharification trials (all studied endomannanases in all doses at 24, 48 and 144 h) showed a clear linear correlation between the release of glucose, mannose, and xylose throughout the degradation (Figure 14). These results support the interpretation that the softwood substrate comprises a complex network with glucomannans and xylans located throughout the lignocellulose matrix (Várnai et al. 2011), and not only on the outer surface. This means that their concurrent hydrolysis is crucial to obtain extensive hydrolysis of cellulose and in turn maximize the overall glucose yields from the softwood substrate. A reason for the low conversion of mannan (approximately 60 %) might be related to the galactose substitutions on galactoglucomannans that hinder the mannosidase in fully degrading the released mannooligosaccharides to mannose. Lower conversion could also be due to lignin carbohydrate complexes that can prevent enzymatic degradation (Eriksson et al. 1980; Ono et al. 2017).

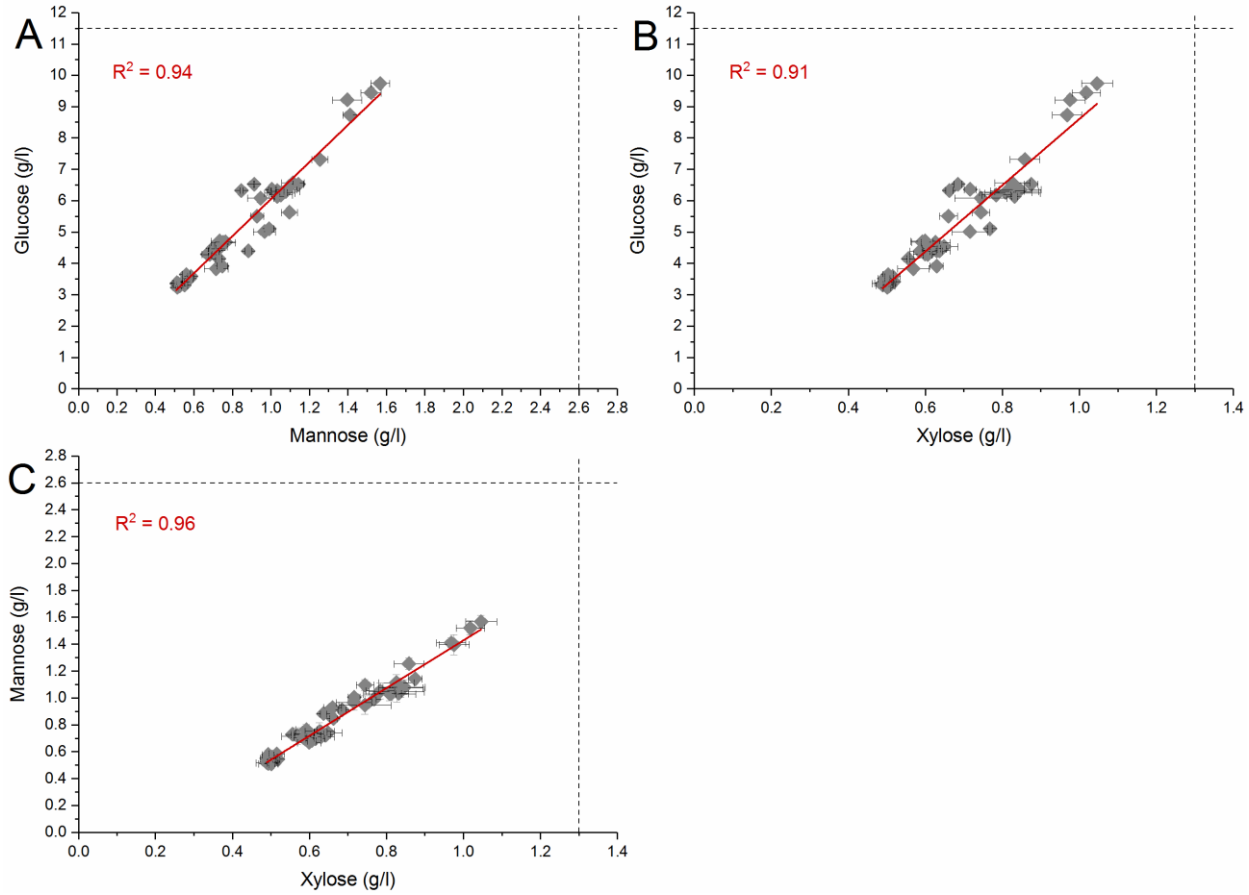


Figure 14: Correlation between released sugars (g/l) during softwood saccharification: glucose and mannose (**A**), glucose and xylose (**B**) and mannose and xylose (**C**). Dashed lines show the theoretical maximal monomeric yield based on the composition of the pretreated softwood substrate (**Paper II**).

5.2.4 Saccharification performance was not predicted by initial rate

Based on the measured activity levels on the pure mannans (Figure 11), it was not possible to predict the efficiency of the enzymes in softwood saccharification (Figure 12). Despite its high boosting capacity, *Tres*Man5A was found to have the lowest initial rate on the pure mannans, and *Ypen*Man26A, with low boosting capacity, had some of the highest initial rates on the pure mannans. Nor did particular substrate preferences observed on the pure mannans, such as having higher rates on glucomannans than on galactomannans, seem to determine the softwood saccharification performance of the individual endomannanases. Likewise, enzyme robustness alone could not explain the observed difference between the boosting capacity of the investigated endomannanases. An example of the latter is that *Tres*Man5A and *Cvir*Man26A had the same stable nature during 48 h incubation at 30 °C (Table 4) but differed in their boosting capacity (Figure 12). However, the low stability of *Ypen*Man26A and *Anid*Man26A (Table 4) may partially explain their poor overall performance in boosting of glucose release from softwood (Figure 12).

Even though the substrate concentration of the pure mannans (2.5 mg/ml) and the mannan present in the softwood saccharification mixtures (2 % DM = 20 g/l at 12 % mannan gives 2.4 mg/ml mannan) was approximately the same, the hydrolysis rate of mannan during softwood saccharification was much lower than the initial hydrolysis rates measured on the pure mannans (see calculation example in **Paper II**). This was probably partly because of limited access to the mannan in the softwood substrate. As indicated by Figure 14, not all mannan in the substrate is available at a given point in time. A coordinated interplay is required between the cellulases and xylanases to expose new mannan to the endomannanases. It is also likely that mannan conformation in the pretreated softwood substrate is different from the conformation of the pure mannans, and may be more difficult for the enzymes to degrade. The association of mannan with cellulose microfibrils may in particular hinder the endomannanase attack. Even though the softwood was pretreated, it cannot be ruled out that at least part of the softwood mannan is closely associated with cellulose, as it is in the biomass before pretreatment (Åkerholm & Salmén 2001). It is therefore also likely that another active site architecture is required to efficiently catalyze the galactoglucomannan removal from the cellulose microfibrils than for hydrolysis of pure mannans.

5.2.5 Explaining the high performance of *Tres*Man5A

*Tres*Man5A has earlier been found to catalyze hydrolysis of softwood galactoglucomannan (Tenkanen et al. 1997). It has also been reported that hemicellulases from *T. reesei* reduce hemicellulose exposed at the cellulose surface of wood materials to a greater extent than hemicellulases from *Aspergillus sp.* (Bombeck et al. 2017). However, the markedly better performance of *Tres*Man5A in softwood saccharification compared to all the other studied GH5 and GH26 fungal endomannanases was unexpected. Our hypothesis to explain the additional boosting effect is that *Tres*Man5A catalyzes a faster or more profound degradation of a certain type of mannan. This type of mannan is not immediately accessible to the other studied endomannanases and moreover, when degraded, allows for a more profound cellulose degradation. The particular portion of the mannan may be a more crystalline part that is more tightly intertwined with the cellulose.

It is likely that the CBM1 in the full-length *Tres*Man5A helps target the cellulose-associated mannan more efficiently than its truncated counterpart which lacks CBM1. This view is supported by a previous study of *Tres*Man5A action on cellulose-mannan complexes and the CBM1 cellulose binding capacity (Hägglund et al. 2003). Since the *Tres*Man5A core without the CBM1 was among the top performers in the softwood saccharification, the core module itself also played a role for the efficient degradation of the insoluble mannans. The reason that the GH26 endomannanases with a CBM1, i.e. *Cvir*Man26A and *Ast*Man26, did not release the same levels of mannose and glucose as *Tres*Man5A could be because their core modules are not as optimal

as *Tres*Man5A for degradation of mannan associated with cellulose. Another reason might be that their CBM1s have slightly different specificities than the *Tres*Man5A CBM1.

5.2.6 Summary

With regard to H2, it was established that fungal endomannanases perform differently in enzymatic cellulose saccharification and that *Tres*Man5A with its CBM1 was found to be the superior endomannanase in this study (Figure 12 and Figure 13). However, the capability to efficiently catalyze the removal of galactoglucomannans from cellulose microfibrils was not found to correlate with activity levels on pure mannan substrates. Therefore, it is most likely not the chemical structure of the mannan itself (whether the mannan contains galactose substitutions, glucose in the backbone or is acetylated) that determines the hydrolysis rate. Rather, the hydrolysis rate seems to be related to the fact that the mannan is found within a complex lignocellulosic matrix. It could be the ability to degrade more crystalline mannan or mannan more tightly intertwined with the cellulose that confers superior *Tres*Man5A performance.

5.3 Substrate interactions in fungal GH26 endomannanases (Paper III)

This section focuses on the interactions between fungal GH26 endomannanases and galactomannan. The following hypothesis and objective are addressed:

H3 The ability to accommodate multiple galactopyranosyl moieties in the active site cleft is a common feature among fungal GH26 endomannanases and semi-conserved residues associated with this feature can be identified.

Obj. 3. To experimentally identify amino acids involved in binding galactomannans in the active site cleft of fungal GH26 endomannanases, but also to evaluate if the substrate binding amino acids are conserved and to analyze how non-conserved amino acids affect the catalytic rate and substrate affinity.

To determine the substrate interactions, it was desirable to solve the crystal structure of a GH26 endomannanase in a complex with a galactomanno-oligosaccharide. Crystalization experiments were set up with the two wild type core GH26 endomannanases investigated in this study: *Anid*Man26A and *Ypen*Man26A. With no structure of a fungal GH26 endomannanase without a CBM, it was tried to solve the structure of these two core enzymes. Furthermore, *Anid*Man26A and *Ypen*Man26A were distinguishable from other analyzed GH26 endomannanases by having the highest initial rates on galactomannans (Figure 11). Crystals with *Anid*Man26A was not obtained within the time course of this PhD project.

5.3.1 The structure of *Ypen*Man26A in complex with MGG

The two catalytic residues of *Ypen*Man26A, Glu165 and Glu257, were identified based on a sequence alignment of *Pans*Man26A and *Ypen*Man26A. A *Ypen*Man26A acid/base substituted variant, E165Q, was made (by Novozymes). The variant was synthesized and expressed in *A. oryzae* (by Novozymes) and deglycosylated and purified to electrophoretic purity (in work performed as part of this PhD study). Crystals of *Ypen*Man26A E165Q were obtained in the presence of MGGMM.

The three-dimensional structure of the *Ypen*Man26A E165Q variant, in complex with MGG, was solved by molecular replacement using the known structure of *Pans*Man26A (3ZM8, Couturier et al. 2013) as template and refined at 1.36 Å. Neither the active *Ypen*Man26A nor the E165Q mutant crystallized as apoenzymes, which suggests that stability and/or conformational changes were obtained upon ligand binding. The *Ypen*Man26A structure forms a $(\beta/\alpha)_8$-barrel fold (Figure 15A) as expected for enzymes belonging to clan-GHA. The active site was identified in the groove with the conserved catalytic residue Glu165 (acid/base) mutated to Gln and the conserved catalytic residue Glu257 (nucleophile) (Figure 15A and C).

Crystals of *Ypen*Man26A E165Q were obtained in the presence of MGGMM with the aim that the oligosaccharide would span the catalytic site. However, the electron density of the ligand was modelled as MGG situated in the -4 to -2 subsites (Figure 15B). This result suggests that residual activity of the E165Q variant caused hydrolysis of the ligand between the backbone monomers in the -1 and +1 subsites after which MGG migrated to span the -4 to -2 subsites, which is indicative of high ligand affinity in these subsites. Cleavage of MGGMM might have been avoided if the nucleophile had been knocked-out instead of the acid/base. The electron density of MGG is clear and unambiguous except for the galactopyranosyl unit in the -3 subsite, which points out from the binding cleft (Figure 15B). All the interactions between the enzyme and the ligand are clearly defined, except for the flexible galactopyranosyl unit. Like *Pans*Man26A, *Ypen*Man26A has eight large loops that form a deep cleft at the active site. These loops are involved in binding of the substrate: loop 1 (36-39), loop 2 (60-73), loop 3 (95-131), loop 4 (166-179), loop 5 (207-211), loop 6 (227-235), loop 7 (259-263), and loop 8(279-291). The -1 and +1 subsites of YpenMan26A are similar to *Pans*Man26A, with the conserved residues His164, Trp170, Phe171, Tyr227, Trp279 (Figure 15C). As described for the homologous enzymes (Hogg et al. 2001; Le Nours et al. 2005; Couturier et al. 2013), *Ypen*Man26A Tyr227 is involved in a hydrogen bond with the catalytic nucleophile Glu257 whilst the aromatic amino acids Trp170 and Trp279 stabilize the mannopyranose rings at the -1 and +1 subsites, respectively (Figure 15). Like *Pans*Man26A, *Ypen*Man26A displays a prominent -4 subsite, with stacking interactions between the mannopyrannose ring and the two aromatic residues W109 and W110, and hydrogen bonds between Asp61, Arg66 and the mannopyrannose ring (Figure 15C). In the -2 subsite the two

aromatic residues, Phe113 and Tyr114, equivalent to Phe248 and Tyr249 in *Pans*Man26A, stabilize the interactions with the mannopyranose unit. Interestingly, because of the captured ligand in the present work, it is possible to identify previously undescribed interactions between the galactopyranose unit and the *Ypen*Man26A in the -2 subsite. Gln36, Asp37, and Asp58 are involved in hydrogen bonds with the galactose residue. Asp37 has a double conformation in the crystal structure, possibly because the amino acid conformation shifts upon ligand binding. *Pans*Man26A has a Glu172 instead of the Asp37 in *Ypen*Man26A, but otherwise the enzymes have an essentially identical environment for interactions with the galactose residue.

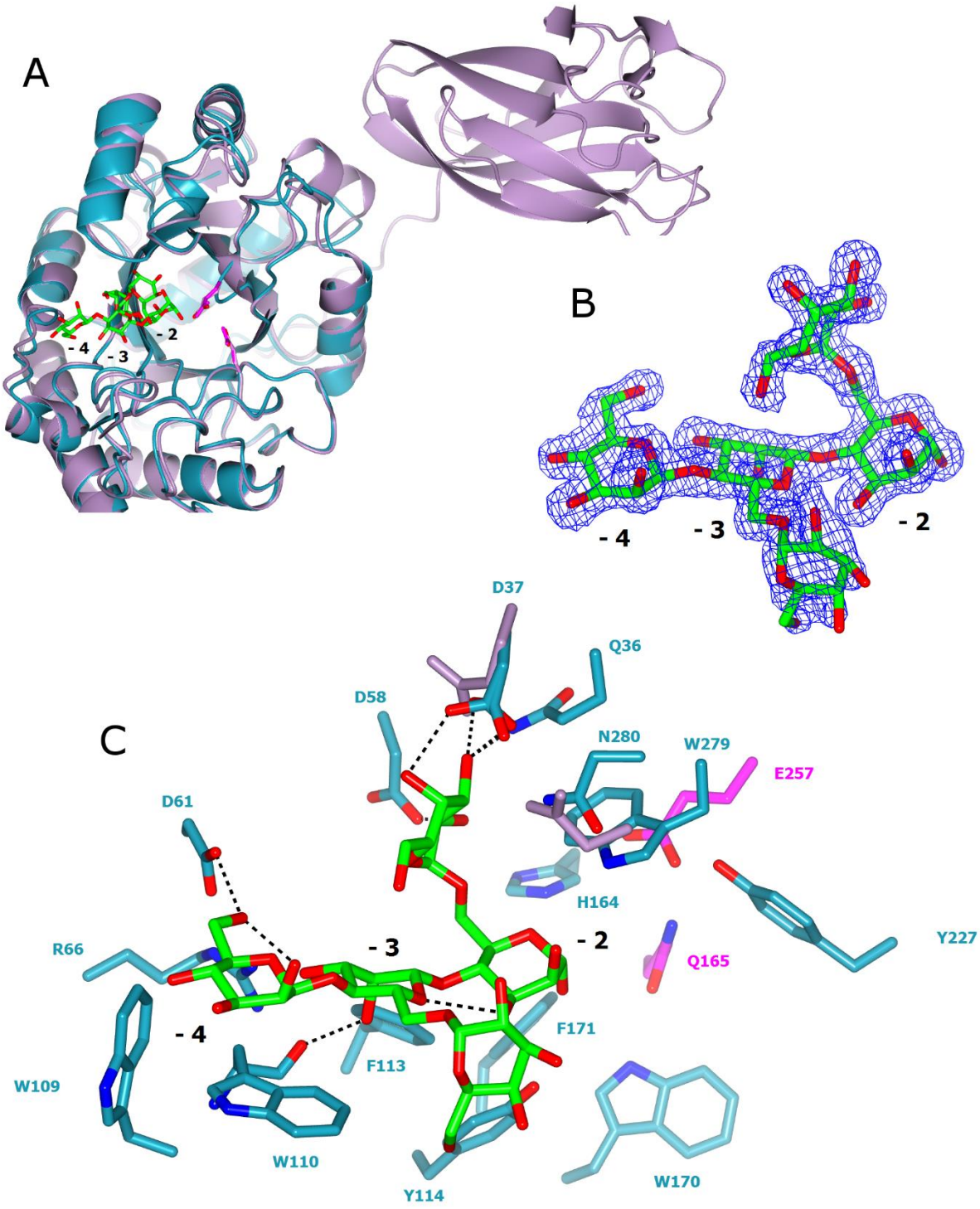


Figure 15: The structure of *Ypen*Man26A (6HPF, dark cyan) superimposed with that of *Pans*Man26A (3ZM8, Couturier et al. 2013, lilac) (**A**). The ligand, α-6^2-6^1-di-galactosyl-mannotriose (MGG), in *Ypen*Man26A (subsites -4 to -2) is shown as green cylinders and the catalytic residues are shown in shades of pink. Observed electron density for MGG in the -4 to -2 subsites (**B**). The positive electron density REFMAC $F_o - F_c$ map, contoured at 3.5 σ (0.37 e Å^{-3}), is shown in blue, with phases calculated prior to the incorporation of any ligand atoms in refinement. The organization of the binding subsites and the MGG ligand in the -4 to -2 subsites of *Ypen*Man26A (**C**). *Pans*Man26A residues are shown (lilac) if they differ from *Ypen*Man26A.

5.3.2 Ligand binding amino acids in fungal GH26 endomannanases

In the previous investigations of the influence of galactopyranosyl residues in the active site cleft of fungal endomannanases, 30-40 % conversion of guar gum was found to be characteristic for *Anid*Man26A and *Pans*Man26 (Figure 6) which were also found to accommodate multiple galactopyranosyl residues in the active site cleft. The conversion of guar gum was also found to reach 30 % for the six other endomannanases characterized in this study (Figure 16), and also they produced G from guar gum hydrolysis (data not shown), which suggests that they have a similar capacity as *Anid*Man26A and *Pans*Man26 to accommodate galactopyranosyl units. A multiple sequence alignment of all eight fungal GH26 endomannanases, and *Mt*Man26A from *Myceliophthora thermophila* previously characterized by Katsimpouras et al (Katsimpouras et al. 2016), showed that the amino acids that take part in ligand binding in *Ypen*Man26A are highly conserved among the examined fungal GH26 endomannanases (Figure 17, blue circles). However, *Wsp.*Man26A has two striking differences compared to *Ypen*Man26A and the other endomannanases. The first difference is in the -2 subsite (*Ypen*Man26A Asp37) where the analyzed endomannanases are predicted to have either an Asp or a Glu, while *Wsp.*Man26A has Thr (Figure 17). The second difference is in the -4 subsite (*Ypen*Man26A Trp110), where the tested endomannanases are predicted to have Trp or Tyr, while *Wsp.*Man26A has His (Figure 17).

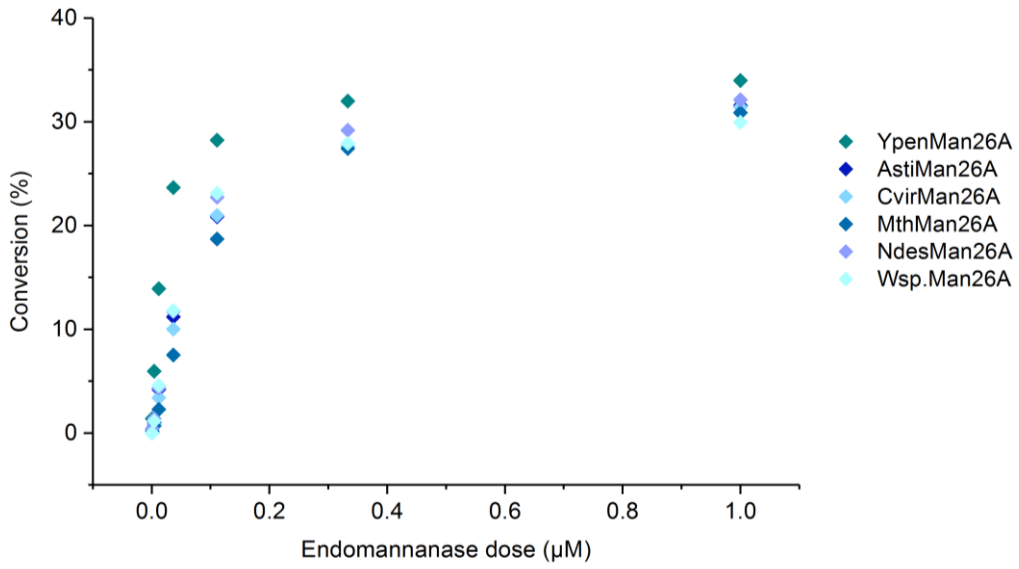


Figure 16: Conversion (%) of guar gum by fungal GH26 endomannanases (37 °C and pH 5). The degree of conversion was defined as released reducing ends relative to the theoretical monomeric yield. Values represent a single replicate.

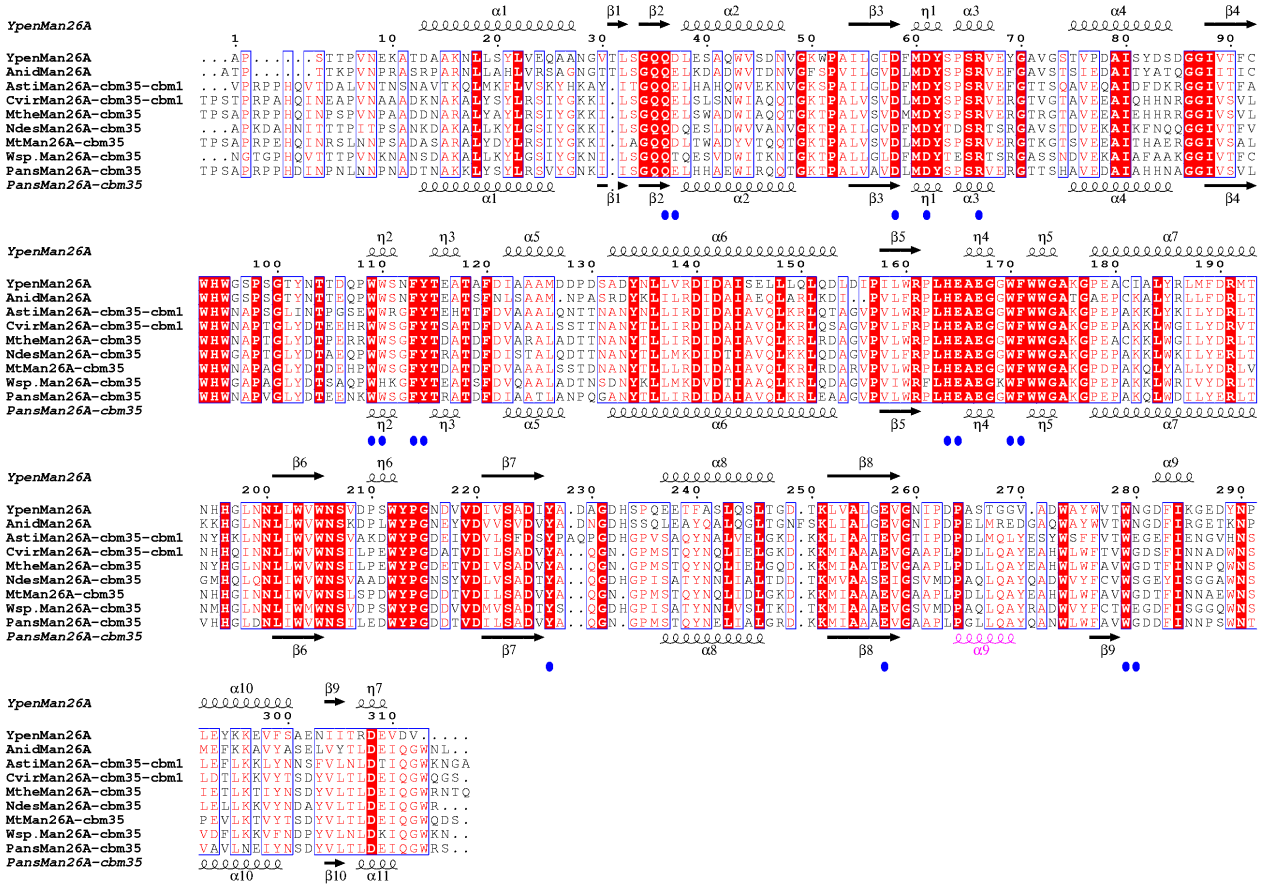


Figure 17: Sequence alignment of the catalytic core region from nine fungal GH26 endomannanases. Secondary structure elements for *Ypen*Man26A and *Pans*Man26A are displayed above and below the alignment, respectively. • Residues involved in ligand binding in the *Ypen*Man26A structure, including the two catalytic residues, are shown as blue circles. The α-helix in *Pans*Man26A (α9), which is nearest the CBM35 and which is a surface loop in *Ypen*Man26A, is colored pink. Identical residues are shown in white on red background. Highly similar residues (when the similarity score assigned to one column is above 0.7) are colored red and framed in a blue box. The GH26 core sequence of *Ypen*Man26A (MH899111), *Anid*Man26A (Q5AWB7), *Asti*Man26A (BBW45412), *Cvir*Man26A (BBW45415), *Mthe*Man26A (MH208368), *Ndes*Man26A (MH208367), *Mt*Man26A (99077) (Katsimpouras et al. 2016), *Wsp.*Man26A (MH208369), *Pans*Man26A (B2AEP0) were aligned by MUSCLE (Edgar 2004) and the figure was generated using ESPript 3 Web server (Robert & Gouet 2014).

5.3.3 Differences in ligand binding amino acids between *Ypen*Man26A and *Wsp*.Man26A affect substrate affinity and binding mode

The identified differences in ligand binding amino acids found in *Wsp*.Man26A did not change its conversion profile on guar gum (Figure 16). However, when screened for activity on pure mannans (Figure 11), *Wsp*.Man26 seemed to be more hindered or have less affinity for the increased amount of galactose substitutions in guar gum compared to the other endomannanases. While there was hardly any difference between activity of *Ypen*Man26A on locust bean gum and guar gum, *Wsp*.Man26A had an approximately four times higher initial hydrolysis rate on locust bean gum than on guar gum. It requires analyses on several galactomannan substrates, to discover these differences in galactose sensitivity between the enzymes. The hydrolysis product profiles from full conversion of guar gum by *Ypen*Man26A and *Wsp*.Man26A were analyzed using the DASH method (Li et al. 2013) (Figure 18). Like *Anid*Man26A and *Pans*Man26A (Figure 7), *Ypen*Man26A produced primarily G (2.10 DE) and MGG (4.10 DE), whereas *Wsp*.Man26A also produced M2 and M3 (Figure 18A). To investigate if the difference in ligand interacting amino acids between *Ypen*Man26A and *Wsp*.Man26A played a role in the observed differences in substrate preference and binding mode, two *Ypen*Man26A mutants, *Ypen*Man26A D37T and *Ypen*Man26A W110H, were designed, expressed (by Novozymes) and purified to electrophoretic purity (Table 6 and Figure 18).

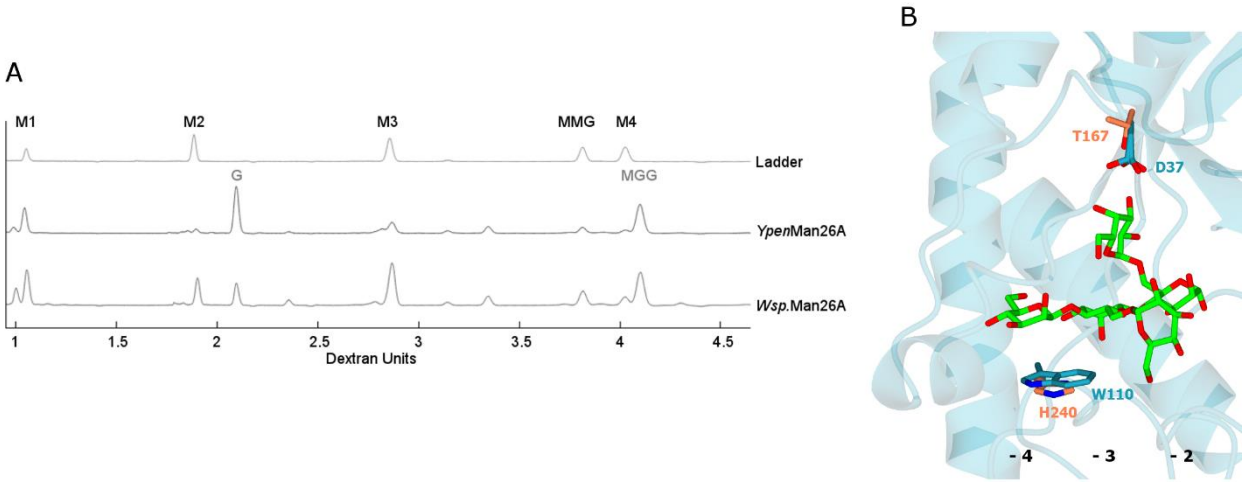


Figure 18: Product profiles from guar gum hydrolysis by *Ypen*Man26A and *Wsp*.Man26A (**A**). Aligned electropherograms of product profiles at 30 % guar gum conversion. Migration of oligosaccharides is given in dextran units (DE). A ladder was run containing: mannose (M1, 0.9 DE), mannobiose (M2, 1.87 DE), mannotriose (M3, 2.85 DE), and α-6^1-galactosyl-mannotriose (MMG, 3.81 DE). Migration of α-galactosyl-mannose (G, 2.10 DE) and α-6^2-6^1-di-galactosyl-mannotriose (MGG, 4.10 DE) was established during the PhD study. The structure of *Ypen*Man26A with MGG in the -4 to -2 subsites (**B**). The two differences in ligand binding amino acids between *Ypen*Man26A and a superimposed homology model of *Wsp*.Man26A are highlighted in dark cyan and orange, respectively.

Table 6: The wild type *Ypen*Man26A and the investigated variants

Enzyme	Domains	Mw [a] (kDa)	*Tm* [b] (°C)
*Ypen*Man26A (MH899111)	GH26	34.5	50
*Ypen*Man26A D37T	GH26	34.5	50
*Ypen*Man26A W110H	GH26	34.4	47

Mannopentaose hydrolysis product analyses using HPAEC combined with solvent isotope labelling and MS analysis (Hekmat et al. 2010; Couturier et al. 2013) was used to estimate the relative frequency of productive binding modes for the *Ypen*Man26A wild type and W110H mutant. These experiments were carried out in Lund by Mathias Wiemann (**Paper III**) and showed that the W110H mutation did in fact change the productive mannopentaose binding mode. While the wild type *Ypen*Man26A had its dominant binding mode from subsites -4 to +1 (80 %), the W110H variant had its dominant binding mode from subsites -3 to + 2 (63 %). This is most likely a consequence of Trp110 in the -4 subsite being changed to His and resulting in a weaker subsite.

Michaelis-Menten kinetic parameters for locust bean gum and guar gum were determined for the *Ypen*Man26A wild type and the two *Ypen*Man26A mutants D37T and W110H using a reducing end assay (Table 7) to evaluate their rate and affinity for the two galactomannan substrates. The wild type had the highest k_{cat} and k_{cat}/K_M on both substrates. The wild type and variant D37T had identical K_M for locust bean gum, but D37T had higher K_M than the wild type on guar gum. This indicates that the D37T mutant has lower affinity for the galactose residues in the highly substituted guar gum than the wild type. Unsubstituted blocks of mannan found in locust bean gum (McCleary 1985) might be the reason for no difference in K_M being observed for this substrate. It is likely that both the wild type and the D37T variant degrade the unsubstituted, more easily accessible part of the substrate first, which is why the initial rate reflects the enzyme affinity for the unsubstituted regions of the substrate. Guar gum is known to have no (or few) blocks without substitutions (McCleary 1985). As the K_M value suggests, the *Ypen*Man26A W110H had very low affinity for locust bean gum when compared to the *Ypen*Man26A wild type and the D37T mutant (Table 7). On guar gum galactomannan it was not possible to determine the kinetic parameters separately because saturation was not reached, but the low k_{cat}/K_M indicates low affinity or low hydrolysis rate.

Table 7: Kinetic parameters on locust bean gum and guar gum for *Ypen*Man26A and the D37T and W110H variants

Enzyme	Locust bean gum			Guar gum		
	k_{cat} (s^{-1})	K_M (mg/ml)	k_{cat}/K_M (ml/(mg*s))	k_{cat} (s^{-1})	K_M (mg/ml)	k_{cat}/K_M (ml/(mg*s))
Wild type	475 ± 5	0.6 ± 0.03	792 ± 40	636 ± 19	2.2 ± 0.2	289 ± 28
D37T	334 ± 6	0.6 ± 0.05	557 ± 47	473 ± 12	2.7 ± 0.2	175 ± 14
W110H	404 ± 18	10 ± 0.8	40 ± 4	n.d.[a]	n.d.[a]	17 ± 0.6

[a] Not determined (n.d.) because saturation was not reached. Linear regression was used to determine k_{cat}/K_M from the initial part of the Michaelis-Menten curve.

k_{cat}/K_M on MGGMM for the *Ypen*Man26A wild type and the D37T mutant was determined by following substrate depletion at low substrate concentration (0.1 mM) by MS (Table 8) (**Paper III**, supplementary material for more detailed MS results). The D37T variant had four times lower k_{cat}/K_M on MGGMM than the wild type enzyme (19 vs 84 $s^{-1}*mM^{-1}$, Table 8), which means that the mutant has lower k_{cat} and/or higher K_M (probably a combination of both as for the individual kinetic parameters determined on guar gum). This result emphasizes that the substitution of Asp37 with Thr seems to decrease the affinity for the galactopyranosyl moiety in the -2 subsite. This is consistent with the expected increase in distance between the galactopyranosyl unit and the amino acid residue when Asp is substituted by Thr (Figure 18B). It is likely that MGGMM binds from the -4 to the +1 subsites in *Ypen*Man26A and therefore accommodates the galactopyranosyl residues in the -3 and -2 subsite, as in the X-ray structure (Figure 15C). This can also be assumed from the dominant mannopentaose productive binding mode for *Ypen*Man26A from subsite – 4 to +1 (**Paper III**). Furthermore, *Anid*Man26A, which is the closest homolog to *Ypen*Man26A (67.5 %), was found to produce MGGM and M from MGGMM (Figure 9).

Table 8: Kinetic efficiency for the *Ypen*Man26A wild type and the D37T variant

Enzyme	k_{cat}/K_M ($s^{-1}*mM^{-1}$) on MGGMM
Wild type	84 ± 5
D37T	19 ± 2

5.3.4 Fungal GH26 endomannanases with and without a CBM35

Most regions are highly conserved between *Ypen*Man26A and *Pans*Man26A (Figure 15A and C). However, unlike *Pans*Man26A and most of the other fungal GH26 endomannanases,

*Ypen*Man26A does not have a N-terminal CBM35 domain. From the superimposition of the two crystal structures (Figure 15A), the main difference in the secondary structure between the core modules is seen to be in the area which approaches the CBM35 of *Pans*Man26A. In this area, *Pans*Man26A has an α-helix and *Ypen*Man26A has a surface loop. Study of the crystal structure of *Pans*Man26A revealed that interactions occur through water between Ala402 and Gln404 in the *Pans*Man26A core domain and Leu58 and the Ser130 in its CBM35 and linker. Couturier et al. (2013) also state that a hydrophobic patch comprising Leu58 and Leu130 on the surface of the CBM35 stands in front of a cluster of hydrophobic residues, Ala402, Tyr403 and Leu399, of the core domain (Couturier et al. 2013). The multiple sequence alignment (Figure 17) confirms variation in the region in and around α9 in *Pans*Man26A (Figure 17, marked pink). The seven enzymes with a CBM35 have identical sequences to *Pans*Man26A (LQAY; for *Asti*Man26A it is MQLY), which forms the α-helix in the *Pans*Man26A structure, while the two enzymes with no CBM35 have a different and seemingly more random sequence (TGGV for *Ypen*Man26A and MRED for *Anid*Man26A). From this analysis, it appears that co-evolution has taken place between the GH26 core domain and the CBM35.

5.3.5 Summary

The findings support the hypothesis H3 in that the ability to accommodate multiple galactopyranosyl units in the active site cleft is conserved among the fungal GH26 endomannanases tested. The galactomannan binding amino acids in the -4 to -2 subsite of *Ypen*Man26A were identified by solving and analyzing the enzyme crystal structure in complex with MGG. The identified ligand binding amino acids were found by sequence alignment to be highly conserved among the investigated GH26 enzymes. In addition, all analyzed GH26 endomannanases reached around 30 % conversion on guar gum. However, *Wsp.*Man26 was found to have a different hydrolysis product profile and lower initial rate of guar gum hydrolysis than *Ypen*Man26A. In agreement with these differences, *Wsp.*Man26A was also found to have a few changes in the ligand binding amino acids when compared to *Ypen*Man26A and the other GH26 endomannanases. *Ypen*Man26A mutants, W110H and D37T, were designed inspired by the variations observed in *Wsp.*Man26A. Analyses of the binding modes and kinetic studies showed that these specific substitutions weakened the -4 subsite interactions with the mannan backbone and interactions with the galactose moiety in the -2 subsite, respectively. Interestingly, it was also found that the core module of the investigated GH26 endomannanases differed depending on whether they were carrying a CBM35 or not. *Ypen*Man26A and *Anid*Man26A, which did not carry a CBM35, had a flexible surface loop instead of the α9-helix found in the enzymes carrying CBM35. The α9-helix interacts with the CBM35 and might help in positioning of the binding domain.

6 General discussion

Relative few entries for fungal GH26 endomannanases are found in the CAZy database (www.cazy.org) (Lombard et al. 2014), when compared to the amount of bacterial GH26 endomannanases, but also when compared to registered fungal endomannanases in GH family 5 and 134. In the work related to this thesis, fungal endomannanases were found predominantly as modular enzymes carrying a CBM35. Two enzymes carrying both a CBM35 and a CBM1 and two core enzymes was also found and characterized. The results obtained in this PhD study add to the current understanding of fungal GH26 endomannanases. These enzymes appear to be conserved in their main known functional characteristics.

6.1 Fungal GH26 endomannanases

Characterized fungal GH26 endomannanases, including *Pans*Man26A and *Ypen*Man26A, have a characteristic binding site with a strong -4 subsite, and a dominant mannopentaose binding mode from the -4 to +1 subsite (**Paper III**) (Couturier et al. 2013; Marchetti et al. 2016). To date, the fungal GH26 endomannanases that have been analyzed with a focus on the accommodation of galactopyranosyl units are able to degrade highly substituted galactomannans by allowing accommodation of galactose substitutions at least in the -3, -2, -1 and +1 subsites, evaluated by biochemical data and crystal structures (**Paper I** and **III**). The biochemical data include the observations that *Pans*Man26A and *Anid*Man26A produce α-galactosyl-mannose (G) as their dominant hydrolysis product from guar gum galactomannan and that *Anid*Man26A catalyzes the hydrolysis of MGGMM to MGGM and mannose (**Paper I**). The structural data include the crystal structure of *Pans*Man26A (Couturier et al. 2013) and *Ypen*Man26A (**Paper III**) and the homology model of *Anid*Man26A, which all show an open active site cleft with space for galactose substitutions in multiple subsites (**Paper I** and **III**). The observation that the amino acids that participate in MGG binding in *Ypen*Man26A (-4 to -2 subsites) are highly conserved between the studied GH26 endomannanases (Figure 17) supports the view that the ability to accommodate multiple galactopyranosyl moieties in the active site may be conserved among fungal GH26 endomannanases. However, since single mutations can change affinities or binding modes, as was also seen in this PhD study (**Paper III**), it is not unlikely that a fungal GH26 endomannanase with altered characteristics may be found. It is also possible that GH26 endomannanases from fungi other than ascomycetes may have altered characteristics.

The characterized fungal GH26 endomannanases were all found to have lower melting temperatures than the investigated fungal GH5 endomannanases (50 – 68 °C vs 70 – 87 °C). The melting temperatures were not directly comparable to stability at 30 °C. However, *Anid*Man26A

and *Ypen*Man26A, which were the two GH26 endomannanases not carrying a CBM, had the lowest melting temperatures of the studied GH26 endomannanases (53 and 50 °C, respectively, vs 57 – 68 °C) and were also the least stable enzymes at 30 °C, with half-lives of 10 and 21 h, respectively. By comparison, the other tested enzymes had half-lives of 90 h or above. The *Pans*Man26A enzyme, both in its full length and as a truncated variant without its CBM35, had the same stability, both when assessing the melting temperature (57 and 58 °C) and half-live (90 and 103 h). These results suggest that it is not the CBM35 that adds extra stability to the core enzyme, but rather that it could be the α9-helix, found in both versions of *Pans*Man26A and all the other tested endomannanases with a CBM35 but not in *Anid*Man26A and *Ypen*Man26A which instead had a surface loop. Extra surface loops can lower protein stability and decrease the enzyme temperature optimum (Kim et al. 2014). Whether *Anid*Man26A and *Ypen*Man26A are particularly suited for functioning under cold conditions still needs to be investigated.

6.1.1 The CBM35

Based on the enzymes characterized in this study and in public databases, most fungal GH26 endomannanases carry a CBM35 (Table 3 and Table 4). CBM35s with different ligand specificities have been found and a classification system has been suggested, which divides the family into four subfamilies with specificity for (I) α-galactopyranosyl residues, (II) internal regions of β-mannans, (III) gluco-configurations (less defined), or (IV) uronic acids (Correia et al. 2010). Initial rate measurements with *Pans*Man26A, with and without its CBM35 on a variety of mannans, suggest that *Pans*CBM35 binds the mannan backbone and not the galactose substitutions (Figure 6 and Figure 11). A phylogenetic tree constructed using the CBM35s from the endomannanases studied here and other, better characterized CBM35s used by Correia et al. (2010) in the subfamily classification, suggests that all the CBMs from the current study belong in the same subfamily and cluster with CBM35s known to bind mannan backbones (data not shown). However, more focused characterization of the CBM35s carried by GH26 endomannanases is needed to establish their specificity.

The results of the present PhD study indicate that the GH26 core modules of the endomannanases carrying a CBM35 have evolved to harbor this binding domain (15 kDa) in close proximity to the core with the aid of an α-helix (**Paper III**). Interactions were found between this α-helix and the CBM35 in the crystal structure of *Pans*Man26A (Couturier et al. 2013). Furthermore, the two core enzymes with no CBM, *Anid*Man26A and *Ypen*Man26A, had a surface loop, where the CBM35 carrying enzymes had the α-helix. It is possible that this α-helix plays an important role in positioning the CBM35. It is also possible that the position we see in the crystal structure of *Pans*Man26A (Couturier et al. 2013) is not the position of the CBM35 in solution. Possibly, the core domain and the CBM35 can come into even closer contact, maybe facilitated by ligand binding. Examples of possibly similar inter-module interactions have been reported

for processive GH9 endoglucanases whose CBM3c modules were shown to align in continuation with the catalytic cleft of the GH9 module, presumably to form one functional entity (Petkun et al. 2015). The linker in these GH9 endoglucanases is wrapped around the core domain (Petkun et al. 2015), similarly to the linker in *Pans*Man26A (Couturier et al. 2013), and contributes significantly to the positioning of the CBM3c.

6.1.2 Comparisons to other endomannanases

The fungal GH26 endomannanases were found to be efficient in their degradation of highly substituted galactomannans, both in terms of high rate and particularly in extent of degradation, when compared to the investigated fungal GH5 endomannanases. This is also evident when these GH26 endomannanases were compared to known fungal GH134 endomannanases. The latter have a strong preference for unsubstituted linear mannans over galactomannans (Jin et al. 2016; Sakai et al. 2017; You, Qin, Li et al. 2018).

The bacterial GH26 endomannanases, as exemplified by *Bova*Man26A and *Bova*Man26B from *Bacteroides ovatus*, demonstrate greater variation in the ability to accommodate galactopyranosyl residues in the active site cleft. *Bova*Man26A appear to be severely restricted by galactopyranosyl side groups, while *Bova*Man26B is much less so. Furthermore, *Bova*Man26A produces small (DP 2-6) oligosaccharides from guar gum hydrolysis, whereas *Bova*Man26B produced larger galactomannooligosaccharides (Bågenholm et al. 2017). This is an example of fine-tuned functional differences within the bacterial GH26 endomannanases (Bågenholm et al. 2017).

Enzyme interactions with a galactopyranosyl substituent in the -1 subsite of *Cjap*Man26C have previously been described (Cartmell et al. 2008). A surface view of *Ypen*Man26A and *Cjap*Man26C (2VX6) with their ligands superimposed over one another (the MGG from *Ypen*Man26A and a bound α-6^3-galactosyl-mannotetraose (MGMM) in the -2 to +2 subsite of *Cjap*Man26C) shows that the ligands nicely overlap and hence gives an indication of the accommodation of galactopyranosyl residues in the -3, -2 and -1 subsites of both enzymes (Figure 19). These superimpositions show that while *Ypen*Man26A seems to accommodate the galactopyranosyl moieties in the -3, -2 and -1 subsites, as was also seen in the current study, *Cjap*Man26C does not accommodate the galactopyranosyl unit in the – 2 subsite where the moiety is pointing into the enzyme structure (Figure 19). The same is seen when *Bova*Man26A (4ZXO) is superimposed on *Ypen*Man26A and *Cjap*Man26A to visualize their ligands in its active site (Figure 19). Since there is no crystal structure of *Bova*Man26B which seems better at degrading galactomannan polymers than *Bova*Man26A (Bågenholm et al. 2017), it is not known if this enzyme accommodates galactose units in the -2 subsite.

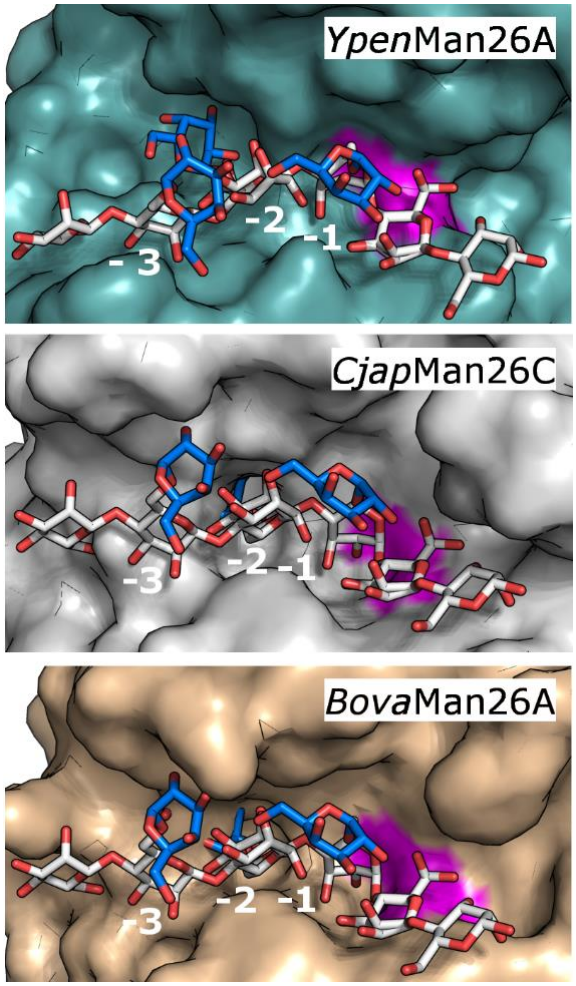


Figure 19: Surface views of the active site clefts of *Ypen*Man26A (6HPF), *Cjap*Man26C (2VX6), and *Bova*Man26A (4ZXO) showing the architecture of their active site cleft. Superimposition of the three structures allowed visualization of ligands from *Ypen*Man26A and *Cja*pMan26C in all three structures: α-6^2-6^1-di-galactosyl-mannotriose (MGG) binding from the -4 to -2 subsites in *Ypen*Man26A and α-6^3-galactosyl-mannotetraose (MGMM) binding from the -2 to +2 subsites in *Cjap*Man26C. Mannose units are colored white and the galactose substitutions are colored blue. Catalytic residues are shown in magenta.

6.1.3 The biological role

Some organisms are known to express several endomannanases. A well-studied example is the bacterium *C. japonicus* which expresses both GH5 and GH26 endomannanases (Hogg et al. 2003; Tailford et al. 2009; Zhang et al. 2014). Studies describe modular and specificity differences between the GH5 and GH26 endomannanases from *C. japonicus*, which point to differences in their biological role in mannan degradation (Tailford et al. 2009; Zhang et al. 2014). The GH26 endomannanases in *C. japonicus* are lipoproteins without CBMs (Hogg et al. 2003; Zhang et al. 2014). These endomannanases have tight specificity for mannose (not glucose) in the -2 subsite (Tailford et al. 2009) and seem to be particularly efficient in hydrolyzing mannooligosaccharides but not plant cell wall associated mannans. These results suggest that their likely biological role

is hydrolysis of mannooligosaccharides and mannan fragments released from the plant cell wall (Zhang et al. 2014). In contrast, the *C. japonicus* GH5 endomannanases contain CBMs and are secreted into the culture medium (Hogg et al. 2003; Zhang et al. 2014). These endomannanases can bind both mannose and glucose at the -2 subsite and was found to be superior, both in rate and extent, in the degradation of plant cell wall mannans, but less efficient in degradation of small mannooligosaccharides than their GH26 counterparts (Zhang et al. 2014). These results suggest that the GH5 endomannanases from *C. japonicus* was adapted to target cell wall mannans and/or glucomannans (Zhang et al. 2014).

The present studies of GH5 and GH26 endomannanases indicate that GH family categorization with regard to biological role is not as clear cut for the fungal endomannanases (**Paper II**). A GH26 endomannanase was found to perform best on the pure and soluble mannans and a GH5 was the most effective on the pretreated cell wall substrate (from lodgepole pine). However, most of the tested GH26 endomannanases performed on par with or even better than the *Anig*Man5A in softwood saccharification (Figure 12). The fact that some fungal GH26 endomannanases were found with a CBM1 and that they were secreted into the culture medium, also indicates that these enzymes could participate in plant cell wall degradation in nature. However, it should be emphasized that the present PhD study was not designed to assess the biological role of the enzymes. As an example, it has not been addressed how the investigated fungal endomannanases differ in the degradation of small mannooligosaccharides. Furthermore, it is not known how the pretreatment affects the structure and composition of the softwood and it is therefore possible that other results would be obtained on a non-pretreated softwood substrate.

Studies with *Pans*Man26A and *Pans*Man5A from *P. anserina* and the multiple endomannanases expressed by *A. nidulans* (the *Anid*Man26A, *Anid*Man5A and *Anid*Man5C studied in this work but also *Anid*Man5B (Rosengren et al. 2014), and *Anid*Man134 (Shimizu et al. 2015)) also suggest that these organisms use another strategy for mannan degradation than *C. japonicus* (Couturier et al. 2013; Zhang et al. 2014). It has been suggested that these ascomycetes have several endomannanases with complementary substrate preferences to create synergy in the deconstruction of mannans or to be able to induce a specific enzyme in a situation when its characteristics are needed (Dilokpimol et al. 2011; Couturier et al. 2013; Rosengren et al. 2014). The *A. nidulans* example highlights that differences in substrate specificity are not seen only between endomannanases from different GH families, but are also found within one family, as is the case with *Anid*Man5A, *Aidn*Man5B and *Aidn*Man5C.

6.2 Evaluation of selected methods

6.2.1 DASH

DASH was used for the assessment of product profiles from endomannanase catalyzed hydrolysis of galactomannans. The method was advantageous, because obtained peaks were narrow and precise, and similar oligosaccharides with identical Mw could be separated as was shown for mannobiose (1.87 DE) and G (2.1 DE) (Figure 7). Some oligosaccharides were also found to co-elute, as was observed for mannotriose (2.85 DE) and MG (2.85 DE) (Figure 8). It is likely that separation would be possible if some of the data acquisition settings of the ABI 3730xl sequencer were changed, e.g. the temperature or the applied voltage. Another advantage of DASH was that the migration of a specific oligosaccharide was the same in every run after alignment of the electropherograms using the applied mobility markers. This precise and reproducible migration makes it possible to obtain libraries with known migration of specific oligosaccharides at specific acquisition settings. Such libraries could even be shared between laboratories that use the method. A further advantage is that oligosaccharides with high DP remain widely spaced in the electropherograms unlike in LC separations. Li et al. (2013) even show that they can detect dextran oligosaccharides with DP up to 35 by extending the run time to 90 min using DASH (Li et al. 2013) (in the current PhD study a run time of 27 min was used).

However, the DASH method also has some draw-backs. The signal strength measured by the ABI 3730xl sequencer was found to vary, both between samples in the same run (the same microtiter plate) and between different experiments (different microtiter plates run at different time spots). The observed variation seemed to be random and was not explained. The variation in signal made it impossible to quantitatively compare different samples without having an internal standard (or even standard row) in each sample. Furthermore, the variation made it difficult to "guess" the amount of labeled sample to analyze in order to obtain a response within the detection limits. Uncertainty over which concentration to use was a problem due to the long and labor-intensive sample preparation procedure, with long labelling and evaporation steps. The labelling reaction itself also introduced uncertainty, because this reaction can be difficult to control if all of a sample is derivatized, and especially if a sample contains an unknown concentration of oligosaccharides with different reducing end sugars that have different labelling efficiencies. In the work of this PhD study, only mannose was found as the reducing end sugar on the analyzed saccharides and the concentration of reducing ends in a sample was known. It was thus possible to keep the APTS concentration higher than the reducing ends to be analyzed (200 nmole APTS was used for 50 nmole reducing ends).

6.2.2 Kinetic measurements on MGGMM by MS

A novel, MS based method was developed to allow kinetic measurements on MGGMM. The results showed that MS in MRM mode can be used for studying endomannanase activity on oligosaccharides. Detection of product formation or substrate depletion by MS has several advantages compared to more traditional methods such as reducing end assays or HPAEC methods. For assessing kinetics on oligosaccharide like MGGMM, reducing end assays are bad alternatives because of the high background of the MGGMM substrate. The applied MS method allowed fast online analysis in real time, which means that the reaction is monitored with minimal disturbance as it proceeds. A negligible amount of sample volume (4 µl out of 500 µl) was drawn from the reaction at a given time point, by the MS autosampler, and was immediately quenched when injected into the flow path (mobile phase, pH 2.4). Another advantage was that the integrity of the sample was maintained since no artifacts such as heat or basic inactivation were introduced by extensive sample handling. If the MS had been coupled to an HPLC, the method could also have been descriptive in terms of product profiling (Perna et al. 2018). Other advantages of MS detection are high sensitivity, low substrate concentration requirement and low noise, particularly in the MRM mode. The MS approach opens for the possibility to study kinetics of other GHs on other oligosaccharides.

The applied MS method also had some drawbacks. When no separation is applied before detection, the method becomes sensitive to the ionic strength in the buffer or the reaction. This is due to ion suppression effects. In the present study 1 mM sodium acetate was used. It was not a problem to use low buffer concentration because neither MGGMM nor the products affect pH. However, to ensure that no change in pH was observed during the reaction, the pH of the reaction mixture was measured just before and after the MS measurements. In order to follow possible ion suppression and signal stability issues, the assay was performed with an internal standard. Another drawback of choosing to skip the HPLC separation before the MS detection was that no simultaneous information about hydrolysis products were obtained. The reason that the separation step was not included, was the increase in time span between injected samples, that this step would give. An increased time span between samples was unacceptable because we were interested in the initial part of the reaction before high levels of products, which could also be new substrates, had accumulated. Furthermore, testing chromatographic elution of MGGMM revealed an unexpected fragmentation pattern, different from the fragmentation pattern observed in the direct injection mode. The unexpected fragmentation pattern was not explained within the time frame of this PhD work.

7 Perspectives and outlook

This PhD study has investigated the boosting capacity of fungal endomannanases in softwood saccharification. Performance studies were carried out at 30 °C to test if the novel specificity of fungal GH26 endomannanases was beneficial in this application. The low melting temperature was chosen to ensure that the results obtained were caused by enzyme specificity rather than stability. However, despite low saccharification temperature, the utilization of fungal GH26 endomannanases in softwood saccharification was not very promising when compared to the yields obtained with *Tres*Man5A. Even the GH26 enzymes that carried a CBM1, demonstrated to be important for hydrolysis of cellulose-containing substrates (Hägglund et al. 2003) such as pretreated softwood, did not perform at the same level as *Tres*Man5A with its CBM1. It is likely that the open active site topology of the characterized fungal GH26 endomannanases is not optimal for attacking the galactoglucomannans when present on the cellulose microfibrils and/or more crystalline parts of the mannan, also because these regions may be less substituted.

The utilization of fungal GH26 endomannanases still has potential in other applications, particularly those using galactomannans. Interesting examples could be in the production of prebiotics (Jian et al. 2013), or partially hydrolyzed guar gum. Partially hydrolyzed guar gum is a supplemental dietary fiber used in some foods. This partially hydrolyzed fiber has the same thickening effect as guar gum but the produced galactomanno-oligosaccharides might also have health benefits (Li et al. 2017). It is possible that some of the galactomanno-oligosaccharides produced by hydrolysis of galactomannans by fungal GH26 endomannanase are particularly beneficial for human or animal health when compared to galactomannao-oligosaccharides produced by other endomannanases. The different binding modes and affinities for galactose in the -2 subsite, as was observed between *Ypen*Man26A and *Wsp.*Man26A, and the knowledge about the involved amino acids (**Paper III**), may be used for designing enzymes with altered product profiles. Galactomannans are found as additives in our foods so it is also likely that fungal GH26 endomannanases could be used as stain removers in the laundry detergent industry. However, such use in industrial applications may require protein engineering to increase the thermostability of the enzymes.

Our understanding of the specificity and the biological role of fungal GH26 endomannanases will be further increased if the specificity of the CBM35s on these enzymes is explored. This could be done by assessing the binding of different ligands to the isolated CBMs (Montanier et al. 2009) to avoid influence by the core domains. In addition, it would be beneficial to study the interactions and positions of the GH26 core module and the CBM35 in a liquid environment. For engineering of CBM35 domains to natural enzymes that do not carry a CBM, it is important to

know if the linker and/or the α-helix interacting with the CBM35 in *Pans*Man26A needs to be incorporated for full functionality of the binding domain. It is also highly important to confirm if the CBM35s on fungal GH26 endomannanases target the mannan backbone, and to explore when this binding domain is advantageous. Another important feature to understand is whether the binding domain itself or the α-helix (α9 in *Pans*Man26A) adds stability to the enzymes.

In order to use endomannanases in softwood saccharification, it will be advantageous to test the hypothesis that T*res*Man5A catalyzes a faster or more profound degradation of a certain type of mannan (maybe a more crystalline part that is more tightly intertwined with the cellulose) than the other endomannanases. Also, we need to understand why the enzyme is capable of attacking this part of the mannan. This is a complex task, given the complex substrate, the multiple enzymes working together, and the long reaction times. A way forward may be to employ pure ivory nut mannan, or the complex of ivory nut mannan and cellulose that can be found in the residue of real ivory nut mannan after removal of low molecular weight mannan by alkali treatment (Hägglund et al. 2003). Both the ivory nut mannan and the mannan/cellulose complex are insoluble, similar to a lignocellulosic matrix. Studies with such a mannan/cellulose complex have been carried out by Hägglund et al 2003 to assess the role of the CBM1 on *Tres*Man5A (Hägglund et al. 2003). However, it is not known to what extent this complex correlates with softwood substrates which most likely also differ depending on the type of softwood and pretreatment method. A cheaper option might be to adsorb mannan on Avicel as done by Wang et al (2017) (Wang et al. 2017). In this PhD study we used low solids to allow assessment of multiple enzymes on the limited amount of available substrate. To address industrial softwood saccharification it will also be relevant to test enzyme performance on higher solids, and in such trials inhibition might play an important role.

This study also highlights the problem of selecting enzymes for some industrial applications based on characterization on pure and well-defined substrates. Even though several enzymes were tested on several pure mannans, including extracted spruce galactoglucomannan, neither rates nor substrate preferences observed in the initial characterization could be correlated with the enzymes performance in softwood saccharification. If a large screening campaign were set up to measure the initial rate of multiple endomannanases e.g. on locust bean gum, and only the best performing candidates were tested in the longer and more expensive softwood saccharification trials, *Tres*Man5A would most likely be deselected before the saccharification trials. Again, this leads to the fact that we need to know which obstacle or obstacles are rate limiting. If it is difficult for the enzymes to attack the mannans when in contact with the cellulose microfibrils, this is ideally what such a screening assay should assess. In other applications than softwood saccharification, the simpler and pure mannan substrates may still be valuable as screening substrates.

8 Conclusion

Endomannanases catalyze the degradation of β-mannans which are an abundant class of plant polysaccharides. This class includes the galactomannans widely used in our food and feed industry and the acetylated galactoglucomannans which constitute up to 25 % of softwood dry matter. Endomannanases are attracting increased attention not only from a fundamental perspective but also due to their industrial importance in biomass conversion processes. Softwood has significant potential as feedstock for renewable energy production and biorefining due to its abundance and low cost. The high amount of galactoglucomannan in softwood, requires efficient mannan-degrading enzymes to unlock its full potential. The need for efficient endomannanases has become even more important as new pretreatment methods are emerging to maximize mannan recovery. To enhance the industrial performance of endomannanases, both for softwood saccharification as well as in other industries, it is important to understand the molecular background of their natural specificity and substrate interactions. This PhD study aimed at elucidating aspects of the natural specificity of selected fungal GH5 and GH26 endomannanases, but also to evaluate how these specificities affected enzyme performance in softwood saccharification.

One obstacle when using endomannanases in mannan conversion is that they are restricted by galactose substitutions on the mannan backbone, which limits their use on highly substituted galactomannans. This study identified fungal GH26 endomannanases with a novel degradation pattern on galactomannans and the capability to achieve full conversion of the highly substituted guar gum galactomannan. By analyzing product profiles from galactomannan hydrolysis using DASH and by structural assessment of the enzymes active site clefts, the GH26 endomannanases were found to accommodate galactopyranosyl moieties at least in the -2, -1 and +1 subsites, in contrast to the investigated GH5 endomannanases which allowed galactopyranosyl units in the -1 subsite but not in the -2 and +1 subsites. The fungal GH26 endomannanases were also found to have lower melting temperatures than the analyzed GH5 enzymes.

Endomannanases have been shown to boost saccharification of softwood to fermentable monomers. However, little is known about which specificities are valuable for catalyzing degradation of galactoglucomannans in the complex lignocellulosic matrix. By assessing the activity levels of ten fungal endomannanases on a variety of pure mannans (Figure 11) and their boosting effect on enzymatic saccharification of softwood (Figure 12 and Figure 13), it was found that their saccharification performance varied significantly (increasing glucose yield by 3 – 30 % after 24 h) but that this performance was not predicted by the initial rates on the pure mannans. The results emphasize that it is enzyme ability to act on the mannan in complex with the

lignocellulosic matrix that determines enzyme performance, rather than enzyme specificity towards galactoglucomannans as such. The best performing endomannanase in softwood saccharification was *Tres*Man5A from *Trichoderma reesei*. Based on the data, we hypothesized that *Tres*Man5A attacks a portion of mannan in the lignocellulosic matrix that is inaccessible for other endomannanases, which thus allows for better cellulose degradation. Both the catalytic efficiency of the core module and the presence of the CBM1 played important roles in the superior performance of *Tres*Man5A. These findings are highly relevant to the search for endomannanases for use in softwood saccharification. Yet even more important is the recognition of the importance of the employed screening assays and strategies.

The molecular background was assessed in order to exploit the novel specificity of fungal GH26 endomannanases. The substrate binding amino acids in the -4 to -2 subsites in one of these enzymes, *Ypen*Man26A from *Yunnania penicillata*, was identified by solving and analyzing its crystal structure in complex with MGG. The -2 subsite in particular was found to have multiple interactions with the galactopyranosyl unit, whereas the side group in the -3 subsite pointed out from the active site. Ligand binding amino acids in the active site cleft were highly conserved among the investigated GH26 endomannanases, which strongly indicates that the capability of accommodating multiple galactopyranosyl side-groups in the binding cleft is conserved among fungal enzymes in the GH26 family. However, a single amino acid substitution in the – 2 subsite, which loosens the interactions with the galactopyranosyl unit in this subsite, lowered enzyme affinity as well as the catalytic rate on highly substituted galactomannans. A more pronounced structural difference between the investigated GH26 endomannanases was found in the area of the core module approaching the CBM35. The enzymes carrying a CBM35 all seem to have an α-helix that permits ordered interactions with the binding domain, whereas the enzymes without a CBM had a flexible surface loop.

This research work of this PhD thesis has significantly contributed to the elucidation of fungal GH26 endomannanase specificity and to the understanding of endomannanase performance in softwood saccharification. The capability to accommodate galactopyranosyl moieties in the -3, -2, -1 and +1 subsites constitutes a novel endomannanase specificity. The results are based on the assessed enzymes and do not give information on all fungal GH5 and GH26 endomannanases found in nature. Since single mutations can cause changes in binding mode, affinity and likely also in the accommodation of galactopyranosyl units in a specific subsite, it is expected that fungal GH26 endomannanases with altered tolerance towards galactose substitutions exist. None the less, it is my belief that this novel endomannanase specificity can increase the potential use of endomannanases in industrial processes; this specificity provides a new tool for degradation of highly substituted galactomannans and production of other galactomanno-oligosaccharides than is possible with previously characterized endomannanases.

9 References

Ademark, P.; Lundqvist, J.; Hägglund, P.; Tenkanen, M.; Torto, N.; Tjerneld, F. and Stålbrand, H. (1999) Hydrolytic properties of a *β*-mannosidase purified from *Aspergillus niger*. *Journal of Biotechnology*, 75(2–3), pp.281–289.

Åkerholm, M. and Salmén, L. (2001) Interactions between wood polymers studied by dynamic FT-IR spectroscopy. *Polymer*, 42, pp.963–969.

Akita, M.; Takeda, N.; Hirasawa, K.; Sakai, H.; Kawamoto, M.; Yamamoto, M.; Grant, W.D.; Hatada, Y.; Ito, S. and Horikoshi, K. (2004) Crystallization and preliminary X-ray study of alkaline mannanase from an alkaliphilic *Bacillus* isolate. *Acta Crystallographica Section D: Biological crystallography*, 60(Pt 8), pp.1490–1492.

Al-Ghazzewi, F.H.; Khanna, S.; Tester, R.F. and Piggott, J. (2007) The potential use of hydrolysed konjac glucomannan as a prebiotic. *Journal of the Science of Food and Agriculture*, 87, pp.1758–1766.

Anderson, L. and Hunter, C.L. (2006) Quantitative mass spectrometric multiple reaction monitoring assays for major plasma proteins. *Molecular & Cellular Proteomics*, 5(4), pp.573–588.

Andersson, A.; Persson, T.; Zacchi, G.; Stålbrand, H. and Jönsson, A.S. (2007) Comparison of diafiltration and size-exclusion chromatography to recover hemicelluloses from process water from thermomechanical pulping of spruce. *Applied Biochemistry and Biotechnology*, 137, pp.971–983.

Asensio, J.L.; Ardá, A.; Cañada, F.J. and Jiménez-Barbero, J. (2013) Carbohydrate-aromatic interactions. *Accounts of Chemical Research*, 46(4), pp.946–954.

Asherie, N. (2004) Protein crystallization and phase diagrams. *Methods*, 34(3), pp.266–272.

Aspeborg, H.; Coutinho, P.M.; Wang, Y.; Brumer, H. and Henrissat, B. (2012) Evolution, substrate specificity and subfamily classification of glycoside hydrolase family 5 (GH5). *BMC evolutionary biology*, 12(186).

Aspinall, G.O.; Hirst, E.L.; Percival, E.G. V and Williamson, I.R. (1953) The mannans of ivory nut (*Phytelephas macrocarpa*). Part I. The methylation of mannan A and mannan B. *Journal of the Chemical Society*, (Part III), pp.3184–3188.

Bååth, J.A.; Abad, A.M.; Berglund, J.; Larsbrink, J.; Vilaplana, F. and Olsson, L. (2018) Mannanase hydrolysis of spruce galactoglucomannan focusing on the influence of acetylation on enzymatic mannan degradation. *Biotechnology for Biofuels*, 11, p.114.

Bågenholm, V.; Reddy, S.K.; Bouraoui, H.; Morrill, J.; Kulcinskaja, E.; Bahr, C.M.; Aurelius, O.; Rogers, T.; Xiao, Y.; Logan, D.T.; Martens, E.C.; Koropatkin, N.M. and Stålbrand, H. (2017) Galactomannan catabolism conferred by a polysaccharide utilization locus of *Bacteroides ovatus*. *The Journal of Biological Chemistry*, 292(1), pp.229–243.

Bayer, E.A.; Chanzy, H.; Lamed, R. and Shoham, Y. (1998) Cellulose, cellulases and cellulosomes. *Current Opinion in Structural Biology*, 8(5), pp.548–557.

Berlin, A.; Gilkes, N.; Kilburn, D.; Bura, R.; Markov, A.; Skomarovsky, A.; Okunev, O.; Gusakov, A.; Maximenko, V.; Gregg, D.; Sinitsyn, A. and Saddler, J. (2005) Evaluation of novel fungal cellulase preparations for ability to hydrolyze softwood substrates - Evidence for the role of accessory enzymes. *Enzyme and Microbial Technology*, 37(2), pp.175–184.

Bombeck, P.L.; Khatri, V.; Meddeb-Mouelhi, F.; Montplaisir, D.; Richel, A. and Beauregard, M. (2017) Predicting the most appropriate wood biomass for selected industrial applications: comparison of wood, pulping, and enzymatic treatments using fluorescent-tagged carbohydrate-binding modules. *Biotechnology for Biofuels*, 10, p.293.

Boraston, A.B.; Bolam, D.N.; Gilbert, H.J. and Davies, G.J. (2004) Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochemical Journal*, 382, pp.769–781.

Broekaert, W.F.; Courtin, C.M.; Verbeke, K.; van de Wiele, T.; Verstraete, W. and Delcour, J.A. (2011) Prebiotic and other health-related effects of cereal-derived arabinoxylans, arabinoxylan-oligosaccharides, and xylooligosaccharides. *Critical Reviews in Food Science and Nutrition*, 51, pp.178–194.

Buchert, J.; Salminen, J.; Siika-aho, M.; Ranua, M. and Viikari, L. (1993) The role of *Trichoderma reesei* xylanase and mannanase in the treatment of softwood kraft pulp prior to bleaching. *Holzforschung*, 47(6), pp.473–478.

Busch, A.; Kunert, G.; Heckel, D.G. and Pauchet, Y. (2017) Evolution and functional characterization of CAZymes belonging to subfamily 10 of glycoside hydrolase family 5 (GH5_10) in two species of phytophagous beetles. *PLoS ONE*, 12(8), pp.1–23.

Carpita, N.C. and Gibeaut, D.M. (1993) Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth. *The Plant Journal: for cell and molecular biology*, 3(1), pp.1–30.

Cartmell, A.; Topakas, E.; Ducros, V.M.A.; Suits, M.D.L.; Davies, G.J. and Gilbert, H.J. (2008) The *Cellvibrio japonicus* mannanase *Cj*Man26C displays a unique exo-mode of action that is conferred by subtle changes to the distal region of the active site. *Journal of Biological Chemistry*, 283(49), pp.34403–34413.

Chandra, R.P.; Chu, Q.L.; Hu, J.; Zhong, N.; Lin, M.; Lee, J.S. and Saddler, J. (2016) The influence of lignin on steam pretreatment and mechanical pulping of poplar to achieve high sugar recovery and ease of enzymatic hydrolysis. *Bioresource Technology*, 199, pp.135–141.

Chandra, R.P.; Gourlay, K.; Kim, C.S. and Saddler, J.N. (2015) Enhancing hemicellulose recovery and the enzymatic hydrolysis of cellulose by adding lignosulfonates during the two-stage steam pretreatment of poplar. *ACS Sustainable Chemistry and Engineering*, 3(5), pp.986–991.

Charalampopoulos, D.; Wang, R.; Pandiella, S.S. and Webb, C. (2002) Application of cereals and cereal components in functional foods: a review. *International Journal of Food Microbiology*, 79(1–2), pp.131–141.

Chauhan, P.S.; Puri, N.; Sharma, P. and Gupta, N. (2012) Mannanases: microbial sources, production, properties and potential biotechnological applications. *Applied Microbiology and Biotechnology*, 93(5), pp.1817–1830.

Chayen, N.E. (2005) Methods for separating nucleation and growth in protein crystallisation. *Progress in Biophysics and Molecular Biology*, 88(3), pp.329–337.

Cherubini, F. (2010) The biorefinery concept: Using biomass instead of oil for producing energy and chemicals. *Energy Conversion and Management*, 51(7), pp.1412–1421.

Correia, M.A.S.; Abbott, D.W.; Gloster, T.M.; Fernandes, V.O.; Prates, J.A.M.; Montanier, C.; Dumon, C.; Williamson, M.P.; Tunnicliffe, R.B.; Liu, Z.; Flint, J.E.; Davies, G.J.; Henrissat, B.; Coutinho, P.M.; Fontes, C.M.G.A. and Gilbert, H.J. (2010) Signature active site architectures illuminate the molecular basis for ligand specificity in family 35 carbohydrate binding module. *Biochemistry*, 49(29), pp.6193–6205.

Coutinho, P.M.; Andersen, M.R.; Kolenova, K.; Van Kuyk, P.A.; Benoit, I.; Gruben, B.S.; Trejo-Aguilar, B.; Visser, H.; van Solingen, P.; Pakula, T.; Seiboth, B.; Battaglia, E.; Aguilar-Osorio, G.; de Jong, J.F.; Ohm, R.A.; Aguilar, M.; Henrissat, B.; Nielsen, J.; Stålbrand, H. and de Vries, R.P. (2009) Post-genomic insights into the plant polysaccharide degradation potential of *Aspergillus nidulans* and comparison to *Aspergillus niger* and *Aspergillus oryzae*. *Fungal Genetics and Biology*, 46(1), pp.S161–S169.

Couturier, M.; Haon, M.; Coutinho, P.M.; Henrissat, B.; Lesage-Meessen, L. and Berrin, J. (2011) *Podospora anserina* hemicellulases potentiate the *Trichoderma reesei* secretome for saccharification of lignocellulosic biomass. *Applied and Environmental Microbiology*, 77(1), pp.237–246.

Couturier, M.; Roussel, A.; Rosengren, A.; Leone, P.; Stålbrand, H. and Berrin, J. (2013) Structural and biochemical analyses of glycoside hydrolase families 5 and 26 β-(1,4)-mannanases from *Podospora anserina* reveal differences upon manno-oligosaccharide catalysis. *The Journal of Biological Chemistry*, 288(20), pp.14624–14635.

Cox, J. and Mann, M. (2011) Quantitative, high-resolution proteomics for data-driven systems biology. *Annual Review of Biochemistry*, 80, pp.273–299.

Crestini, C.; Crucianelli, M.; Orlandi, M. and Saladino, R. (2010) Oxidative strategies in lignin chemistry: A new environmental friendly approach for the functionalisation of lignin and lignocellulosic fibers. *Catalysis Today*, 156, pp.8–22.

Davies, G. and Henrissat, B. (1995) Structures and mechanisms of glycosyl hydrolases. *Structure*, 3(9), pp.853–859.

Davies, G.J.; Wilson, K.S. and Henrissat, B. (1997) Nomenclature for sugar-binding subsites in glycosyl hydrolases. *Biochemical Journal*, 321, pp.557–559.

Dhawan, S. and Kaur, J. (2007) Microbial mannanases: An overview of production and applications. *Critical Reviews in Biotechnology*, 27, pp.197–216.

Dias, F.M. V; Vincent, F.; Pell, G.; Prates, J.A.M.; Centeno, M.S.J.; Tailford, L.E.; Ferreira, L.M.A.; Fontes, C.M.G.A.; Davies, G.J. and Gilbert, H.J. (2004) Insights into the molecular determinants of substrate specificity in glycoside hydrolase family 5 revealed by the crystal structure and kinetics of *Cellvibrio mixtus* mannosidase 5A. *The Journal of Biological Chemistry*, 279(24), pp.25517–25526.

Dilokpimol, A.; Nakai, H.; Gotfredsen, C.H.; Baumann, M.J.; Nakai, N.; Hachem, M.A. and Svensson, B. (2011) Recombinant production and characterisation of two related GH5 endo-β-1,4-mannanases from *Aspergillus nidulans* FGSC A4 showing distinctly different transglycosylation capacity. *Biochimica et Biophysica Acta*, 1814, pp.1720–1729.

Ducros, V.M.-A.; Zechel, D.L.; Murshudov, G.N.; Gilbert, H.J.; Szabó, L.; Stoll, D.; Withers, S.G. and Davies, G.J. (2002) Substrate distortion by a β-mannanase: Snapshots of the Michaelis and covalent-intermediate complexes suggest a $B_{(2,5)}$ conformation for the transition state. *Angewandte Chemie International Edition*, 41(15), pp.2824–2827.

Ebringerová, A. (2006) Structural diversity and application potential of hemicelluloses. *Macromolecular Symposia*, 232(1), pp.1–12.

Ebringerová, A. and Heinze, T. (2000) Xylan and xylan derivatives - Biopolymers with valuable properties, 1: Naturally occurring xylans structures, isolation procedures and properties. *Macromolecular Rapid Communications*, 21(9), pp.542–556.

Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, 32(5), pp.1792–1797.

Emsley, P.; Lohkamp, B.; Scott, W.G. and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallographica Section D: Biological Crystallography*, 66(4), pp.486–501.

Eriksson, Ö.; Goring, D.A.I. and Lindgren, B.O. (1980) Structural studies on the chemical bonds between lignins and carbohydrates in spruce wood. *Wood Science and Technology*, 14(4), pp.267–279.

Eriksson, T.; Börjesson, J. and Tjerneld, F. (2002) Mechanism of surfactant effect in enzymatic hydrolysis of lignocellulose. *Enzyme and Microbial Technology*, 31, pp.353–364.

Espiñeira, J.M.; Novo Uzal, E.; Gómez Ros, L. V.; Carrión, J.S.; Merino, F.; Ros Barceló, A. and Pomar, F. (2011) Distribution of lignin monomers and the evolution of lignification among lower plants. *Plant Biology*, 13(1), pp.59–68.

Evangelista, R.A.; Liu, M.S. and Chen, F.T.A. (1995) Characterization of 9-aminopyrene-1,4,6-trisulfonate-derivatized sugars by capillary electrophoresis with laser-induced fluorescence detection. *Analytical Chemistry*, 67(13), pp.2239–2245.

Fanutti, C.; Ponyi, T.; Black, G.W.; Hazlewood, G.P. and Gilbert, H.J. (1995) The conserved noncatalytic 40-residue sequence in cellulases and hemicellulases from anaerobic fungi functions as a protein docking domain. *The Journal of Biological Chemistry*, 270(49), pp.29314–29322.

Fry, S.C.; Nesselrode, B.H.W.A.; Miller, J.G. and Mewburn, B.R. (2008) Mixed-linkage (1–3,1–4)-β-D-glucan is a major hemicellulose of Equisetum (horsetail) cell walls. *New Phytologist*, 179(1), pp.104–115.

Fry, S.C.; York, W.S.; Albersheim, P.; Darvill, A.; Hayashi, T.; Joseleau, J.; Kato, Y.; Lorences, E.P.; Maclachlan, G.A.; McNeil, M.; Mort, A.J.; Reid, J.S.G.; Seitz, H.U.; Selvendran, R.R.; Voragen, A.G.J. and White, A.R. (1993) An unambiguous nomenclature for xyloglucan-derived oligosaccharides. *Physiologia Plantarum*, 89(1), pp.1–3.

Fujimoto, Z.; Kuno, A.; Kaneko, S.; Kobayashi, H.; Kusakabe, I. and Mizuno, H. (2002) Crystal structures of the sugar complexes of *Streptomyces olivaceoviridis* E-86 xylanase: Sugar binding structure of the family 13 carbohydrate binding module. *Journal of Molecular Biology*, 316(1), pp.65–78.

Gibson, L.J. (2012) The hierarchical structure and mechanics of plant materials. *Journal of the Royal Society Interface*, 9(76), pp.2749–2766.

Gilbert, H.J.; Knox, J.P. and Boraston, A.B. (2013) Advances in understanding the molecular basis of plant cell wall polysaccharide recognition by carbohydrate-binding modules. *Current Opinion in Structural Biology*, 23(5), pp.669–677.

Gilbert, H.J.; Stålbrand, H. and Brumer, H. (2008) How the walls come crumbling down: recent structural biochemistry of plant polysaccharide degradation. *Current opinion in plant biology*, 11(3), pp.338–48.

Gonçalves, A.M.D.; Silva, C.S.; Madeira, T.I.; Coelho, R.; de Sanctis, D.; San Romão, M. V and Bento, I. (2012) Endo-*β*-D-1,4-mannanase from *Chrysonilia sitophila* displays a novel loop arrangement for substrate selectivity. *Acta crystallographica Section D: Biological Crystallography*, 68, pp.1468–1478.

Goubet, F.; Jackson, P.; Deery, M.J. and Dupree, P. (2002) Polysaccharide analysis using carbohydrate gel electrophoresis: A method to study plant cell wall polysaccharides and polysaccharide hydrolases. *Analytical Biochemistry*, 300(1), pp.53–68.

Grantham, N.J.; Wurman-Rodrich, J.; Terrett, O.M.; Lyczakowski, J.J.; Stott, K.; Iuga, D.; Simmons, T.J.; Durand-Tardif, M.; Brown, S.P.; Dupree, R.; Busse-Wicher, M. and Dupree, P. (2017) An even pattern of xylan substitution is critical for interaction with cellulose in plant cell walls. *Nature Plants*, 3, pp.859–865.

Hägglund, P.; Eriksson, T.; Collén, A.; Nerinckx, W.; Claeyssens, M. and Stålbrand, H. (2003) A cellulose-binding module of the *Trichoderma reesei β*-mannanase Man5A increases the mannan-hydrolysis of complex substrates. *Journal of Biotechnology*, 101(1), pp.37–48.

Hekmat, O.; Lo Leggio, L.; Rosengren, A.; Kamarauskaite, J.; Kolenova, K. and Stålbrand, H. (2010) Rational engineering of mannosyl binding in the distal glycone subsites of *Cellulomonas fimi* endo-*β*-1,4-mannanase: Mannosyl binding promoted at subsite -2 and demoted at subsite -3. *Biochemistry*, 49(23), pp.4884–4896.

Henrissat, B. (1991) A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochemical Journal*, 280(2), pp.309–316.

Henrissat, B. and Bairoch, A. (1996) Updating the sequence-based classification of glycosyl hydrolases. *Biochemical Journal*, 316, pp.695–696.

Henrissat, B.; Callebaut, I.; Fabrega, S.; Lehn, P.; Mornon, J. and Davies, G. (1995) Conserved catalytic machinery and the prediction of a common fold for several families of glycosyl hydrolases. *Proceedings of the National Academy of Sciences of the USA*, 92(15), pp.7090–7094.

Henrissat, B. and Davies, G. (1997) Structural and sequence-based classification of glycoside hydrolases. *Current Opinion in Structural Biology*, 7(5), pp.637–644.

Henrissat, B.; Teeri, T.T. and Warren, R.A.J. (1998) A scheme for designating enzymes that hydrolyse the polysaccharides in the cell walls of plants. *FEBS letters*, 425(2), pp.352–354.

Hilge, M.; Gloor, S.M.; Rypniewski, W.; Sauer, O.; Heightman, T.D.; Zimmermann, W.; Winterhalter, K. and Piontek, K. (1998) High-resolution native and complex structures of thermostable *β*-mannanase from *Thermomonospora fusca* - susbstrate specificity in glycosyl hydrolase family 5. *Structure*, 6(11), pp.1433–1444.

de Hoffmann, E. and Stroobant, V. (2007) *Mass spectrometry principles and applications* Third edit., John Wiley & Sons Inc.

Hogg, D.; Pell, G.; Dupree, P.; Goubet, F.; Martín-Orúe, S.M.; Armand, S. and Gilbert, H.J. (2003) The modular architecture of *Cellvibrio japonicus* mannanases in glycoside hydrolase families 5 and 26 points to differences in their role in mannan degradation. *Biochemical Journal*, 371(3), pp.1027–1043.

Hogg, D.; Woo, E.J.; Bolam, D.N.; McKie, V.A.; Gilbert, H.J. and Pickersgill, R.W. (2001) Crystal structure of mannanase 26A from *Pseudomonas cellulosa* and analysis of residues involved in substrate binding. *Journal of Biological Chemistry*, 276(33), pp.31186–31192.

Inoue, H.; Yano, S. and Sawayama, S. (2015) Effect of *β*-mannanase and *β*-mannosidase supplementation on the total hydrolysis of softwood polysaccharides by the *Talaromyces cellulolyticus* cellulase system. *Applied Biochemistry and Biotechnology*, 176(6), pp.1673–1686.

Izydorczyk, M.S. and Biliaderis, C.G. (1995) Cereal arabinoxylans: advances in structure and physicochemical properties. *Carbohydrate Polymers*, 28(1), pp.33–48.

Jian, H.; Zhu, L.; Zhang, W.; Sun, D. and Jiang, J. (2013) Enzymatic production and characterization of manno-oligosaccharides from *Gleditsia sinensis* galactomannan gum. *International Journal of Biological Macromolecules*, 55, pp.282–288.

Jin, Y.; Petricevic, M.; John, A.; Jenkins, H.; Souza, L.P.D.; Cuskin, F.; Gilbert, H.J.; Rovira, C.; Goddard-borger, E.D.; Williams, S.J. and Davies, G.J. (2016) A β-mannanase with a lysozyme-like fold and a novel molecular catalytic mechanism. *American Chemical Society Publications*, 2, pp.896–903.

Jørgensen, H.; Sanadi, A.R.; Felby, C.; Lange, N.E.K.; Fischer, M. and Ernst, S. (2010) Production of ethanol and feed by high dry matter hydrolysis and fermentation of palm kernel press cake. *Applied Biochemistry and Biotechnology*, 161(1–8), pp.318–332.

Katsimpouras, C.; Dimarogona, M. and Petropoulos, P. (2016) A thermostable GH26 endo-β-mannanase from *Myceliophthora thermophila* capable of enhancing lignocellulose degradation. *Applied Microbiology and Biotechnology*, 100, pp.8385–8397.

Katsuraya, K.; Okuyama, K.; Hatanaka, K.; Oshima, R.; Sato, T. and Matsuzaki, K. (2003) Constitution of konjac glucomannan: chemical analysis and 13C NMR spectroscopy. *Carbohydrate Polymers*, 53(2), pp.183–189.

Khatri, V.; Meddeb-Mouelhi, F. and Beauregard, M. (2018) New insights into the enzymatic hydrolysis of lignocellulosic polymers by using fluorescent tagged carbohydrate-binding modules. *Sustainable Energy & Fuels*, 2, pp.479–491.

Kim, M.-K.; An, Y.J.; Song, J.M.; Jeong, C.-S.; Kang, M.H.; Kwon, K.K.; Lee, Y.-H. and Cha, S.-S. (2014) Structure-based investigation into the functional roles of the extended loop and substrate-recognition sites in an endo-β-1,4-D-mannanase from the Antarctic springtail, *Cryptopygus antarcticus*. *Proteins: Structure, Function and Bioinformatics*, 82(11), pp.3217–3223.

Kishida, N.; Okimasu, S. and Kamata, T. (1978) Molecular weight and intrinsic viscosity of konjac gluco-mannan. *Agricultural and Biological Chemistry*, 42(9), pp.1645–1650.

Klemm, D.; Heublein, B.; Fink, H.P. and Bohn, A. (2005) Cellulose: Fascinating biopolymer and sustainable raw material. *Angewandte Chemie - International Edition*, 44, pp.3358–3393.

Kök, M.S. (2010) Rheological study of galactomannan depolymerisation at elevated temperatures: Effect of varying pH and addition of antioxidants. *Carbohydrate Polymers*, 81(3), pp.567–571.

Kulcinskaja, E.; Rosengren, A.; Ibrahim, R.; Kolenová, K. and Stålbrand, H. (2013) Expression and characterization of a *Bifidobacterium adolescentis* β-mannanase carrying mannan-binding and cell association motifs. *Applied and Environmental Microbiology*, 79(1), pp.133–140.

Larsson, A.M.; Anderson, L.; Xu, B.; Muñoz, I.G.; Usón, I.; Janson, J.; Stålbrand, H. and Ståhlberg, J. (2006) Three-dimensional crystal structure and enzymic characterization of β-mannanase Man5A from blue mussel *Mytilus edulis*. *Journal of Molecular Biology*, 357(5), pp.1500–1510.

Li, X.; Jackson, P.; Rubtsov, D. V; Faria-Blanc, N.; Mortimer, J.C.; Turner, S.R.; Krogh, K.B.; Johansen, K.S. and Dupree, P. (2013) Development and application of a high throughput carbohydrate profiling technique for analyzing plant cell wall polysaccharides and carbohydrate active enzymes. *Biotechnology for Biofuels*, 6, p.94.

Li, Y.; Yi, P.; Wang, N.; Liu, J.; Liu, X.; Yan, Q. and Jiang, Z. (2017) High level expression of β-mannanase (*Rm*Man5A) in *Pichia pastoris* for partially hydrolyzed guar gum production. *International Journal of Biological Macromolecules*, 105, pp.1171–1179.

Lombard, V.; Ramulu, H.G.; Drula, E.; Coutinho, P.M. and Henrissat, B. (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Research*, 42, pp.D490–D495.

Lundqvist, J.; Jacobs, A.; Palm, M.; Zacchi, G.; Dahlman, O. and Stålbrand, H. (2003) Characterization of galactoglucomannan extracted from spruce (*Picea abies*) by heat-fractionation at different conditions. *Carbohydrate Polymers*, 51(2), pp.203–211.

Lundqvist, J.; Teleman, A.; Junel, L.; Zacchi, G.; Dahlman, O.; Tjerneld, F. and Stålbrand, H. (2002) Isolation and characterization of galactoglucomannan from spruce (*Picea abies*). *Carbohydrate Polymers*, 48(1), pp.29–39.

Malgas, S.; van Dyk, J.S. and Pletschke, B.I. (2015a) A review of the enzymatic hydrolysis of mannans and synergistic interactions between β-mannanase, β-mannosidase and α-galactosidase. *World Journal of Microbiology and Biotechnology*, 31(8), pp.1167–1175.

Malgas, S.; van Dyk, J.S. and Pletschke, B.I. (2015b) β-Mannanase (Man26A) and α-galactosidase (Aga27A) synergism – A key factor for the hydrolysis of galactomannan substrates. *Enzyme and Microbial Technology*, 70, pp.1–8.

Marchetti, R.; Berrin, J.; Couturier, M.; Qader, S.A.U.; Molinaro, A. and Silipo, A. (2016) NMR analysis of the binding mode of two fungal endo-β-1,4-mannanases from GH5 and GH26 families. *Organic & Biomolecular Chemistry*, 14, pp.314–322.

McCleary, B. V (1979) Modes of action of β-mannanase enzymes of diverse origin on legume seed galactomannans. *Phytochemistry*, 18(5), pp.757–763.

McCleary, B. V (1985) The fine structures of carob and guar galactomannans. *Carbohydrate Research*, 139, pp.237–260.

McCleary, B. V and Matheson, N.K. (1983) Action patterns and substrate-binding requirements of β-D-mannanase with mannosaccharides and mannan-type polysaccharides. *Carbohydrate Research*, 119, pp.191–219.

McCleary, B. V; Nurthen, E.; Taravel, F.R. and Joseleau, J. (1983) Characterisation of the oligosaccharides produced on hydrolysis of galactomannan with β-D-mannanase. *Carbohydrate Research*, 118, pp.91–109.

Mikkelson, A.; Maaheimo, H. and Hakala, T.K. (2013) Hydrolysis of konjac glucomannan by *Trichoderma reesei* mannanase and endoglucanases Cel7B and Cel5A for the production of glucomannooligosaccharides. *Carbohydrate Research*, 372, pp.60–68.

Millward-Sadler, S.J.; Hall, J.; Black, G.W.; Hazlewood, G.P. and Gilbert, H.J. (1996) Evidence that the *Piromyces* gene family encoding endo-1,4-mannanases arose through gene duplication. *FEMS Microbiology Letters*, 141(2–3), pp.183–188.

Mizutani, K.; Tsuchiya, S.; Toyoda, M.; Nanbu, Y.; Tominaga, K.; Yuasa, K.; Takahashi, N.; Tsuji, A. and Mikami, B. (2012) Structure of β-1,4-mannanase from the common sea hare *Aplysia kurodai* at 1.05 Å resolution. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications*, 68(10), pp.1164–1168.

Montanier, C.; van Bueren, A.L.; Dumon, C.; Flint, J.E.; Correia, M.A.; Prates, J.A.; Firbank, S.J.; Lewis, R.J.; Grondin, G.G.; Ghinet, M.G.; Gloster, T.M.; Herve, C.; Knox, J.P.; Talbot, B.G.; Turkenburg, J.P.; Kerovuo, J.; Brzezinski, R.; Fontes, C.M.G.A.; Davies, G.J.; Boraston, A.B. and Gilbert, H.J. (2009) Evidence that family 35 carbohydrate binding modules display conserved specificity but divergent function. *Proceedings of the National Academy of Sciences of the USA*, 106(9), pp.3065–70.

Moreira, L.R.S. and Filho, E.X.F. (2008) An overview of mannan structure and mannan-degrading enzyme systems. *Applied Microbiology and Biotechnology*, 79, pp.165–78.

Morrill, J.; Månberger, A.; Rosengren, A.; Naidjonoka, P.; von Freiesleben, P.; Krogh, K.B.R.M.; Bergquist, K.E.; Nylander, T.; Karlsson, E.N.; Adlercreutz, P. and Stålbrand, H. (2018) β-mannanase-catalyzed synthesis of alkyl mannooligosides. *Applied Microbiology and Biotechnology*, 102(12), pp.5149–5163.

Mortimer, J.C.; Faria-Blanc, N.; Yu, X.; Tryfona, T.; Sorieul, M.; Ng, Y.Z.; Zhang, Z.; Stott, K.; Anders, N. and Dupree, P. (2015) An unusual xylan in *Arabidopsis* primary cell walls is synthesised by GUX3, IRX9L, IRX10L and IRX14. *The Plant Journal*, 83(3), pp.413–426.

Murshudov, G.N.; Skubák, P.; Lebedev, A.A.; Pannu, N.S.; Steiner, R.A.; Nicholls, R.A.; Winn, M.D.; Long, F. and Vagin, A.A. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallographica Section D: Biological Crystallography*, 67(4), pp.355–367.

Nadaroglu, H.; Adiguzel, A. and Adiguzel, G. (2015) Purification and characterisation of β-mannanase from *Lactobacillus plantarum* (M24) and its applications in some fruit juices. *International Journal of Food Science and Technology*, 50(5), pp.1158–1165.

Le Nours, J.; Anderson, L.; Stoll, D.; Stålbrand, H. and Lo Leggio, L. (2005) The structure and characterization of a modular endo-β-1,4-mannanase from *Cellulomonas fimi*. *Biochemistry*, 44(38), pp.12700–12708.

Oh, Y.H.; Eom, I.Y.; Joo, J.C.; Yu, J.H.; Song, B.K.; Lee, S.H.; Hong, S.H. and Park, S.J. (2015) Recent advances in development of biomass pretreatment technologies used in biorefinery for the production of bio-based fuels, chemicals and polymers. *Korean Journal of Chemical Engineering*, 32(10), pp.1945–1959.

Ono, Y.; Saito, T. and Isogai, A. (2017) Branched structures of softwood celluloses: proof based on size-exclusion chromatography and multi-angle laser-light scattering. In *Nanocelluloses: Their Preparation, Properties, and Applications*. Washington: American Chemical Society, pp. 151–169.

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

Pell, G.; Szabo, L.; Charnock, S.J.; Xie, H.; Gloster, T.M.; Davies, G.J. and Gilbert, H.J. (2004) Structural and biochemical analysis of *Cellvibrio japonicus* xylanase 10C: How variation in substrate-binding cleft influences the catalytic profile of family GH-10 xylanases. *Journal of Biological Chemistry*, 279(12), pp.11777–11788.

Perna, V.; Agger, J.W.; Holck, J. and Meyer, A.S. (2018) Multiple Reaction Monitoring for quantitative laccase kinetics by LC-MS. *Scientific Reports*, 8(1), pp.1–9.

Petkun, S.; Grinberg, I.R.; Lamed, R.; Jindou, S.; Burstein, T.; Yaniv, O.; Shoham, Y.; Shimon, L.J.W.; Bayer, E.A. and Frolow, F. (2015) Reassembly and co-crystallization of a family 9 processive endoglucanase from its component parts: structural and functional significance of the intermodular linker. *PeerJ*, 3, p.e1126.

Pham, T.A.; Berrin, J.G.; Record, E.; To, K.A. and Sigoillot, J.C. (2010) Hydrolysis of softwood by *Aspergillus* mannanase: Role of a carbohydrate-binding module. *Journal of Biotechnology*, 148, pp.163–170.

Ragauskas, A.J.; Beckham, G.T.; Biddy, M.J.; Chandra, R.; Chen, F.; Davis, M.F.; Davison, B.H.; Dixon, R.A.; Gilna, P.; Keller, M.; Langan, P.; Naskar, A.K.; Saddler, J.N.; Tschaplinski, T.J.; Tuskan, G.A. and Wyman, C.E. (2014) Lignin valorization: Improving lignin processing in the biorefinery. *Science*, 344(6185).

Rahar, S.; Swami, G.; Nagpal, N.; Nagpal, M.A. and Singh, G.S. (2011) Preparation, characterization, and biological properties of β-glucans. *Journal of Advanced Pharmaceutical Technology & Research*, 2(2), p.94.

Raman, R.; Venkataraman, M.; Ramakrishnan, S.; Lang, W.; Raguram, S. and Sasisekharan, R. (2006) Advancing glycomics: Implementation strategies at the consortium for functional glycomics. *Glycobiology*, 16(5), p.82R–90R.

Ratcliffe, I.; Williams, P.A.; Viebke, C. and Meadows, J. (2005) Physicochemical characterization of konjac glucomannan. *Biomacromolecules*, 6(4), pp.1977–1986.

Rättö, M.; Siika-aho, M.; Buchert, J.; Valkeajävi, A. and Viikari, L. (1993) Enzymatic hydrolosis of isolated and fibre-bound galactoglucomannans from pine-wood and pine kraft pulp. *Applied Microbiology and Biotechnology*, 40, pp.449–454.

Receveur, V.; Czjzek, M.; Schülein, M.; Panine, P. and Henrissat, B. (2002) Dimension, shape, and conformational flexibility of a two domain fungal cellulase in solution probed by small angle X-ray scattering. *Journal of Biological Chemistry*, 277(43), pp.40887–40892.

Reddy, S.K.; Bågenholm, V.; Pudlo, N.A.; Bouraoui, H.; Koropatkin, N.M.; Martens, E.C. and Stålbrand, H. (2016) A β-mannan utilization locus in *Bacteroides ovatus* involves a GH36 α-galactosidase active on galactomannans. *FEBS Letters*, 590, pp.2106–2118.

Reddy, S.K.; Rosengren, A.; Klaubauf, S.; Kulkarni, T.; Karlsson, E.N.; De Vries, R.P. and Stålbrand, H. (2013) Phylogenetic analysis and substrate specificity of GH2 β-mannosidases from Aspergillus species. *FEBS Letters*, 587(21), pp.3444–3449.

Robert, X. and Gouet, P. (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Research*, 42, pp.320–324.

Rosengren, A.; Hägglund, P.; Anderson, L.; Pavon-Orozco, P.; Peterson-Wulff, R.; Nerinckx, W. and Stålbrand, H. (2012) The role of subsite +2 of the *Trichoderma reesei* β-mannanase *Tr*Man5A in hydrolysis and transglycosylation. *Biocatalysis and Biotransformation*, 30(3), pp.338–352.

Rosengren, A.; Reddy, S.K.; Sjöberg, J.S.; Aurelius, O.; Logan, D.T.; Kolenová, K. and Stålbrand, H. (2014) An *Aspergillus nidulans* β-mannanase with high transglycosylation capacity revealed through comparative studies within glycosidase family 5. *Applied Microbiology and Biotechnology*, 98(24), pp.10091–104.

Rupp, B. (2009) *Biomolecular crystallography. Principles, Practice, and Application to Structural Biology* 1st ed., Garland Science. Taylor and Francis Group.

Sabini, E.; Schubert, H.; Murshudov, G.; Wilson, K.S.; Siika-Aho, M. and Penttilä, M. (2000) The three-dimensional structure of a *Trichoderma reesei* β-mannanase from glycoside hydrolase family 5. *Acta Crystallographica Section D Biological Crystallography*, 56(1), pp.3–13.

Sachslehner, A.; Foidl, G.; Foidl, N.; Gübitz, G. and Haltrich, D. (2000) Hydrolysis of isolated coffee mannan and coffee extract by mannanases of *Sclerotium rolfsii*. *Journal of Biotechnology*, 80(2), pp.127–34.

Sakai, K.; Kimoto, S.; Shinzawa, Y.; Minezawa, M.; Suzuki, K.; Jindou, S.; Kato, M. and Shimizu, M. (2018) Characterization of pH-tolerant and thermostable GH 134 β-1,4-mannanase *Ss*GH134 possessing carbohydrate binding module 10 from *Streptomyces sp*. NRRL B-24484. *Journal of Bioscience and Bioengineering*, 125(3), pp.346–352.

Sakai, K.; Mochizuki, M.; Yamada, M.; Shinzawa, Y.; Minezawa, M.; Kimoto, S.; Murata, S.; Kaneko, Y.; Ishihara, S.; Jindou, S.; Kobayashi, T.; Kato, M. and Shimizu, M. (2017) Biochemical characterization of thermostable β-1,4-mannanase belonging to the glycoside hydrolase family 134 from *Aspergillus oryzae*. *Applied Microbiology and Biotechnology*, 101(8), pp.3237–3245.

Scheller, H. V and Ulvskov, P. (2010) Hemicelluloses. *Annual Review of Plant Biology*, 61, pp.263–289.

Schröder, R.; Atkinson, R.G. and Redgwell, R.J. (2009) Re-interpreting the role of endo-β-mannanases as mannan endotransglycosylase/hydrolases in the plant cell wall. *Annals of botany*, 104(2), pp.197–204.

Shaw Stewart, P.D.; Kolek, S.A.; Briggs, R.A.; Chayen, N.E. and Baldock, P.F.M. (2011) Random microseeding: A theoretical and practical exploration of seed stability and seeding techniques for successful protein crystallization. *Crystal Growth and Design*, 11(8), pp.3432–3441.

Shimizu, M.; Kaneko, Y.; Ishihara, S.; Mochizuki, M.; Sakai, K.; Yamada, M.; Murata, S.; Itoh, E.; Yamamoto, T.; Sugimura, Y.; Hirano, T.; Takaya, N.; Kobayashi, T. and Kato, M. (2015) Novel β-1,4-mannanase belonging to a new glycoside hydrolase family in *Aspergillus nidulans*. *Journal of Biological Chemistry* , 290(46), pp.27914–27927.

Sinnott, M.L. (1990) Catalytic mechanisms of enzymic glycosyl transfer. *Chemical Reviews*, 90(7), pp.1171–1202.

Sørensen, T.H.; Cruys-Bagger, N.; Windahl, M.S.; Badino, S.F.; Borch, K. and Westh, P. (2015) Temperature effects on kinetic parameters and substrate affinity of Cel7A cellobiohydrolases. *Journal of Biological Chemistry*, 290(36), pp.22193–22202.

Stålbrand, H.; Saloheimo, A.; Vehmaanperä, J.; Henrissat, B. and Penttilä, M. (1995) Cloning and expression in *Saccharomyces cerevisiae* of a *Trichoderma reesei* β-mannanase gene containing a cellulose binding domain. *Applied and Environmental Microbiology*, 61(3), pp.1090–1097.

Stålbrand, H.; Siika-aho, M.; Tenkanen, M. and Viikari, L. (1993) Purification and characterization of two β-mannanases from *Trichoderma reesei*. *Journal of Biotechnology*, 29, pp.229–242.

Sunna, A. (2010) Modular organisation and functional analysis of dissected modular β-mannanase *Cs*Man26 from *Caldicellulosiruptor* Rt8B.4. *Applied Microbiology and Biotechnology*, 86(1), pp.189–200.

Tailford, L.E.; Ducros, V.M.A.; Flint, J.E.; Roberts, S.M.; Morland, C.; Zechel, D.L.; Smith, N.; Bjørnvad, M.E.; Borchert, T. V; Wilson, K.S.; Davies, G.J. and Gilbert, H.J. (2009) Understanding how diverse β-mannanases recognize heterogeneous substrates. *Biochemistry*, 48(29), pp.7009–7018.

Tailford, L.E.; Offen, W.A.; Smith, N.L.; Dumon, C.; Morland, C.; Gratien, J.; Heck, M.P.; Stick, R. V.; Blériot, Y.; Vasella, A.; Gilbert, H.J. and Davies, G.J. (2008) Structural and biochemical evidence for a boat-like transition state in β-mannosidases. *Nature Chemical Biology*, 4(5), pp.306–312.

Teleman, A.; Lundqvist, J.; Tjerneld, F.; Stålbrand, H. and Dahlman, O. (2000) Characterization of acetylated 4-O-methylglucuronoxylan isolated from aspen employing 1H and13C NMR spectroscopy. *Carbohydrate Research*, 329(4), pp.807–815.

Tenkanen, M.; Makkonen, M.; Perttula, M.; Viikari, L. and Teleman, A. (1997) Action of *Trichoderma reesei* mannanase on galactoglucomannan in pine kraft pulp. *Journal of Biotechnology*, 57(1–3), pp.191–204.

Tenkanen, M.; Puls, J.; Rättö, M. and Viikari, L. (1993) Enzymatic deacetylation of galactoglucomannans. *Applied Microbiology and Biotechnology*, 39(2), pp.159–165.

Thorsheim, K.; Siegbahn, A.; Johnsson, R.E.; Stålbrand, H.; Manner, S.; Widmalm, G. and Ellervik, U. (2015) Chemistry of xylopyranosides. *Carbohydrate Research*, 418, pp.65–88.

Timell, T.E. (1967) Recent progress in the chemistry of wood hemicelluloses. *Wood Science and Technology*, 1, pp.45–70.

Timell, T.E. (1957) Vegetable ivory as source of mannan polysaccharide. *Canadian journal of chemistry*, 35, pp.333–338.

Tsukagoshi, H.; Nakamura, A.; Ishida, T.; Touhara, K.K.; Otagiri, M.; Moriya, S.; Samejima, M.; Igarashi, K.; Fushinobu, S.; Kitamoto, K. and Arioka, M. (2014) Structural and biochemical analyses of glycoside hydrolase family 26 β-mannanase from a symbiotic protist of the termite *Reticulitermes speratus*. *The Journal of Biological Chemistry*, 289(15), pp.10843–10852.

Ueda, M.; Hirano, Y.; Fukuhara, H.; Naka, Y.; Nakazawa, M.; Sakamoto, T.; Ogata, Y. and Tamada, T. (2018) Gene cloning, expression, and X-ray crystallographic analysis of a β-mannanase from *Eisenia fetida*. *Enzyme and Microbial Technology*, 117(May), pp.15–22.

Vanholme, R.; Demedts, B.; Morreel, K.; Ralph, J. and Boerjan, W. (2010) Lignin biosynthesis and structure. *Plant Physiology*, 153(3), pp.895–905.

Várnai, A.; Huikko, L.; Pere, J.; Siika-Aho, M. and Viikari, L. (2011) Synergistic action of xylanase and mannanase improves the total hydrolysis of softwood. *Bioresource Technology*, 102(19), pp.9096–9104.

Várnai, A.; Siika-aho, M. and Viikari, L. (2013) Carbohydrate-binding modules (CBMs) revisited: reduced amount of water counterbalances the need for CBMs. *Biotechnology for biofuels*, 6(30).

Vinzant, T.B.; Adney, W.S.; Decker, S.R.; Baker, J.O.; Kinter, M.T.; Sherman, N.E.; Fox, J.W. and Himmel, M.E. (2001) Fingerprinting *Trichoderma reesei* hydrolases in a commercial cellulase preparation. *Applied Biochemistry and Biotechnology*, 91–93, pp.99–107.

Wang, X.; Li, K.; Yang, M.; Wang, J. and Zhang, J. (2017) Hydrolyzability of mannan after adsorption on cellulose. *Cellulose*, 24(1), pp.35–47.

Wang, Y.; Azhar, S.; Gandini, R.; Divne, C.; Ezcurra, I. and Aspeborg, H. (2015) Biochemical characterization of the novel endo-β-mannanase *At*Man5-2 from *Arabidopsis thaliana*. *Plant Science*, 241, pp.151–163.

Wang, Y.; Vilaplana, F.; Brumer, H. and Aspeborg, H. (2014) Enzymatic characterization of a glycoside hydrolase family 5 subfamily 7 (GH5_7) mannanase from *Arabidopsis thaliana*. *Planta*, 239(3), pp.653–665.

Westh, P.; Borch, K.; Sørensen, T.; Tokin, R.; Kari, J.; Badino, S.; Cavaleiro, M.A.; Røjel, N.; Christensen, S.; Vesterager, C.S. and Schiano-di-Cola, C. (2018) Thermoactivation of a cellobiohydrolase. *Biotechnology and Bioengineering*, 115, pp.831–838.

Willför, S.; Sjöholm, R.; Laine, C.; Roslund, M.; Hemming, J. and Holmbom, B. (2003) Characterisation of water - soluble galactoglucomannans from Norway spruce wood and thermomechanical pulp. *Carbohydrate Polymers*, 52, pp.175–187.

Williamson, M. (2012) *How Proteins Work*, Garland Science. Taylor and Francis Group, LLC, an informa business.

Winn, M.D.; Ballard, C.C.; Cowtan, K.D.; Dodson, E.J.; Emsley, P.; Evans, P.R.; Keegan, R.M.; Krissinel, E.B.; Leslie, A.G.W.; McCoy, A.; McNicholas, S.J.; Murshudov, G.N.; Pannu, N.S.; Potterton, E.A.; Powell, H.R.; Read, R.J.; Vagin, A. and Wilson, K.S. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallographica Section D: Biological Crystallography*, 67(4), pp.235–242.

Withers, S.G. (2001) Mechanisms of glycosyl transferases and hydrolases. *Carbohydrate Polymers*, 44, pp.325–337.

Wlodawer, A.; Minor, W.; Dauter, Z. and Jaskolski, M. (2008) Protein crystallography for non-crystallographers, or how to get the best (but not more) from published macromolecular structures. *The FEBS Journal*, 275(1), pp.1–21.

Wolfenden, R. and Snider, M.J. (2001) The depth of chemical time and the power of enzymes as catalysts. *Accounts of Chemical Research*, 34, pp.938–945.

Wolfrom, M.L.; Laver, M.L. and Patin, D.L. (1961) Carbohydrates of the coffee bean II. Isolation and characterization of a mannan. *The Journal of Organic Chemistry*, 26(11), pp.4533–4535.

Xia, W.; Lu, H.; Xia, M.; Cui, Y.; Bai, Y.; Qian, L.; Shi, P.; Luo, H. and Yao, B. (2016) A novel glycoside hydrolase family 113 endo-β-1,4-mannanase from *Alicyclobacillus sp.* strain A4 and insight into the substrate recognition and catalytic mechanism of this family. *Applied and Environmental Microbiology*, 82(9), pp.2718–2727.

Xu, C.; Leppänen, A.S.; Eklund, P.; Holmlund, P.; Sjöholm, R.; Sundberg, K. and Willför, S. (2010) Acetylation and characterization of spruce (*Picea abies*) galactoglucomannans. *Carbohydrate Research*, 345(6), pp.810–816.

Yamabhai, M.; Sak-Ubol, S.; Srila, W. and Haltrich, D. (2016) Mannan biotechnology: from biofuels to health. *Critical Reviews in Biotechnology*, 36(1), pp.32–42.

You, X.; Qin, Z.; Li, Y.-X.; Yan, Q.-J.; Li, B. and Jiang, Z.-Q. (2018) Structural and biochemical insights into the substrate-binding mechanism of a novel glycoside hydrolase family 134 β-mannanase. *Biochimica et Biophysica Acta - General Subjects*, 1862(6), pp.1376–1388.

You, X.; Qin, Z.; Yan, Q.; Yang, S.; Li, Y. and Jiang, Z. (2018) Structural insights into the catalytic mechanism of a novel glycoside hydrolase family 113 β-1,4-mannanase from *Amphibacillus xylanus*. *Journal of Biological Chemistry*, 293(30), pp.11746–11757.

Zechel, D.L. and Withers, S.G. (2000) Glycosidase mechanisms:Anatomy of a finely tuned catalyst. *Acc. Chem. Res.*, 33(1), pp.11–18.

Zhang, X.; Rogowski, A.; Zhao, L.; Hahn, M.G.; Avci, U.; Knox, J.P. and Gilbert, H.J. (2014) Understanding how the complex molecular architecture of mannan-degrading hydrolases contributes to plant cell wall degradation. *The Journal of Biological Chemistry*, 289(4), pp.2002–2012.

Zhang, Y.; Ju, J.; Peng, H.; Gao, F.; Zhou, C.; Zeng, Y.; Xue, Y.; Li, Y.; Henrissat, B.; Gao, G.F. and Ma, Y. (2008) Biochemical and structural characterization of the intracellular mannanase *Aa*ManA of *Alicyclobacillus acidocaldarius* reveals a novel glycoside hydrolase family belonging to Clan GH-A. *The Journal of Biological Chemistry*, 283(46), pp.31551–31558.

van Zyl, W.H.; Rose, S.H.; Trollope, K. and Görgens, J.F. (2010) Fungal β-mannanases: Mannan hydrolysis, heterologous production and biotechnological applications. *Process Biochemistry*, 45(8), pp.1203–1213.

PAPERS

Paper I

An *Aspergillus nidulans* GH26 endo-*β*-mannanase with a novel degradation pattern on highly substituted galactomannans

Enzyme and Microbial Technology 83 (2016) 68–77

Contents lists available at ScienceDirect

Enzyme and Microbial Technology

journal homepage: www.elsevier.com/locate/emt

An *Aspergillus nidulans* GH26 endo-β-mannanase with a novel degradation pattern on highly substituted galactomannans

Pernille von Freiesleben[a,b], Nikolaj Spodsberg[a], Thomas Holberg Blicher[a], Lars Anderson[a], Henning Jørgensen[b], Henrik Stålbrand[c], Anne S. Meyer[b,*], Kristian B.R.M. Krogh[a]

[a] Novozymes A/S, Krogshøjvej 36, 2880 Bagsværd, Denmark
[b] Center for Bioprocess Engineering, Department of Chemical and Biochemical Engineering, Building 229, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark
[c] Department of Biochemistry and Structural Biology, Center for Molecular Protein Science, Lund University, PO Box 124, SE-221 00 Lund, Sweden

ARTICLE INFO

Article history:
Received 10 July 2015
Received in revised form 4 October 2015
Accepted 31 October 2015
Available online 4 November 2015

Keywords:
Endo-β-(1 → 4)-mannanase
Galactomannan
DASH
Substrate degradation pattern
Glycoside hydrolase family

ABSTRACT

The activity and substrate degradation pattern of a novel *Aspergillus nidulans* GH26 endo-β-mannanase (*An*Man26A) was investigated using two galactomannan substrates with varying amounts of galactopyranosyl residues. The *An*Man26A was characterized in parallel with the GH26 endomannanase from *Podospora anserina* (*Pa*Man26A) and three GH5 endomannanases from *A. nidulans* and *Trichoderma reesei* (*An*Man5A, *An*Man5C and *Tr*Man5A). The initial rates and the maximal degree of enzymatically catalyzed conversion of locust bean gum and guar gum galactomannans were determined. The hydrolysis product profile at maximal degree of conversion was determined using DNA sequencer-Assisted Saccharide analysis in High throughput (DASH). This is the first reported use of this method for analyzing galactomannooligosaccharides. *An*Man26A and *Pa*Man26A were found to have a novel substrate degradation pattern on the two galactomannan substrates. On the highly substituted guar gum *An*Man26A and *Pa*Man26A reached 35–40% as their maximal degree of conversion whereas the three tested GH5 endomannanases only reached 8–10% as their maximal degree of conversion. α-Galactosyl-mannose was identified as the dominant degradation product resulting from *An*Man26A and *Pa*Man26A action on guar gum, strongly indicating that these two enzymes can accommodate galactopyranosyl residues in the −1 and in the +1 subsite. The degradation of α-6^4-6^3-di-galactosyl-mannopentaose by *An*Man26A revealed accommodation of galactopyranosyl residues in the −2, −1 and +1 subsite of the enzyme. Accommodation of galactopyranosyl residues in subsites −2 and +1 has not been observed for other characterized endomannanases to date. Docking analysis of galactomannooligosaccharides in available crystal structures and homology models supported the conclusions drawn from the experimental results. This newly discovered diversity of substrate degradation patterns demonstrates an expanded functionality of fungal endomannanases, than hitherto reported.

1. Introduction

β-Mannan (hereafter mannan) is the second most abundant hemicellulose in nature. Mannans are composed of a linear backbone containing D-mannopyranosyl residues (linear mannans) or D-mannopyranosyl and D-glucopyranosyl residues in an alternating manner (glucomannans) linked together by β-(1 → 4)-linkages. The backbone can be decorated with α-(1 → 6)-linked D-galactopyranosyl residues (galactomannans or galactoglucomannans) and acetyl groups [1–3]. Large amounts of mannans are found in the secondary plant cell wall of softwood (coniferous trees), where acetylated glucomannans and galactoglucomannans, comprise 25% of the wood dry matter [2,4]. The significance of softwood galactoglucomannans in biomass processing is a main reason for the interest in mannan degrading enzymes: endo-β(1 → 4)-mannanases (endomannanases, EC 3.2.1.78), β-mannosidases (EC 3.2.1.25) and α-galactosidases (EC 3.2.1.22), and their synergistic action [5,6]. Pure galactoglucomannan is not widely available, which is why galactomannans are often used to study enzymatic degradation of mannans containing galactopyranosyl substitutions [7]. Guar gum from the seeds of the guar plant (*Cyamopsis tetragonolobus*) and locust bean gum, from the carob tree

* Corresponding author. Fax.: +45 45882258.
E-mail address: am@kt.dtu.dk (A.S. Meyer).

http://dx.doi.org/10.1016/j.enzmictec.2015.10.011
0141-0229/

(*Ceretonia siliqua*) are significant sources of galactomannans. Guar gum contains more galactopyranosyl residues (Gal:Man, 1:2) than locust bean gum (Gal:Man, 1:4) [2,8]. In locust bean gum, the distribution of galactopyranosyl residues is irregular with a high proportion of unsubstituted blocks whereas in guar gum, the galactopyranosyl residues are more ordered and found mainly in pairs and triplets with few non-substituted regions [9].

A variety of bacteria, yeasts and filamentous fungi express mannan degrading enzymes [10]. Endomannanases are classified into three glycosyl hydrolase (GH) families: 5, 26 and 113, based on sequence similarity [11]. Endomannanases from these families belong to clan GH-A, share the $(\beta/\alpha)_8$-TIM barrel fold, and catalyze the hydrolysis of the *O*-glycosidic bonds in the mannan backbone with retention of the anomeric configuration [12–14]. "Endo-type" enzymes often use an open active site cleft in contrast to exo enzymes, e.g. β-mannosidases, using a pocket shaped active site region [15]. Characterized endo-mannanases show higher initial rates on locust bean gum compared to guar gum, and it has been proposed that the lower activity on guar gum is caused by the larger amount of galactopyranosyl residues [7,16,17]. Reported fungal endomannanases are predominantly found in GH5. Limited knowledge is available relating the sequence similarity classification (GH5 or GH26) to structural and functional differences among the two types of enzymes. Genome analysis has revealed that some organisms have endomannanases from both GH5 and GH26. The potentially different biological roles (substrate preferences) have been addressed for the bacterium *Cellvibrio japonicus* [18,19] and the fungus *Podospora anserina* [20,21]. A subdivision of fungal endomannanases from GH5 based on sequence similarity and functional characteristics has been proposed [22]. This phenomenon has been studied in *Aspergillus nidulans*, having a variety of different GH5 endomannanases (*An*Man5A, *An*Man5B, and *An*Man5C) [17,23,24]. *An*Man5B has been reported to have a high transglycosylation capacity and significantly lower k_{cat} and k_{cat}/k_m on locust bean gum compared with *An*Man5A and *An*Man5C [24]. Some fungal GH5 endo-mannanases are modular, typically having a carbohydrate-binding module from family 1 (CBM1), known to confer cellulose binding and to increase the mannan hydrolysis of complex substrates [25,26]. Fungal GH26 endomannanases may be fused to a CBM35 [21], a CBM family known to bind β-mannans, uronic acids and α-D-galactopyranosyl residues on carbohydrate polymers [27,28].

The hypothesis of this study is that fungal endomannanases of GH5 and GH26 have different degradation patterns on galactomannans due to structural differences, which affect their ability to accommodate galactopyranosyl residues in the active site cleft. With only a few fungal GH26 endomannanases studied, this work focused on the functional characterization of a novel GH26 endomannanase from *A. nidulans* (*An*Man26A). For comparison, well characterized fungal GH5 and GH26 endomannanases were analyzed as well, including *An*Man5A and *An*Man5C. The functional characterization included determination of the initial hydrolysis rate and maximal degree of conversion of locust bean gum and guar gum. The DNA sequencer-Assisted Saccharide analysis in High throughput (DASH) method [29] was used to characterize hydrolysis end product profiles on the two substrates. This method has not previously been reported for the analysis of galactomanno-oligo-saccharides. A dominant degradation product on guar gum was elucidated by obtaining the same compound in the degradation of α-6^1-galactosyl-mannotriose with a known β-mannosidase. Further knowledge about the accommodation of galactopyranosyl residues in the active site cleft was obtained by degradation of α-6^4-6^3-di-galactosyl-manno-pentaose by *An*Man26A, and by computational methods.

2. Materials and methods

2.1. Materials

Locust bean gum (low viscosity; borohydride reduced), guar gum (high viscosity), mannobiose, mannotriose, mannotetraose, mannopentaose, α-6^1-galactosyl-mannotriose, and α-6^4-6^3-di-galactosyl-mannopentaose were purchased from Megazyme (Ireland). All other chemicals were from Sigma (Germany). Mobility markers, dextran ladder, and the DASHboard software for DASH analyses were kindly donated by Prof. Paul Dupree (University of Cambridge, UK).

2.2. Structural comparison, homology modeling, and ligand docking

Homology models of *An*Man26A, *An*Man5A and *An*Man5C were generated using the HHpred-Homology server (http://toolkit.tuebingen.mpg.de/hhpred) [30] with *Pa*Man26A as template for *An*Man26A (PDB ID: 3ZM8, [21], 47% sequence identity), an *Aspergillus niger* GH5 endomannanase as template for *An*Man5A (PDB ID: 3WH9, not published, 68% sequence identity), and the *Chrysonilia sitophila* GH5 endomannanase as template for *An*Man5C (PDB ID: 4AWE, [31] 71% sequence identity). Model quality was evaluated using a Ramachandran analysis in MolProbity (http://molprobity.biochem.duke.edu/) [32]. *An*Man26A had 97.6% (323/331) of all residues in allowed regions, *An*Man5A had 99.1% (335/338) of all residues in allowed regions, and *An*Man5C had 98.7% (376/381) of all residues in allowed regions. Mannooligosaccharides with relevant α-(1 → 6)-linked D-galacto-pyranosyl residues were docked into the binding cleft of the homology models (*An*Man26A, *An*Man5A and *An*Man5C) as well as available crystal structures (*Pa*Man26A and *Tr*Man5A) as described below. First, β-(1 → 4)-linked D-manno-oligosaccharide models were designed with all residues in regular chair conformation except for the mannopyrano-syl unit in the -1 subsite, which was arranged in the same skewed boat conformation as observed in the ligand in a *C. japonicus* GH26C mannanase structure (PDB ID: 2VX6, [33]). This conformation is believed to be a prerequisite for catalysis to take place. The positioning of the structural oligomer units was guided by the location of mannoses, mannobioses, mannan analogues and other small molecule compounds found in the active site of various available mannanase crystal structures, see PDB IDs: 1QNR (*Tr*Man5A; mannobiose + glycerol in active site, [34]), 3WH9 (*An*Man5A; Tris in active site, unpublished), 4AWE (*C. sitophila* GH5; Tris + acetate in active site, [31]), 4CD6, 4CD7, 4CD8 (all *Aspergillus aculeatus* GH113A; various substrate analogues including mannobiose, [35]), 2VX6, 4CD4, 4CD5 (all *C. japonicus* GH26C; various substrate analogues including mannobiose, [33,35]), 3ZM8 (*Pa*Man26A; tartaric acid in active site, [21]). In all cases, the structures with ligands were aligned to the models and to the crystal structures of *Tr*Man5A and *Pa*Man26A such that the two catalytic glutamates were exactly overlapping and the surrounding residues in approximate (close) agreement. This alignment is preferred over a global structural alignment, because larger structural differences between (distantly related) mannanases away from the active site means that global alignment will lead to misalignment of the active sites. Based on active site alignment, the mannooligomeric ligands (with relevant galactopyranosyl substitutions) were positioned with C1 of the -1 manno-pyranosyl residue ~3 Å from the relevant glutamate nucleophile (as seen in PDB ID 2VX6). By adjusting the torsion angles in the neighboring mannose residues one by one, nearly perfect alignment with experimentally determined manno-pyranosyl units and other fragments was possible without obvious structural overlaps. Notice that only the five central subsites (−3 to +2) have verified crystal structure ligands, the extra mannopyranosyl residues were added to visualize the possibility of extending the substrate in both ends (except in *An*Man5C). Following the positioning of the mannanooligosaccharides, each model was individually energy-minimized to optimize details of the local geometry with only one restriction. Models were analyzed with PyMOL v1.7.2.0 (DeLano Scientific LLC, San Carlos, CA).

2.3. Expression and purification

Three *A. nidulans* endomannanases (*An*Man26A, *An*Man5A and *An*Man5C), a *P. anserina* GH26 endomannanase with and without its N-terminal CBM35 (*Pa*Man26A and *Pa*Man26A core), a *Trichoderma reesei* GH5 endomannanase with and without its C-terminal CBM1 (*Tr*Man5A and TrMan5A core) and an *A. niger* GH2 β-mannosidase (*Aniger*BM2, UNIPROT:A2QWU9) were recombinantly expressed in *Aspergillus oryzae* MT3568an amdS [36]. The β-mannosidase phylogenetically belongs to clade GH2 of β-mannosidases [37]. The enzymes were purified to electrophoretic purity using hydrophobic interaction and ion exchange chromatography (SDS-PAGE gels shown in Supplemental Fig. S1). The identity of the purified endomannanases was validated with mass spectrometry analyzing a tryptic digest of the protein band excised from a SDS-PAGE gel.

2.4. Protein determination

Protein concentrations were determined spectrophotometrically at 280 nm on an 8453-UV–vis Spectrophotometer (Aglient Technologies), using 1 ml cuvettes with 1 cm light path, and the molar extinction coefficient (ε) of a given protein. All measurements were made in triplicates. ε at 280 nm of all proteins were estimated

by GPMAW 9.20 (Lighthouse Data), and were based on mature proteins without modifications.

2.5. Thermal stability

The thermal stability was investigated with Differential Scanning Calorimetry (DSC) using a VP-Capillary Differential Scanning Calorimeter (MicroCal Inc., Piscataway, NJ, USA). DSC measures heat capacity as a function of temperature. In the transition between folded and unfolded state the heat capacity of a protein increases (it absorbs more energy). The Thermal midpoint (*Tm*) was determined as the top of the denaturation peak obtained with a constant heating rate of 200 °C/h (denaturation peaks shown in Supplemental Fig. S2). All endomannanases were buffer exchanged to 0.5 mg/ml in 50 mM sodium acetate buffer, pH 5 on illustra NAP-5 columns (GE Healtcare) before the scan. Denaturation temperatures were determined at an accuracy of +/− 1 °C.

2.6. pH optimum

The hydrolytic activity was determined on locust bean gum at 37 °C, after 15 min, over a pH range from 2.0 to 8.0, with 0.5 pH unit intervals. The hydrolysis volume was 200 μl, with 5 mg/ml locust bean gum in a Britton–Robinson buffer: 50 mM phosphoric acid, 50 mM acetic acid, 0.01% Trition X-100, 50 mM potassium chloride and 1 mM calcium chloride. The buffer pH was adjusted with sodium hydroxide from pH 2.0 to 8.0. Released reducing sugars were measured with the 4-hydroxybenzoic acid hydrazide (PAHBAH) method described by Lever [38], with mannose as standard.

2.7. Initial rates and maximal degree of conversion

The initial rate and the degree of conversion of locust bean gum and guar gum by the endomannanases were determined. The hydrolysis volume was 200 μl, with 2.5 mg/ml substrate in 50 mM sodium acetate buffer, pH 5. Hydrolysis was carried out at 37 °C for 15 min. This substrate concentration resulted in substrate solutions having sufficiently low viscosity to allow determination of initial rates of both locust bean gum and guar gum. Released reducing sugars were measured with the PAHBAH method as described above. All hydrolysis assays were carried out at 7 different endomannanase doses, in duplicates. Initial rates were calculated in the initial linear range of the hydrolysis. One unit (U) was defined as the amount of endomannanase required to release 1 μmole of reducing ends per minute, under the assay conditions specified. The degree of conversion was defined as released reducing ends relative to the theoretical monomeric yield of the substrate (including both mannopyranosyl and galactopyranosyl residues). The calculation procedure is explained under supplemental example S1. Hydrolysis end product profiles were analyzed by DASH after inactivation by heating at 95 °C for 15 min.

2.8. Degradation of galactomannooligosaccharides by mannan degrading enzymes

α-6^1-Galactosyl-mannotriose was degraded using *Aniger*BM2. Mannopentaose and α-6^4-6^3-di-galactosyl-mannopentaose was degraded using *An*Man26A. The hydrolysis volume was 200 μl, with 1 mM substrate in 50 mM sodium acetate buffer, pH 5. Hydrolysis was carried out in 96 well flat bottomed microtiter plates, (Nunc™) at 37 °C for 15 min. Released reducing sugars were measured with the PAHBAH method as described above. Background absorbance, caused by the reducing ends of the substrate before enzyme treatment, was subtracted from the absorbance measured after hydrolysis. All hydrolysis assays were carried out in duplicates. Hydrolysis products were analyzed by DASH after inactivation by heating at 95 °C for 15 min.

2.9. DASH—APTS labeling and analysis of the labeled oligosaccharides

APTS (9-aminopyrene-1,4,6-trisulfonate) labeling and analysis of the labeled saccharides were carried out as described by Li et al. [29]. More than one amount of labelled saccharides was tested for all samples. Samples were analyzed with an ABI 3730xl 96-sample DNA sequencer using the standard DNA analysis buffer system, and settings described in supplemental table S1. DASHboard software was used to align the peak profiles using the mobility markers that were present in every sample. The ratios between peaks in a sample were highly reproducible as confirmed by analyzing 12 samples run in parallel containing the same labeled dextran ladder. The peak area ratios between all adjacent peaks in the ladder (P_{n-1}/P_n) had a coefficient of variation (CV) below 5% as long as peaks were above 24,000 RFU^2 and not out of scale, which confer with the data obtained by Li et al. [29].

3. Results and discussion

The fungal endomannanases tested in this study (Table 1) were recombinantly expressed in the fungal host *A. oryzae* and purified to electrophoretic purity.

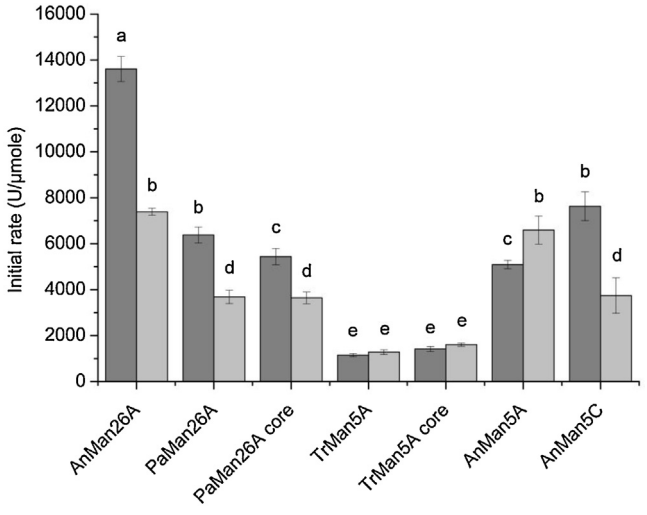


Fig. 1. Initial rates of hydrolysis given in U per μmole of endomannanase on galactomannans. The initial reaction rates (U/μmole) were determined at 37 °C and pH 5 on locust bean gum (dark grey) and guar gum (light grey). Multiple doses were tested for each enzyme, but only doses which resulted in hydrolysis rates within the linear initial rate of hydrolysis were used. Initial rates were calculated at substrate conversion levels ranging from 5 to 20% of the maximal degree of conversion by the specific enzyme on the specific substrate. A one-way ANOVA analysis validated significant difference between means. Values are given as mean values ± SD ($n = 2–6$). The letters (a–e) each represents a group of initial rates which are significantly different from initial rates belonging to all other groups (ratings are assigned with a 95% confidence interval for means based on a pooled SD of 411).

3.1. Physicochemical properties

The endomannanases were found to have pH optima in the range of 3.5–5.5, with *An*Man26A having the least acidic optimum of them all (pH 5.5). The *Tm* of the tested fungal GH26 endomannanases were found to be 53 °C and 57 °C, with *An*Man26A having the lowest *Tm* at 53 °C, while the *Tm* of the tested fungal GH5 endomannanases were between 70 and 81 °C (Table 1). Based on the *Tm*'s fungal GH5 endomannanases were more thermostable than the fungal GH26 endomannanases. The same pattern was observed within one organism, *A. nidulans*: *An*Man26A had a *Tm* of 53 °C compared with *An*Man5A and *An*Man5C of 70 °C. However a fungal GH26 endomannanase from *A. niger* has been reported to have an optimum temperature at 80 °C at pH 5 in a one hour hydrolysis [7], suggesting that more thermostable fungal GH26 endomannanases do exist. The thermal stability did not seem to be influenced by the presence of the CBM, irrespective of the CBM belonging to CBM1 (*Tr*Man5A) or CBM35 (*Pa*Man26A), with CBM1 being 40 amino acids and having an S/T-linker, and CBM35 being 100 amino acids and having a P-rich linker.

3.2. Initial rates of hydrolysis

The initial rate of hydrolysis of locust bean gum and guar gum by the endomannanases was determined (Fig. 1) at 37 °C and pH 5, since these conditions assured stability (Table 1, *Tm*) and close to optimal activity (Table 1, pH_5/pH_{opt}) of the enzymes. *An*Man26A had the highest initial rate on locust bean gum (13,600 U/μmole) and on guar gum (7400 U/μmole) of all tested endomannanases. The lowest initial rates were observed for *Tr*Man5A on locust bean gum (1150 U/μmole) and on guar gum (1300 U/μmole) irrespective of the presence of the CBM1. This is in accordance with *Tr*Man5A CBM1 being known to bind to cellulose but not to mannan [26]. *Pa*Man26A containing a CBM35 had a significantly higher initial rate on locust bean gum (6400 U/μmole) compared to *Pa*Man26A core (5400 U/μmole). An explanation could be that CBM35 interacts

Table 1
Properties of the studied endomannanases.

Origin	Enzyme name	GH family	CBMs	Mw[a] (kDa)	pH_{opt}[b]	Relative activity (pH_5/pH_{opt})	Tm[c] (°C)	Swissprot ID
A. nidulans	*An*Man26A	26	No	35.2	5.5 (4.5–7)	0.93	53	Q5AWB7
P. anserina	*Pa*Man26A	26	CBM35	49.8	nd	nd	57	B2AEP0
P. anserina	*Pa*Man26A core	26	No	34.4	nd	nd	58	na[d]
T. reesei	*Tr*Man5A	5	CBM1	45.2	3.5 (3–6.5)	0.93	81	Q99036
T. reesei	*Tr*Man5A core	5	No	38.8	4.0 (3–6)	0.88	78	na[d]
A. nidulans	*An*Man5A	5	No	40.7	4.0 (3.5–7)	0.99	70	Q5B7X2
A. nidulans	*An*Man5C	5	No	43.5	3.5 (3-6)	0.86	70	Q5AZ53

[a] Theoretical protein weight.
[b] pH optimum and pH interval with 80 % retained activity in brackets, estimated at 37 °C, nd = not determined.
[c] The Thermal midpoint (*Tm*) was estimated for each enzyme at pH 5. The denaturation peaks can be seen in supplemental Fig. S2.
[d] na = Not applicable, here because the core construct has the same ID as the wild type (full length) version of the enzyme.

with locust bean gum through binding to β-mannans or α-D-galactopyranosyl residues, as it has been reported for other CBMs from family 35 [27,28]. *An*Man26A and *Pa*Man26A showed significantly higher initial rates on locust bean gum than on guar gum at the same substrate concentration. These data are in accordance with previous studies by Malgas et al. [7] showing reduced specific activity on guar gum compared to locust bean gum for a fungal GH26 endomannanase from *A. niger*. The same tendency was observed for *An*Man5C whereas *Tr*Man5A, *Tr*Man5A core and *An*Man5A seemed to discriminate less between the two substrates with similar initial rates on guar gum compared to locust bean gum. In an earlier study, comparing *An*Man5A and *An*Man5C, both endomannanases had higher initial rates on locust bean gum compared to guar gum, but also here *An*Man5A discriminated much less between the two substrates than *An*Man5C [17]. These data are in agreement with the present work, with the only exception being that *An*Man5A in the present study exhibited a higher initial rate on guar gum compared to locust bean gum.

3.3. Maximal degree of conversion

A dose-response study to determine the maximal degree of conversion of locust bean gum and guar gum by the endomannanases was carried out (Fig. 2). Maximal degree of conversion was defined as reached when the conversion degree did not increase with increasing enzyme dose. Hydrolyses were carried out at stable conditions for the enzymes: 37 °C, pH 5 for 15 min. All tested endomannanases exhibited a linear dose-response relation at low doses and reached a plateau at high doses. Significant differences were observed between the maximal degrees of conversion reached by the individual endomannanases. This was especially the case at the highly galactopyranosyl substituted mannan—guar gum (Fig. 2, bottom). For the less substituted locust bean gum the differences were less pronounced as *An*Man26A, *Pa*Man26A and *Pa*Man26A core reached 38–40% conversion and *An*Man5A, AnMan5C, *Tr*Man5A and *Tr*Man5A core reached 26–29% conversion. On guar gum the tested GH26 endomannanases, *An*Man26A, *Pa*Man26A and *Pa*Man26A core, reached 35–40% conversion, whereas the tested GH5 endomannanases: *An*Man5A, *Tr*Man5A and *Tr*Man5A reached 8-10% and *An*Man5C only 3 % conversion of guar gum. The maximal degree of conversion reached by *An*Man26A, *Pa*Man26A and *Pa*Man26A core was found to be significantly different from the maximal degree of conversion reached by *An*Man5A, AnMan5C, *Tr*Man5A and *Tr*Man5A core on both substrates ($p < 0.01$). The presence of a CBM (either CBM1 or CBM35) did not seem to affect the maximal degree of conversion on locust bean gum or guar gum. The presented data for the three tested GH5 endomannanases are in agreement with a study carried out by McCleary in which the maximal degree of conversion of locust bean gum and guar gum by three endomannanases was found to be approximately 30% and 10%, respectively [39]. To the best of our

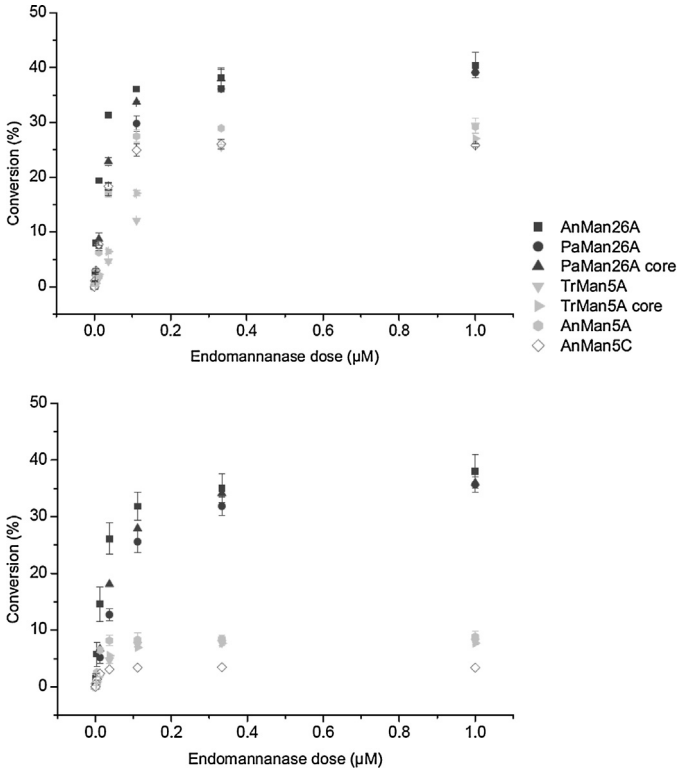


Fig. 2. Conversion of galactomannans by fungal endomannanases. Conversion (%) of locust bean gum (top) and guar gum (bottom). Hydrolyses were performed at 37 °C, pH 5 for 15 min. The degree of conversion was defined as released reducing ends relative to the theoretical monomeric yield. A calculation example can be seen in supplemental example S1. A one-way ANOVA analysis validated that the maximal degree of conversion (%) (measured with 1 μM endomannanase) for *An*Man26A and *Pa*Man26A were significantly different from the maximal degree of conversion (%) of *Tr*Man5A, *An*Man5A and *An*Man5C ($p < 0.01$) for both locust bean gum and guar gum. Values are represented as mean values ± SD ($n = 2$).

knowledge endomannanase catalyzed conversion yields of 35–40% of guar gum, as presented in the present study, have not been reported before.

The measured initial rates (Fig. 1) and the maximal degrees of conversion (Fig. 2), present two very different pictures of the characterized endomannanases. Despite differences in initial rates between endomannanases belonging to one GH family (Fig. 1), they reach almost the same degree of conversion. This difference demonstrates that initial rates obtained by enzymatic degradation of heterogeneous substrates only give information about the cleavage of the accessible bonds in the substrate and that the amount of accessible bonds may differ for the different types of enzymes. An enzyme with a relatively high initial rate on a given substrate might thus only be able to degrade a limited part of the substrate, as is

the case for *An*Man5A on guar gum. Locust bean gum and guar gum were consequently characterized to have easily accessible areas for endomannanases and areas being more difficult to degrade. Possibly *An*Man26A and *Pa*Man26A can degrade parts of the substrates which are inaccessible for *An*Man5A, AnMan5C, and *Tr*Man5A. The difference between the two groups of endomannanases are more pronounced on the highly substituted guar gum than on locust bean gum, thus the inaccessibility is likely due to the galactopyranosyl residues.

3.4. Hydrolysis product profiles

At the maximal degree of conversion the product profile was analyzed using the DASH method (Fig. 3). The analysis was run with a ladder containing known mannooligo- and galactomannooligosaccharides, in order to identify these products. The nomenclature used for galactomannooligosaccharides in this study is similar as that suggested for xyloglucan by Fry et al. [40]. Oligosaccharides are written from the non-reducing to the reducing end and backbone units bearing a substitution are only represented by the substitution. With this nomenclature α-6^1-galactosyl-mannotriose is named MMG and α-6^4-6^3-di-galactosyl-mannopentaose is named MGGMM. On both locust bean gum (Fig. 3A) and guar gum (Fig. 3B) the product profiles had differences and similarities among endomannanases. Compounds observed in between the known mannooligosaccharides in the DASH profiles must be galactomannooligosaccharides. However it is possible that some galactomanno-oligosaccharides have the same migration as the known unsubstituted mannooligosaccharides represented in the ladder. The product profiles for *An*Man5A, AnMan5C, and *Tr*Man5A were generally simpler than the product profiles for *An*Man26A and *Pa*Man26A and contained fewer galactomanno-oligosaccharides. No unsubstituted mannooligosaccharides above mannotriose were observed in the product profiles for *An*Man5A, AnMan5C, and *Tr*Man5A strongly suggesting that they cleave mannotetraose or longer mannooligosaccharides in accordance with previous studies [17,41]. Mannotetraose was observed in the product profiles for *An*Man26A and *Pa*Man26A which is in agreement with studies reporting that *Pa*Man26A needs binding of at least five subsites for efficient hydrolytic activity [21]. *An*Man26A and *Pa*Man26A produced mostly mannose and an unidentified galactomannooligosachharide which migrated to 2.1 dextran units (DE), in-between mannobiose and mannotriose. On guar gum this galactomanno-oligosaccharide was the dominant product for the two GH26 endomannanases. However, it was not observed in any of the product profiles for the tested GH5 endomannanases on both substrates. The few products observed in the profile of *An*Man5C on both locust bean gum and guar gum were unexpected and suggest a different substrate degradation pattern than the other tested endomannanases. The complex product profiles for *An*Man26A and *Pa*Man26A can be explained by an ability to degrade regions of the substrates with higher substitution. Such ability will also explain why additional products were observed for these two enzymes after hydrolysis of the highly substituted guar gum compared with locust bean gum and why they reached a higher maximal degree of conversion compared with *An*Man5A, AnMan5C, and *Tr*Man5 (Fig. 2).

3.5. Elucidation of the dominant hydrolysis product for An*Man26A and* Pa*Man26A on guar gum*

The product profiles from the hydrolysis of guar gum by *An*Man26A and *Pa*Man26A had the same unidentified dominant compound, migrating to 2.1 DE. The compound was expected to be a galactomannooligosaccharide containing two or three sugar units, as it migrated after mannobiose but before mannotriose. Elucidation of this compound was crucial for understanding the degradation pattern and the accommodation of galactopyranosyl residues in the active site cleft of these two endomannanases. To determine the migration pattern of some well-defined small galactomannooligosaccharides, MMG was treated with GH2 β-mannosidase from *A. niger* (*Aniger*BM2) and the resulting products analyzed by DASH (Fig. 4). The products formed from MMG were those expected for a GH2 clade A β-mannosidase [37,42]. Due to the nature of the DASH analysis, the size of peaks can be compared relatively within each sample but not across samples. The degradation of 1 mM MMG with 0.001 and 0.01 μM *Aniger*BM2 resulted in the production of mannose and α-6^1-galactosyl-mannobiose (MG). MG migrated to 2.87 DE, which is the same migration as observed for M3. At 1 μM *Aniger*BM2, the β-mannosidase converted MG into mannose and α-galactosyl-mannose (G), migrating to 2.1 DE. Because of identical migration, this product was assumed to be the same as the dominant oligosaccharide in the product profiles of *An*Man26A and *Pa*Man26A on guar gum. The assumption is supported by the knowledge that the migration of other manno- and galactomannooligosaccharides with similar molecular weight has been established to other DE: M1 at 1 DE, M2 at 1.87 DE and MG and M3 at 2.87 DE. To date the shortest reported galactomannooligosaccharide produced from galactomannan or galactoglucomannan hydrolysis by an endomannanase is MG. MG has previously been found among the hydrolysis products of *Tr*Man5A and an *A. niger* GH5 endomannanase [43,44]. It has been speculated that the ability to produce this product was related to the sequence and structure of the fungal GH5 endomannanases, with the reasoning that a bacterial GH26 endomannanase from *Bacillus subtilis* had MMG as its smallest produced galactomannooligosaccharides [44]. To produce G as the dominant end product, *An*Man26A and *Pa*Man26A must be able to accept galactopyranosyl residues both in the −1 and +1 subsite in the active site cleft. Earlier studies investigating the accommodation of galactopyranosyl residues in the active site clefts of endomannanases, including a fungal endomannanase from *A. niger*, found that they could be accommodated in subsite −1, but not in subsite −2 and +1 [43]. In fact it has been highlighted that galactopyranosyl residues are absent from the non-reducing, terminal mannopyranosyl group in earlier identified hydrolysis products [39,43,45]. With the accommodation of galactopyranosyl residues in the +1 subsite, *An*Man26A and *Pa*Man26A seem to possess a novel characteristic, which likely contributes to the differentiated degradation pattern of galactomannans.

3.6. Further elucidation of the accommodation of galactopyranosyl residues in the active site of An*Man26A*

Additional knowledge about the accommodation of galactopyranosyl residues in the active site cleft of *An*Man26A was gathered by degrading mannopentaose and MGGMM. The resulting products were analyzed by DASH (Fig. 5). Degradation of mannopentaose and MGGMM by *An*Man26A was accomplished with similar rates based on reducing ends generated. Degradation of mannopentaose by *An*Man26A at 0.03 μM resulted predominantly in the production of mannose and mannotetraose and to a smaller extent mannobiose and mannotriose (Fig. 5B). With increasing enzyme concentration, all of the initial substrate was degraded and at 3 μM *An*Man26A even mannotetraose was consumed. These data confer with an earlier study reporting that *Pa*Man26A (47% identity with *An*Man26A) degrades mannopentaose to mannose and mannotetraose with preferred binding from the −4 to +1 subsite [21]. Degradation of MGGMM by 0.3 μM *An*Man26A resulted in the production of mannose and an unidentified galactomannooligosaccharide migrating to 5.06 DE and to a smaller extent mannobiose and another unidentified galactomannooligosaccharide migrating to 4.10 DE. With the preferred binding of mannopentaose to *Pa*Man26A from subsite

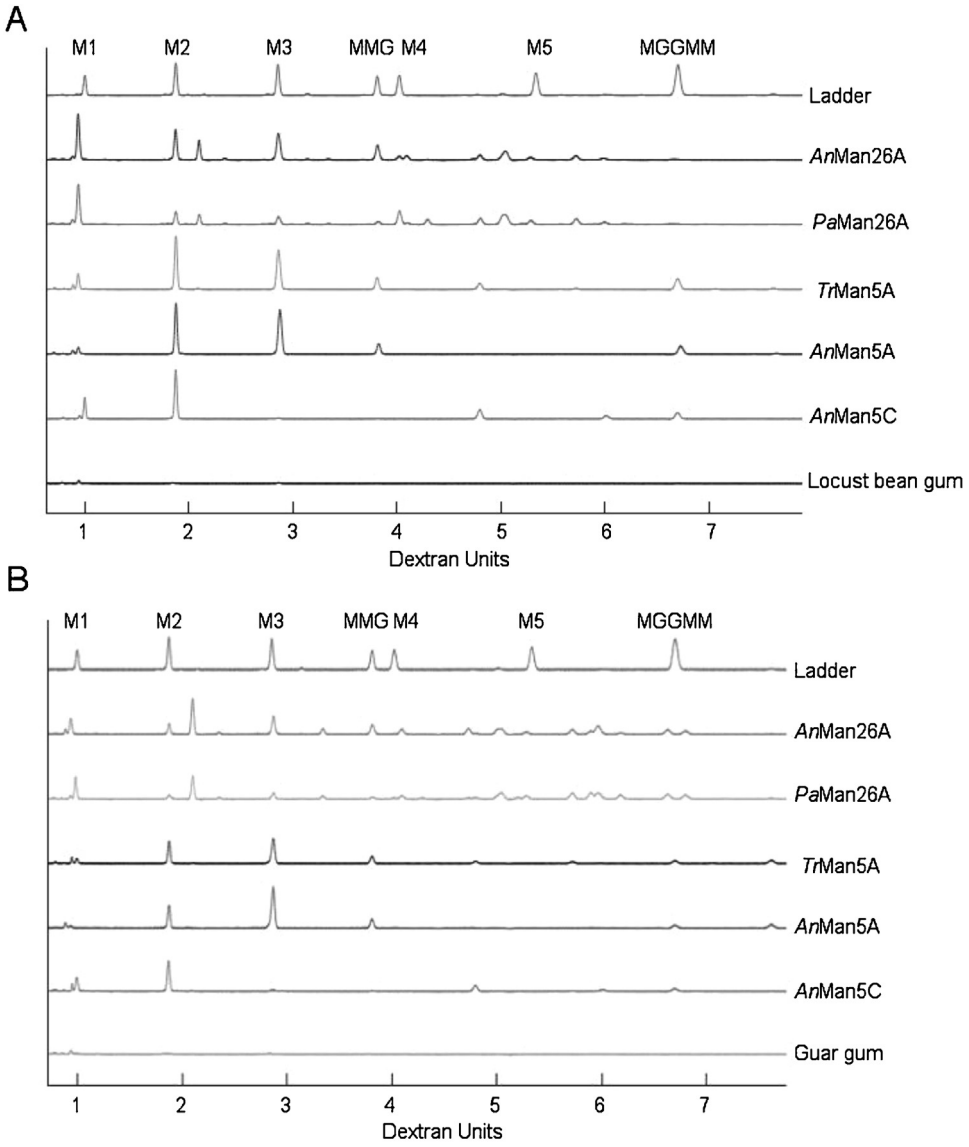


Fig. 3. Product profiles from galactomannan hydrolysis by fungal endo-mannanases. Aligned electropherograms of enzymatic hydrolysis on locust bean gum (A) and guar gum (B) at the maximal degree of conversion. The migration of the oligosaccharides is in dextran units (DE). A ladder containing mannooligosaccha-rides (mannose to mannopentaose) and two galactomannooligosaccharides were run: mannose (M1, migrating to 0.9 DE), mannobiose (M2 migrating to 1.87 DE), mannotriose (M3, migrating to 2.85 DE), α-6^1-galactosyl-mannotriose (MMG, migrating to 3.81 DE), mannotetraose (M4, migrating to 4.02 DE), mannopentaose (M5, migrating to 5.33 DE), α-6^4-6^3-di-galactosyl-mannopentaose (MGGMM, migrating to 6.70 DE). Substrates subjected to the hydrolysis conditions as controls.

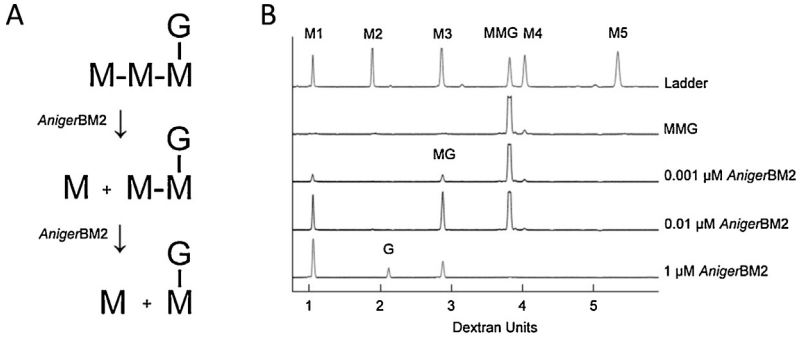


Fig. 4. Degradation of MMG by *A. niger*BM2. Degradation scheme (A) and aligned electropherograms (B) of the degradation of MMG by an *A. niger* GH2 β-manno-sidase (*Aniger*BM2). The migration of the oligosaccharides is in dextran units (DE). A ladder containing mannooligosaccharides from mannose to mannopentaose and galactomannooligosaccharides were run: mannose (M1, migrating to 0.9 DE), mannobiose (M2 migrating to 1.87 DE), mannotriose (M3, migrating to 2.85 DE), α-6^1-galactosyl-mannotriose (MMG, migrating to 3.81 DE), mannotetraose (M4, migrating to 4.02 DE) and mannopentaose (M5, migrating to 5.33 DE). MMG treated at hydrolysis conditions was also tested. Other abbreviations: α-6^1-galactosyl-mannobiose (MG) and α-galactosyl-mannose (G).

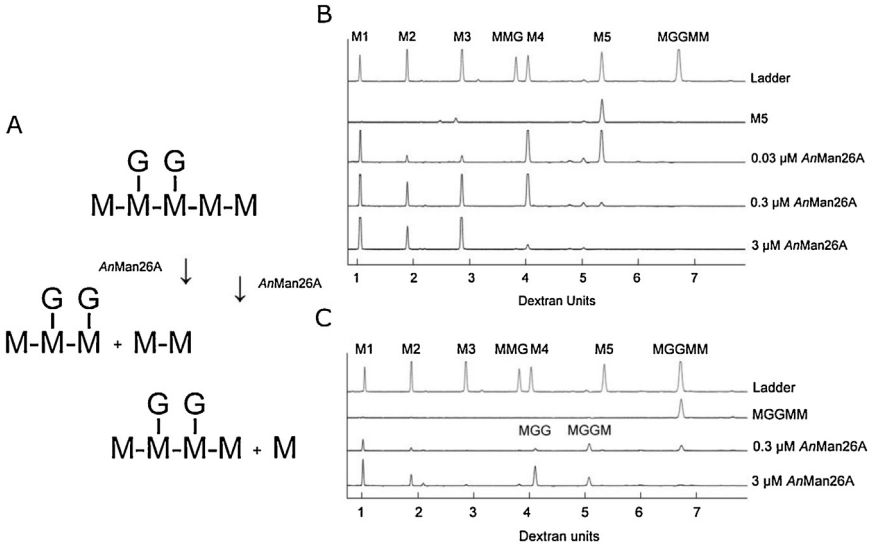


Fig. 5. Degradation of mannopentaose and MGGMM by *An*Man26A. Proposed scheme of MGGMM degradation by an *Aspergillus nidulans* GH26 endomannanase (*An*Man26A) (A) and aligned electropherograms of the degradation of manno-pentaose (B) and MGGMM (C) by *An*Man26A. The migration of the oligosaccharides is in dextran units (DE). A ladder containing mannooligosaccharides from mannose to mannopentaose and two galactomannooligosaccharides were run: mannose (M1, migrating to 0.9 DE), mannobiose (M2 migrating to 1.87 DE), mannotriose (M3, migrating to 2.85 DE), α-6^1-galactosyl-mannotriose (MMG, migrating to 3.81 DE), mannotetraose (M4, migrating to 4.02 DE), mannopentaose (M5, migrating to 5.33 DE), α-6^4-6^3-di-galactosyl-mannopentaose (MGGMM, migrating to 6.70 DE). Substrates treated with hydrolysis conditions were also tested.

−4 to +1 [21], the close sequence homology of *Pa*Man26A and *An*Man26A, and mannose and mannotetraose being the dominant products for both enzymes, similar preferences for binding can be expected. A comparison of a model of *An*Man26A and the crystal structure of *Pa*Man26 (Fig. 6D–F), showing highly conserved organization of the binding residues in the active site clefts of these two enzymes, supported this assumption. In subsite −4 the *An*Man26A has two aromatic residues: W150 and W151, positioned as W244 and W245 in *Pa*Man26A; these residues have been reported to mediate the strong binding in this subsite [21]. It is therefore likely that *An*Man26A predominantly binds MGGMM from subsite −4 to 1 producing mannose and MGGM which is expected to be to the oligosaccharide observed at 5.06 DE. This cleavage will allow accommodation of galactopyranosyl residues in subsite −3 and −2. To produce mannobiose *An*Man26A must bind MGGMM from subsite −3 to +2, implying that the galactopyranosyl residues will be accommodated in subsite −2 and −1. This cleavage will create mannobiose and MGG, which is expected to be the oligosaccharide seen at 4.10 DE (Fig. 5C). At this enzyme concentration binding of MGGMM from −2 to +3, creating MG and GMM, is not observed, since no MG (migrating to 2.87 DE) is detected. At 3 μM *An*Man26A, MGGMM is fully degraded and also MGGM is partly degraded. With this high enzyme concentration small amounts of G, MG (or GM) and MMG (or GMM) seemed to appear among the products (Fig. 5C). These products might have been produced by cleavage of MGG to MG and G, cleavage of MGGM to MG and GM or even cleavage of MGGMM to MG and GMM. It is strongly believed that these products only arise at very high enzyme to substrate ratio because *An*Man26A, like *Pa*Man26A, needs to bind in more than four subsites for effective binding [21].

3.7. Ligand docking and structural comparison of the active site clefts

As assessment of the structural background for the observed product profiles, a comparison of the active site cleft of the tested endomannanases with docked galactomannooligosaccharides was performed using available crystal structures and homology models (Fig. 6). Notice how the mannobiose in the binding cleft of the *Tr*Man5A crystal structure (1QNR) align with the docked galactomanno-oligosaccharide (Fig. 6B). The structural comparison revealed differences between the tested endomannanases. The docking analysis strongly indicated that *An*Man5A (Fig. 6A) and *Tr*Man5A (Fig. 6B) could both accommodate a galactopyranosyl residue in subsite −1, but not in subsite −2 or subsite +1. For *Tr*Man5A productive accommodation of a galactopyranosyl unit in the −1 subsite has been shown previously [44]. The model of *An*Man5A appears to have a pocket evolved to accommodate the −1 galactopyranosyl residue. A similar pocket was not observed in the crystal structure of *Tr*Man5A. A dedicated pocket for binding of a galactopyranosyl residue in subsite −1 could contribute to the significantly higher initial rates on locust bean gum and guar gum obtained by *An*Man5A compared to *Tr*Man5A (Fig. 1). The model of *An*Man5C (Fig. 6C) exhibits a narrower active site cleft with no room for galactopyranosyl residues in either subsite −2, −1 or +1. The glycone region of the active site cleft in *An*Man5C appears to be closed off by a loop, preventing binding beyond subsite −2. A similar loop arrangement is observed in the GH5 endomannanase from *C. sitophila* (*Cs*Man5) [31], which has been used as template for the *An*Man5C model. The authors solving the *Cs*Man5 structure suggests that the loop arrangement may confer exo-activity and that *Cs*Man5 could be a 2-mannobiohydrolase [31]. These observations are in perfect agreement with the *An*Man5C product profiles on locust bean gum and guar gum, both consisting predominantly of mannobiose. Together these results corroborate that *An*Man5C possesses exo-activity, with preference for cleaving off two mannopyranosyl residues at a time. Compared to the fungal GH5 endomannanases, the active site cleft in *An*Man26A (Fig. 6D) and *Pa*Man26A (Fig. 6E) is more open and the docking results reveal how galactopyranosyl residues can be accommodated in subsite −2, −1 and +1. These observations agree with the observed product profiles of locust bean gum and guar gum obtained by DASH (Fig. 3), which showed that only *An*Man26A and *Pa*Man26A produced G, which require accommodation of a galactopyranosyl residue in the +1 subsite and −1. In addition, by the degradation of MGGMM it was shown that *An*Man26A could accommodate a

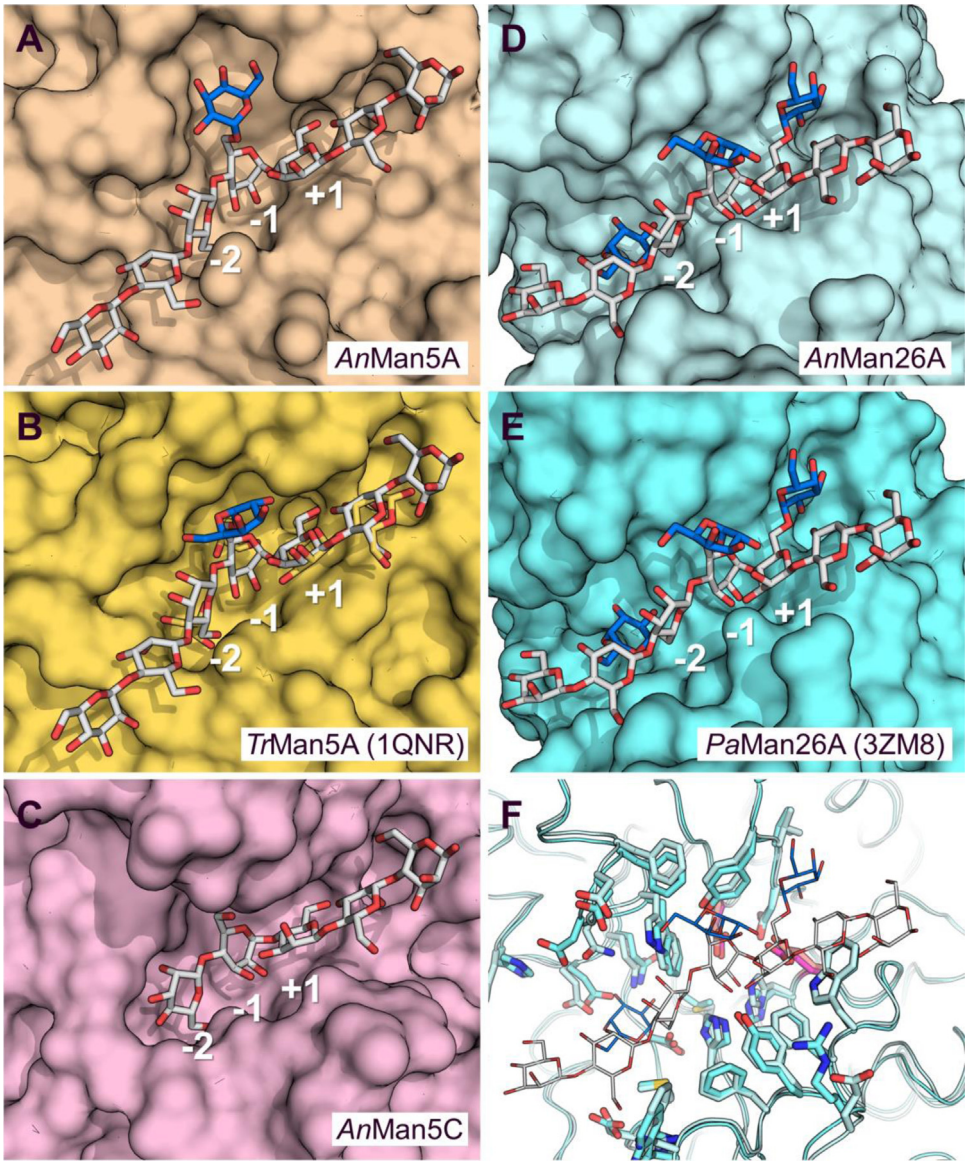


Fig. 6. Structural comparison of the active site cleft of fungal endomannanases. The *An*Man5A model (A) appears to have a dedicated pocket to accommodate a galactopyranosyl residue in subsite −1, but no room for substitutions in subsite −2 and +1. Although *Tr*Man5A (B; PDB ID: 1QNR, [34]) does not have a similar pocket there is still room for a galactopyranosyl residue in subsite −1, but not in −2 and +1. Notice how the experimentally determined mannobiose (in yellow) in the binding site align with the docked galactomannooligosaccharide. The *An*Man5C model (C) does not appear to have room for substituents in the −2, −1 and +1 subsites, and the binding pocket appears to be closed off preventing binding beyond the −2 subsite. The model of *An*Man26A (D) and the crystal structure of *Pa*Man26A (E; PDB ID: 3ZM8, [21]) both have a wide binding cleft with room for galactopyranosyl residues in subsite −2, −1, and +1. Panel (F) shows the conservation of substrate binding residues between *An*Man26A (light blue) and PaMan26A (cyan). The active site glutamate residues are highlighted in shades of pink.

galactopyranosyl residue in subsite −2 (while this remains to be assessed for *Pa*Man26A). Interestingly, the docking results with *Pa*Man26A and *An*Man26A also indicate that there is room for galactopyranosyl substitutions on the mannan backbone beyond the −2, −1 and +1 subsites, and thus both enzymes might be able to work on extensively galactopyranosyl-substituted mannans. Based on the presented data it seems that the ability to accommodate galactopyranosyl residues in both subsite −2, −1 and +1 and therefore reach higher degradation of the highly substituted guar gum is a unique GH26 feature. However, it will require examination of more endomannanases belonging to the two GH families to truly verify this hypothesis.

Some organisms express several endomannanases, a well-studied example being the bacterium *C. japonicus* expressing both GH5 and GH26 mannanases [18]. Fungal examples are *A. nidulans* expressing the studied *An*Man26A, *An*Man5A and *An*Man5C and the enzyme *An*Man5B which is known for its transglycosylation capacity [24], and *P. anserina* expressing two GH5 endomannanases and *Pa*Man26A [21]. In these organisms, the individual enzymes possess different characteristics and therefore support the idea that a single organism may have several endomannanases with complementary substrate preferences to create synergy in the deconstruction of mannans or to be induced at different situations when their characteristics are needed [17,24]. These examples highlight that differences in substrate degradation patterns is not only seen between endomannanases from different GH families but can also be found within one family as is the case with *An*Man5A, *An*Man5C and *An*Man5B.

4. Conclusions

A novel GH26 endomannanase from *A. nidulans* was successfully cloned, purified, and characterized in parallel with known fungal endomannanases from GH5 and GH26. The analyzed GH5 endomannanases were found to be more thermostable than the GH26 endomannanases and neither presence of CBM1 nor CBM35 influenced the thermal stability of the endomannanases. The DASH method was successfully adapted for characterization of endomannanases and their hydrolysis products on galactomannans. The method can be a great tool in the process of screening endomannanases and other enzymes degrading heterogeneous substrates. *An*Man26A and *Pa*Man26A were found to have a novel substrate degradation pattern on guar gum and locust bean gum making these endomannanases superior in their capability to degrade highly substituted galactomannans in comparison to the tested fungal GH5 endomannanases. This capability seems to be due to the possibility of accommodating galactopyranosyl residues in both the −1, −2 and +1 subsite. The functional characterization of the enzymes was supported by ligand docking and structural comparison of the 5 fungal endomannanases. It is a novel characteristic for endomannanases to allow galactopyranosyl residues in subsite −2 and +1, whereas the accommodation in subsite −1 has been reported before. This novel characteristic provided by the assessed GH26 endomannanases expands the functional diversity of fungal endomannanases and may open for new applications of these enzymes.

Authors contribution

Each author has materially participated in the research and the article preparation.

Pernille von Freiesleben, PvF: Participated in the design of the work. Conducted experiments, analysed the data, and co-wrote the manuscript.

Nikolaj Spodsberg, NS: Cloned and recombinantly expressed the enzymes.

Thomas Holberg Blicher, THB: Helped making the docking analysis of the enzymes.

Henning Jørgensen, HJ, Lars Anderson, LA, Henrik Stålbrand, HS, Anne S. Meyer, AM, Kristian B. R. M. Krogh, KBK: Designed and supervised the work, co-wrote the manuscript with input from all authors.

All authors have approved the final revised (R1) article.

Acknowledgements

We would like to thank Prof. Paul Dupree (University of Cambridge, UK) for help with implementation and optimization of the DASH method. This study was partially financed by The Danish Agency for Science, Technology and Innovation. Henrik Stålbrand was partially funded by FORMAS (213-2014-1254) and the Swedish Foundation for Strategic Research (14-0046).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.enzmictec.2015.10.011.

References

[1] A. Ebringerová, Structural diversity and application potential of hemicelluloses, Macromol. Symp. 232 (2006) 1–12.
[2] L.R.S. Moreira, E.X.F. Filho, An overview of mannan structure and mannan-degrading enzyme systems, Appl. Microbiol. Biotechnol. 79 (2008) 165–178.
[3] H.V. Scheller, P. Ulvskov, Hemicelluloses, Annu. Rev. Plant Biol. 61 (2010) 263–289.
[4] A. Várnai, L. Huikko, J. Pere, M. Siika-Aho, L. Viikari, Synergistic action of xylanase and mannanase improves the total hydrolysis of softwood, Bioresour. Technol. 102 (2011) 9096–9104.
[5] H.J. Gilbert, H. Stålbrand, H. Brumer, How the walls come crumbling down: recent structural biochemistry of plant polysaccharide degradation, Curr. Opin. Plant Biol. 11 (2008) 338–348.
[6] S. Malgas, J.S. van Dyk, B.I. Pletschke, A review of the enzymatic hydrolysis of mannans and synergistic interactions between β-mannanase, β-mannosidase and α-galactosidase, World J. Microbiol. Biotechnol. 31 (2015) 1167–1175.
[7] S. Malgas, S.J. van Dyk, B.I. Pletschke, β-Mannanase (Man26A) and α-galactosidase (Aga27A) synergism—a key factor for the hydrolysis of galactomannan substrates, Enzyme Microb. Technol. 70 (2015) 1–8.
[8] K.S. Parvathy, N.S. Susheelamma, R.N. Tharanathan, A.K. Gaonkar, A simple non-aqueous method for carboxymethylation of galactomannans, Carbohydr. Polym. 62 (2005) 137–141.
[9] B.V. McCleary, The fine structures of carob and guar galactomannans, Carbohydr. Res. 139 (1985) 237–260.
[10] S. Dhawan, J. Kaur, Microbial mannanases: an overview of production and applications, Crit. Rev. Biotechnol. 27 (2007) 197–216.
[11] B. Henrissat, G. Davies, Structural and sequence-based classification of glycoside hydrolases, Curr. Opin. Struct. Biol. 7 (1997) 637–644.
[12] M.L. Sinnott, Catalytic mechanisms of enzymic glycosyl transfer, Chem. Rev. 90 (1990) 1171–1202.
[13] B. Henrissat, I. Callebaut, S. Fabrega, P. Lehn, J. Mornon, G. Davies, Conserved catalytic machinery and the prediction of a common fold for several families of glycosyl hydrolases, Proc. Natl. Acad. Sci. U. S. A. 92 (1995) 7090–7094.
[14] S.G. Withers, Mechanisms of glycosyl transferases and hydrolases, Carbohydr. Polym. 44 (2001) 325–337.
[15] G. Davies, B. Henrissat, Structures and mechanisms of glycosyl hydrolases, Structure 3 (1995) 853–859.
[16] J. Le Nours, L. Anderson, D. Stoll, H. Stålbrand, L. Lo Leggio, The structure and characterization of a modular endo-β-1,4-mannanase from *Cellulomonas fimi*, Biochemistry 44 (2005) 12700–12708.
[17] A. Dilokpimol, H. Nakai, C.H. Gotfredsen, M.J. Baumann, N. Nakai, M.A. Hachem, et al., Recombinant production and characterisation of two related GH5 endo-β-1,4-mannanases from *Aspergillus nidulans* FGSC A4 showing distinctly different transglycosylation capacity, Biochim Biophys Acta 1814 (2011) 1720–1729.
[18] L.E. Tailford, V.M.A. Ducros, J.E. Flint, S.M. Roberts, C. Morland, D.L. Zechel, et al., Understanding how diverse β-mannanases recognize heterogeneous substrates, Biochemistry 48 (2009) 7009–7018.
[19] X. Zhang, A. Rogowski, L. Zhao, M.G. Hahn, U. Avci, J.P. Knox, et al., Understanding how the complex molecular architecture of mannan-degrading hydrolases contributes to plant cell wall degradation, J. Biol. Chem. 289 (2014) 2002–2012.
[20] M. Couturier, J. Féliu, S. Bozonnet, A. Roussel, J. Berrin, Molecular engineering of fungal GH5 and GH26 β-(1,4)-mannanases toward improvement of enzyme activity, PLoS One 8 (2013) e79800.
[21] M. Couturier, A. Roussel, A. Rosengren, P. Leone, H. Stålbrand, J. Berrin, Structural and biochemical analyses of glycoside hydrolase families 5 and 26 β-(1,4)-mannanases from Podospora anserina reveal differences upon manno-oligosaccharide catalysis, J. Biol. Chem. 288 (2013) 14624–14635.
[22] H. Aspeborg, P.M. Coutinho, Y. Wang, H. Brumer, B. Henrissat, Evolution, substrate specificity and subfamily classification of glycoside hydrolase family 5 (GH5), BMC Evol. Biol. 12 (2012).
[23] P.M. Coutinho, M.R. Andersen, K. Kolenova, K. Van, P.A. Van Kuyk, I. Benoit, B.S. Gruben, et al., Post-genomic insights into the plant polysaccharide degradation potential of *Aspergillus nidulans* and comparison to *Aspergillus niger* and *Aspergillus oryzae*, Fungal Genet. Biol. 46 (2009) S161–S169.
[24] A. Rosengren, S.K. Reddy, J.S. Sjöberg, O. Aurelius, D.T. Logan, K. Kolenová, et al., An *Aspergillus nidulans* β-mannanase with high transglycosylation capacity revealed through comparative studies within glycosidase family 5, Appl. Microbiol. Biotechnol. 98 (2014) 10091–10104.
[25] H. Stålbrand, A. Saloheimo, J. Vehmaanperä, B. Henrissat, M. Penttilä, Cloning and expression in *Saccharomyces cerevisiae* of a *Trichoderma reesei* β-mannanase gene containing a cellulose binding domain, Appl. Environ. Microbiol. 61 (1995) 1090–1097.
[26] P. Hägglund, T. Eriksson, A. Collén, W. Nerinckx, M. Claeyssens, H. Stålbrand, A cellulose-binding module of the *Trichoderma reesei* β-mannanase Man5A increases the mannan-hydrolysis of complex substrates, J. Biotechnol. 101 (2003) 37–48.
[27] C. Montanier, B. van, A.L. ueren, C. Dumon, J.E. Flint, M.A. Correia, J.A. Prates, et al., Evidence that family 35 carbohydrate binding modules display conserved specificity but divergent function, Proc. Natl. Acad. Sci. U. S. A. 106 (2009) 3065–3070.
[28] M.A.S. Correia, D.W. Abbott, T.M. Gloster, V.O. Fernandes, J.A.M. Prates, C. Montanier, et al., Signature active site architectures illuminate the molecular basis for ligand specificity in family 35 carbohydrate binding module, Biochemistry 49 (2010) 6193–6205.
[29] X. Li, P. Jackson, D. Rubtsov V, N. Faria-Blanc, J.C. Mortimer, S.R. Turner, et al., Development and application of a high throughput carbohydrate profiling technique for analyzing plant cell wall polysaccharides and carbohydrate active enzymes, Biotechnol. Biofuels 6 (2013) 94.

[30] J. Söding, A. Biegert, A.N. Lupas, The HHpred interactive server for protein homology detection and structure prediction, Nucleic Acids Res. 33 (2005) 244–248.
[31] A.M.D. Gonçalves, C.S. Silva, T.I. Madeira, R. Coelho, S. de, D. anctis, R. San, M. omao, V, et al. Endo-β-D-1,4-mannanase from *Chrysonilia sitophila* displays a novel loop arrangement for substrate selectivity, Acta Crystallogr. Sect. D Biol. Crystallogr. 68 (2012) 1468–1478.
[32] I.W. Davis, A. Leaver-Fay, V.B. Chen, J.N. Block, G.J. Kapral, X. Wang, et al., MolProbity: all-atom contacts and structure validation for proteins and nucleic acids, Nucleic Acids Res. 35 (2007) 375–383.
[33] A. Cartmell, E. Topakas, V.M.A. Ducros, M.D.L. Suits, G.J. Davies, H.J. Gilbert, The *Cellvibrio japonicus* mannanase CjMan26C displays a unique exo-mode of action that is conferred by subtle changes to the distal region of the active site, J. Biol. Chem. 283 (2008) 34403–34413.
[34] E. Sabini, H. Schubert, G. Murshudov, K.S. Wilson, M. Siika-Aho, M. Penttilä, The three-dimensional structure of a *Trichoderma reesei* β-mannanase from glycoside hydrolase family 5, Acta Crystallogr. Sect. D Biol. Crystallogr. 56 (2000) 3–13.
[35] R.J. Williams, J. Iglesias-Fernández, J. Stepper, A. Jackson, A.J. Thompson, E.C. Lowe, et al., Combined inhibitor free-energy landscape and structural analysis reports on the mannosidase conformational coordinate, Angew. Chemie Int. Ed. 53 (2014) 1087–1091.
[36] J. Lehmbeck, F. Wahlbom. Production of a monoclonal antibody in a heterokaryon fungus or in a fungal host cell. WO2005070962 A1, 2005.
[37] S.K. Reddy, A. Rosengren, S. Klaubauf, T. Kulkarni, E.N. Karlsson, R.P. De Vries, et al., Phylogenetic analysis and substrate specificity of GH2 β-mannosidases from *Aspergillus species*, FEBS Lett. 587 (2013) 3444–3449.
[38] M. Lever, A new reaction for colorimetric determination of carbohydrates, Anal Biochem. 47 (1972) 273–279.
[39] B.V. McCleary, Modes of action of β-mannanase enzymes of diverse origin on legume seed galactomannans, Phytochemistry 18 (1979) 757–763.
[40] S.C. Fry, W.S. York, P. Albersheim, A. Darvill, T. Hayashi, J. Joseleau, et al., An unambiguous nomenclature for xyloglucan-derived oligosaccharides, Physiol. Plant 89 (1993) 1–3.
[41] A. Rosengren, P. Hägglund, L. Anderson, P. Pavon-Orozco, R. Peterson-Wulff, W. Nerinckx, et al., The role of subsite +2 of the *Trichoderma reesei* β-mannanase *Tr*Man5A in hydrolysis and transglycosylation, Biocatal. Biotransform. 30 (2012) 338–352.
[42] P. Ademark, J. Lundqvist, P. Hägglund, M. Tenkanen, N. Torto, F. Tjerneld, et al., Hydrolytic properties of a β-mannosidase purified from *Aspergillus niger*, J. Biotechnol. 75 (1999) 281–289.
[43] B.V. McCleary, N.K. Matheson, Action patterns and substrate-binding requirements of β-D-mannanase with mannosaccharides and mannan-type polysaccharides, Carbohydr. Res. 119 (1983) 191–219.
[44] M. Tenkanen, M. Makkonen, M. Perttula, L. Viikari, A. Teleman, Action of *Trichoderma reesei* mannanase on galactoglucomannan in pine kraft pulp, J. Biotechnol. 57 (1997) 191–204.
[45] B.V. McCleary, E. Nurthen, F.R. Taravel, J. Joseleau, Characterisation of the oligosaccharides produced on hydrolysis of galactomannan with β-D-mannanase, Carbohydr. Res. 118 (1983) 91–109.

Paper I

An *Aspergillus nidulans* GH26 endo-*β*-mannanase with a novel degradation pattern on highly substituted galactomannans

---

Supplementary data

Supplementary material:

An *Aspergillus nidulans* GH26 endo-β-mannanase with novel degradation pattern on highly substituted galactomannans

von Freiesleben et al.

Examples of calculations

Example S1

Mw of mannose (standard): 180.16 g/mol
Mannose concentration in stock: 36 g/L
Dilution in assay: 10
Substrate concentration in assay: 2.5 g/L

Standard name	MiliQ water µL	Mannose stock µL	Dilution	Assay mannose g/L	Conversion %
S1	500	500	2	1.8	65
S2	625	375	2.7	1.35	49
S3	750	250	4	0.9	32
S4	875	125	8	0.45	16
S5	937	63	15.9	0.23	8
S6	969	31	32.3	0.11	4
S7	1000	0	0	0	0

Example of calculation of degree of conversion (%) using standard S1:
Standard S1 contains 500 µl MQ water and 500 µl mannose stock with 36 g/L. S1 is diluted 2 times with respect to the stock and additionally 10 times when added to the assay. The amount of mannose in the assay thus is: 36 (g/L)/(2·10) = 1.8 g/L. Degree of conversion is then calcuated as actual yield / theoretical yield ·100 %.
The Theoretical yield is defined as: the substrate concentration (2.5 g/L) corrected for the monomeric units weight when assembled in the polymeric chains (180/162). This will give 1.8/(2.5·(180/162))·100 = 65 %.
By measuring the absorbance of S1 – S7 in the assay at 405 nm a standard curve can be made. The degree of conversion of other samples can be determined by this standard curve. With this method all sugars in the substrate, both mannose and galactose are used. Therefore it will never be possible to reach 100 % conversion.

Tables

Table S2

Settings used to separate APTS labeled saccharides with DNA sequencer-Assisted Saccharide analysis in High throughput (DASH).

Oven temperature	30 °C
Current stability	30 µA
Prerun voltage	15 kV
Prerun time	180 s
Injection voltage	1.2 kV
Injection time	23 s
Voltage number of steps	10
Voltage step interval	20 s
Data delay time	500 s
Run voltage	15 kV
Run time	26.67 min (1600 s)
Ramp delay	1 s

Figures

Lane	Sample
1	*An*Man26A – 5µl
2	LMW – 10µl
3	*Pa*Man26A – 5µl
4	LMW – 10µl
5	*Pa*Man26A core – 5µl
6	LMW – 10µl
7	*Tr*Man5A – 5µl
8	LMW – 10µl
9	*Tr*Man5A core – 5µl
10	LMW – 10µl
11	*An*Man5A – 5µl
12	LMW – 10µl
13	*An*Man5C – 5µl
14	LMW – 10µl

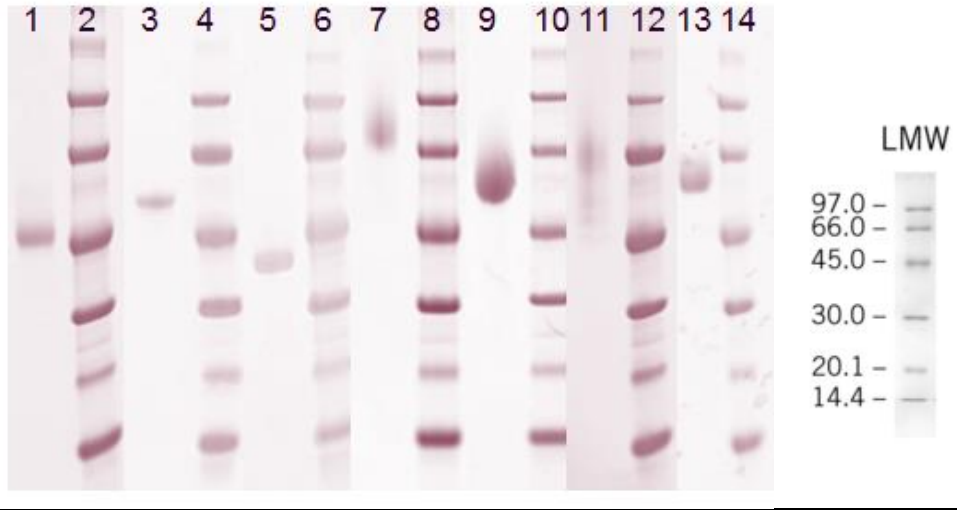


Figure S1. SDS-PAGE gels of the purified enzymes. The protein concentration in the sample was 0.5 mg/ml and prior to loading on the gel the sample was diluted 1:1 with the loading mix. The loading mix was prepared as a 9:1 mix of Novex ® Tris-Glycine SDS Sample Buffer (“X) (Life Technologies and Nupage ® Sample Reducing Agent (10X) (Life Technologies). Please consult the manuscript text for enzyme abbreviations.

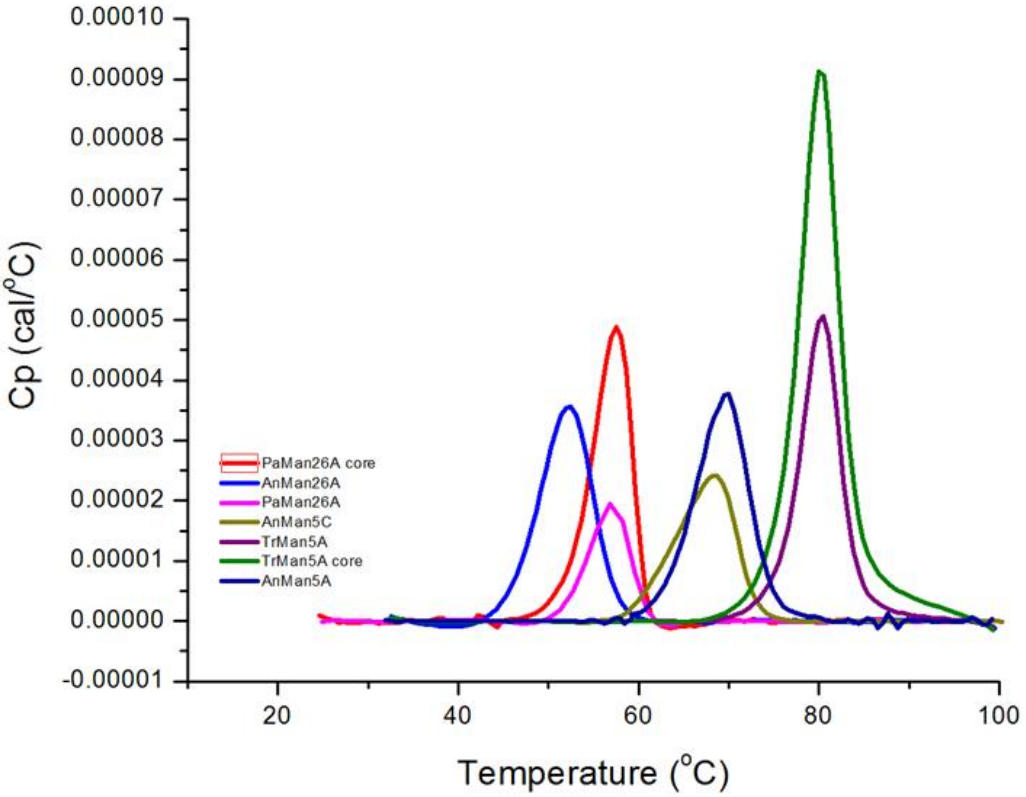


Figure S2. Denaturation peaks illustrating the thermal stability of the purified enzymes. An overlay of denaturation peaks from Differential Scanning Calorimetry (DSC) scans obtained with a constant heating rate of 200 °C/hr. All endomannanases were buffer exchanged to 0.5 mg/ml in 50 mM acetic acid/sodium hydroxide pH 5 before each scan.

Paper II

Boosting of enzymatic softwood saccharification by fungal GH5 and GH26 endomannanases

Biotechnology for Biofuels

RESEARCH Open Access

Boosting of enzymatic softwood saccharification by fungal GH5 and GH26 endomannanases

Pernille von Freiesleben[1,2], Nikolaj Spodsberg[1], Anne Stenbæk[1], Henrik Stålbrand[3], Kristian B. R. M. Krogh[1] and Anne S. Meyer[2*]

Abstract

Background: Softwood is a promising feedstock for lignocellulosic biorefineries, but as it contains galactoglucomannan efficient mannan-degrading enzymes are required to unlock its full potential.

Results: Boosting of the saccharification of pretreated softwood (Canadian lodgepole pine) was investigated for 10 fungal endo-β(1→4)-mannanases (endomannanases) from GH5 and GH26, including 6 novel GH26 enzymes. The endomannanases from *Trichoderma reesei* (*Tres*Man5A) and *Podospora anserina* (*Pans*Man26) were investigated with and without their carbohydrate-binding module (CBM). The pH optimum and initial rates of enzyme catalysed hydrolysis were determined on pure β-mannans, including acetylated and deacetylated spruce galactoglucomannan. Melting temperature (Tm) and stability of the endomannanases during prolonged incubations were also assessed. The highest initial rates on the pure mannans were attained by GH26 endomannanases. Acetylation tended to decrease the enzymatic rates to different extents depending on the enzyme. Despite exhibiting low rates on the pure mannan substrates, *Tres*Man5A with CBM1 catalysed highest release among the endomannanases of both mannose and glucose during softwood saccharification. The presence of the CBM1 as well as the catalytic capability of the *Tres*Man5A core module itself seemed to allow fast and more profound degradation of portions of the mannan that led to better cellulose degradation. In contrast, the presence of the CBM35 did not change the performance of *Pans*Man26 in softwood saccharification.

Conclusions: This study identified *Tres*Man5A as the best endomannanase for increasing cellulase catalysed glucose release from softwood. Except for the superior performance of *Tres*Man5A, the fungal GH5 and GH26 endomannanases generally performed on par on the lignocellulosic matrix. The work also illustrated the importance of using genuine lignocellulosic substrates rather than simple model substrates when selecting enzymes for industrial biomass applications.

Keywords: Glucose release, Endo-β(1→4)-mannanases, GH5, GH26, CBM1 (carbohydrate-binding module 1), CBM35, Galactoglucomannan, Acetylation

*Correspondence: asme@dtu.dk
[2] Protein Chemistry & Enzyme Technology, DTU Bioengineering, Technical University of Denmark, Building 221, 2800 Kgs. Lyngby, Denmark
Full list of author information is available at the end of the article

Background

Softwood has significant potential as feedstock for renewable energy production and biorefining, due to its abundancy, low cost, and lack of competition with the food and feed industry. Enzymatic degradation of softwood to fermentable monomeric sugars is, however, still challenging due to its complex composition and inhomogeneous architecture [1]. Not only lignin but also hemicellulose, β-1,4 mannan and β-1,4 xylan (hereafter named mannan and xylan), prevents enzymatic hydrolysis of cellulose in the absence of relevant accessory enzymes [2, 3]. The hemicelluloses are closely associated with the cellulose fibrils and with lignin [1, 3–5]. The main hemicellulose in softwood is *O*-acetylated galactoglucomannan (Fig. 1), accounting for 15–25% of the wood dry matter [6, 7]. Galactoglucomannan consists of a β-1,4 linked backbone of D-mannopyranosyl and D-glucopyranosyl units. The mannopyranosyl units can be decorated with single α-1,6 linked D-galactopyranosyl residues at C-6 and be *O*-acetylated at C-2 and C-3 [6, 8]. The typical Man:Glc:Gal ratio in Norway spruce galactoglucomannan has been reported to be 3.5–4.5:1:0.5–1.1 with the mannopyranosyl residues being *O*-acetylated to an approximate degree of 0.2–0.3 [9–11]. Variations in the ratios depend on the raw material, but also extraction methods and pretreatment can reduce the amount of backbone decorations [11]. Mannans are not only found as structural units in plant cell walls, but also serve as storage polysaccharides in certain species, e.g., guar gum from the seeds of the guar plant (*Cyamopsis tetragonolobus*, Man:Gal, 2:1), locust bean gum, from the carob tree (*Ceratonia siliqua,* Man:Gal, 4:1) and the glucomannan from the konjac plant (*Amorphophallus konjac,* Man:Glc, 1.6:1) [7, 12].

It requires a coordinated interplay of different types of enzymes to degrade the *O*-acetylated galactoglucomannan found in softwood (Fig. 1). A variety of bacteria, yeasts and filamentous fungi express these mannan-degrading enzymes [13]: endo-β(1→4)-mannanases (endomannanases, EC 3.2.1.78), β-mannosidases (EC 3.2.1.25), β-glucosidases (EC 3.2.1.21), α-galactosidases (EC 3.2.1.22) and acetyl mannan or glucomannan esterases (EC 3.1.1.-) [14, 15]. Also, certain β(1→4)-glucanases (EC 3.2.1.4), primarily attacking the cellulose fraction of the softwood, have been shown to have activity on glucomannans [16].

Endomannanases are important enzymes for facilitating the solubilization and release of mannan from the substrate matrix [3, 17]. Endomannanases are classified into four glycosyl hydrolase (GH) families: 5, 26, 113 and 134 based on sequence similarity [18]. The endomannanases from family 5, 26, and 113, all belonging to clan GH-A, share a $(\beta/\alpha)_8$-TIM barrel fold in their structure, and catalyse the cleavage of the *O*-glycosidic bonds with retention of the anomeric configuration [19–21]. Based on studies of bacterial *Cellvibrio* mannanases, it has been proposed that GH26 enzymes may primarily attack soluble mannans, whilst the GH5 counterparts primarily

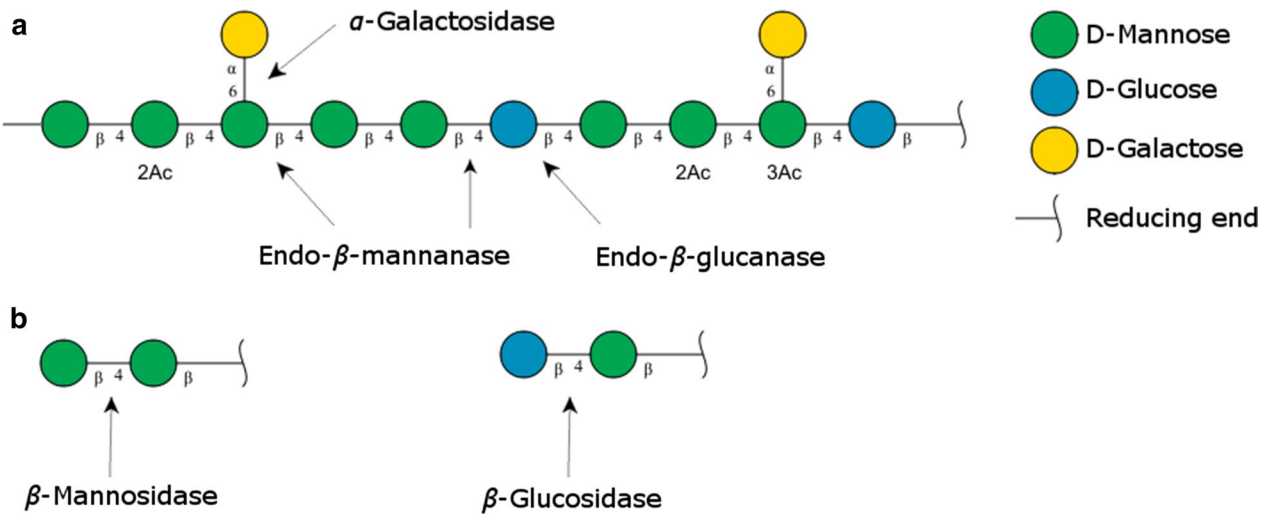


Fig. 1 Schematic illustration of *O*-acetylated galactoglucomannan and enzymes involved in degradation of the backbone (**a**) and released oligosaccharides (**b**). Sugars shown using the Consortium for Functional Glycomics notation [48]. The figure shows general glycoside linkage specificity of each type of enzyme and a given enzyme may be restricted by neighbouring backbone sugar monomers and/or substitutions, exemplified by the varying influence of galactosyl substituents and potentially backbone glucosyl units on mannanase activity [17, 26]. Dual linkage specificity is known to occur among some of the illustrated enzymes. As example, some endo-β(1→4)-glucanases may hydrolyse within the glucomannan backbones, either by action on the glucopyranosyl units or by being unspecific, i.e., hydrolysing at mannopyranosyl units [16]

attack insoluble mannans [22, 23]. However, it is unclear whether this perception is valid for the fungal GH26 endomannanases [24]. Regarding fungal endomannanases, different substrate binding modes were observed for the two *Podospora anserina* endomannanases, *Pans*Man26A and *Pans*Man5A [24, 25]. The *Pans*Man26A together with the GH26 endomannanase from *Aspergillus nidulans*, *Anid*Man26A, were also found to accommodate more galactopyranosyl residues in the active site pocket than their GH5 counterparts [26].

Many fungal GH5 endomannanases are modular, typically having a carbohydrate-binding module from family 1 (CBM1) as part of their structure. CBM1 is known to confer cellulose binding and increase the mannan hydrolysis of complex substrates such as softwood and ivory nut extractions containing both mannan and cellulose [27, 28]. Fungal GH26 endomannanases may have a CBM35 [24, 29], a CBM family known to bind to β-mannans, uronic acids and α-D-galactopyranosyl residues on carbohydrate polymers [30, 31].

The capacity of endomannanases to boost saccharification of softwood to fermentable monomers has been demonstrated and studied mostly with selected fungal GH5 endomannanases. The available literature in particular includes several studies with the *Trichoderma reesei* GH5 endomannanase, (*Tres*Man5A) [3, 4, 17, 32], but also of other endomannanases [33]. A few studies have shown increased glucose release from wood substrates when cellulase (and xylanase) cocktails have been supplemented with fungal GH26 endomannanases [29, 34], but a comparison of the performance of several different endomannanases on the same softwood substrate is not available in the literature. Severe pretreatment methods, leaving only small amounts of mannan in the pretreated substrate, are generally used on softwood to allow enzymatic saccharification. However, as the quest for efficient, yet sustainable utilization of plant biomass increases, new tailor-made pretreatment methods that also maximize the hemicellulose recovery, including mannan recovery, have appeared [35].

Based on the hypothesis that fungal endomannanases differ in their capacity to catalyse removal of galactoglucomannans from cellulose fibrils, and thus in turn may have different effects on enzymatic cellulose saccharification, this study compares 10 fungal endomannanases and their boosting effect on enzymatic cellulose degradation from softwood (the softwood being pretreated lodgepole pine, *Pinus contorta*) with 12% mannan left after pretreatment. The saccharification studies were performed at low temperature (30 °C) to focus on comparing activity of the enzymes, and at 50 °C to mimic industrial saccharification conditions. A subsidiary aim was to address the importance of the CBM35 in softwood saccharification by testing the *Podospora anserina* GH26 endomannanase with and without its CBM35. To touch upon any possible differences in the biological role of the fungal GH5 and GH26 endomannanases, and to assess if any of the substrate preferences on pure mannans could help explain performance differences on the softwood substrate, the initial rate of the studied endomannanases on soluble mannans, including acetylated and deacetylated spruce galactoglucomannan were also determined.

Results and discussion

Based on a phylogenetic sequence comparison of more than 50 fungal GH26 endomannanases, and subsequent recombinant expression assessment, 8 wild type ascomycete GH26 endomannanases were selected for investigation, 6 of them previously uncharacterised (Table 1). Two of the GH26 endomannanases carry both a N-terminal CBM35, a common module among fungal GH26 enzymes [24, 29], as well as a C-terminal CBM1, previously only found in fungal GH5 endomannanases. In addition, two previously characterized GH5 endomannanases from *A. niger* [36] and *T. reesei* [27], respectively, were included in the study. The selected enzymes are all expressed well in the fungal host *Aspergillus oryzae.* The enzymes were all expressed using their native gene sequence and natural signal peptide and purified from the culture broth, the latter indicating that they function as secreted enzymes in nature. To address the influence of their CBMs, the *P. anserina* GH26 and *T. reesei* GH5 enzymes were expressed both with and without their CBM, i.e., CBM35 and CBM1, respectively (Table 1).

Physicochemical properties of the enzymes

The GH26 endomannanases had pH optima in the range of 5–7 and Tm between 50 and 68 °C, with the two wild type core enzymes, *Anid*Man26A and *Ypen*Man26A, having lowest Tm of 53 and 50 °C, respectively (Table 1). Despite the high Tm values, *Ypen*Man26A and *Anid*Man26A had surprisingly low half-lives during prolonged incubation at 30 °C (Table 1). Tm is considered the temperature at which the protein molecule unfolds. However, according to the classical van't Hoff equation and the Arrhenius equation, the equilibrium constant for protein denaturation and the rate of the enzymatic reaction vary with temperature. In practice, this implies that inactivation and rate constant changes caused by altered conformation of enzymes may occur (gradually and slowly) at lower temperatures than the Tm. The net effect is that altered conformation of the enzyme proteins may cause gradual activity loss during prolonged incubation as can be seen in the 30 °C stability data. The GH5 endomannanases had pH optima in the range of 3–6 and appeared more thermally robust than the GH26 endomannanases

Table 1 Properties of the studied endomannanases

Origin	Domains	Mw[a] kDa	pH^{b}_{opt}	Relative activity pH_5/pH_{opt}	Tm[c] °C	$t½$[d] h	Sequence ID[e]	Identity[f] %
GH26								
Collariella virescens (*Cvir*Man26A)	CBM35-GH26-CBM1	57.9	6 (5–7)	0.97	62	–	BBW45415	76
Mycothermus thermophiles (*Mthe*Man26A)	CBM35-GH26	52.1	5 (5–8)	1.00	68	91 ± 0.3	MH208368	76
Podospora anserina (PansMan26A)	CBM35-GH26	49.8	6 (5–7)	0.97	57	90 ± 5.5	B2AEP0, [24]	100
Podospora anserina (*Pans*Man26A core)	GH26	34.4	5 (5–7)	1.00	58	103 ± 2.2	(B2AEP0)	100
Neoascochyta desmazieri (*Ndes*Man26A)	CBM35-GH26	48.7	5 (4–7)	1.00	65	267 ± 24.1	MH208367	60
Westerdykella sp. (*Wsp.*Man26A)	CBM35-GH26	50.4	6 (6–7)	0.79	58	59 ± 4.3	MH208369	55
Ascobolus stictoideus (*Asti*Man26A)	CBM35-GH26-CBM1	59.4	7 (5–7)	0.80	61	81 ± 7.7	BBW45412	55
Aspergillus nidulans (*Anid*Man26A)	GH26	35.2	6 (5–7)	0.93	53	10 ± 0.1	Q5AWB7, [26]	52
Yunnania penicillata (*Ypen*Man26A)	GH26	34.5	6 (5–8)	0.87	50	21 ± 0.1	BDN98740	48
GH5								
Trichoderma reesei (*Tres*Man5A)	GH5-CBM1	45.2	4 (4–5)	0.93	81	–	Q99036, [27]	36
Trichoderma reesei (*Tres*Man5A core)	GH5	38.8	4 (3–6)	0.88	78	2390 ± 360	(Q99036), [27]	36
Aspergillus niger (*Anig*Man5A)	GH5	39.8	4 (3–5)	0.85	87	137 ± 10.0	BCK48306, [36]	30

[a] Theoretical (non-glycosylated protein)

[b] pH optimum at 37 °C and pH interval with 80% relative activity in brackets

[c] The thermal midpoint (Tm) at pH 5

[d] Half-lives ($t½$) at 30 °C. No decay was observed for *Cvir*Man26A and *Tres*Man5A during the 48 h incubation period (Additional file 1: Figure S1)

[e] A reference is given when the enzyme is previously characterized

[f] Homology to *Pans*Man26A for which the structure is known (PDB ID: 3ZM8)

with Tm values ranging from 78 to 87 °C and half-lives above 137 h. For the truncated enzymes, the thermal stability did not seem to be drastically influenced by the lack of the CBM, neither the CBM1 (*Tres*Man5A) or CBM35 (*Pans*Man26A) (Table 1). When compared at 37 °C, all the endomannanases had 79–100% relative activity at pH 5 compared to the activity at their pH optimum.

Initial rates of enzymatic hydrolysis on pure mannans

The initial rates of endomannanase catalysed hydrolysis of locust bean gum, guar gum, konjac glucomannan, as well as acetylated and deacetylated spruce galactoglucomannan were determined at pH 5 and at 37 °C, to assess the activity without confounding effects of differential thermostability of the enzymes (Fig. 2). No β-glucanase cross-activity was found for any of the tested endomannanases, *Cvir*Man26A, *Pans*Man26A, *Pans*Man26A core, *Anig*Man5A, *Tres*Man5A and *Tres*Man5A core (assessed on barley β-glucan and carboxymethyl cellulose).

The endomannanases showed different activity levels on the pure mannan substrates.

GH5 endomannanases tended to exhibit lower initial rates than the GH26, but no consistent discrimination between the enzymes' substrate preferences was evident (Fig. 2). *Ypen*Man26A had a significantly higher initial rate than all the other enzymes on the two galactomannans: locust bean gum (15,050 U/μmole) and guar gum (13,850 U/μmole). In contrast, *Wsp*Man26A had the highest initial rates of all tested endomannanases on the glucomannans: konjac glucomannan (12,550 U/μmole), acetylated galactoglucomannan (8650 U/μmole), and deacetylated galactoglucomannan (9150 U/μmole) (Fig. 2). Deacetylation of galactoglucomannan doubled the rate for a few GH26 endomannanases (*Pans*Man26A and *Cvir*Man26A), but did not generally affect rates or caused rate increase. The lowest initial rates were observed for *Tres*Man5A on the acetylated spruce galactoglucomannan and on konjac glucomannan (on both substrates 600 U/μmole). In general, *Tres*Man5A had lowest initial rates of all the endomannanases on all the tested mannans, irrespective of the presence of the CBM1. The lack of significance of the CBM1 presence is in accordance with *Tres*Man5A CBM1 being known to bind to cellulose and not to mannan [27]. In contrast, the initial rates of *Pans*Man26A containing a CBM35 tended to be higher than those for the *Pans*Man26A core, especially on locust bean gum and konjac glucomannan. A positive effect of the CBM35 may be related to the reported interaction of CBM35 with β-mannans or α-D-galactopyranosyl residues [30, 31]. It seems more likely that the *Pans*Man26A CBM35 interacts with the β-mannan backbone than with the α-D-galactopyranosyl

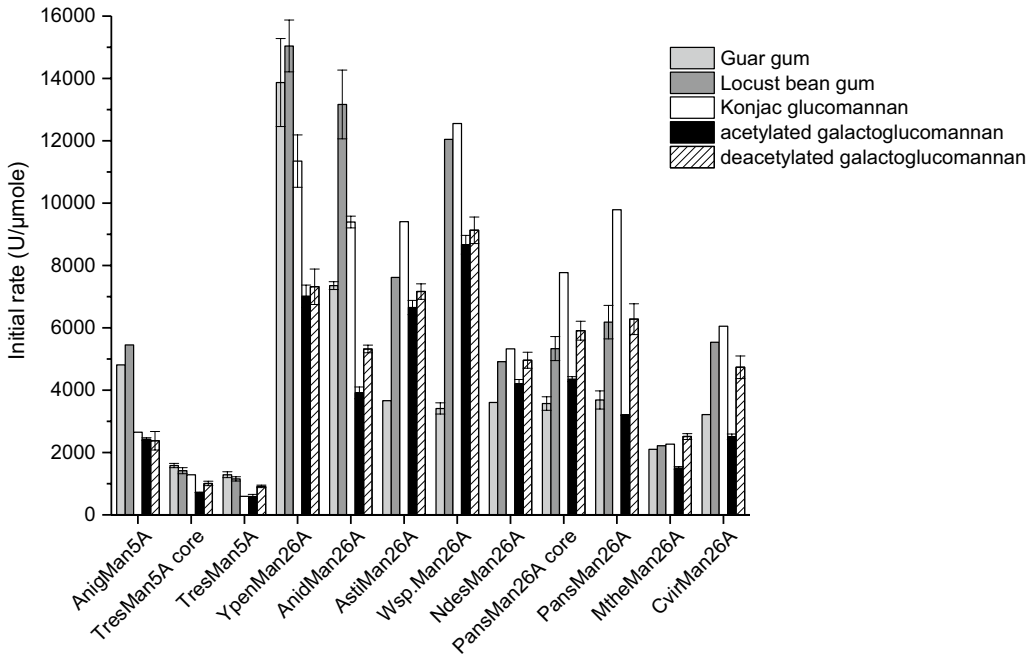


Fig. 2 Initial reaction rates (U/µmole) by endomannanases on pure mannans (37 °C and pH 5). Guar gum (light grey), locust bean gum (dark grey), konjac glucomannan (white), acetylated spruce galactoglucomannan (black) and deacetylated spruce galactoglucomannan (lined). Values are given as mean values ± SD ($n = 1–7$). One-way ANOVA analyses can be seen in Additional file 1: Tables S1 and S2

substitutions, since the positive effect of the CBM was not found on the highly substituted guar gum.

Differences in substrate preferences

The wild type core GH26 endomannanases, *Ypen*Man26A and *Anid*Man26A, had a significantly higher initial rate on locust bean gum (15,050 and 13,150 U/µmole) compared to on konjac glucomannan (11,350 and 9400 U/µmole). All other tested GH26 endomannanases had higher (or similar) initial rates on konjac glucomannan than on locust bean gum, including *Pans*Man26A core for which the CBM35 were removed artificially. *Ypen*Man26A and *Anid*Man26A do not have a CBM, hence the substrate preferences exhibited by these enzymes as compared to those with a CBM35 must be tied to the enzyme core properties rather than to presence of the CBM35 domain. Interestingly, Katsimpouras et al. [29] reported that the GH26 endomannanase with a CBM35 from *Myceliophthora thermophila* had similar substrate preference trends, i.e., showing higher activity on konjac glucomannan compared to locust bean gum.

Like the wild type core GH26 endomannanases, *Tres*Man5A and *Anig*Man5A, had significantly higher initial rates on locust bean gum (1150 and 5450 U/µmole) compared to on konjac glucomannan (600 and 2650 U/µmole). The data for the GH5 enzymes correspond with the substrate preferences reported for *A. nidulans* GH5 endomannanases [37].

Endomannanase performance in softwood saccharification

The efficiency of the 10 endomannanases for saccharification of softwood was assessed by adding equal molar amounts of each endomannanase on top of Cellic® CTec3, where in each case the Cellic® CTec3 had been supplied with a pure GH2 β-mannosidase from *Aspergillus niger* (BM2). When assessed on locust bean gum Cellic® CTec3 itself exerted weak mannan-degrading activity. The endomannanase addition levels were ten times higher than this background activity. The release of glucose, mannose and xylose, respectively, was quantified at 24, 48 and 144 h (Figs. 3, 4 and Additional file 1: Figure S2). BSA was added as a protein control to assure that any differences in release of monosaccharides were not due to increased levels of protein as sometimes observed in lignocellulose hydrolysis (BSA binds non-productively) [38]. In a direct comparison of Cellic® CTec3 with Cellic® CTec3 plus BM2 plus BSA after 24 h hydrolysis, the β-mannosidase itself profoundly increased the release of mannose from 0.07 to 0.5 g/l (0.43 g/l increase) and the release of glucose from 3.06 to 3.18 g/l (0.12 g/l increase) (Fig. 3).

The increased mannose and glucose release is most likely due to BM2 activity on soluble galactoglucomannan

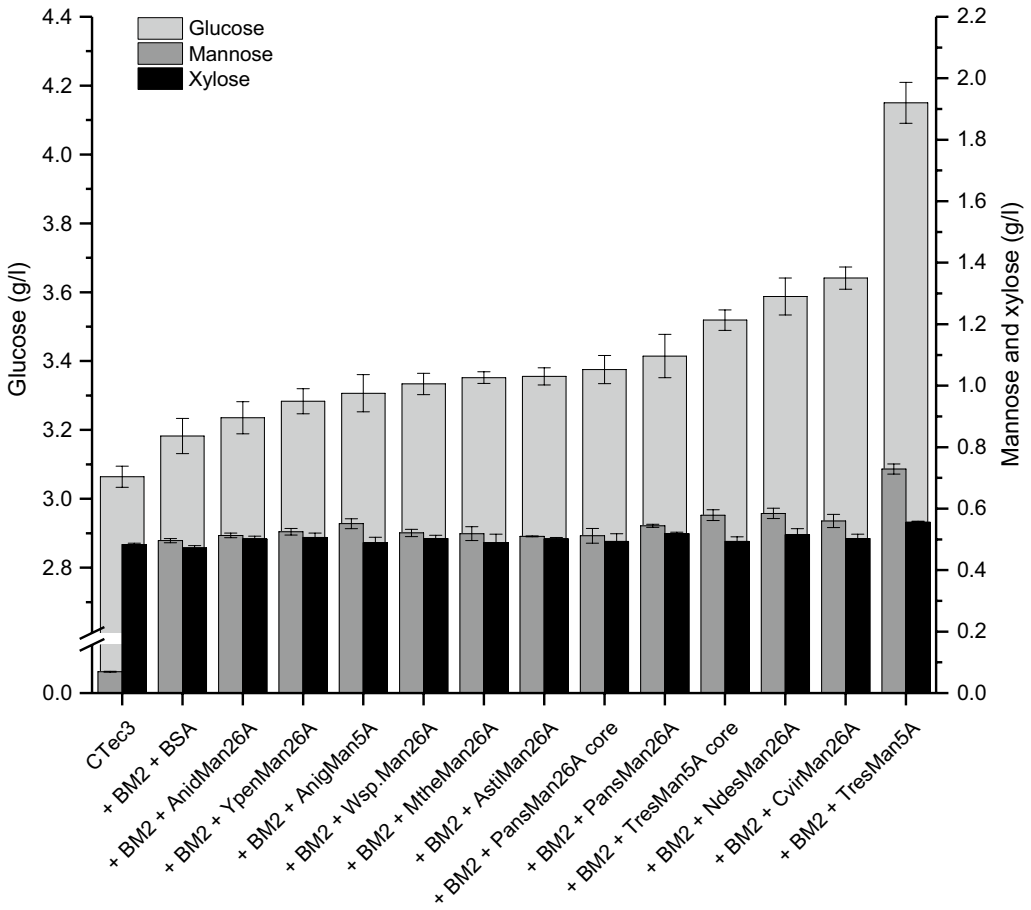


Fig. 3 Softwood saccharification. Endomannanases or BSA were added in equal molar amounts on top of Cellic ® CTec3 plus an *A. niger* GH2 *β*-mannosidase (BM2). Samples were taken after 24 h saccharification at 30 °C. Glucose (g/l, light grey), mannose (dark grey) and xylose (black) yields are given as mean values ± SD ($n = 3$). One-way ANOVA analyses can be seen in Additional file 1: Table S3. Yields from 48 and 144 h can be seen in Additional file 1: Figure S2

oligosaccharides in the mixtures. By removing mannopyranosyl units from the nonreducing end of these oligosaccharides, the BM2 will expose glucopyranosyl residues in the nonreducing end, which can be released by *β*-glucosidase activity from the Cellic® CTec3. The released amount of mannose and glucose upon addition of BM2 corresponds to a Man:Glc ratio of 3.6:1. This ratio is in agreement with reported Man:Glc ratios in softwood (spruce) galactoglucomannans [11], suggesting that no additional cellulose was degraded. The galactoglucomannan oligosaccharides were most likely released by low endomannanase activity present in the Cellic® CTec3 preparation and/or by weak glucomannan degrading capacity by some endoglucanases of this enzyme cocktail (Fig. 1).

Supplementation of Cellic® CTec3 with endomannanase significantly increased the release of glucose for all tested endomannanases, with *Tres*Man5A being the best performing candidate. After 24 h of enzyme treatment, the release of glucose and mannose obtained with the *Tres*Man5A addition was 30% (increase in 1 g/l glucose) and 15% (increase in 0.23 g/l mannose) higher than that of the control (Cellic® CTec3 + BM2 + BSA, Fig. 3), and much higher than those obtained with any of the other endomannanases. The relative amount of released glucose and mannose (Man:Glc, 0.2:1) infer that the released glucose did not derive solely from hydrolysed galactoglucomannan, but also from the cellulose fraction. The increased cellulose degradation obtained with *Tres*Man5A after 24 h, was apparently accompanied by a slightly increased xylose release as well (Additional file 1: Table S3). As discussed below, the overall impression was that the mannose and xylose and the glucose and xylose release were linearly correlated.

The second-best enzyme was *Cvir*Man26A, containing both CBM35 and CBM1, with a glucose yield of 88%

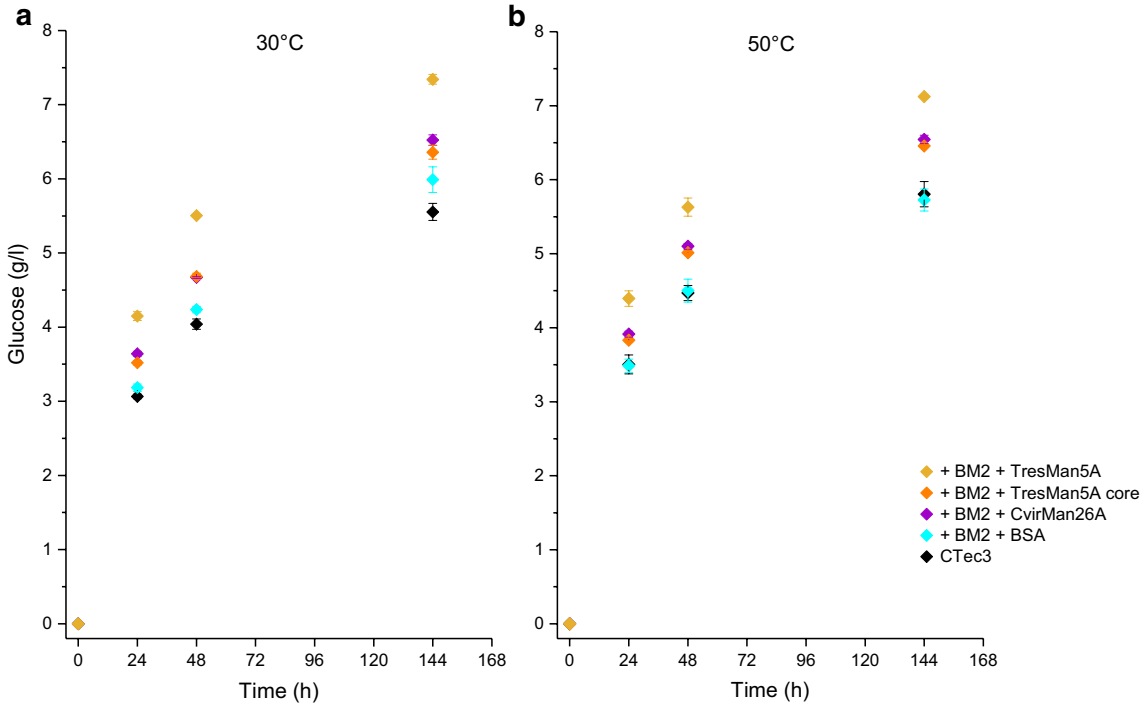


Fig. 4 Softwood saccharification at 30 °C (**a**) and 50 °C (**b**) with *Tres*Man5A, *Tres*Man5A core, *Cvir*Man26A or BSA. Endomannanases or BSA were added in equal molar amounts on top of Cellic ® CTec3 plus an *A. niger* GH2 *β*-mannosidase (BM2). Glucose yields (g/l) are given as mean values ± SD ($n = 3$). One-way ANOVA analyses can be seen in Additional file 1: Table S4

of that obtained by *Tres*Man5A (Fig. 3). The release of glucose and mannose continued throughout the 144 h hydrolysis (Additional file 1: Figure S2).

No obvious trends in the effect of GH5 versus GH26 endomannanases could be discerned. *Tres*Man5A was the superior enzyme, but glucose yields obtained with the other GH5 endomannanase, *Anig*Man5A, were in the low–middle range. For *Tres*Man5A, the presence of CBM1 improved the release of both mannose and glucose. However, both *Cvir*Man26A and *Asti*Man26 with a CBM1 caused release of medium levels of glucose, but there were no evident differences in their mannose release compared to the other endomannanases. No significant effect of the presence of CBM35 was observed for *Pans*Man26A.

Enzyme robustness does not explain the observed difference between the boosting capacity of the investigated endomannanases (Table 1 and Fig. 3). For example, *Tres*Man5A and *Cvir*Man26A had the same stable nature during 48 h incubation at 30 °C (Table 1 and Additional file 1: Figure S1), but differed in their boosting capacity. However, enzyme robustness possibly contributed to some of the observed differences in the enzymes' boosting capacities. For example, the low stability of *Ypen*Man26A and *Anid*Man26A (Table 1 and Additional file 1: Figure S1) may partially explain their poor overall performance in boosting of glucose release from softwood (Fig. 3).

Saccharification at 50 °C

When the two endomannanases, *Tres*Man5A and *Cvir*Man26A, were assessed at 50 °C, it was confirmed that they both boosted the glucose release catalysed by Cellic® CTec3 and that the complete *Tres*Man5A with the CBM1 induced release of significantly more glucose than the *Cvir*Man26A and the *Tres*Man5A core. The time curves of the enzymatic glucose release at 50 and 30 °C were in complete agreement (Fig. 4), and the ranking of the performance of the enzymes was similar at the two reaction temperatures. The lack of increase in hydrolytic rate by cellulases in Cellic® CTec3 with temperature (Q_{10} close to 1 between 30 and 50 °C) is in agreement with data published by Westh et al. [39, 40] who showed that at low Avicel concentrations reduction in substrate affinity caused by heating (increase in K_M) cancels thermoactivation (increase in k_{cat}). In the present study, the substrate concentration was low (2% DM ~ 1% cellulose). The effective accessible substrate concentration may have been lower, because not all cellulose is equally accessible. Since industrial lignocellulose conversion is usually

executed at 50 °C, the data strongly indicate that addition of *Tres*Man5A to commercial cellulase preparations can efficiently boost glucose yields in industrial softwood saccharification reactions.

Correlation between release of glucose, mannose and xylose

With the Cellic ® CTec3 and endomannanase doses used (Table 3), the maximal degree of conversion was approximately 60% of glucose (7.3 g/l of the available 11.5 g/l of glucose were released) after 144 h (Additional file 1: Figure S2). To assess the softwood saccharification at a higher degree of conversion, the enzyme loadings were increased, i.e., addition levels of *Tres*Man5A and *Cvir*-Man26A, respectively, and the Cellic ® CTec3 dose were increased (Table 3). With the high enzyme doses, 85% cellulose, 60% mannan, and 81% xylan conversion were obtained at 30 °C after 144 h.

The data obtained, plus the saccharification results presented in Figs. 3, 4, and Additional file 1: Figure S2, showed a clear linear correlation between the release of glucose, mannose, and xylose throughout the degradation (Fig. 5). These results support the comprehension that the softwood substrate comprises a complex network with glucomannans and xylans located throughout the lignocellulose matrix [3], and not on the outer surface. Their concurrent hydrolysis is crucial to obtain extensive hydrolysis of cellulose and in turn maximize the overall glucose yields from the softwood substrate. A reason for the lower conversion of mannan (approximately 60%) than glucan and xylan (approximately 80%) might be related to the galactose substitutions on galactoglucomannans that hinder the β-mannosidase in fully degrading the released mannooligosaccharides to mannose (Fig. 1), which in turn would explain why not all solubilized galactoglucomannan was analysed as monomers.

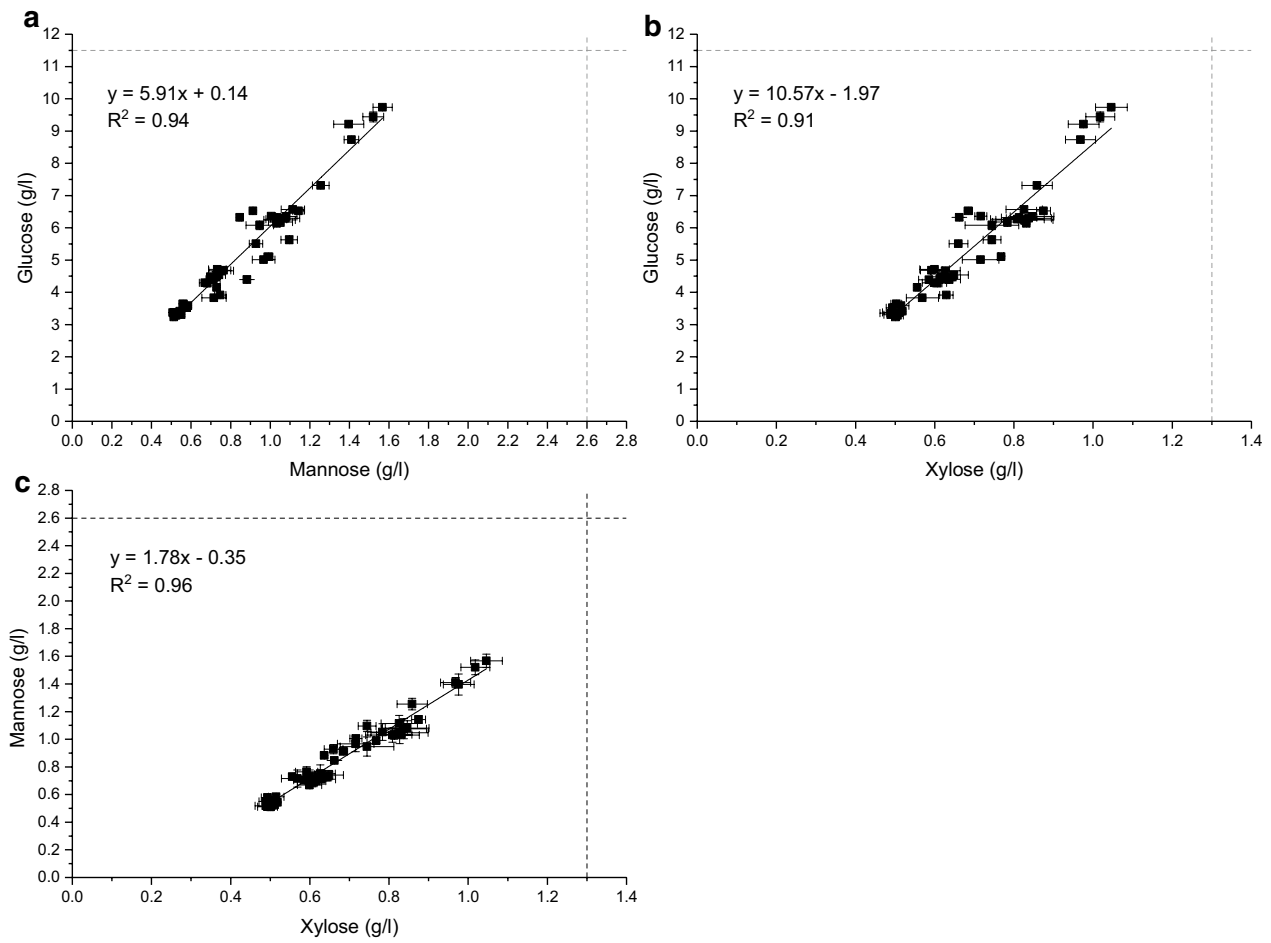


Fig. 5 Correlation between the release (g/l) of glucose and mannose (**a**), glucose and xylose (**b**), and mannose and xylose (**c**) during softwood saccharification. Dashed lines show the theoretical monomeric yield of the substrate based on the chemical composition analysis (Table 2)

Performance in softwood saccharification was not predicted by initial rate

When comparing the initial rates on the soluble model substrates and the boosting effect, the general trend is that the GH5 enzymes show low initial rates but in the case of *Tres*Man5A comparably high boosting effect. For the GH26 enzymes the situation appears more complex, although high initial rate (e.g., *Ypen*Man26A) is not correlated to high boosting effect. Despite its high boosting capacity, *Tres*Man5A was found to have the lowest initial rate on the pure mannans (Fig. 2), demonstrating that performance comparison on these substrates could not predict the efficiency of the enzymes in softwood saccharification. Earlier studies have proposed differences in the biological role for bacterial GH5 and GH26 endomannanases, with the GH5 being optimal for degradation of cell wall mannans [22, 23]. The present study suggests that the GH family categorization and the biological role is not as clear cut for the fungal endomannanases. It was a GH26 endomannanase that performed best on the pure and soluble mannans and a GH5 on the cell wall substrate. However, most of the tested GH26 endomannanases performed on par with, or even better than the *Anig*Man5A. The fact that some fungal GH26 endomannanases are found with a CBM1 and that we have seen no sign of cell wall association for these enzymes, also indicate that the fungal GH26 endomannanases participate in cell wall degradation in nature.

In general, the hydrolysis rate of mannan during softwood saccharification was much lower than the initial hydrolysis rates measured on the pure mannans. During the first 24 h, the mannose release from the softwood substrate by *Tres*Man5A and BM2 reached 0.73 g/l (Fig. 3). If averaging over 24 h, this amount corresponds to a hydrolysis rate on 2.8 U/l (approximately 11 U/µmole *Tres*Man5A). This is likely an overestimated rate for the endomannanase since it attacks the mannan backbone, while it is the surplus of BM2 that causes the release of mannose. On the other hand, the rate is probably not constant throughout the first 24 h, but higher during the initial reaction period. Even if the initial reaction rate for *Tres*Man5A on the softwood substrate was 20 U/µmole, this rate is still 45 times lower than the initial rates obtained for *Tres*Man5A on extracted deacetylated spruce galactoglucomannan at 900 U/µmole (the acetyl moieties are expected to be lost during pretreatment of the softwood). On this insoluble lignocellulosic matrix, the hydrolysis rate of the endomannanases is comparable with the rate of cellulases working on insoluble cellulose [39]. Neither the initial rate itself nor particular substrate preferences of the individually endomannanases with regard to galactose substitutions or acetylation, seemed to determine their performance in softwood saccharification. This in turn means that other properties of the enzymes must be considered.

Explaining the high boosting effect of *Tres*Man5A in softwood saccharification

Our results clearly show that *Tres*Man5A with its CBM1 was the most efficient for softwood saccharification among the tested endomannanases (when added to Cellic ® CTec3). It has previously been observed that hemicellulases from *T. reesei* reduce hemicellulose exposed at the cellulose surface of wood materials to a greater extent than hemicellulases from *Aspergillus* sp. [41], and that *Tres*Man5A catalyses hydrolysis of softwood galactoglucomannan [17]. Our hypothesis for a mechanistic explanation about the additional boosting effect is that *Tres*Man5A catalyses a faster or more profound degradation of a certain type of mannan that is not immediately accessible for the other endomannanases and which moreover, when degraded, allows for a more profound cellulose degradation. Since at least part of softwood mannan is closely associated with cellulose [42], it cannot be ruled out that this is the case also with the pretreated material. The particular portion of the mannan may be a more crystalline part that is more tightly intertwined with the cellulose. It is likely, that the CBM1 in the full-length *Tres*Man5A helps target the cellulose-associated mannan more efficiently than its truncated counterpart lacking CBM1. This view is supported by a previous study of *Tres*Man5A action on cellulose-mannan complexes and the CBM1 cellulose binding capacity [27]. When plant cell wall material are enzymatically degraded by endomannanases and other glycoside hydrolases, CBMs are in general important for ensuring correct positioning and close proximity between enzymes and glycan substrates in turn facilitating enzymatic hydrolysis by the catalytic core modules [22].

Since the *Tres*Man5A core without the CBM1 was among the top performers in the softwood saccharification, the core module itself also played a role for the efficient degradation of the insoluble mannans. The reason that the other GH26 endomannanases with a CBM1, i.e., *Cvir*Man26A and *Ast*Man26, were not releasing the same levels of mannose and glucose as *Tres*Man5A, could be because their core modules are not as optimal as *Tres*Man5A for degradation of mannan associated to cellulose. Another reason might be that their CBM1s have slightly different specificities than the *Tres*Man5A CBM1.

Conclusion

This study strongly confirms that fungal endomannanases differ in their capacity to degrade galactoglucomannan in softwood, and that this degradation contributes significantly to obtain increased enzymatic cellulose

saccharification of softwood. Apart from this main finding, the key novel result was that the well-studied *Tres*Man5A was superior to all other tested GH5 and GH26 endomannanses (also with CBM1 modules) in genuine softwood saccharification, despite being among the slowest on purified mannan substrates. Two GH5 and eight GH26 endomannanases (including six novel endomannanases) were successfully recombinantly expressed, purified, and characterized with focus on their performance in softwood saccharification. The fungal GH26 endomannanases from *Yunnania penicillata* and *Westerdykella* sp. were found to have highest initial rates among the tested enzymes on pure soluble galactomannans and glucomannans, respectively. The acetyl groups on extracted spruce galactoglucomannan tended to decrease the initial enzymatic rates when compared to the initial rates on the deacetylated substrate. However, these initial rates on the pure mannans did not correlate with the results obtained in the extended softwood saccharification reactions. All tested endomannanases caused increased glucose release during softwood saccharification when compared to the glucose release catalysed by Cellic ® CTec3 plus the *A. niger* β-mannosidase alone. However, the GH5 endomannanase from *T. reesei* with a CBM1 produced a markedly higher mannose and glucose release at all time points than all the other tested endomannanases. Based on the data, our hypothesis is that *Tres*Man5A is able to attack an additional portion of mannan in the lignocellulosic matrix, allowing for better cellulose degradation. Both the catalytic efficiency of the core module and the presence of the CBM1 play important roles in the superior performance of this enzyme on softwood. The other nine GH5 and GH26 endomannanases performed on par with the softwood lignocellulosic matrix, giving no clear signs of different biological roles for these fungal endomannanases. Presence of the CBM35 did not change the performance of *Pans*Man26 in softwood saccharification.

The data obtained highlight the problematic strategy of selecting enzymes for industrial applications based on basic characterisation on pure and well-defined substrates. Neither rates nor substrate preferences observed in the basic characterisation correlated with efficient softwood saccharification. Evidently, the data of this study have implications for the selection and use of endomannanases in industrial softwood saccharification applications, especially as new pretreatment methods leaves more hemicellulose in the lignocellulosic matrix after pretreatment.

Methods

Materials

Locust bean gum (low viscosity; borohydride reduced), guar gum (high viscosity), konjac glucomannan (high viscosity), β-glucan (barley; high viscosity), and carboxymethyl cellulose, were purchased from Megazyme (Ireland). Spruce galactoglucomannan (Man:Glc:Gal:Ac, 3.3:1:0.83:1.32) was prepared as described previously [43]. The *O*-acetyl moieties of the *O*-acetylated spruce galactoglucomannan were removed by alkaline hydrolysis in the presence of ammonium hydroxide as described by Jacobs et al. [44]. All other chemicals were from Sigma (Germany).

Expression and purification

The fungal GH26 endomannanases from *Collariella virescens* (*Cvir*Man26A), *Mycothermus thermophiles* (*Mthe*Man26A), *Neoascochyta desmazieri* (*Ndes*Man26A), *Ascobolus stictoideus* (*Asti*Man26A), *Westerdykella* sp. (*Wsp.*Man26A), *Aspergillus nidulans* (*Anid*Man26A), *Yunnania penicillata* (*Ypen*Man26A), *Podospora anserina* (*Pans*Man26A and *Pans*Man26A core), the fungal GH5 endomannanases from *Aspergillus niger* (*Anig*Man5A), and *Trichoderma reesei* (*Tres*Man5A and *Tres*Man5A core) were recombinantly expressed in *Aspergillus oryzae* MT3568an amdS [45]. *Pans*Man26 core and *Tres*Man5A core were expressed without the linker and the N-terminal CBM35 and the C-terminal CBM1, respectively. The enzymes were purified to electrophoretic purity using hydrophobic interaction and ion exchange chromatography (SDS-PAGE gels shown in Additional file 1: Figure S3). The identity of the purified endomannanases was validated with mass spectrometry analysing a tryptic digest of the protein band excised from a SDS-PAGE gel. Protein concentrations were determined by UV absorption at 280 nm using theoretical extinction coefficients (ε). ε at 280 nm of all proteins were estimated by GPMAW 9.20 (Lighthouse Data), and were based on mature proteins without modifications.

pH optimum

The hydrolytic activity was determined at 37 °C, after 15 min, over a pH range from 2.0 to 12.0, with 1 pH unit intervals. The hydrolysis volume was 200 µl, with 2.5 mg/ml locust bean gum in a buffer with 50 mM acetic acid, 50 mM HEPES, 50 mM glycine, 0.01% Trition X-100, 50 mM potassium chloride and 1 mM calcium chloride. The buffer pH was adjusted with sodium hydroxide from pH 2.0–12.0. Released reducing sugars were measured with the 4-hydroxybenzoic acid hydrazide (PAHBAH) method described by Lever [46], with mannose as standard.

Initial rates

The initial rate on locust bean gum, guar gum, konjac glucomannan, acetylated and deacetylated spruce galactoglucomannan, β-glucan, and carboxymethyl cellulose by the endomannanases were determined with 2.5 mg/ml substrate in 50 mM sodium acetate pH 5 at 37 °C. All substrates, except the deacetylated spruce galactoglucomannan and the carboxymethyl cellulose, were soluble at the concentration employed. The deacetylated spruce galactoglucomannan and the carboxymethyl cellulose were only partly soluble, but these substrates were kept in suspension during reaction and sampling via vigorous mixing. Released reducing sugars were measured with the PAHBAH method as described above, except that glucose was used as standard for measurements on β-glucan and carboxymethyl cellulose. All hydrolysis assays were carried out at seven different endomannanase doses as described elsewhere [26]. Initial rates were calculated in the initial linear range of the hydrolysis. To validate that the slope calculation was reproducible, up to seven replicates were done for selected enzymes on selected substrates. The CV was below 10%. The initial rate by the other enzymes was calculated from one slope. One unit (U) was defined as the amount of endomannanase required to release 1 µmole of reducing ends per minute, under the assay conditions specified.

Thermal stability

The thermal stability at pH 5 was investigated with differential scanning calorimetry (DSC) as described elsewhere [26]. The thermal midpoint (Tm) was determined as the top of the protein denaturation peak, and was determined at an accuracy of ± 1 °C. To assay enzyme stability at 30 °C, which was the temperature used in the softwood saccharification reactions, the purified enzymes were incubated individually in triplicates at 30 °C, in 50 mM sodium acetate pH 5, for 24 h and 48 h. Residual activity was determined at 30 °C, pH 5, on locust bean gum with the assay conditions described above.

Biomass and pretreatment

Commercially available grey-stage beetle killed lodgepole pinewood (LPP) chips (*Pinus contorta*) were pretreated, and the chemical composition was analyzed (pretreatment and compositional analysis was done by University of British Colombia, Vancouver, Canada). Wood chips were screened and a size fraction between 2.5 × 2.5 and 5.0 × 5.0 cm was collected and used as feedstock for pretreatment. The pretreatment was performed in a similar manner to the procedure developed by Chandra et al. [35] which was shown to preserve the hemicellulose component in the water insoluble substrate. Prior to steam pretreatment, 200 g of LPP chips (8% moisture) were placed in thermal plastic bags, and mixed with 200 ml of water containing 6% sodium sulfite and 4% sodium carbonate (w/w based on dry wood). The bag of chips was sealed and submerged in a water bath at 60 °C for 12 h. The wet chemical impregnated biomass was then loaded to a 2-l Stake Tech II steam gun (Stake Tech II batch reactor, SunOpta of Norval, ON, Canada) and pretreated at 130 °C for 30 min. After steaming, the biomass remained as chips which were filtered, suspended in 20 l of water and then subjected to mechanical size reduction using a commercial juicer (Angel model 8500). After this process, the sample was filtered with the water insoluble fraction subsequently characterized for its chemical composition (Table 2) by the NREL method [47] and then used for enzymatic hydrolysis experiments.

Enzymes for softwood hydrolysis

The applied enzymes were the purified endomannanases (see section about "Expression and purification"), a purified *A. niger* GH2 β-mannosidase (BM2, UNIPROT A2QWU9), and a commercially available cellulase- and xylanase-rich enzyme cocktail (Cellic® CTec3). Except for the endomannanases, which were purified in this study, the enzymes were kindly provided by Novozymes. The applied enzyme doses can be seen in Table 3. Cellic®

Table 2 Chemical composition of the softwood dry matter (DM)

Softwood component	Percent of DM (%)
Arabinose	1.3
Galactose	1.3
Glucose	52
Xylose	5.5
Mannose	11.7
Lignin	27

Table 3 Doses of enzymes, enzyme cocktail and BSA in softwood saccharification

Set-up[a]	CTec3[b]	BM2	Endomannanase or BSA
	mg EP/g DM	mg EP/g DM	mol/g DM
(1)	10	1	1.26×10^{-8}
(2)	50	1	1.26×10^{-7}

[a] Two set-ups were used: comparing endomannanases at 30 and 50 °C (1). Increased enzyme doses to evaluate saccharification at a higher degree of conversion (2)

[b] Cellic® CTec3 and the *A. niger* GH2 β-mannosidase (BM2) doses are given as mg enzyme protein (EP)/g dry matter (DM) and not as mg product

CTec3 is a *Trichoderma*-based product with different recombinant enzymes. One can assume that the mannan-degrading activity from Cellic® CTec3 is equivalent to *Tres*Man5A. Cellic® CTec3 supplies approximately 1.3×10^{-9} mol/g dry matter (DM) endomannanase (*Tres*Man5A) to the softwood hydrolysis, when added in a concentration of 10 mg Cellic® CTec3/g DM. The lowest addition of purified endomannanase was 1.26×10^{-8} mol/g DM.

Enzymatic hydrolysis of pretreated softwood

Enzymatic hydrolysis of the pretreated softwood was performed in 50 ml falcon tubes with two metal balls (9 mm). Softwood was added to give 0.4 g dry matter per tube (resulting in 2% dry matter) along with sodium acetate buffer, pH 5, and proxel to give concentrations of 50 mM and 0.2%, respectively, in the final mixture. Milli-Q water was added to give a total reaction mass of 20 g after addition of the required amount of enzyme. The tubes were incubated in a heated (30 or 50 °C) 20 cm diameter drum, rotating at 20 rpm. All experiments were performed in triplicates and run for 144 h. At sampling, 2 ml representative whole slurry sample was transferred to Eppendorf tubes and centrifuged 18,213*g* for 10 min. The liquid was decanted, filtered through a 0.45 and a 0.22 µm filter and kept for sugar analysis on high-performance liquid chromatography (HPLC) for separation and quantification of glucose amount and by 1-phenyl-3-methyl-5-pyrazolone (PMP) derivatization followed by reverse phase ultra-performance liquid chromatography (UPLC) for separation and quantification of mannose and xylose amounts.

Analysis of sugar release

The amount of glucose was analysed by a Waters HPLC system coupled with a refractive index detector and equipped with a Aminex HPX 87H column (300 by 7.8 mm). 10 µl sample was injected and separation was performed at 65 °C with a flow rate of 0.6 ml/min 5 mM H_2SO_4. Calibration curves for glucose were plotted (Empower) and used to estimate the amount of glucose released.

The derivatization reaction, for analysing the amount of mannose and xylose, was performed with 200 µl sample in an appropriate concentration, 20 µl 6-deoxy-D-glucose as internal standard, 20 µl 4 M NaOH, and 200 µl 0.5 M PMP in aqueous methanol. The mixtures were mixed well and incubated at 70 °C for 30 min, cooled to room temperature and mixed with 20 µl 4 M HCL and 400 µl methanol. The separation and quantification of mannose and xylose was analysed by a Waters Acquity UPLC system coupled with a UV (245 nm) detector and equipped with a Waters Acquity CSH C18 column (dimensions: 150 × 2.1 mm, particle size: 1.7 µm and pore size: 130 Å). 3 µl sample was injected and separation was performed at 65 °C with a flow rate of 0.5 ml/min. A two-eluent system was used, (A) 0.15% formic acid in MiliQ water and (B) 0.15% formic acid in ACN with the following gradient: 0 min, 83:17 (% A:B); 1 min, 83:17 (% A:B); 10 min, 77.2:22.8 (% A:B); 10.5 min, 5:95 (% A:B); 11 min, 83:17 (% A:B). The total run time per injection was 13 min. Calibration curves for mannose and xylose were plotted (Empower3) and used to estimate the amount of released mannose and xylose. All results are expressed in g/l.

Additional file

Additional file 1: Table S1 and S2. Ratings from one-way ANOVA analyses of initial reaction rates by endomannanases on pure mannans (Fig. 2). A one-way ANOVA analysis was made for all endomannanases on each substrate (top, Table S1) and for each enzyme across all substrates (bottom, Table S2). Ratings are assigned with a 95 % confidence interval with the Tukey–Kramer method in SASjmp. Please consult the manuscript for enzyme abbreviations. **Table S3.** Ratings from one-way ANOVA analyses of sugar release during softwood saccharification at 30 °C (Fig. 3 and Additional file 1: Figure S2). A one-way ANOVA analysis was made for each sugar at each time point. Ratings are assigned with a 95 % confidence interval with the Tukey–Kramer method in SASjmp. Please consult the manuscript for enzyme abbreviations. **Table S4.** Ratings from one-way ANOVA analyses of glucose release during softwood saccharification at 30 and 50 °C (Fig. 4). A one-way ANOVA analysis was made for each temperature at each time point. Ratings are assigned with a 95 % confidence interval with the Tukey–Kramer method in SASjmp. Please consult the manuscript for enzyme abbreviations. **Figure S1.** Stability of endomannanases at 30 °C. The relative activity (%) was determined on locust bean gum at 30 °C, pH 5 and are given as mean values ± SD (n=3). Please consult the manuscript for enzyme abbreviations. **Figure S2.** Softwood saccharification. Endomannanases or BSA were added in equal molar amounts on top of Cellic ® CTec3 plus an *A. niger* GH2 *β*-mannosidase (BM2). Samples were taken after 24, 48 and 144 h saccharification at 30 °C. Glucose (g/l, light grey), mannose (dark grey) and xylose (black) yields are given as mean values ± SD (n=3). One-way ANOVA analyses can be seen in Additional file 1: Table S3. **Figure S3.** SDS-PAGE gels of the purified enzymes. The protein concentration in the samples was 0.5 mg/ml. Prior to gel loading, samples were diluted 1:1 with loading mix. Loading mix was prepared as a 9:1 mix of Novex ® Tris-Glycine SDS Sample Buffer (2X) (Life Technologies) and Nupage ® Sample Reducing Agent (10X) (Life Technologies). Please consult the manuscript for enzyme abbreviations. Samples with *Ndes*Man26A and *Mthe*Man26A both contain molecules with and without the CBM35.

Authors' contributions

PvF, HS, KBK, and AM planned the study, analysed the results, and wrote the manuscript. NS did the phylogenetic sequence comparison, cloning, and expression of the GH26 and GH5 endomannanases. PvF performed the experiments and drafted the manuscript. AS participated in the data analysis. All authors read and approved the final manuscript.

Author details

[1] Novozymes A/S, Krogshøjvej 36, 2880 Bagsværd, Denmark. [2] Protein Chemistry & Enzyme Technology, DTU Bioengineering, Technical University of Denmark, Building 221, 2800 Kgs. Lyngby, Denmark. [3] Department of Biochemistry and Structural Biology, Center for Molecular Protein Science, Lund University, PO Box 124, 221 00 Lund, Sweden.

Acknowledgements
We thank Richard Chandra and Prof. Jack Saddler at UBC, Vancouver, Canada for providing the pretreated lodgepole pine.

Competing interests
PvF, NS, AS, and KBK are employed at Novozymes A/S. All authors declare that they have no competing interests.

Availability of data and materials
All data generated and analysed during this study are included in the manuscript and its additional file in form of graphs and tables.

Consent for publication
All authors agree to the submission.

Ethics approval and consent to participate
Not applicable.

Funding
This study was partially funded by the BioValue SPIR Strategic Platform for Innovation and Research on value added products from biomass, which is co-funded by The Innovation Fund Denmark, Case No: 0603-00522B. Henrik Stålbrand was partially funded by FORMAS (213-2014-1254) and the Swedish Foundation for Strategic Research (RBP 14-0046).

Publisher's Note

Received: 29 April 2018 Accepted: 21 June 2018
Published online: 17 July 2018

References

1. Khatri V, Meddeb-Mouelhi F, Beauregard M. New insights into the enzymatic hydrolysis of lignocellulosic polymers by using fluorescent tagged carbohydrate-binding modules. Sustain Energy Fuels. 2018;2:479–91.
2. Berlin A, Gilkes N, Kilburn D, Bura R, Markov A, Skomarovsky A, et al. Evaluation of novel fungal cellulase preparations for ability to hydrolyze softwood substrates—evidence for the role of accessory enzymes. Enzyme Microb Technol. 2005;37:175–84.
3. Várnai A, Huikko L, Pere J, Siika-Aho M, Viikari L. Synergistic action of xylanase and mannanase improves the total hydrolysis of softwood. Bioresour Technol. 2011;102:9096–104.
4. Eriksson Ö, Goring DAI, Lindgren BO. Structural studies on the chemical bonds between lignins and carbohydrates in spruce wood. Wood Sci Technol. 1980;14:267–79.
5. Barros J, Serk H, Granlund I, Pesquet I. The cell biology of lignification in higher plants. Ann Bot. 2015;115:1053–74.
6. Timell TE. Recent progress in the chemistry of wood hemicelluloses. Wood Sci Technol. 1967;1:45–70.
7. Moreira LRS, Filho EXF. An overview of mannan structure and mannan-degrading enzyme systems. Appl Microbiol Biotechnol. 2008;79:165–78.
8. Xu C, Leppänen AS, Eklund P, Holmlund P, Sjöholm R, Sundberg K, et al. Acetylation and characterization of spruce (*Picea abies*) galactoglucomannans. Carbohydr Res. 2010;345:810–6.
9. Lundqvist J, Teleman A, Junel L, Zacchi G, Dahlman O, Tjerneld F, et al. Isolation and characterization of galactoglucomannan from spruce (*Picea abies*). Carbohydr Polym. 2002;48:29–39.
10. Willför S, Sjöholm R, Laine C, Roslund M, Hemming J, Holmbom B. Characterisation of water—soluble galactoglucomannans from Norway spruce wood and thermomechanical pulp. Carbohydr Polym. 2003;52:175–87.
11. Bååth JA, Abad AM, Berglund J, Larsbrink J, Vilaplana F, Olsson L. Mannanase hydrolysis of spruce galactoglucomannan focusing on the influence of acetylation on enzymatic mannan degradation. Biotechnol Biofuels. 2018;11:114.
12. Katsuraya K, Okuyama K, Hatanaka K, Oshima R, Sato T, Matsuzaki K. Constitution of konjac glucomannan: chemical analysis and 13C NMR spectroscopy. Carbohydr Polym. 2003;53:183–9.
13. Srivastava PK, Kapoor M. Production, properties, and applications of endo-β-mannanases. Biotechnol Adv. 2017;35:1–19.
14. Gilbert HJ, Stålbrand H, Brumer H. How the walls come crumbling down: recent structural biochemistry of plant polysaccharide degradation. Curr Opin Plant Biol. 2008;11:338–48.
15. Malgas S, van 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. 2015;31:1167–75.
16. Mikkelson A, Maaheimo H, Hakala TK. Hydrolysis of konjac glucomannan by *Trichoderma reesei* mannanase and endoglucanases Cel7B and Cel5A for the production of glucomannooligosaccharides. Carbohydr Res. 2013;372:60–8.
17. Tenkanen M, Makkonen M, Perttula M, Viikari L, Teleman A. Action of *Trichoderma reesei* mannanase on galactoglucomannan in pine kraft pulp. J Biotechnol. 1997;57:191–204.
18. Lombard V, Ramulu HG, Drula E, Coutinho PM, Henrissat B. The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res. 2014;42:D490–5.
19. Sinnott ML. Catalytic mechanisms of enzymic glycosyl transfer. Chem Rev. 1990;90:1171–202.
20. Henrissat B, Callebaut I, Fabrega S, Lehn P, Mornon J, Davies G. Conserved catalytic machinery and the prediction of a common fold for several families of glycosyl hydrolases. Proc Natl Acad Sci USA. 1995;92:7090–4.
21. Withers SG. Mechanisms of glycosyl transferases and hydrolases. Carbohydr Polym. 2001;44:325–37.
22. Zhang X, Rogowski A, Zhao L, Hahn MG, Avci U, Knox JP, et al. Understanding how the complex molecular architecture of mannan-degrading hydrolases contributes to plant cell wall degradation. J Biol Chem. 2014;289:2002–12.
23. Tailford LE, Ducros VMA, Flint JE, Roberts SM, Morland C, Zechel DL, et al. Understanding how diverse β-mannanases recognize heterogeneous substrates. Biochemistry. 2009;48:7009–18.
24. Couturier M, Roussel A, Rosengren A, Leone P, Stålbrand H, Berrin J. Structural and biochemical analyses of glycoside hydrolase families 5 and 26 β-(1,4)-mannanases from *Podospora anserina* reveal differences upon manno-oligosaccharide catalysis. J Biol Chem. 2013;288:14624–35.
25. Marchetti R, Berrin J, Couturier M, Qader SAU, Molinaro A, Silipo A. NMR analysis of the binding mode of two fungal endo-β-1,4-mannanases from GH5 and GH26 families. Org Biomol Chem. 2016;14:314–22.
26. von Freiesleben P, Spodsberg N, Blicher TH, Anderson L, Jørgensen H, Stålbrand H, et al. An *Aspergillus nidulans* GH26 endo-β-mannanase with a novel degradation pattern on highly substituted galactomannans. Enzyme Microb Technol. 2016;83:68–77.
27. Hägglund P, Eriksson T, Collén A, Nerinckx W, Claeyssens M, Stålbrand H. A cellulose-binding module of the *Trichoderma reesei* β-mannanase Man5A increases the mannan-hydrolysis of complex substrates. J Biotechnol. 2003;101:37–48.
28. Pham TA, Berrin JG, Record E, To KA, Sigoillot JC. Hydrolysis of softwood by *Aspergillus* mannanase: role of a carbohydrate-binding module. J Biotechnol. 2010;148:163–70.
29. Katsimpouras C, Dimarogona M, Petropoulos P. A thermostable GH26 endo-β-mannanase from *Myceliophthora thermophila* capable of enhancing lignocellulose degradation. Appl Microbiol Biotechnol. 2016;100:8385–97.
30. Montanier C, van Bueren AL, Dumon C, Flint JE, Correia MA, Prates JA, et al. Evidence that family 35 carbohydrate binding modules display conserved specificity but divergent function. Proc Natl Acad Sci USA. 2009;106:3065–70.
31. Correia MAS, Abbott DW, Gloster TM, Fernandes VO, Prates JAM, Montanier C, et al. Signature active site architectures illuminate the molecular basis for ligand specificity in family 35 carbohydrate binding module. Biochemistry. 2010;49:6193–205.
32. Rättö M, Siika-aho M, Buchert J, Valkeajävi A, Viikari L. Enzymatic hydrolysis of isolated and fibre-bound galactoglucomannans from pine-wood and pine kraft pulp. Appl Microbiol Biotechnol. 1993;40:449–54.
33. Inoue H, Yano S, Sawayama S. Effect of β-mannanase and β-mannosidase supple-mentation on the total hydrolysis of softwood polysaccharides by the *Talaromyces cellulolyticus* cellulase system. Appl Biochem Biotechnol. 2015;176:1673–86.

34. Couturier M, Haon M, Coutinho PM, Henrissat B, Lesage-Meessen L, Berrin J. *Podospora anserina* hemicellulases potentiate the *Trichoderma reesei* secretome for saccharification of lignocellulosic biomass. Appl Environ Microbiol. 2011;77:237–46.
35. Chandra RP, Chu QL, Hu J, Zhong N, Lin M, Lee JS, et al. The influence of lignin on steam pretreatment and mechanical pulping of poplar to achieve high sugar recovery and ease of enzymatic hydrolysis. Bioresour Technol. 2016;199:135–41.
36. Jørgensen H, Sanadi AR, Felby C, Lange NEK, Fischer M, Ernst S. Production of ethanol and feed by high dry matter hydrolysis and fermentation of palm kernel press cake. Appl Biochem Biotechnol. 2010;161:318–32.
37. Dilokpimol A, Nakai H, Gotfredsen CH, Baumann MJ, Nakai N, Hachem MA, et al. Recombinant production and characterisation of two related GH5 endo-β-1,4-mannanases from *Aspergillus nidulans* FGSC A4 showing distinctly different transglycosylation capacity. Biochim Biophys Acta. 2011;1814:1720–9.
38. Eriksson T, Börjesson J, Tjerneld F. Mechanism of surfactant effect in enzymatic hydrolysis of lignocellulose. Enzyme Microb Technol. 2002;31:353–64.
39. Sørensen TH, Cruys-Bagger N, Windahl MS, Badino SF, Borch K, Westh P. Temperature effects on kinetic parameters and substrate affinity of Cel7A cellobiohydrolases. J Biol Chem. 2015;290:22193–202.
40. Westh P, Borch K, Sørensen T, Tokin R, Kari J, Badino S, et al. Thermoactivation of a cellobiohydrolase. Biotechnol Bioeng. 2018;115:831–8.
41. Bombeck PL, Khatri V, Meddeb-Mouelhi F, Montplaisir D, Richel A, Beauregard M. Predicting the most appropriate wood biomass for selected industrial applications: comparison of wood, pulping, and enzymatic treatments using fluorescent-tagged carbohydrate-binding modules. Biotechnol Biofuels. 2017;10:293.
42. Åkerholm M, Salmén L. Interactions between wood polymers studied by dynamic FT-IR spectroscopy. Polymer (Guildf). 2001;42:963–9.
43. Andersson A, Persson T, Zacchi G, Stålbrand H, Jönsson A-S. Comparison of diafiltration and size-exclusion chromatography to recover hemicelluloses from process water from thermomechanical pulping of spruce. Appl Biochem Biotechnol. 2007;137:971–83.
44. Jacobs A, Lundqvist J, Stålbrand H, Tjerneld F, Dahlman O. Characterization of water-soluble hemicelluloses from spruce and aspen employing SEC/MALDI mass spectroscopy. Carbohydr Res. 2002;337:711–7.
45. Lehmbeck J, Wahlbom F. Production of a monoclonal antibody in a heterokaryon fungus or in a fungal host cell. WO2005070962 A1; 2005.
46. Lever M. A new reaction for colorimetric determination of carbohydrates. Anal Biochem. 1972;47:273–9.
47. Sluiter A, Hames B, Ruiz R, Scarlata C, Sluiter J, Templeton D, et al. Determination of structural carbohydrates and lignin in Biomass. Technical Report: NREL/TP-510-42618; 2008.
48. Raman R, Venkataraman M, Ramakrishnan S, Lang W, Raguram S, Sasisekharan R. Advancing glycomics: implementation strategies at the consortium for functional glycomics. Glycobiology. 2006;16:82R–90R.

Paper II

Boosting of enzymatic softwood saccharification by fungal GH5 and GH26 endomannanases

---

Supplementary data

Additional material

Boosting of enzymatic softwood saccharification by fungal GH5 and GH26 **endomannanases**
von Freiesleben et al.

Table S1 and S2 Ratings from one-way ANOVA analyses of initial reaction rates by endomannanases on pure mannans (Figure 2). A one-way ANOVA analysis was made for all endomannanases on each substrate (top, Table S1) and for each enzyme across all substrates (bottom, Table S2). Ratings are assigned with a 95 % confidence interval with the Turkey-Kramer method in SASjmp. Please consult the manuscript for enzyme abbreviations.

Table S1

Endomannanases	Guar gum	Locust bean gum	Konjac glucomannan	Acetylated GGM [a]	Deacetylated GGM [a]
*Ypen*Man26A	A	A	A	B	B
*Wsp.*Man26A	C, D	B	A	A	A
*Anid*Man26A	B	B	A, B	C	D, E
*Asti*Man26A	C, D	C	A, B, C, D	B	B, C
PansMan26A	C	C	A, B, C	C, D	B, C, D
*Pans*Man26A core	C, D	C	A, B, C, D	C	C, D, E
*Ndes*Man26A	C, D	C, D	A, B, C, D	C	E
CvirMan26A	C, D	C	A, B, C, D	D, E	E
*Anig*Man5A	C	C	B, C, D	D, E	F
*Mthe*Man26A	D, E	D, E	B, C, D	E, F	F
*Tres*Man5A core	E	E	C, D	F	G
TresMan5A	E	E	D	F	G

Table S2

Endomannanases	Guar gum	Locust bean gum	Konjac glucomannan	Acetylated GGM [a]	Deacetylated GGM [a]
*Ypen*Man26A	A, B	A	B	C	C
*Wsp.*Man26A	C	A	A	B	B
*Anid*Man26A	B, C	A	B	D	C, D
*Asti*Man26A	C	A, B	A	B	A, B
PansMan26A	C	B	A	C	B
*Pans*Man26A core	C	B	A	B, C	B
*Ndes*Man26A	A	A	A	A	A
CvirMan26A	A, B	A	A	B	A
*Anig*Man5A	A, B	A	B, C	C	C
*Mthe*Man26A	A, B	A	A	B	A
*Tres*Man5A core	A	A, B	B, C	D	C
TresMan5A	A	A	B, C	C	B

[a] spruce galactoglucomannan (GGM)

Table S3 Ratings from one-way ANOVA analyses of sugar release during softwood saccharification at 30 °C (Figure 3 and Figure S2). A one-way ANOVA analysis was made for each sugar at each time point. Ratings are assigned with a 95 % confidence interval with the Turkey-Kramer method in SASjmp. Please consult the manuscript for enzyme abbreviations.

Samples	24h			48h			144h		
	Glucose	Mannose	Xylose	Glucose	Mannose	Xylose	Glucose	Mannose	Xylose
+ *BM2* + *Tres*Man5A	A	A	A	A	A	A	A	A	A
+ *BM2* + *Cvir*Man26A	B	B, C	B, C	B, C	B	A, B	B, C	A, B	A
+ *BM2* + *Ndes*Man26A	B	B	A, B	B	B	A, B	B	A, B, C	A, B
+ *BM2* + *Tres*Man5A core	B, C	B	B, C	B	B	A, B	B, C, D	B, C, D	B
+ *BM2* + PansMan26A	C, D	B, C, D	A, B	B, C, D	B	A	B, C, D, E	B, C, D	A, B
+ *BM2* + *Pans*Man26A core	D	D, E	B, C	C, D	B	A, B	B, C, D, E	B, C, D	A, B
+ *BM2* + *Asti*Man26A	D, E	D, E	B, C	D, E	B	A, B	B, C, D	B, C, D	A, B
+ *BM2* + *Mthe*Man26A	D, E	C, D, E	B, C	D, E, F	B	A, B	B, C, D, E	B, C, D	A, B
+ *BM2* + *Wsp.*Man26A	D, E	C, D, E	B, C	D, E, F, G	B	A, B	B, C, D, E	B, C, D	A, B
+ *BM2* + *Anig*Man5A	D, E	B, C, D	B, C	D, E, F	B	A, B	C, D, E	B, C, D	A, B
+ *BM2* + *Ypen*Man26A	D, E, F	C, D, E	B, C	E, F, G	B	A, B	C, D, E	B, C, D	A, B
+ *BM2* + *Anid*Man26A	E, F	C, D, E	B, C	F, G	B	A, B	D, E	D	A, B
+ *BM2* + *BSA*	F	E	C	G	B	A, B	E	C, D	A, B
CTec3	G	F	B, C	H	C	B	F	E	A, B

Table S4 Ratings from one-way ANOVA analyses of glucose release during softwood saccharification at 30 and 50 °C (Figure 4). A one-way ANOVA analysis was made for each temperature at each time point. Ratings are assigned with a 95 % confidence interval with the Turkey-Kramer method in SASjmp. Please consult the manuscript for enzyme abbreviations.

Samples	24 h		48 h		144 h	
	30 °C	50 °C	30 °C	50 °C	30 °C	50 °C
+ *BM2* + *TresMan5A*	A	A	A	A	A	A
+ *BM2* + *CvirMan26A*	B	B	B	B	B	B
+ *BM2* + *TresMan5A core*	B	B	B	B	B	B
+ *BM2* + *BSA*	C	C	C	C	C	C
CTec3	D	C	D	C	D	C

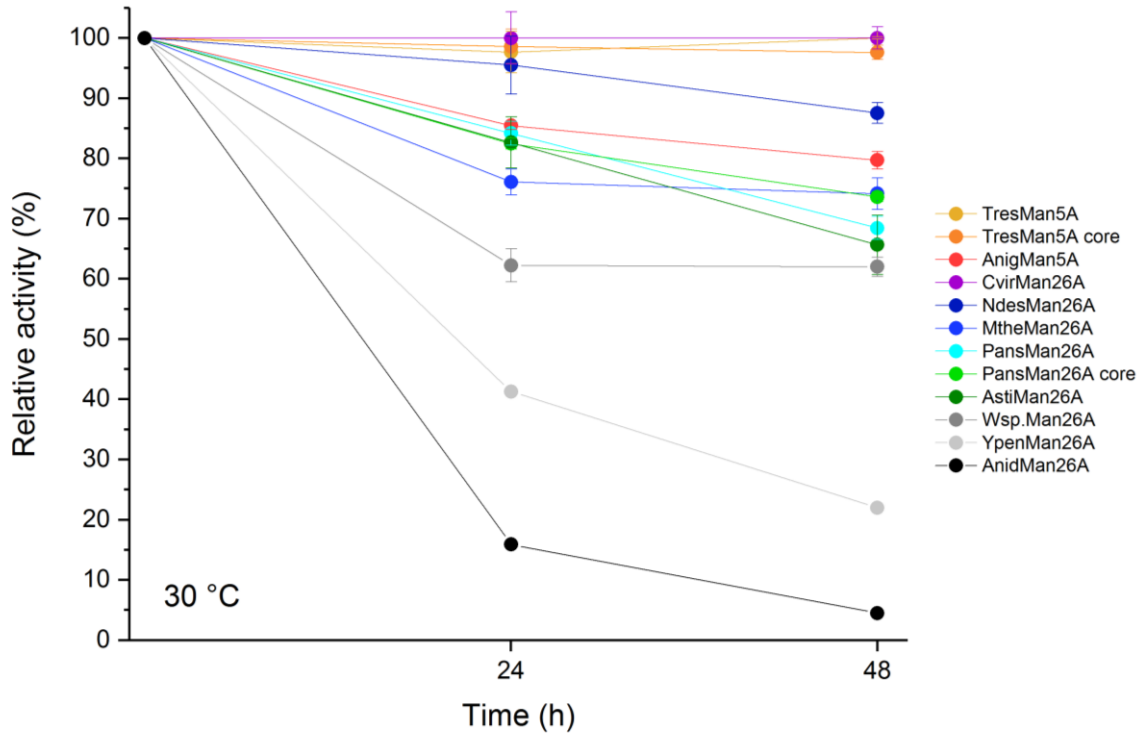


Figure S1 Stability of endomannanases at 30 °C. The relative activity (%) was determined on locust bean gum at 30 °C, pH 5 and are given as mean values ± SD (n=3). Please consult the manuscript for enzyme abbreviations.

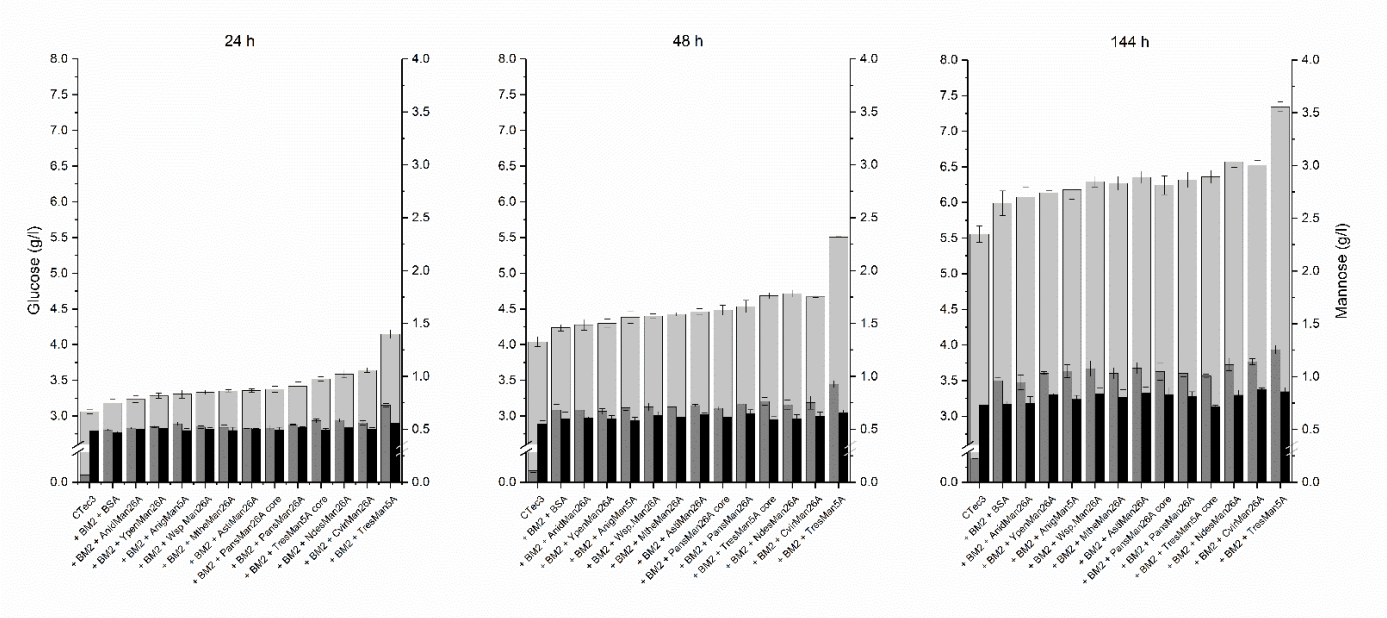


Figure S2 Softwood saccharification. Endomannanases or BSA were added in equal molar amounts on top of Cellic ® CTec3 plus an *A. niger* GH2 β-mannosidase (BM2). Samples were taken after 24, 48 and 144 h saccharification at 30 °C. Glucose (g/l, light grey), mannose (dark grey) and xylose (black) yields are given as mean values ± SD (n=3). One-way ANOVA analyses can be seen in Table S3.

Lane	Sample
1	*Anid*Man26A – 5µl
2	LMW – 10µl
3	*Pans*Man26A – 5µl
4	LMW – 10µl
5	*Pans*Man26A core – 5µl
6	LMW – 10µl
7	*Tres*Man5A – 5µl
8	LMW – 10µl
9	*Tres*Man5A core – 5µl
10	LMW – 10µl
11	*Anig*Man5A – 5 µl
12	LMW – 10µl

Lane	Sample
1	*Cvir*Man26A – 5µl
2	LMW – 10µl
3	*Asti*Man26A – 5µl
4	LMW – 10µl
5	*Ypen*Man26A – 5µl
6	LMW – 10µl
7	*Wsp*Man26A – 5µl
8	Mark 12 – 10µl
9	*Ndes*Man26A – 5µl
10	Mark 12 – 10µl
11	*Mthe*Man5A – 5 µl
12	Mark 12 – 10µl

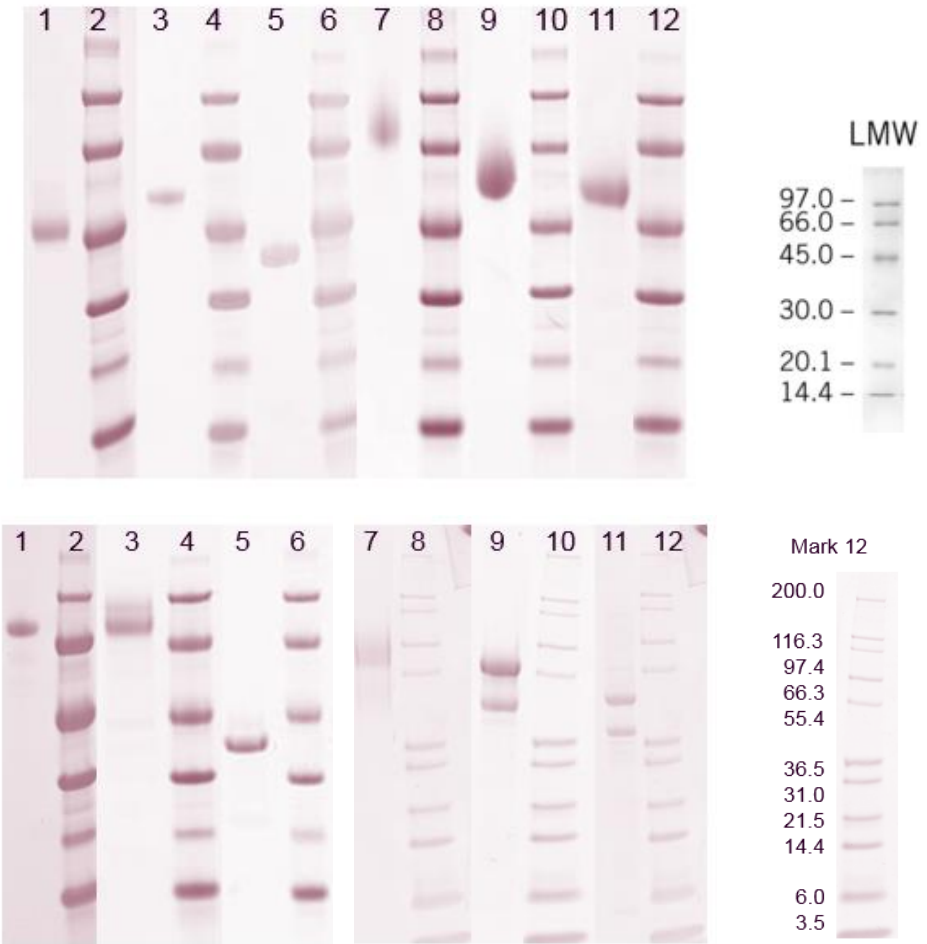


Figure S3 SDS-PAGE gels of the purified enzymes. The protein concentration in the samples was 0.5 mg/ml. Prior to gel loading, samples were diluted 1:1 with loading mix. Loading mix was prepared as a 9:1 mix of Novex ® Tris-Glycine SDS Sample Buffer (2X) (Life Technologies) and Nupage ® Sample Reducing Agent (10X) (Life Technologies). Please consult the manuscript for enzyme abbreviations. Samples with *Ndes*Man26A and *Mthe*Man26A both contain molecules with and without the CBM35.

Paper III

Crystal structure and substrate interactions of an unusual fungal non-CBM carrying GH26 endo-β-mannanase from *Yunnania penicillata*

Crystal structure and substrate interactions of an unusual fungal non-CBM carrying GH26 endo-β-mannanase from *Yunnania penicillata*

Pernille von Freiesleben [a, b], Olga V. Moroz [c], Elena Blagova [c], Mathias Wiemann [d], Nikolaj Spodsberg [a], Jane W. Agger [b], Gideon J. Davies [c], Keith S. Wilson [c, *], Henrik Stålbrand [d], Anne S. Meyer [b, *] and Kristian B. R. M. Krogh [a]

[a] Novozymes A/S, Krogshøjvej 36, 2880 Bagsværd, Denmark

[b] DTU Bioengineering, Department of Biotechnology and Biomedicine, Building 221, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark

[c] York Structural Biology Laboratory, Department of Chemistry, University of York, York YO10 5DD. U.K.

[d] Department of Biochemistry and Structural Biology, Center for Molecular Protein Science, Lund University, PO Box 124, SE-221 00 Lund, Sweden

* Corresponding authors. *E-mail:* asme@dtu.dk (Anne S. Meyer) and keith.wilson@york.ac.uk (Keith S. Wilson)

Abstract

Endo-β(1→4)-mannanases (endomannanases) catalyse degradation of β-mannans, an abundant class of plant polysaccharides. This study investigates structural features and substrate binding of *Ypen*Man26A, a non-CBM carrying endo-mannanase from *Yunnania penicillata.* Structural and sequence comparisons to other fungal family GH26 endomannanases showed high sequence similarities and conserved binding residues, indicating that fungal GH26 endomannanases accommodate galactopyranosyl units in the -3 and -2 subsite. Two striking amino acid differences in the active site were found when the *Ypen*Man26A structure was compared to a homology model of *Wsp.*Man26A from *Westerdykella* and the sequences of nine other fungal GH26 endo-mannanases. Two *Ypen*Man26A mutants W110H and D37T, inspired by differences observed in the *Wsp.*Man26A, produced a shift in how mannopentaose bound across the active site cleft and a decreased affinity for galactose in the -2 subsite, respectively, compared to the *Ypen*Man26A. The *Ypen*Man26A was moreover found to have a flexible surface loop in the position where *Pans*Man26A from *Podospora anserina* has an α-helix (α9) which interacts with its CBM of family 35. A sequence alignment inferred that the core structure of fungal GH26 endomannanases differ depending on the natural presence of this type of CBM. These new findings have implications for selecting and optimising enzymes for galactomannan-degradation.

Keywords: GH26, kinetics, substrate binding, galactomannan, CBM35, mass spectrometry

Introduction

Endo-β(1→4)-mannanases (endomannanases, EC 3.2.1.78) are important enzymes, catalysing the degradation of abundant plant β-mannans (hereafter mannan) in nature. Endomannanases are also industrially relevant enzymes used for various applications including plant biomass conversion [1], food and feed [2,3], detergent formulations [4] and oil drilling [5]. An understanding of the intimate interactions between endomannanases and their substrates is key to optimising their utilisation and industrial performance. Mannan is an abundant hemicellulose in nature, found as a structural unit, primarily in the cell wall of softwood. Mannans also serve as storage polysaccharides in some seeds [6]. Mannans are composed of a linear backbone containing D-mannopyranosyl residues (linear mannans) or D-mannopyranosyl and D-glucopyranosyl residues in an alternating manner (glucomannans) linked by β-(1→4)-linkages. The backbone can be decorated with α-(1→6)-linked D-galactopyranosyl groups (galactomannans or galactoglucomannans) and acetyl groups [6–8] (examples of galactomannans are shown in Figure 1). Large amounts of mannans are found in the secondary plant cell wall of softwood (coniferous trees), where acetylated galactoglucomannans comprise approximately 25 % of the wood dry matter [9–11]. Guar gum, produced from the seeds of the guar plant (*Cyamopsis tetragonolobus*) and locust bean gum, produced from the seeds of the carob tree (*Ceretonia siliqua*) are significant sources of galactomannans. Guar gum contains more galactopyranosyl groups (Gal:Man, 1:2) than locust bean gum (Gal:Man, 1:4) [6]. In locust bean gum, the distribution of galactopyranosyl side-groups is irregular with a high proportion of unsubstituted blocks, whereas in guar gum, the

galactopyranosyl groups are more ordered and found mainly in pairs and triplets with few non-substituted regions [12] (Figure 1).

(Please insert Figure 1 around here)

Endomannanases are the main enzymes which catalyse depolymerisation of mannan. They produce mannooligosaccharides which are further processed by e.g. the exo-acting β-mannosidases and α-galactosidases. Soluble substrates are often accessible to all these enzymes, but the mannan attack by endomannanases may also occur on water-insoluble substrate matrixes [1,2,14]. Endomannanases are classified into four glycoside hydrolase (GH) families: 5, 26, 113 and 134 based on sequence similarity [15]. Endomannanases from families 5, 26, and 113 belong to clan GH-A and share the $(\beta/\alpha)_8$-TIM barrel fold, catalytic machinery and catalyse the cleavage of the *O*-glycosidic bonds in the mannan backbone with net retention of the anomeric configuration [16–18]. The newly identified GH134 endomannanases have a lysozyme-like fold and catalyse the hydrolysis of the mannan backbone via an inverting mechanism [19]. Fungal endomannanases known to date are predominantly categorised as GH5 with a few as GH26. Although several GH26 endomannanases from different organisms have been characterised (e.g. *Cfim*Man26A from *Cellulomonas fimi* (2BVY) [20], *Cjap*Man26A (1J9Y) [21] and *Cjap*Man26C (2VX6) [22] from *Cellvibrio japonicus*, *Bova*Man26A (4ZXO) and *Bova*Man26B from *Bacteroides ovatus* [23] and *Rspe*Man26A from a symbiotic protist of the termite *Recticulitermes speratus* (3WDR) [24]), fewer studies have focused on the GH26 fungal enzymes and only one crystal structure is available, namely that of *Pans*Man26A from *Podospora anserina*, 3ZM8 [25], which carries a family 35 carbohydrate-binding module (CBM35). *Pans*Man26A and the GH26 endomannanase from *Aspergillus nidulans*, *Anid*Man26A, were

shown to have a significant -4 subsite, and to accommodate galactopyranosyl units not only in the -1 subsite, but also in the -2 and +1 subsites, in contrast to the GH5 counterparts from *A. nidulans Anid*Man5A and *Anid*Man5C [25–27]. Several fungal GH26 endomannanases were found to have higher initial rates on soluble galactomannans than the tested GH5 endomannanases, with the GH26 endomannanase from *Yunnania penicillata, Ypen*Man26A, having the highest initial hydrolysis rate, closely followed by *Anid*Man26A and the GH26 endomannanase from *Westerdykella sp*, *Wsp.*Man26A [1]. However, the tested fungal GH26 endomannanases discriminated differently between the soluble mannans [1], exemplified by the *Ypen*Man26A and the *Wsp.*Man26A which both had high initial hydrolysis rates on locust bean gum, but behaved differently on the more heavily substituted galactomannan. While *Ypen*Man26A also showed high hydrolysis rate on guar gum, *Wsp.*Man26A appeared more restricted by the extra galactose substitutions.

Most fungal GH26 endomannanases have a CBM35 [25,27,28]; a CBM family known to include members that bind β-mannans, uronic acids, β-1,3-galactan or α-1,6-galactopyranosyl residues on carbohydrate polymers [29,30]. The binding site of CBM35s has been reported to be located in between the loops connecting the β-strands and not on the concave surface presented by the β-strands as observed in CBM1[29,30].

In the present study, the Michaelis-Menten kinetic parameters for *Ypen*Man26A was determined, the crystal structure in complex with a galactomanno-oligosaccharide was solved and the amino acids involved in substrate interactions identified. The structure of this unusual fungal wild type enzyme with no CBM35 was compared to the known *Pans*Man26A structure harbouring a CBM35 and by sequence alignment to seven other fungal GH26

endomannanases. The roles of selected substrate binding amino acids, were evaluated from two *Ypen*Man26A mutants, D37T and W110H. The mutations were inspired by the sequence of *Wsp.*Man26A, an endomannanase seemingly more restricted by galactose substitutions than *Ypen*Man26A.

Results

Yunnania penicillata possesses at least one protein with endomannanase activity [1] (Genbank ID: MH899111). This enzyme, studied in the current paper, has a signal peptide and a GH26 catalytic domain, but no CBM, in contrast to most known fungal GH26 endomannanases which carries a CBM35 [1,25,28]. A gene encoding the catalytic domain, named *Ypen*Man26A, was synthesised and expressed in *Aspergillus oryzae*. Based on a sequence alignment with the sequence of *Pans*Man26A, the two catalytic residues (previously identified for GH26 enzymes [31,32]), Glu165 and Glu257 in *Ypen*Man26A were identified, with Glu257 being the nucleophile, performing the nucleophilic attack on an anomeric carbon in the mannan backbone, and Glu165 the acid/base, which serves as proton donor and later deprotonates the glycosyl acceptor in the first and second step of the retaining catalytic mechanism respectively [16,33]. This mechanism is characteristic for Clan-GHA glycosyl hydrolases, such as GH26 endomannanases [16]. The Michaelis-Menten kinetic parameters with locust bean gum and guar gum were determined for *Ypen*Man26A. Interestingly, the k_{cat} on guar gum (636 s^{-1}) was found to be higher than the k_{cat} on locust bean gum (475 s^{-1}). Former studies reports a decrease in hydrolytic rate of endomannanases going from less to more substituted galactomannans, such as from locust bean gum to guar gum [20,23,34]. It is thought

that the galactose substitutions cause steric hindrance, making the mannan backbone less accessible for the enzyme [6,35]. As expected, the K_M was also higher on guar gum (2.2 mg/ml) than on locust bean gum (0.6 mg/ml) and the k_{cat}/K_M therefore lower on guar gum (289 ml/(mg*s)) than on locust bean gum (792 ml/(mg*s)) (the kinetic parameters are listed in Table 4). Motivated by the desire to see how this enzyme accommodates and interacts with the galactopyranosyl groups in galactomannan, we sought to determine the crystal structure of *Ypen*Man26A in complex with a galactomanno-oligosaccharide. A *Ypen*Man26A acid/base substituted variant, E165Q, was made using synthetic oligonucleotides and PCR, replacing the codon GAG at position 165 with CAG. The variant was synthesised and expressed in *Aspergillus oryzae*. *N*-Deglycosylation of the purified wild type and the E165Q *Ypen*Man26A mutant using Endoglycosidase H, resulted in a small shift (~ 5 kDa) in the apparent molecular mass on SDS-PAGE (Figure S1). These results confirm that *Ypen*Man26A is *N*-glycosylated, in agreement with the GPMAW (Lighthouse data) prediction.

Structure of *Ypen*Man26A

The structure of the deglycosylated *Ypen*Man26A acid/base substituted variant E165Q, in complex with a α-6^2-6^1-di-galactosyl-mannotriose (MGG), was solved by molecular replacement using the known structure of *Pans*Man26A [25] as template. The enzyme was refined at 1.36 Å resolution (Table 1). Neither the active *Ypen*Man26A nor the E165Q mutant crystallized as apoenzymes, suggesting that ligand binding resulted in increased stability and/or conformational changes leading to successful crystallogenesis. The *Ypen*Man26A chain can be traced from Ala1 to Val312 without breaks, and forms a $(\beta/\alpha)_8$-barrel fold (Figure 2A) as

expected. The active site was identified in the groove with the conserved catalytic residue Glu165 (acid/base) mutated to Gln, and the conserved catalytic residue Glu257 (nucleophile) [31,32] (Figure 2A), equivalent to those observed in *Pans*Man26A [25]. The activity-crippled *Ypen*Man26A E165Q variant showed an initial rate of hydrolysis of locust bean gum of 40 U/μmole enzyme, roughly 400-fold lower than the wild type enzyme (15050 U/μmole). The low residual activity may be a consequence of the acid/base and not the nucleophile being substituted or the fact that the E165Q variant has made with a single base change from codon GAG (Glu) to CAG (Gln), in which case activity can be translational incorporated. There is a single *N*-glycosylation site at Asn103, located on the external side of the barrel, with a remaining *N*-acetylglucosamine (GlcNac). As expected, *Ypen*Man26A shows the highest structural similarity to other endomannanases (from both fungal, bacterial and protists origin) in family GH26 (Table 2). Judged from the Z-score (used by the DALI protein structure comparison server [36] for ranking of structural matches) *Ypen*Man26A has the greatest structural similarity to *Pans*Man26A (3ZM8) followed by *Rspe*Man26C (3WRD) (Table 2).

(Please insert Figure 2 around here)

Table 1: Data collection and refinement statistics of *Ypen*Man26A

Data set[a]	**MGG - *Yp*Man26 E165Q**
PDB code	6HPF
Data collection	
Beamline	I04, Diamond, 2017.09.18
Space group	$P6_522$
Unit-cell parameters (Å)	a = 98.99, b = 98.99, c = 170.50
Resolution range (Å)	34.22 – 1.36 (1.38 – 1.36)
No. of reflections	1924268
Unique reflections	105928
Completeness (%)	100 (100)
$CC_{1/2}$	1 (0.894)
Multiplicity	18.2 (18.5)
$\langle I/\sigma(I)\rangle$	20.7(1.2)
R_{merge}	0.058 (1.242)
$R_{r.i.m.}$[b]	0.061 (1.313)
Refinement statistics	
Percentage of R_{free} reflections	4.97
(%)R_{cryst}=Σ||Fo| - |Fc||/Σ|Fo|(%)	12.2
Free R factor (%)	14.4
Bond distances (Å)	0.017 (0.020)
Bond angles (°)	1.72 (1.92)
Chiral centres (Å^3)	0.118 (0.200)
Planar groups (Å)	0.014 (0.021)
Average B value protein (Å^2)	18
Average B value ligand (Å^2)	24
Average B value water (Å^2)	35
Molprobity score	0.81
Ramachandran favoured	97.4
Ramachandran outliers	0.37

[a] Values in parentheses correspond to the highest resolution shell. [b] Estimated $R_{r.i.m.} = R_{merge}[N/(N-1)]^{1/2}$, where N is the data multiplicity, and R_{merge} is defined as $\Sigma|I - \langle I\rangle|/\Sigma I$, where I is the intensity of the reflection. [c] CC(1/2) values for Imean are calculated by splitting the data randomly in half. [d] Ramachandran plot analysis was carried out using Molprobity [38].

Table 2: The five closest structural matches to *Ypen*Man26A, calculated using the DALI protein structure comparison server [36] (excluding duplicates).

Enzyme	**PDB code**	**Z-score**	**R.m.s.d. (Å)**	**Sequence identity (%)**	**Residues aligned**
PansMan26A, Podospora anserina [25]	3ZM8	48.1	1.0	46	309/444
*Rspe*Man26C, *Reticulitermes speratus* [24]	3WDR	42.0	1.5	36	298/330
*Bsub*Man26A, *Bacillus subtilis* [39]	2WHK	33.1	2.1	27	276/332
BCMan, *Bacillus subtilis* [40]	2QHA	33.1	2.1	27	276/336
Bacillus subtilis	3CBW	33.1	2.0	27	275/336

Ligand binding to *Ypen*Man26A

Crystals of *Ypen*Man26A E165Q were obtained in the presence of α-6^4-6^3-di-galactosyl-mannopentaose (MGGMM) with the aim that the oligosaccharide would span the catalytic site. However, the electron density of the ligand was modelled as MGG situated in the -4 to -2 subsites (Figure 2C), suggesting that the residual activity of the E165Q variant has caused hydrolysis of the ligand between the backbone monomers in the -1 and +1 subsites, whereafter MGG has migrated to span the subsite -4 to -2, indicating high ligand affinity in these subsites. The electron density of MGG is clear and unambiguous, except for the galactopyranosyl unit in the -3 subsite, which points out of the binding cleft (Figure 2B). The B values for the galactopyranosyl residue in the -3 subsite are also higher (between 34 – 63 Å^2 for the C atoms), than for the galactopyranosyl unit in the -2 subsite (between 17 - 30 for the C atoms) or for the mannopyranosyl moieties (between 14 - 28 for the C atoms). All the interactions between the enzyme and the ligand are clearly defined, except for the flexible galactopyranosyl unit. There is electron density present near the mutated Q165, which is separate from the ligand, and described as acetate, which fits the density well. There was no obvious acetate present in crystallisation conditions, but most probably, it was a contaminant during purification or crystallisation, or it was present in the cell growth media, similar to the unknown ligand described as propionate in 5G4Z [41].

Like *Pans*Man26A, *Ypen*Man26A has eight large loops that form a deep cleft at the active centre and are involved in binding of the substrate: loop 1 (36-39), loop 2 (60-73), loop 3 (95-131), loop 4 (166-179), loop 5 (207-211), loop 6 (227-235), loop 7 (259-263), and loop 8(279-

291). The -1 and +1 subsites of *Ypen*Man26A are similar to other fungal and bacterial GH26 endomannanases (e.g. *Pans*Man26A, *Cjap*Man26A, *Cfim*Man26A [20,21,25]) with the conserved residues His164, Trp170, Phe171, Tyr227, Trp279 (Figure 2C). As described for the homologous enzymes[20,21,25], *Ypen*Man26A Tyr227 is involved in a hydrogen bond with the catalytic nucleophile Glu257 whilst the aromatic amino acids Trp170 and Trp279 stabilise the mannopyranose rings at the -1 and +1 subsites, respectively (Figure 2C). Like *Pans*Man26A, *Ypen*Man26A displays a prominent -4 subsite, with stacking interactions between the mannopyrannose ring and the two aromatic residues W109 and W110 and hydrogen bonds between Asp61, Arg66 and the mannopyrannose ring (Figure 2C). The -3 subsite appears more weakly bound as judged from the ligand enzyme interactions. In the -2 subsite the two aromatic residues, Phe113 and Tyr114, equivalent to Phe248 and Tyr249 in *Pans*Man26A, stabilise the interactions with the mannopyranose unit. Previously, enzyme interactions with a galactopyranosyl substituent attached to a mannopyranosyl unit within the -1 subsite of *Cjap*Man26C has been described [22]. Interestingly, because of the captured ligand in the present work, it is possible to identify interactions between the galactopyranose unit and the *Ypen*Man26A in the -2 subsite not previously described. Gln36, Asp37, and Asp58 are involved in hydrogen bonds with the galactose residue. Asp37 has a double conformation in the crystal structure, possibly because the amino acid conformation shifts upon ligand binding. *Pans*Man26A has a Glu172 instead of the Asp37 in *Ypen*Man26A, but otherwise the enzymes have essentially identical environments for interactions with the galactose residue. Out of the six closest structural matches (Table 2), only *Pans*Man26A (3ZM8) accommodates galactopyranosyl residues in the -2 subsite as *Ypen*Man26A. A surface view of *Ypen*Man26A

and *Cjap*Man26C (2VX6) with their ligands superimposed (the MGG from *Ypen*Man26A and a bound α-6^3-galactosyl-mannotetraose (MGMM) in the -2 to +2 subsite of *Cjap*Man26C) shows that the ligands overlap nicely. The data thus indicate accommodation of galactopyranosyl residues in the -3, -2 and -1 subsites of both enzymes (Figure S2). These superimpositions show that *Cjap*Man26C does not accommodate the galactopyranosyl unit in the – 2 subsite, where the moiety is pointing into the enzyme structure, whereas *Ypen*Man26A accommodates the galactopyranosyl moieties in the -3, -2 and -1 subsites (Figure S2). The data also show that *Ypen*Man26A has a more open active site than *Cjap*Man26C (Figure S2).

Design of two *Ypen*Man26A variants – inspired by *Wsp*.Man26A

A sequence similarity search with the *Ypen*Man26A sequence, using the NCBI protein-protein BLAST (Basic Alignment Search Tool at http://www.ncbi.nlm.nih.gov/ BLAST/, against the non-redundant protein sequences database)[44], identified the *A. nidulans* GH26 endomannanase (Swissprot ID Q5AWB7 [27]) with 67.5 % amino acid identity as the closest characterised enzyme. A multiple sequence alignment of 9 fungal GH26 endomannanases showed that the amino acids that take part in ligand binding in *Ypen*Man26A are highly conserved (Figure 3, red stars) (see later paragraph for discussion of differences between sequences of the GH26 core domains with and without a CBM35). However, *Wsp*.Man26A has two striking differences compared to *Ypen*Man26A and the other endomannanases. The first is in the -2 subsite (*Ypen*Man26A Asp37), where the analysed endomannanases have either an Asp or a Glu, while *Wsp*.Man26A has Thr (Figure 3).

(Please insert Figure 3 around here)

The second is in the -4 subsite (*Ypen*Man26A Trp110), where the tested endomannanases have Trp or Tyr, while *Wsp.*Man26A has His (Figure 3). von Freiesleben et al. 2018 showed that *Ypen*Man26A and *Wsp.*Man26A differ in their substrate preferences for locust bean gum and guar gum. *Ypen*Man26A barely discriminated between the two substrates, whilst *Wsp.*Man26A had approximately four times higher initial hydrolysis rate on locust bean gum than on guar gum, indicating that this enzyme was more hindered or had less affinity for the increased amount of galactose substitutions in guar gum. In the present study, the hydrolysis product profiles from full conversion of guar gum were analysed using the DNA sequencer-Assisted Saccharide analysis in High throughput (DASH) method [27,45] (Figure 4A).

(Please insert Figure 4 around here)

*Ypen*Man26A produced primarily α-galactosyl-mannose (G, 2.10 DE) and α-6^2-6^1-di-galactosyl-mannotriose (MGG, 4.10 DE), whereas *Wsp.*Man26A in addition produced M2 and M3. To investigate if the difference in ligand interacting amino acids between *Ypen*Man26A and *Wsp.*Man26A played a role in the observed differences in substrate preference and binding mode, two *Ypen*Man26A mutants, *Ypen*Man26A D37T and *Ypen*Man26A W110H, were designed, expressed and purified to electrophoretic purity (Table 3 and Figure S3).

Table 3: The wild type *Ypen*Man26A and the investigated variants

Enzyme	**Domains**	**Mw[a] (kDa)**	***Tm*[b] (°C)**
*Ypen*Man26A (sequence ID: MH899111)	GH26	34.5	50
*Ypen*Man26A D37T	GH26	34.5	50
*Ypen*Man26A W110H	GH26	34.4	47

[a] Theoretical. [b] The Thermal midpoint (*Tm*) at pH 5.

Productive binding of M5

M5 hydrolysis product analysis using HPAEC combined with solvent isotope labelling and mass spectrometry (MS) analysis [25,46], was used to estimate the relative frequency of productive binding modes for the *Ypen*Man26A wild type and W110H mutant. The HPAEC product quantification showed a clear difference between the wild type and the W110H variant (Figure S4), with the wild type preferring producing M4 and M1 (89 % relative productive binding frequency) with little formation of M3 and M2 (11 %). For the W110H mutant the major hydrolysis products were M3 and M2 (70 %) as well as some M4 and M1 (30 %). Because two productive binding modes can give rise to the same products, M5 can for example be hydrolysed into M4 and M1 through removal of the reducing end or the non-reducing end mannopyranosyl unit, the HPAEC data was combined with an *in situ* labelling, matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) analysis procedure [25,46] where M5 hydrolysis is performed in ^{18}O-water, to obtain product ratios of ^{18}O-labelled versus ordinary ^{16}O-products. The newly formed reducing end will be labelled with ^{18}O (heavy product) while the "leaving group" saccharide (light) of each catalytic event will not. With the MS analysis, it is thus possible to distinguish between M4 produced by M5 binding from subsite -4 to +1 (generating heavy M4) and M5 binding from subsite -1 to +4 (generating light M4). The heavy versus light product ratios obtained for M3 and M4 were used to calculate the relative binding frequencies of binding modes that generate these products, respectively (Figure 5).

(**Please insert Figure 5 around here**)

The data show that for the *Ypen*Man26A wild type the dominant productive M5 binding mode is from subsite -4 to +1 (80 % binding frequency) (Figure 5), but that this mode is significantly reduced (28 %) for the W110H mutant. Instead the dominant productive M5 binding mode is shifted to cover subsites -3 to +2 (63 % binding frequency). This is most likely a consequence of Trp110 in the -4 subsite being changed to His, resulting in a weaker subsite.

Kinetics with galactomannans and MGGMM

The Michaelis-Menten kinetic parameters with locust bean gum and guar gum were determined for the two *Ypen*Man26A mutants D37T and W110H and compared with the parameters obtained for the wild type (Table 4). The wild type had the highest k_{cat}/K_M and k_{cat} on both substrates. Wild type and variant D37T had identical K_M for locust bean gum, but D37T had higher K_M than wild type on guar gum. This indicates that the D37T mutant has lower affinity for the galactose residues in the highly substituted guar gum than the wild type. The reason that no difference in K_M was observed for locust bean gum might be due to unsubstituted blocks of mannan [12]. It is likely that both the wild type and the D37T variant degrade the unsubstituted, more easily assessable, part of the substrate first, so the initial rate reflects the enzyme affinity for the unsubstituted regions of the substrate. Guar gum is known to have no (or few) blocks without substitutions [12]. As judged from the K_M value, the *Ypen*Man26A W110H had very low affinity for locust bean gum, when compared to the two other constructs. On guar gum galactomannan it was not possible to determine the kinetic parameters separately, because saturation was not reached, but the low k_{cat}/K_M indicates low affinity or low hydrolysis rate. k_{cat}, K_M and k_{cat}/K_M with locust bean gum for the wild type

Wsp.Man26A was 564 ± 26 s^{-1}, 0.8 ± 0.2 mg/ml and 705 ± 179 ml/(mg*s) respectively and on guar gum 271 ± 31 s^{-1}, 3.6 ± 1 mg/ml and 75 ± 23 ml/(mg*s) respectively. Particularly on guar gum, full saturation was not reach for *Wsp*.Man26A, why the standard deviation is relatively high. The R^2 value for the fitted Michaelis-Menten curve for *Wsp*.Man26A was 0.90 and 0.91 on locust bean gum and guar gum respectively.

Table 4 Kinetic parameters on locust bean gum and guar gum of the wild type *Ypen*Man26A and the variants *Ypen*Man26A D37T and *Ypen*Man26A W110H.

Enzyme	**Locust bean gum**			**Guar gum**		
	k_{cat} (s^{-1})	**K_M (mg/ml)**	**k_{cat}/K_M (ml/(mg*s))**	**k_{cat} (s^{-1})**	**K_M (mg/ml)**	**k_{cat}/K_M (ml/(mg*s))**
Wild type	475 ± 5	0.6 ± 0.03	792 ± 40	636 ± 19	2.2 ± 0.2	289 ± 28
D37T	334 ± 6	0.6 ± 0.05	557 ± 47	473 ± 12	2.7 ± 0.2	175 ± 14
W110H	404 ± 18	10 ± 0.8	40 ± 4	n.d[a]	n.d[a]	17 ± 0.6

a Not determined (n.d), because saturation was not reached. Linear regression was used to determine k_{cat}/K_M from the initial part of the Michaelis-Menten curve.

To validate if the increase of K_M for *Ypen*Man26A D37T on the highly substituted guar gum galactomannan, was caused by the change in the -2 subsite, k_{cat}/K_M on MGGMM for the *Ypen*Man26A wild type and the D37T mutant were determined by following substrate depletion at low substrate concentration (0.1 mM) by MS (Table 5). A novel, MS based method with an internal standard was developed to allow these measurements (relevant spectra, extracted ion chromatograms and a standard curve can be seen in Figure S5). The reaction rate of MGGMM depletion could be described by the equation described by Matsui et al. [47] (Figure S6), which was used to determine k_{cat}/K_M. It is likely that MGGMM binds from the -4 to the +1 subsite in *Ypen*Man26A, and therefore accommodates the galactopyranosyl residues in the -3

and -2 subsite, as in the X-ray structure (Figure 2C). This can be assumed because of the dominant M5 productive binding mode for *Ypen*Man26A from subsite – 4 to +1 (Figure 5) and the demonstrated capability of *Ypen*Man26A to accommodate the galactopyranosyl moiety in the -3 and -2 subsites (Figure 2). Furthermore, *Anid*Man26A, which is the closest homolog to *Ypen*Man26A, was found to produce MGGM and M from MGGMM [27].

Table 5 Kinetic efficiency on MGGMM for *Ypen*Man26A wild type and the variant *Ypen*Man26A D37T

Enzyme	**k_{cat}/K_M ($s^{-1}*mM^{-1}$) on MGGMM**
Wild type	84 ± 5
D37T	19 ± 2

The D37T variant had four times lower k_{cat}/K_M on MGGMM than the wild type enzyme (84 vs 19 $s^{-1}*mM^{-1}$, Table 5), meaning that the mutant has lower k_{cat} and /or higher K_M (probably a combination of both as with the individual kinetic parameters determined on guar gum). The observed k_{cat}/K_M for the wild type *Ypen*Man26A and the D37T variant is at the same level as k_{cat}/K_M's reported for other fungal endomannanases on M5, which were found to be between 23 – 163 $s^{-1}*mM^{-1}$ for GH5 endomannanases from *Aspergillus nidulans* and *Trichoderma reesei* [48] and 22 $s^{-1}*mM^{-1}$ for *Pans*Man26A [25]. The bacterial GH26 endomannanase from *B. ovartus*, *Bova*Man26A, had a k_{cat}/K_M on 247 $s^{-1}*mM^{-1}$ on M5 [23]. This result emphasises that the substitution of Asp37 with Thr, seem to decrease the affinity for the galactopyranosyl moiety in the -2 subsite. The lower k_{cat}/K_M on MGGMM for the D37T mutant compared to the wild type, is consistent with the expected increase in distance between the galactopyranosyl unit and the amino acid residue, when substituting Asp with Thr (Figure 4C).

Differences in the catalytic GH26 domain in fungal endomannanases with and without CBM35

Most regions are highly conserved between *Ypen*Man26A and *Pans*Man26A (Figure 2A and C), but *Ypen*Man26A lacks a N-terminal CBM35 domain. From the superimposition of the two crystal structures (Figure 2A), it is seen that the main difference in the secondary structure between the core modules of the two enzymes is in the area which approaches the CBM35 of *Pans*Man26A. In this area, *Pans*Man26A has an α-helix and *Ypen*Man26A a surface loop. Interactions occur through water between the Ala402 and Gln404 in the *Pans*Man26A core domain and the Leu58 and the Ser130 in its CBM35 and linker respectively. Couturier et al. 2013 also state that a hydrophobic patch comprising Leu58 and Leu130 on the surface of the CBM35 stands in front of a cluster of hydrophobic residues, Ala402, Tyr403 and Leu399 of the core domain [25]. These interactions would not be established if the *Pans*CBM35 were appended to the *Ypen*Man26A, because of differences in the amino acid sequence and the flexible nature of the surface loop. The multiple sequence alignment (Figure 3) of the GH26 core domains of 9 fungal GH26 endomannanases (two wild type core enzymes, five with a N-terminal CBM35 and two with a CBM35 and a C-terminal CBM1), confirms variation in the region in and around α9 in *Pans*Man26A (Figure 3, marked blue), the area of the core domain approaching the CBM35. The seven enzymes with a CBM35 have identical sequences to *Pans*Man26A (LQAY, for *Asti*Man26A it is MQLY), which forms an α-helix in the *Pans*Man26A structure, while the two enzymes with no CBM35, have a different and seemingly more variable sequence (TGGV for *Ypen*Man26A and MRED for *Anid*Man26A). From this analysis, it seems

that co-evolution has occurred between the GH26 core domain and the CBM35. It is likely that the core domain evolved to accommodate and maybe help position the CBM35. When the CBM35 is absent, this α-helix (α9) is not needed.

Discussion

Data presented here add to the current understanding of fungal GH26 endomannanases, which appear to be conserved in their known functional characteristics. Characterised fungal GH26 endomannanases, including *Ypen*Man26A, have a characteristic ligand binding site with a strong – 4 subsite, and a dominant M5 binding mode from the -4 to +1 subsite [25,27,28], in contrast to at least some fungal GH5 endomannanases (including *Pans*Man5A) which mainly bind M5 from the -3 to the +2 subsite [25]. To date, the fungal GH26 endomannanases which have been analysed with a focus on the accommodation of galactopyranosyl units, are able to degrade highly substituted galactomannans by allowing accommodation of galactose substitutions at least in the -3, -2, -1 and +1 subsite as judged by biochemical data and crystal structures. The biochemical data include the observations that *Pans*Man26A and *Anid*Man26A produces produce α-galactosyl-mannose (G) as their dominant hydrolysis product from guar gum galactomannan and that *Anid*Man26A catalyse the hydrolysis of MGGMM to MGGM and mannose [27]. The structural data include the crystal structure of *Pans*Man26A [25] and the homology model of *Anid*Man26A that both show an open active site cleft with space for galactose substitutions [27]. Furthermore, our current data with the crystal structure of *Ypen*Man26A with bound MGG from the – 4 to the -2 subsites and the observation that the amino acids participating in MGG binding in *Ypen*Man26A are highly

conserved between studied GH26 endomannanases (Figure 2), further support this hypothesis. Some fungal GH5 endomannanases, e.g. the *Tres*Man5A from *Trichoderma reesei*, have been found to accommodate galactopyranosyl residues in the - 1 subsite [49], but not in the -2 and +1 subsites [27]. Among the bacterial GH26 endomannanases there is a variation in their ability to accommodate multiple galactopyranosyl residues in the active site cleft, exemplified by *Bova*Man26A and *Bova*Man26B from *Bacteroides ovatus* [23].

We show that a single mutation in the substrate binding amino acids can result in altered binding modes or substrate affinity as seen for the *Ypen*Man26 wild type and mutants investigated in the present study. Of the 17 amino acids involved in ligand binding (including the two catalytic residues) only three residues were not conserved between the nine fungal GH26 endomannanases compared in this study (Figure 3). In two of these changes *Wsp*.Man26A differed from the rest of the endomannanases. Mutation studies showed that W110H shifted the dominant productive M5 binding mode of *Ypen*Man26A from covering the -4 to +1 subsites to covering the -3 to +2 subsites, emphasising the importance of Trp110 in the strong -4 subsite. The D37T mutation lowered the affinity for a galactopyranosyl unit in the -2 subsite of *Ypen*Man26A. A third variation in ligand binding amino acids among the studied GH26 endomannanases was found where *Ypen*Man26A has Asn280 (Figure3). This residue is not conserved between the nine fungal GH26 endomannanases, which might indicate that this residue is not important for ligand binding or it could contribute to different affinity for galactose in the -2 subsite similar to the D37T mutation investigated in the present study. Indeed fungal GH26 endomannanases was shown to have different ratios between their initial rate on locust bean gum vs guar gum [1], indicating variations in galactose affinity and/or

tolerance, which perhaps can be explained by variations at this position (Asn280 in *Ypen*Man26A, Figure 3). Detailed knowledge about binding mode and affinity for substitutions in different subsites is important when using these enzymes to produce specific oligosaccharides e.g. for prebiotics or alkyl mannooligosides.

As seen from the superimposition of *Ypen*Man26A and *Pans*Man26A (Figure 2A) and the sequence alignment of nine fungal GH26 endomannanases (Figure 3), the main difference in their catalytic domains appears to be in the area approaching the CBM35 (if present). The GH26 core module of the enzymes with a CBM35 seems to have evolved to harbour this big binding domain (15 kDa) in close proximity to the core, by aid of an α-helix (α9) whereas the wild type enzymes with no CBM35, *Ypen*Man26A and *Anid*Man26A, have a less structured surface loop in this area. The α9-helix in *Pans*Man26A is situated with the end of the helix pointing directly into the site where the linker is attached to the CBM35. It is possible that this α-helix plays an important role in positioning of the CBM35. It is also possible that the position we see in the crystal structure of *Pans*Man26A is not the position of the CBM35 in solution. It is likely that the core domain and the CBM35 can come in even closer contact, perhaps facilitated by ligand binding. Examples of a similar event have been reported for processive GH9 endoglucanases the closest CBM3c module were shown to align with the catalytic cleft of the GH9 module, presumably forming one functional entity [50]. The linker in these GH9 cellulases is wrapped around the core domain, similar to the linker in *Pans*Man26A[25], and contributes significantly to the positioning of the CBM3c.

Conclusions

This study identified important amino acids for binding galactomannan in the – 4 to – 2 subsites of *Ypen*Man26A, by solving and analysing its crystal structure in complex with MGG. Particularly the -2 subsite has multiple interactions with the galactopyranosyl side group. The study also highlights the high sequence similarity of known fungal GH26 endomannanases, with conserved ligand binding amino acids in the active site cleft. These results strongly indicate that the capability of accommodating multiple galactopyranosyl side-groups in the binding cleft is conserved among the fungal enzymes in the GH26 family. The two *Ypen*Man26A variants, W110H and D37T, showed that these changes shifted the dominant M5 binding mode from covering the -4 to +1 subsite to cover the -3 to +2 subsite and lowered the affinity for galactopyranosyl residues in the -2 subsite. The crystal structure of *Ypen*Man26A had a unique surface loop when compared to the crystal structure of *Pans*Man26A, which appear to be a consequence of the enzyme not harbouring a CBM35. Known fungal GH26 endomannanases, including *Ypen*Man26A, seem tailored for hydrolysing highly substituted galactomannans. Understanding the intimate enzyme-substrate interactions and the possibilities of changing product profiles and substrate affinities are important for fine-tuned optimization and utilization of these enzymes in industrial applications.

Methods

Materials

Locust bean gum (low viscosity; sodium borohydride reduced), guar gum (medium viscosity), mannobiose (M2), mannotriose (M3), mannotetraose (M4), mannopentaose (M5), α-6^1-galactosyl-mannotriose (MMG), α-6^4-6^3-di-galactosyl-mannopentaose (MGGMM), and α-6^2-6^3-6^4-tri-xylosyl-glucotetraose (XXXG) were purchased from Megazyme (Ireland). All other chemicals were purchased from Sigma (Germany), unless otherwise stated. Mobility markers, dextran ladder, and the DASHboard software for DASH analyses were kindly donated by Prof. Paul Dupree (University of Cambridge, UK).

Construction of variants

The gene sequence encoding *Ypen*Man26A (GENESEQP: MH899111) was used to make the mutated constructs. E165Q was introduced into the gene sequence by PCR using synthetic oligonucleotides replacing the codon GAG position 165 of the mature peptide with CAG. PCR was conducted for the 5′ fragment and 3′ fragment separately using Phusion High-Fidelity DNA Polymerase (ThermoFisher Scientific) under the following conditions: 98 °C 2 min, 35 cycles at 98 °C for 10 sec, 72 °C for 150 sec, followed by 72 °C for 10 min. The PCR products were gel purified and used as template for a second round of PCR, using the gene flanking primers to amplify the full-length gene with the native signal peptide. The full-length PCR product was cloned into pDAu222 [51], an *Aspergillus* expression vector under the control of a NA2-tpi double promoter using the BamHI and XhoI restriction sites, and its sequence was

determined. The resulting pDAu222-*Ypen*Man26A-E165Q expression vector was transformed into *A. oryzae* MT3568. MT3568 is an amdS (acetamidase) disrupted derivative of *A. oryzae* Jal_355 [52] in which pyrG auxotrophy was restored in the process of inactivating the *A. oryzae* amdS gene. Secretion of *Ypen*Man26A E165Q in the culture supernatant of the recombinant MT3568 clones was confirmed by SDS-page.

Mutants containing the D37T and W110H substitutions respectively were made as synthetic full-length cDNA constructs with the native signal peptide (ThermoFisher Scientific) cloned into pDAu222 using the BamHI and XhoI restriction sites. For D37T the codon GAC of position 37 of the mature peptide was replaced with ACC. For W110H the codon TGG of position 110 of the mature peptide was replaced with CAC. The constructs were verified by sequencing and the resulting pDAu222 expression vectors were transformed into *A. oryzae* MT3568. Secretion of mutants in the culture supernatant of recombinant MT3568 clones was confirmed by SDS-page.

Expression and purification

The fungal wild type GH26 endomannanases *Westerdykella sp.* (*Wsp.*Man26A) and *Yunnania penicillata* (*Ypen*Man26A), as well as the *Ypen*Man26A mutants D37T, W110H and E165Q were recombinantly expressed in *A. oryzae* MT3568 an amdS [52]. The enzymes, wild types and variants, were purified to electrophoretic purity using hydrophobic interaction and ion exchange chromatography. The inactive *Ypen*Man26A E165Q variant, used for crystallisation, was further purified using size-exclusion chromatography and deglycosylated with Endoglycosidase H (Roche). The identity of the purified endomannanases was validated

with mass spectrometry analysing a tryptic digest of the protein band excised from a SDS-PAGE gel. Protein concentrations were determined by UV absorption at 280 nm using theoretical extinction coefficients (ε). ε at 280 nm of all proteins were estimated by GPMAW 9.20 (Lighthouse Data), and were based on mature proteins without modifications.

Crystallisation

The inactive *Ypen*Man26A mutant E165Q was concentrated to 48 mg/ml, in 20 mM MES, 125 mM NaCl, pH 6 and aliquoted into 50 µl samples. Aliquots not used for immediate crystallisation trials were flash-frozen in liquid nitrogen and stored at -80 °C. Initial crystallisation screening was carried out using sitting-drop vapour-diffusion with drops set up using a *Mosquito Crystal* liquid handling robot (TTP LabTech, UK) with 150 nl protein solution plus 150 nl reservoir solution in 96-well format plates (MRC 2-well crystallisation microplate, Swissci, Switzerland) equilibrated against 54 µl reservoir solution. Experiments were carried out at room temperature with several commercial screens, for the protein on its own and in the presence of 5 mM MGGMM. The best hits were obtained in the $AmSO_4$ suite (QIAGEN), for the ligand complex. The conditions were manually optimised in a 24-well Linbro dish, in hanging drop format. The final crystallisation conditions were 2.6-2.8 M ammonium sulphate, 0.1 M Hepes pH 7.0.

Data collection, structure solution and refinement

All computation was carried out using programs from the CCP4 suite v. 7.0 [53]. For the MGGMM-*Ypen*Man26A complex, data were collected at the Diamond Light Source beamline

I04 to 1.36 Å resolution and processed using *xia2* [54]. The structure was solved using *MOLREP* [55] with *Pans*Man26A (PDB entry: 3zm8; Couturier et al., 2013; sequence identity: 47.7 %) as template. The structure was refined using REFMAC5 [56] iterated with manual model building/correction in Coot [57]. The final model was validated using Molprobity [38] as part of the Phenix package [58]. Data-processing and refinement statistics are given in Table 1. Structure figures were prepared using CCP4*mg* [37] or PyMOL v 1.7.20 (DeLano Scientific LLC, San Carlos, CA). The sequence alignments were created with MUSCLE [42] and ESPript [43].

Homology modelling

The homology model of *Wsp*.Man26A was generated using HHPred-Homology server (https://toolkit.tuebingen.mpg.de/#/tools/hhpred) [59] with *Pans*Man26A as template, (PDB ID: 3ZM8 [25], 54% sequence identity). Model quality was evaluated using a Ramachandran analysis in MolProbity (http://molprobity.biochem.duke.edu/) [38]. The model of *Wsp.*Man26A had 96.4 % (430/437) of all residues in allowed regions. The model was only used to visualise the mutated amino acids in *Ypen*Man26A, which were inspired by *Wsp.*Man26A (Figure 3).

Thermal stability

The thermal stability at pH 5 was investigated with Differential Scanning Calorimetry (DSC) following an established protocol [27]. The Thermal midpoint (*Tm*) was determined as the top of the protein denaturation peak, with an accuracy of +/- 1 °C.

Initial rates and analysis of product profiles by DASH

The initial rate on locust bean gum and guar gum by the endomannanases were determined with 2.5 mg/ml substrate in 50 mM sodium acetate pH 5 at 37 °C. The hydrolytic activity was determined after 15 min in a 200 µl hydrolysis volume. Released reducing sugars were measured with the 4-hydroxybenzoic acid hydrazide (PAHBAH) method described by Lever (1972) [60], with mannose as standard. All hydrolysis assays were carried out at 7 different endomannanase doses as described elsewhere [27]. Initial rates were calculated in the initial linear range of the hydrolysis. One unit (U) was defined as the amount of endomannanase required to release 1 µmole of reducing ends per minute, under the assay conditions specified. Guar gum hydrolysis product profiles at high conversion (26-36%) were analysed by DASH after inactivation by heating at 95 °C for 15 min. APTS (9-aminopyrene-1,4,6-trisulfonate) labelling and analysis of the labelled saccharides were carried out as described elsewhere [27,45].

Productive M5 binding modes

The hydrolytic cleavage pattern of M5 was determined for the *Ypen*Man26A wild type and the W110H variant, by the previously established ^{18}O-water product labelling methodology [25,46]. First, M5 hydrolysis products were analysed and quantified by high performance anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD) using a Dionex ICS-5000 with a Carbo-Pac PA-200 column and guard column. For this, double incubations of 1 mM M5 and 50 nM wild type enzyme or 200 nM W110H mutant in 1.5 mM sodium acetate buffer, pH 5 were stopped by boiling at timed intervals (30 min to 3 h). Data after 30 min incubation for *Ypen*Man26A or 2 h for the W110H mutant (approximately 30

% hydrolysis) were used. The quantification allowed distinguishing between productive M5 binding modes that generated M4 and M1 versus those that generate M3 and M2. However, HPAEC alone cannot distinguish between the two possible binding modes generating M4 and M1 (i.e. binding either from subsite -4 to +1 or from -1 to +4), neither the two binding modes that generate M3 and M2 (i.e. binding from subsite -3 to +2 or -2 to +3). Therefore, incubations as above were also set up at 8 ˚C using 97 % H_2 ^{18}O as stock solvent reaching 92 % ^{18}O-water in the reactions. Duplicate reactions were stopped after 30 min (for wild type) and 2 h (for W110H) by directly spotting 0.5 µl samples with 0.5 ml matrix (10 mg/ml 2,5-dihydroxybenzoic acid) on a stainless-steel plate, followed by immediate drying with warm air. Spectra were then obtained by MALDI-TOF MS and used to calculate the ^{18}O over ^{16}O product ratios using the monoisotopic peak areas as previously described [25,46]. Since M5 hydrolysis in ^{18}O-water generates products where the newly formed reducing end becomes ^{18}O-labelled (and other chain ends do not), the ^{18}O over ^{16}O product ratios can be used to calculate the relative frequency of the productive binding modes mentioned above (i.e. M5 binding from subsite -4 to +1 versus subsite -1 to +4 or binding from subsite -3 to +2 versus subsite -2 to +3) [25,46]. The procedure involves two calculated corrections for the product ratio determination. One correction for the (M+2) natural isotope peak of the light (^{16}O) species which overlaps with the heavy (^{18}O) peak and second for the presence of 8 % ordinary $H_2{}^{16}O$ in the hydrolysis reaction.

Kinetics with locust bean gum and guar gum

The kinetic constants for locust bean gum and guar gum hydrolysis were determined by assaying the initial endomannanase rates at different substrate concentrations (10 to 0.1 mg/ml) using the PAHBAH assay as described above. The enzyme concentrations used for the locust bean gum hydrolysis were 4 nM *Ypen*Man26A wild type, 4 nM *Wsp.*Man26A, 4 nM *Ypen*Man26A D37T, and 18 nM *Ypen*Man26A W110H and for the guar gum hydrolysis were 4 nM *Ypen*Man26A, 10 nM *Wsp.*Man26A, 6 nM *Ypen*Man26A D37T, and 44 nM *Ypen*Man26A W110H. The initial hydrolysis rate, V_i, was plotted as a function of the substrate concentration, [S]. Non-linear regression using the Michaelis-Menten equation was used to determine the values for k_{cat}, K_M and k_{cat}/K_M.

Kinetics with MGGMM

k_{cat}/K_M was determined by following MGGMM depletion over time at low substrate concentration (0.1 mM), pH 5 and 37 °C, with an online, direct injection, mass spectrometry based assay. Duplicate samples were analysed using a HPLC-MS system with a Dionex Ultimate 3000RS HPLC connected to an ESI-iontrap (AmaZon SL, Bruker Daltonics). The HPLC provided a constant flow of 0.1 ml/min of 50/50 vol-% acetonitrile and 0.1 % formic acid. The electrospray was operated in positive ultrascan mode with Multiple Reaction Monitoring (MRM) using a target mass of *m/z* 800. MRM mode was chosen to selectively follow substrate depletion and an internal standard (XXXG). 100 % reaction amplitude was used to ensure fragmentation of the precursor ion. The capillary voltage was set at 4.5 kV, end plate offset was 0.5 kV, nebulizer pressure 3.0 bar, dry gas flow 12.0 l/min, and dry gas

temperature was set to 280 °C. Buffer concentration, 1 mM sodium acetate pH 5, was set as low as possible to minimize ion suppression without compromising pH in the reaction. The total reaction volume was 500 µl and the sample was incubated directly in an HPLC-vial in the HPLC-autosampler. The reaction was started by adding enzyme in 2 nM and 6 nM for the wild type *Ypen*Man26A and the D37T variant respectively. Two min after enzyme addition, the first sample was taken. Thereafter sampling was performed every 5.4 min (including sampling procedure), when the autosampler injected 4 µL sample directly into the flow leading to the MS. The enzyme reaction was immediately quenched when entering the flow path because the mobile phase was pH 2.7 and detection occurred approx. 0.5 min after injection. Total acquisition time was set to 4 min. The enzyme reactions were followed for a maximum time period of 50 min, but only data describing the initial phase of the reaction (less than 25 % conversion of substrate) were used for estimating k_{cat}/K_M. Details on extracted ion chromatograms used for quantification of MGGMM and XXXG can be seen in Figure S5. Data was analysed and quantified using Compass DataAnalysis 4.2 and Compass QuantAnalysis 2.2 provided by Bruker Daltonics. Ln (S_0/S_t) was plotted as function of time (t) (Figure S6) and k_{cat}/K_M was calculated as described by Matsui et al [47]; k = Ln (S_0/S_t), where k = ((k_{cat}/K_M)*[enzyme])*t, S_0 = substrate concentration at time zero and S_t = substrate concentration at time t.

References

1. von Freiesleben, P. *et al.* Boosting of enzymatic softwood saccharification by fungal GH5 and GH26 endomannanases. *Biotechnol. Biofuels* **11,** 194 (2018).

2. Jørgensen, H. *et al.* Production of ethanol and feed by high dry matter hydrolysis and fermentation of palm kernel press cake. *Appl. Biochem. Biotechnol.* **161,** 318–332 (2010).

3. Li, Y. *et al.* High level expression of β-mannanase (RmMan5A) in *Pichia pastoris* for partially hydrolyzed guar gum production. *Int. J. Biol. Macromol.* **105,** 1171–1179 (2017).

4. Morrill, J. *et al.* β-Mannanase-catalyzed synthesis of alkyl mannooligosides. *Appl. Microbiol. Biotechnol.* **102,** 5149–5163 (2018).

5. Srivastava, P. K. & Kapoor, M. Production, properties, and applications of endo-β-mannanases. *Biotechnol. Adv.* **35,** 1–19 (2017).

6. Moreira, L. R. S. & Filho, E. X. F. An overview of mannan structure and mannan-degrading enzyme systems. *Appl. Microbiol. Biotechnol.* **79,** 165–78 (2008).

7. Ebringerová, A. Structural diversity and application potential of hemicelluloses. *Macromol. Symp.* **232,** 1–12 (2006).

8. Scheller, H. V & Ulvskov, P. Hemicelluloses. *Annu. Rev. Plant Biol.* **61,** 263–289 (2010).

9. Timell, T. E. Recent progress in the chemistry of wood hemicelluloses. *Wood Sci. Technol.* **1,** 45–70 (1967).

10. Lundqvist, J. *et al.* Characterization of galactoglucomannan extracted from spruce (*Picea abies*) by heat-fractionation at different conditions. *Carbohydr. Polym.* **51,** 203–211 (2003).

11. Willför, S. *et al.* Characterisation of water - soluble galactoglucomannans from Norway spruce wood and thermomechanical pulp. *Carbohydr. Polym.* **52,** 175–187 (2003).

12. McCleary, B. V. The fine structures of carob and guar galactomannans. *Carbohydr. Res.* **139,** 237–260 (1985).

13. Raman, R. *et al.* Advancing glycomics: Implementation strategies at the consortium for functional

glycomics. *Glycobiology* **16,** 82R–90R (2006).

14. Couturier, M. *et al. Podospora anserina* hemicellulases potentiate the *Trichoderma reesei* secretome for saccharification of lignocellulosic biomass. *Appl. Environ. Microbiol.* **77,** 237–246 (2011).

15. Lombard, V., Ramulu, H. G., Drula, E., Coutinho, P. M. & Henrissat, B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* **42,** D490–D495 (2014).

16. Sinnott, M. L. Catalytic mechanisms of enzymic glycosyl transfer. *Chem. Rev.* **90,** 1171–1202 (1990).

17. Henrissat, B. *et al.* Conserved catalytic machinery and the prediction of a common fold for several families of glycosyl hydrolases. *Proc. Natl. Acad. Sci. USA* **92,** 7090–7094 (1995).

18. Withers, S. G. Mechanisms of glycosyl transferases and hydrolases. *Carbohydr. Polym.* **44,** 325–337 (2001).

19. Jin, Y. *et al.* A β‑Mannanase with a lysozyme-like fold and a novel molecular catalytic mechanism. *Am. Chem. Soc. Publ.* **2,** 896–903 (2016).

20. Le Nours, J., Anderson, L., Stoll, D., Stålbrand, H. & Lo Leggio, L. The structure and characterization of a modular endo-β-1,4-mannanase from *Cellulomonas fimi*. *Biochemistry* **44,** 12700–12708 (2005).

21. Hogg, D. *et al.* Crystal structure of mannanase 26A from *Pseudomonas cellulosa* and analysis of residues involved in substrate binding. *J. Biol. Chem.* **276,** 31186–31192 (2001).

22. Cartmell, A. *et al.* The *Cellvibrio japonicus* mannanase CjMan26C displays a unique exo-mode of action that is conferred by subtle changes to the distal region of the active site. *J. Biol. Chem.* **283,** 34403–34413 (2008).

23. Bågenholm, V. *et al.* Galactomannan catabolism conferred by a polysaccharide utilization locus of *Bacteroides ovatus*. *J. Biol. Chem.* **292,** 229–243 (2017).

24. Tsukagoshi, H. *et al.* Structural and biochemical analyses of glycoside hydrolase family 26 β-Mannanase from a symbiotic protist of the termite *Reticulitermes speratus*. *J. Biol. Chem.* **289,** 10843–10852 (2014).

25. Couturier, M. *et al.* Structural and biochemical analyses of glycoside hydrolase families 5 and 26 β-(1,4)-mannanases from *Podospora anserina* reveal differences upon manno-oligosaccharide catalysis. *J. Biol. Chem.* **288,** 14624–14635 (2013).

26. Marchetti, R. *et al.* NMR analysis of the binding mode of two fungal endo-β-1,4-mannanases from GH5 and GH26 families. *Org. Biomol. Chem.* **14,** 314–322 (2016).

27. von Freiesleben, P. *et al.* An *Aspergillus nidulans* GH26 endo-β-mannanase with a novel degradation pattern on highly substituted galactomannans. *Enzyme Microb. Technol.* **83,** 68–77 (2016).

28. Katsimpouras, C., Dimarogona, M. & Petropoulos, P. A thermostable GH26 endo-β-mannanase from *Myceliophthora thermophila* capable of enhancing lignocellulose degradation. *Appl. Microbiol. Biotechnol.* **100,** 8385–8397 (2016).

29. Montanier, C. *et al.* Evidence that family 35 carbohydrate binding modules display conserved specificity but divergent function. *Proc. Natl. Acad. Sci. USA* **106,** 3065–70 (2009).

30. Correia, M. A. S. *et al.* Signature active site architectures illuminate the molecular basis for ligands specificity in family 35 carbohydrate binding module. *Biochemistry* **49,** 6193–6205 (2010).

31. Bolam, D. N. *et al.* Mannanase A from *Pseudomonas fluorescens* ssp. cellulosa is a retaining glycosyl hydrolase in which E212 and E320 are the putative catalytic residues. *Biochemistry* **35,** 16195–16204 (1996).

32. Ducros, V. M.-A. *et al.* Substrate distortion by a β-mannanase: Snapshots of the Michaelis and covalent-intermediate complexes suggest a B(2,5) conformation for the transition state. *Angew. Chemie Int. Ed.* **41,** 2824–2827 (2002).

33. Zechel, D. L. & Withers, S. G. Glycosidase mechanisms:Anatomy of a finely tuned catalyst. *Acc. Chem. Res.* **33,** 11–18 (2000).

34. Dilokpimol, A. *et al.* Recombinant production and characterisation of two related GH5 endo-β-1,4-mannanases from *Aspergillus nidulans* FGSC A4 showing distinctly different transglycosylation capacity. *Biochim. Biophys. Acta* **1814,** 1720–1729 (2011).

35. Dhawan, S. & Kaur, J. Microbial mannanases: An overview of production and applications. *Crit. Rev. Biotechnol.* **27,** 197–216 (2007).

36. Holm, L. & Laakso, L. M. Dali server update. *Nucleic Acids Res.* **44,** W351–W355 (2016).

37. McNicholas, S., Potterton, E., Wilson, K. S. & Noble, M. E. M. Presenting your structures: The CCP4mg

molecular-graphics software. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67,** 386–394 (2011).

38. Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66,** 12–21 (2010).

39. Tailford, L. E. *et al.* Understanding how diverse β-mannanases recognize heterogeneous substrates. *Biochemistry* **48,** 7009–7018 (2009).

40. Yan, X. X., An, X. M., Gui, L. L. & Liang, D. C. From structure to function: insights into the catalytic substrate specificity and thermostability displayed by *Bacillus subtilis* mannanase BCman. *J. Mol. Biol.* **379,** 535–544 (2008).

41. Brewster, J. L. *et al.* Structural basis for ligand recognition by a Cache chemosensory domain that mediates carboxylate sensing in *Pseudomonas syringae*. *Sci. Rep.* **6,** 35198 (2016).

42. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32,** 1792–1797 (2004).

43. Robert, X. & Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42,** 320–324 (2014).

44. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215,** 403–410 (1990).

45. Li, X. *et al.* Development and application of a high throughput carbohydrate profiling technique for analyzing plant cell wall polysaccharides and carbohydrate active enzymes. *Biotechnol. Biofuels* **6,** 94 (2013).

46. Hekmat, O. *et al.* Rational engineering of mannosyl binding in the distal glycone subsites of *Cellulomonas fimi* endo-β-1,4-mannanase: Mannosyl binding promoted at subsite -2 and demoted at subsite -3. *Biochemistry* **49,** 4884–4896 (2010).

47. Matsui, I. *et al.* Subsite structure of Saccharomycopsis α-amylase secreted from *Saccharomyces cerevisiae*. *J. Biochem.* **109,** 566–569 (1991).

48. Rosengren, A. *et al.* An *Aspergillus nidulans* β-mannanase with high transglycosylation capacity revealed through comparative studies within glycosidase family 5. *Appl. Microbiol. Biotechnol.* **98,** 10091–104

(2014).

49. Tenkanen, M., Makkonen, M., Perttula, M., Viikari, L. & Teleman, A. Action of *Trichoderma reesei* mannanase on galactoglucomannan in pine kraft pulp. *J. Biotechnol.* **57,** 191–204 (1997).

50. Petkun, S. *et al.* Reassembly and co-crystallization of a family 9 processive endoglucanase from its component parts: structural and functional significance of the intermodular linker. *PeerJ* **3,** e1126 (2015).

51. Schnorr, K. M., Anderson, L., Da Fonseca, M. L. Q. C. & Leite, R. Expression constructs comprising a Terebella lapidaria nucleic acid encoding a cellulase, host cells, and methods of making the cellulase. (2015).

52. Lehmbeck, J. & Wahlbom, F. Production of a monoclonal antibody in a heterokaryon fungus or in a fungal host cell. (2005).

53. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67,** 235–242 (2011).

54. Winter, G., Lobley, C. M. C. & Prince, S. M. Decision making in xia2. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **69,** 1260–1273 (2013).

55. Vagin, A. & Teplyakov, A. Molecular replacement with MOLREP. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66,** 22–25 (2010).

56. Murshudov, G. N. *et al.* REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67,** 355–367 (2011).

57. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66,** 486–501 (2010).

58. Adams, P. D. *et al.* The Phenix software for automated determination of macromolecular structures. *Methods* **55,** 94–106 (2011).

59. Zimmermann, L. *et al.* A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. *J. Mol. Biol.* **430,** 2237–2243 (2018).

60. Lever, M. A new reaction for colorimetric determination of carbohydrates. *Anal. Biochem.* **47,** 273–279 (1972).

Acknowledgements

This study was partially financed by the BioValue SPIR, Strategic Platform for Innovation and Research on value added products from biomass, which is co-funded by The Innovation Fund Denmark, case no: 0603 – 00522B. HS and MW were partly funded by FORMAS (942-2016-117) and the Swedish Foundation for Strategic Research (RBP 14-0046). We thank Diamond Light Source for access to beamline I04 (proposal no. mx-13587) that contributed to the results presented here. We thank Johan Turkenburg and Sam Hart for support with the X-ray data collection.

Author contributions

P.v.F, H.S., K.B.K, and A.S.M. designed and supervised the research. N.S. constructed and expressed the mutants. P.v.F., O.V.M., E.B., G.J.D., and K.S.W. conducted the crystallisation and structural resolution experiments and the structural data analysis. M.W. and H.S. did the $H_2{}^{18}O$ experiments. J.W.A. and P.v.F. did the MGGMM kinetics analysis using MRM LC-MS. P.v.F. did all other experiments. P.v.F, H.S., K.B.K, and A.S.M analysed and interpreted the data, and prepared the manuscript. All authors have read and approved the final manuscript.

Competing interests statement

All authors declare no conflicts of financial or non-financial interest.

Data availability All data generated or analysed during this study are included in this article and its Supplementary Information file.

Figure legends

Figure 1: Schematic illustration of the two galactomannans (A) guar gum and (B) locust bean gum, with different amount and pattern of galactose substitutions on the β-mannan backbone [12]. Sugars shown using the Consortium for Functional Glycomics notation [13]. Both polymers are continuing towards the reducing end, having a degree of polymerization around 1500 for locust bean gum and 900 for guar gum [12].

Figure 2. (**A**) The structure of *Ypen*Man26A (blue) superimposed with that of *Pans*Man26A (PDB ID: 3ZM8 [25], gold). The α-6^2-6^1-di-galactosyl-mannotriose (MGG) ligand in *Ypen*Man26A (subsites -4 to -2) is shown as green cylinders and the active residues are shown in shades of pink (**B**) Observed electron density for MGG in the -4 to -2 subsites. The positive electron density REFMAC $F_o - F_c$ map, contoured at 3.5 σ (0.37 e Å^{-3}) is shown in blue, with phases calculated prior to the incorporation of any ligand atoms in refinement. (**C**) The organisation of binding subsites and the MGG ligand in the -4 to +2 subsites of *Ypen*Man26A (blue) compared with *Pans*Man26A (gold). *Pans*Man26A residues are only shown for the residues which differ from *Ypen*Man26A. All panels were drawn using *CCP4mg* [37].

Figure 3. Sequence alignment of the catalytic GH26 core region from 9 fungal GH26 endomannanases. Secondary structure elements for *Ypen*Man26A and *Pans*Man26A are displayed above and below the alignment respectively. * Residues involved in ligand binding in the *Ypen*Man26A structure including the two catalytic residues. The α-helix in *Pans*Man26A (α9) which is nearest the CBM35 and which is a surface loop in *Ypen*Man26A is coloured blue.

Identical residues are shown in white on red background. Highly similar residues (when the similarity score assigned to one column is above 0.7) are coloured red and framed in a blue box. The GH26 core sequence of *Ypen*Man26A (MH899111), *Anid*Man26A (Q5AWB7), *Ascobolus stictoideus Asti*Man26A (BBW45412), *Collariella virescens Cvir*Man26A (BBW45415), *Mycothermus thermophiles Mthe*Man26A (MH208368), *Neoascochyta desmazieri Ndes*Man26A (MH208367), *Myceliophthora thermophila Mt*Man26A (99077), *Westerdykella sp. Wsp.*Man26A (MH208369), *Pans*Man26A (B2AEP0) were aligned by MUSCLE [42] and the figure was generated using ESPript 3 Web server [43].

Figure 4. (**A**) Product profiles from guar gum hydrolysis by *Ypen*Man26A and *Wsp.*Man26A. Aligned electropherograms of product profiles at 30 % guar gum conversion (max conversion). Migration of oligosaccharides is given in dextran units (DE). A ladder was run containing: mannose (M1, 0.9 DE), mannobiose (M2, 1.87 DE), mannotriose (M3, 2.85 DE), and α-6^1-galactosyl-mannotriose (MMG, 3.81 DE). Migration of α-galactosyl-mannose (G, 2.10 DE), and α-6^2-6^1-di-galactosyl-mannotriose (MGG, 4.10 DE) was determined by von Freiesleben et al. 2016. (**B**) Initial reaction rates (U/µmole) by *Ypen*Man26A and *Wsp.*Man26A on galactomannans. Data are from von Freiesleben et al. 2016. Hydrolyses were carried out at 37 °C, pH 5 on guar gum (light grey) and locust bean gum (dark grey). Values are given as mean values ± SD (n=2). (**C**) The structure of *Ypen*Man26A with MGG in the -4 to -2 subsites. The two differences in ligand binding amino acids between *Ypen*Man26A and a superimposed homology model of *Wsp.*Man26A are highlighted in blue and orange, respectively.

Figure 5. (**A**) Relative frequency of the productive binding modes of M5 for the *Ypen*Man26A wild type and the W110H variant. Each circle represents a mannose unit. The dashed line between subsite -1 and +1 represents hydrolytic cleavage. The outmost numbers on respective side represent the total percentage of produced product, i.e. M4 and M1 or M3 and M2, determined by HPAEC-PAD quantification. These numbers were then combined with the individual ratios of labelled (^{18}O) to unlabelled (^{16}O) products (M4- and M3-species, respectively) (see panel B) to calculate the inner numbers which represent the relative frequency of each productive binding mode for the two enzymes. (**B**) Mass spectrometry peaks showing the major labelled (^{18}O) hydrolysis product for *Ypen*Man26A wild type (left) and W110H (right) together with unlabelled (^{16}O) species of the same DP (M4 and M3 for the wild type and W110H, respectively). From these spectra, a M4/M4-^{18}O ratio of 1:8.9 and a M3/M3-^{18}O ratio of 1:9.2 was calculated. The theoretical mass for M3 with a sodium adduct is 527.159 and the theoretical mass for M4 with a sodium adduct is 689.212.

Figures

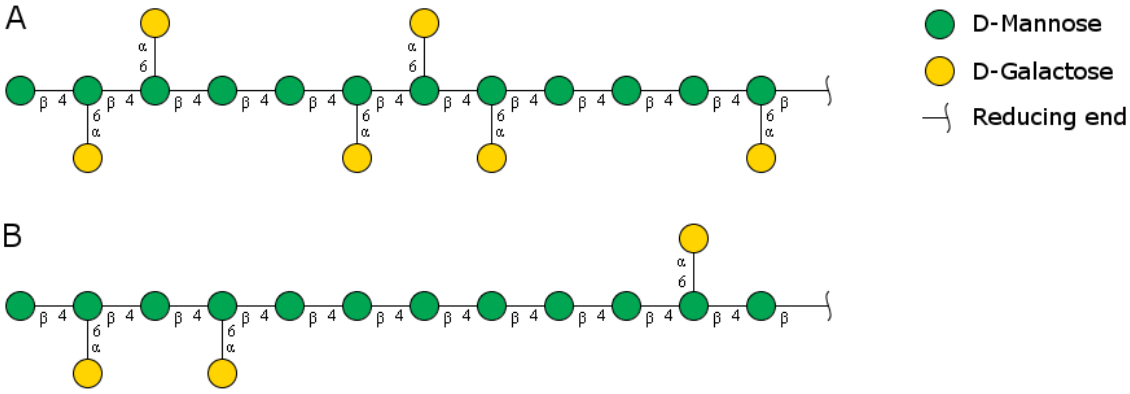


Figure 1.

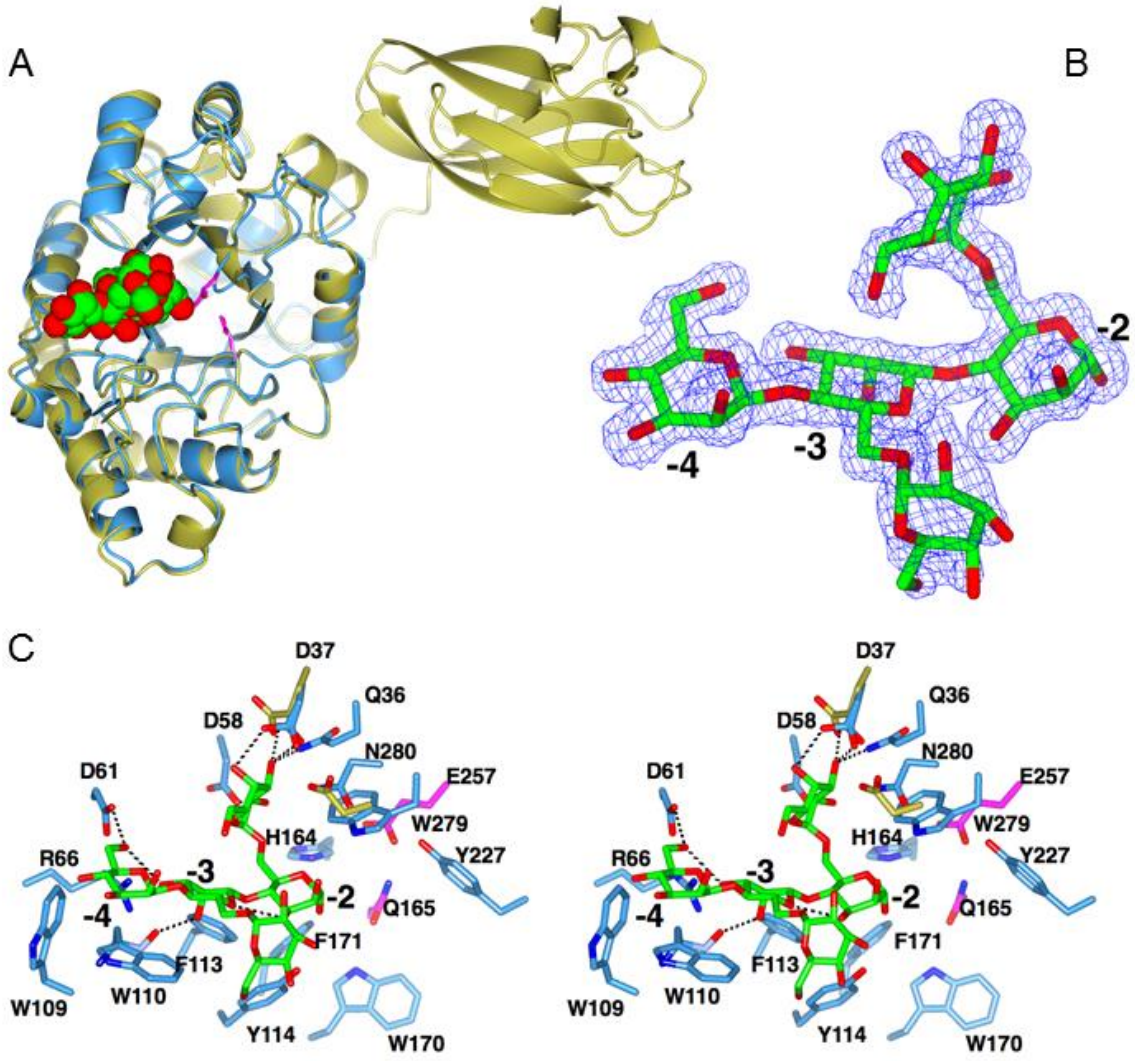

A
B
-2
-3
-4
C
D37
Q36
D58
N280
E257
D61
W279
H164
Y227
R66
-3
-2
Q165
-4
F171
F113
W109
W110
Y114
W170

Figure 2.

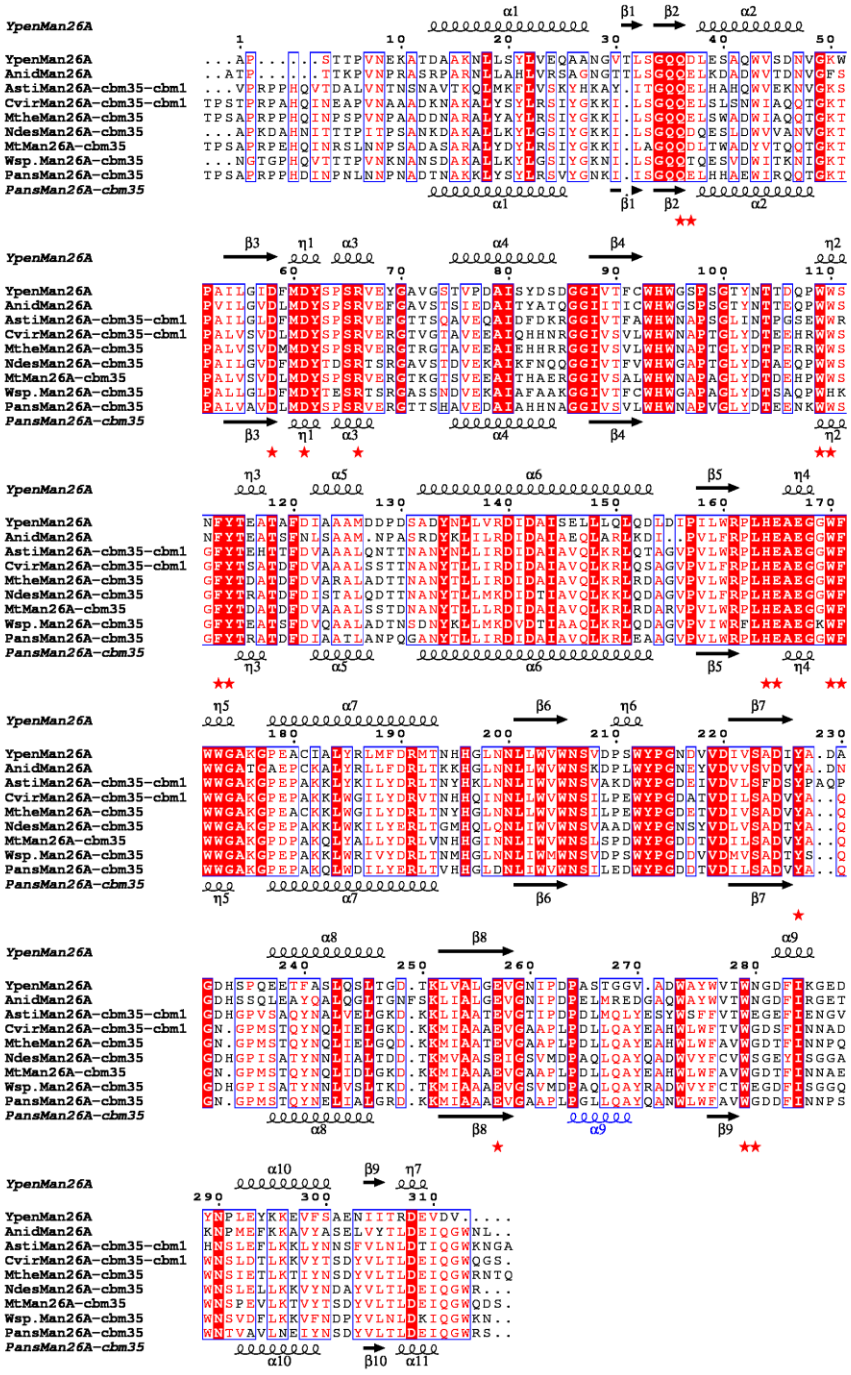

YpenMan26A
AnidMan26A
AstiMan26A-cbm35-cbm1
CvirMan26A-cbm35-cbm1
MtheMan26A-cbm35
NdesMan26A-cbm35
MtMan26A-cbm35
Wsp.Man26A-cbm35
PansMan26A-cbm35
PansMan26A-cbm35

Figure 3.

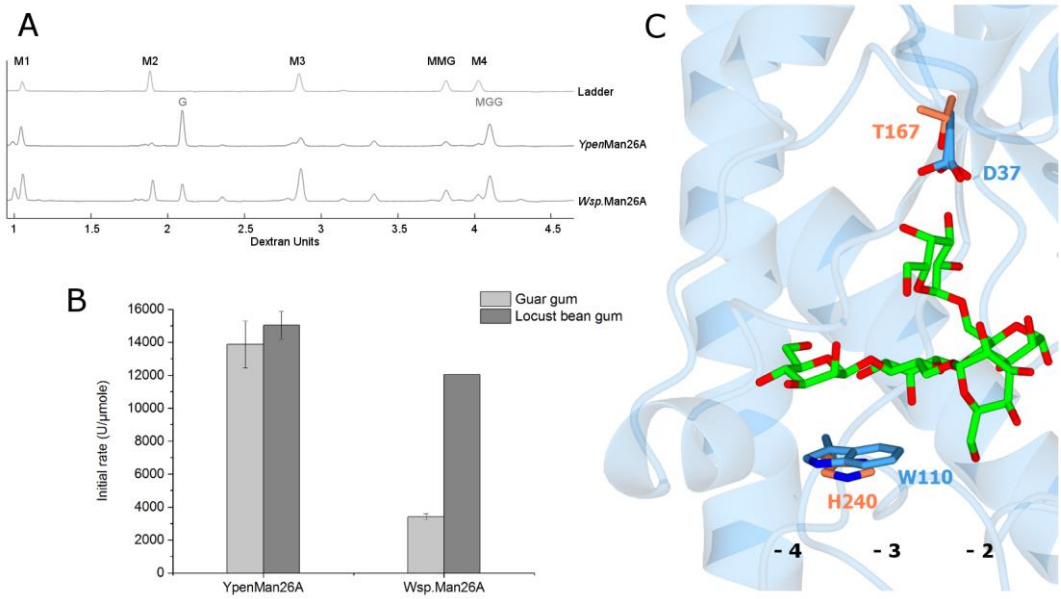


Figure 4.

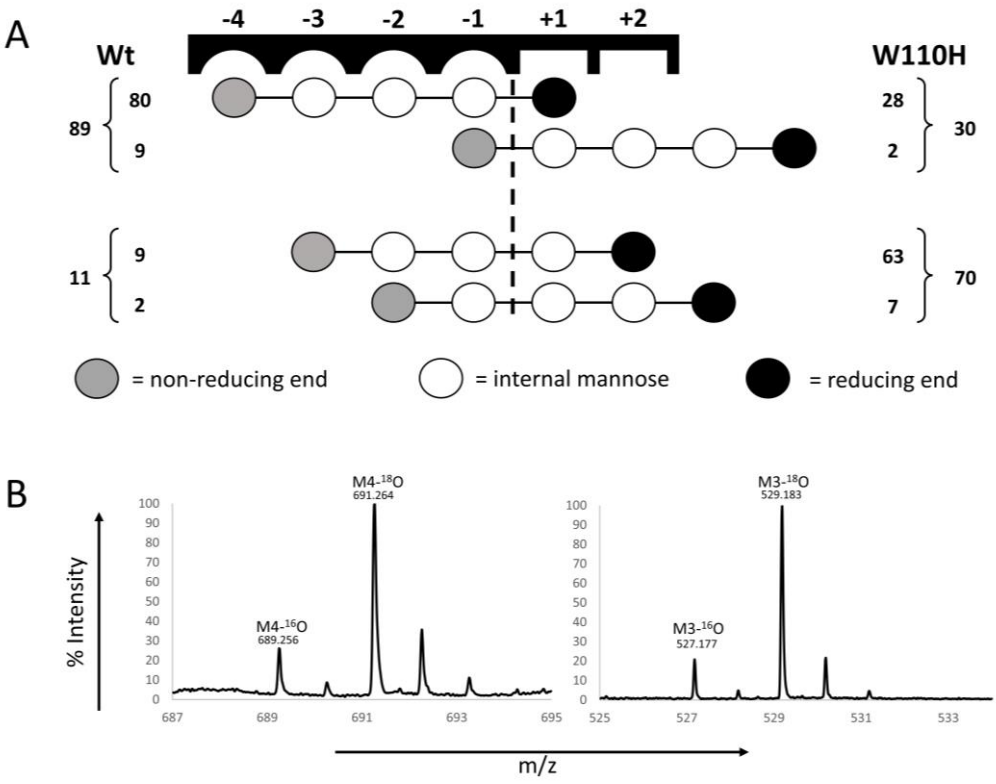


Figure 5.

Paper III

Crystal structure and substrate interactions of an unusual fungal non-CBM carrying GH26 endo-*β*-mannanase from *Yunnania penicillata*

Supplementary data

Supplementary material

Crystal structure and substrate interactions of an unusual fungal non-CBM carrying GH26 endo-β-mannanase from *Yunnania penicillata*

von Freiesleben et al.

Lane	Sample
1	LMW – 10µl
2	*Ypen*Man26A E165Q – 5µl
3	*YpenMan26A E165Q* + *EndoH*– 5µl
4	LMW – 10µl
5	LMW – 10µl
6	*Ypen*Man26A – 5µl
7	*YpenMan26A* + *EndoH*– 5µl
8	LMW – 10µl

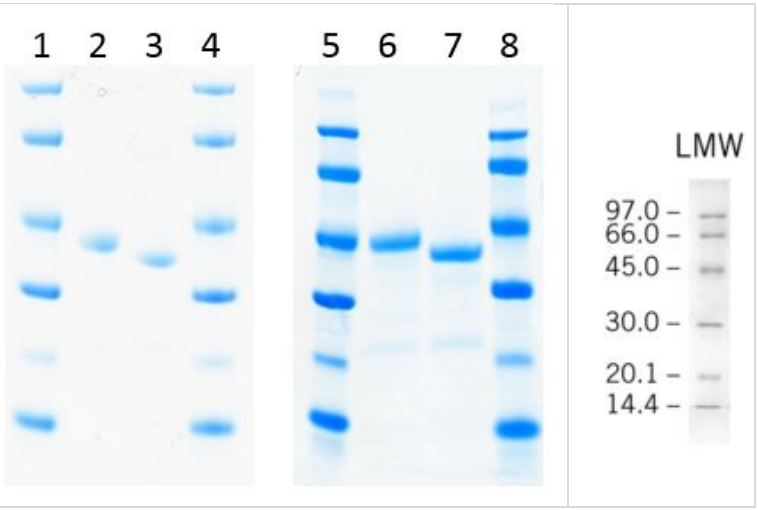


Figure S1. SDS-PAGE before and after EndoH treatment of the active and inactive *Ypen*Man26A. The protein concentration in the samples were 0.5 mg/ml. Prior to gel loading, samples were diluted 1:1 with loading mix. Loading mix was prepared as a 9:1 mix of Novex ® Tris-Glycine SDS Sample Buffer (2X) (Life Technologies) and Nupage ® Sample Reducing Agent (10X) (Life Technologies). Please consult the manuscript for enzyme abbreviations.

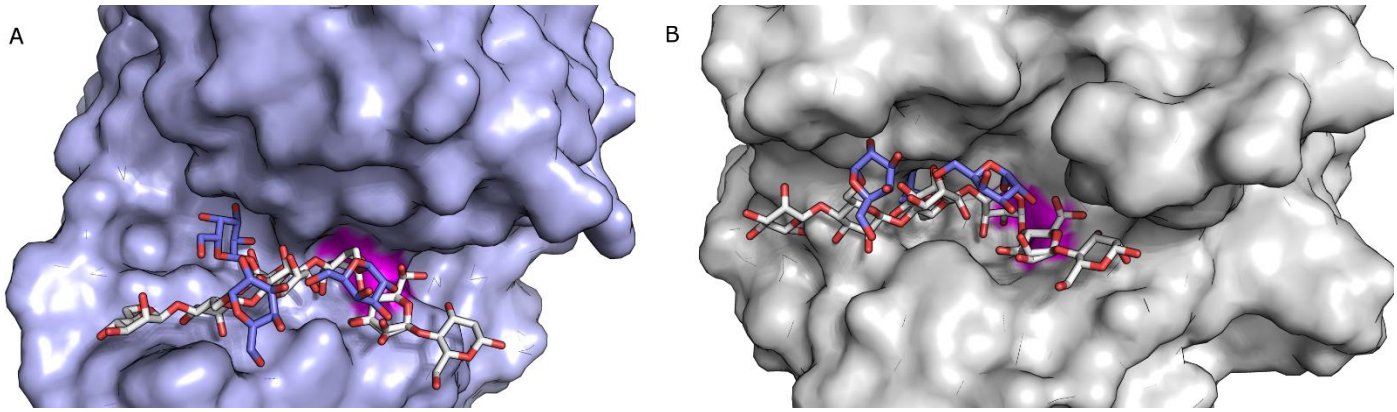


Figure S2. Surface views of the crystal structures of (A) *Ypen*Man26A from *Yunnania penicillata* and (B) *Cjap*Man26C (2VX6) from *Cellvibrio japonicus* showing the architecture of their active site cleft. A superimposition of the two structures allowed visualization of ligands from both crystal structures in each structure: α-6^2-6^1-di-galactosyl-mannotriose (MGG) binding from the -4 to -2 subsites in *Ypen*Man26A and α-6^3-galactosyl-mannotetraose (MGMM) binding from the -2 to +2 subsites in *Cjap*Man26C. Mannose units are coloured white and the galactose substitutions are coloured purple. Catalytic residues are shown in magenta.

Lane	Sample
1	LMW - 10µl
2	*Ypen*Man26A D37T – 10µl
3	*Ypen*Man26A W110H – 10µl

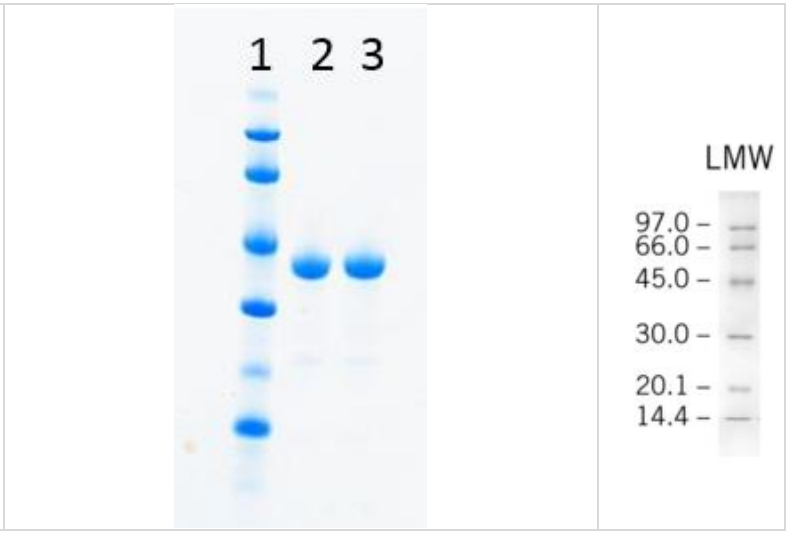


Figure S3. SDS-PAGE of purified *Ypen*Man26A mutants. Prior to gel loading, samples were diluted 1:1 with loading mix. Loading mix was prepared as a 9:1 mix of Novex ® Tris-Glycine SDS Sample Buffer (2X) (Life Technologies) and Nupage ® Sample Reducing Agent (10X) (Life Technologies). Please consult the manuscript for enzyme abbreviations. The purification of *Wsp*.Man26A is described by von Freiesleben et al 2018.

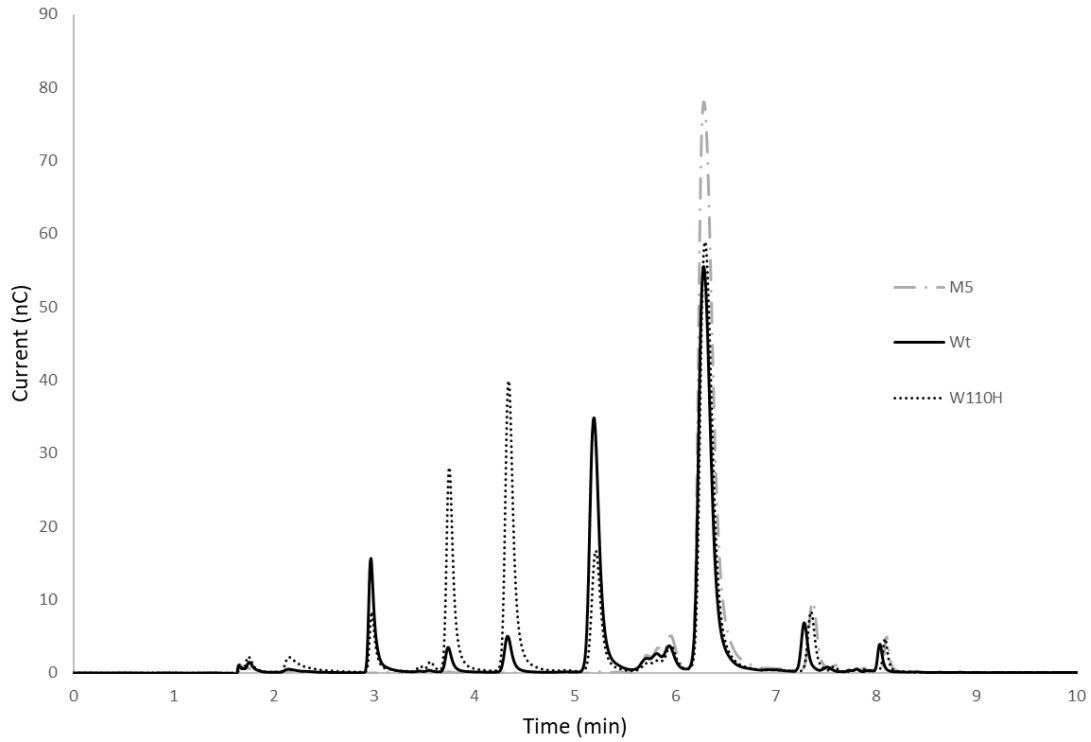


Figure S4. Representative HPAEC chromatogram of M5 hydrolysis by *Ypen*Man26 wild type and W110H mutant. The solid line shows M5 hydrolysis by *Ypen*Man22. The dotted line shows M5 hydrolysis by the W110H mutant. The dotted and dashed line is a control sample containing the M5 substrate without enzyme and incubated under the same conditions.

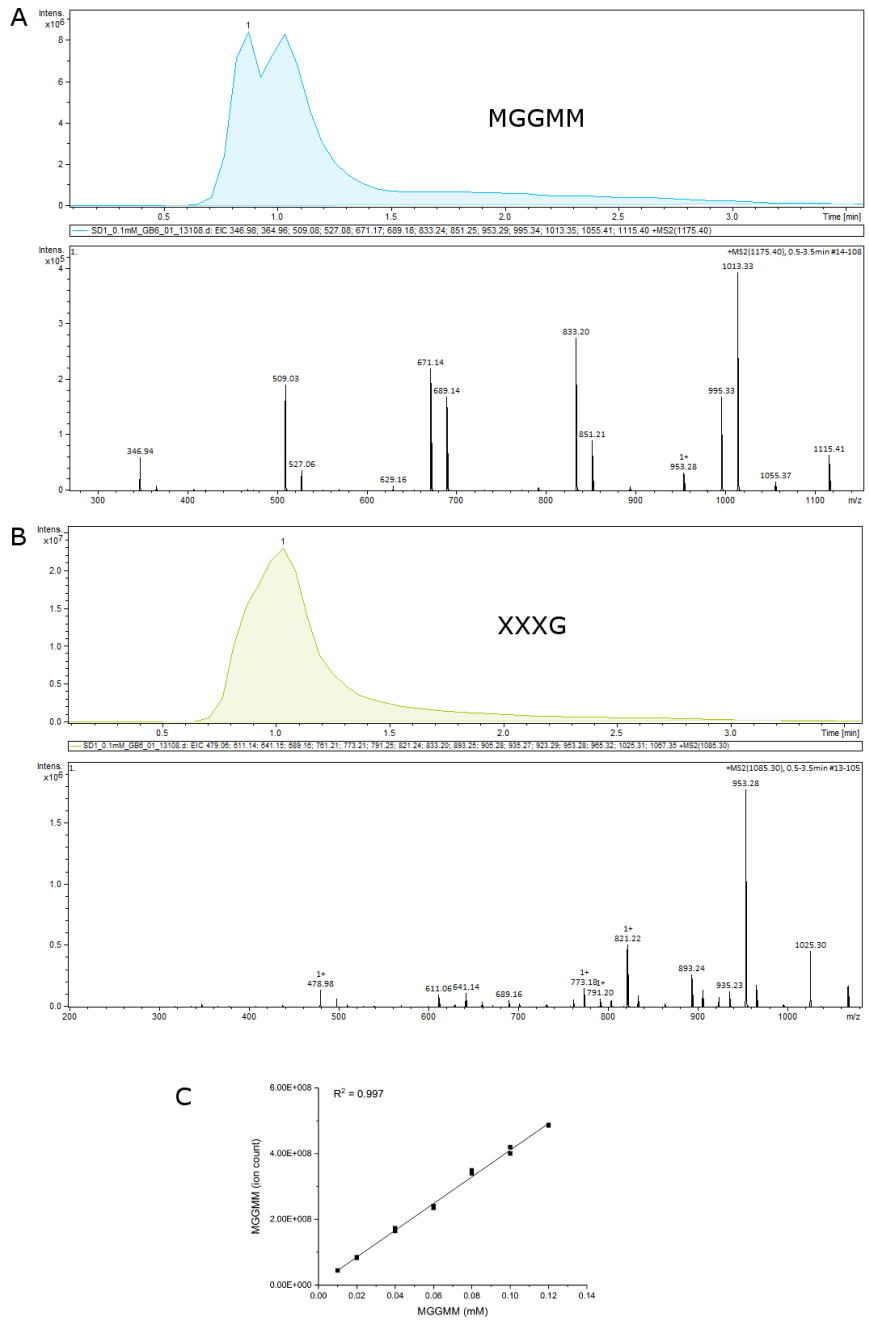


Figure S5. The observed precursor ion of both α-6^4-6^3-di-galactosyl-mannopentaose (MGGMM) and the internal standard (α-6^2-6^3-6^4-tri-xylosyl-glucotetraose) XXXG was the single charged sodium adduct $[M+Na]^+$of *m*/*z* 1175.4 and 1085.3 respectively. Data were collected with a window width of 0.5 amu. XXXG was added to the enzyme reaction as an internal standard to verify consistent signal response. XXXG was chosen because the mass and the branched structure are similar to MGGMM and because the enzymes in question did not have any activity towards it. The extracted ion chromatogram (EIC) and fragmentation pattern of $[M+Na]^+$ is shown for (**A**) 0.1 mM MGGMM (fragmentation ions used for quantification after MS/MS was *m*/*z* 346.9800; 364.9589; 509.0800; 527.0800; 671.1700; 689.1800; 833.2400; 851.2500; 953.2900; 995.3400; 1013.3500; 1055.4100; 1115.4000) and (**B**) 0.1 mM XXXG (fragmentation ions used for quantification after MS/MS was *m*/*z* 479.0600; 611.1400; 641.1500; 689.1600; 761.2052; 773.2100; 791.2500; 821.2400; 833.2031; 893.2500; 905.2779; 935.2700; 923.2869; 953.2800; 965.3159; 1025.3100). Data was obtained with an online, direct injection mass spectrometry based assay in positive ultrascan mode with Multiple Reaction Monitoring (MRM) following only MGGMM and XXXG. Peak integration was performed manually from 0.5 – 3.5 min. (**C**) A standard curve with seven calibration levels for MGGMM (0.01-0.12 mM) using XXXG as internal standard was used for quantification. Values are shown from two individual replicates, with a linear fit.

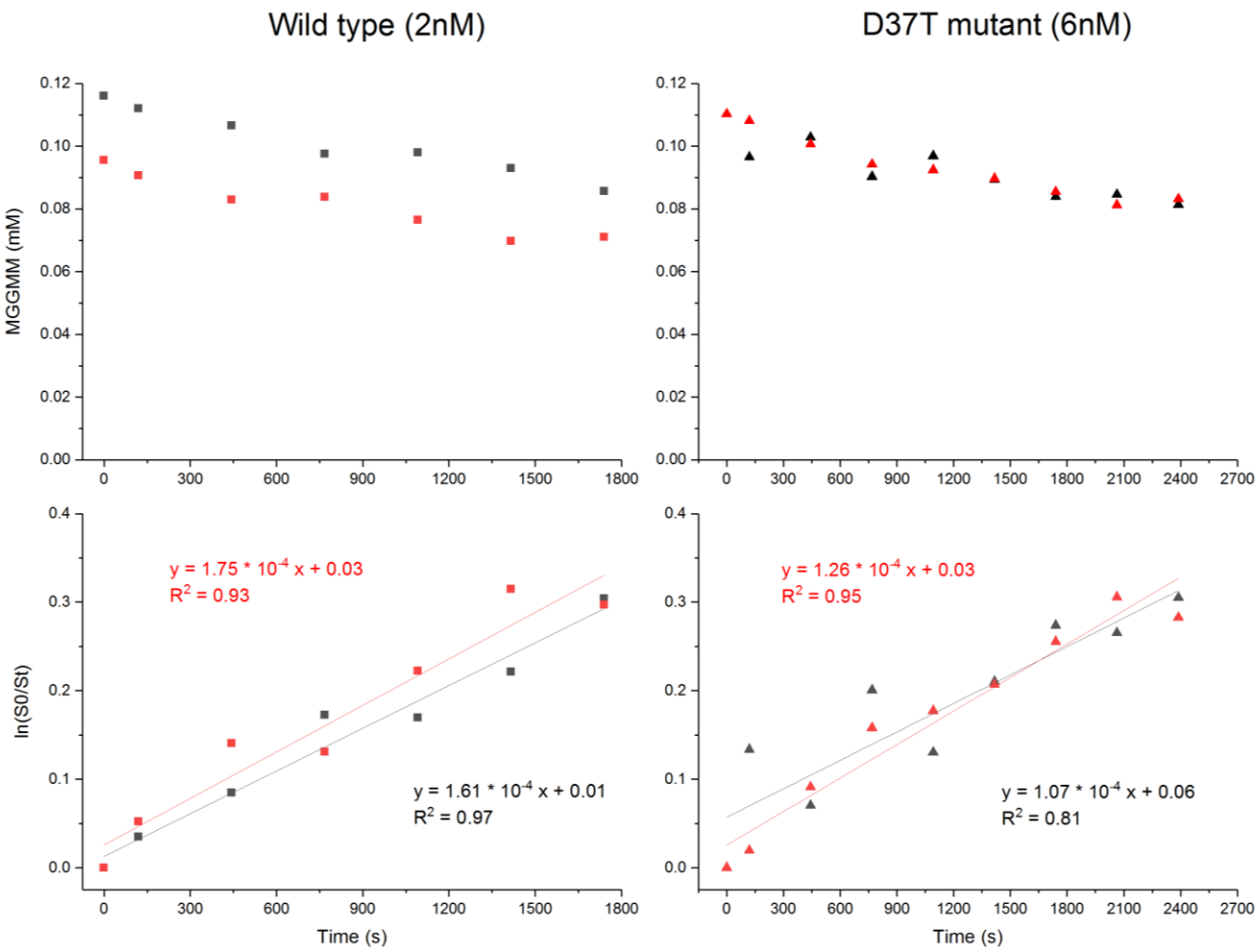


Figure S6. *Ypen*Man26A wild type (**left**) and D37T mutant (**right**) catalysed MGGMM depletion. Data are plotted as **(top)** Substrate (*S*) as a function of time (*t*) to show MGGMM depletion and (**bottom**) as Ln (S_0/S_t) as a function of *t* to illustrate the Matsui equation. Values are shown from two individual replicates (black and red), with a linear fit for each replicate in the Matsui graph (bottom). k_{cat}/K_M can be calculated according to the Matsui equation; k = Ln (S_0/S_t), where *k* = ((k_{cat}/K_M)*[enzyme])*t*, S_0 = substrate concentration at time zero and S_t = substrate concentration at time *t* (Matsui et al. 1991). Enzyme catalysed MGGMM depletion is not necessarily linear and it is not a prerequisite to use only the initial hydrolysis rate for the Matsui method.