The digestive machinery of a human gut bacterium

Structural enzymology of galactomannan utilisation

Bågenholm, Viktoria

2018

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

The digestive machinery of a human gut bacterium

Structural enzymology of galactomannan utilisation

VIKTORIA BÅGENHOLM
DIVISION OF BIOCHEMISTRY AND STRUCTURAL BIOLOGY | LUND UNIVERSITY

Printed by Media-Tryck, Lund 2018 NORDIC SWAN ECOLABEL 3041 0903

LUND UNIVERSITY

Faculty of Science
Department of Chemistry
Division of Biochemistry and Structural Biology

ISBN 978-91-7422-595-2

The digestive machinery of a human gut bacterium

Structural enzymology of galactomannan utilisation

Viktoria Bågenholm

LUND
UNIVERSITY

DOCTORAL DISSERTATION
by due permission of the Faculty of Science, Lund University, Sweden.
To be defended on October 19th 2018, at 13:15 in Hall B, Kemicentrum, Lund.

Faculty opponent
Professor Leila Lo Leggio
University of Copenhagen, Denmark

Organization
LUND UNIVERSITY
Biochemistry and structural biology
P.O. Box 124
Se-221 00 Lund, Sweden

Author(s)
Viktoria Bågenholm

Document name
DOCTORAL DISSERTATION

Date of issue
2018-09-25

Sponsoring organization

Title and subtitle

The digestive machinery of a human gut bacterium: structural enzymology of galactomannan utilisation

Abstract

Human gut bacteria utilise different types of polysaccharides present in our diet. One of these polysaccharides is galactomannan. Many organisms in the phylum Bacteroidetes have gene clusters encoding for all proteins required for hydrolysis, binding and transport of one type of polysaccharide, called polysaccharide utilisation loci (PULs). A PUL from the human gut bacterium *Bacteroides ovatus* was previously indicated to be involved in galactomannan utilisation (*Bo*ManPUL) and codes for one glycosde hydrolase (GH) family 36 α-galactosidase (*Bo*Gal36A, **Paper I**) and two GH26 β-mannanases (*Bo*Man26A and *Bo*Man26B, **Papers II-IV**).

In **Paper I** *B. ovatus* was shown to be able to grow on galactomannan, an ability which was lost upon knockout of *Bo*ManPUL, indicating that it was primarily responsible for galactomanna utilisation in this human gut bacteria. **Papers I-III** revealed a pathway for galactomanna utilisation in which *Bo*Man26B initially cleaves polysaccharide substrates outside the cell, the products of which are transported into the periplasm and there further processed by *Bo*Gal36A, then *Bo*Man26A.

*Bo*Man26B is outer membrane attached, preferntially hydrolyses longer substrates (**Paper II**) and is one of the enzymes in GH26 least restricted by galactose substitutions (**Papers II** and **III**). Crystal structures revealed a long, open active site cleft, only being restricted by galactose substitutions in one subsite -2 and possibly favouring a substitution in subsite -4 (**Paper III**). *Bo*Man26B provides sequential synergy to *Bo*Gal36A (**Paper III**), which preferentially hydrolyses internal galactose substitutions from galactomanno-oligosaccharides (**Paper I**). This hydrolysis of internal galactosyl units is unusual in GH36, which was shown to be caused by the absence of a loop that is present in other GH36 subgroup I members (**Paper I**). *Bo*Gal36A in turn provides sequential synergy to the endo-capable mannobiohydrolase *Bo*Man26A, which preferentially cleaves unsubstituted mannooligosaccharides (**Paper II**). The structure of *Bo*Man26A revealed a narrow active site cleft where at least two subsites are restricted by galactose substitutions, and which was blocked beyond subsite -2 by two loops: 2 and 8 (**Paper II**). NMR assignment of the *Bo*Man26A backbone was carried out in **Paper IV**, to perform further NMR-based studies of substarte binding and protein dynamics relating to loops 2 and 8.

*Bo*Man26A and *Bo*Man26B differ both in structure and biochemistry and in a phylogenetic analysis of selected GH26 sequences they were shown to cluster in different branches (**Paper III**). The level of conservation around the active site cleft was generally low, with the exceptions of the -1 and +1 subsites (**Paper III**).

This thesis reveals a model for galactomannan PUL utilisation in *B. ovatus*, increasing the understanding of our gut microbiota. It also delves into the structure-function relationship of substrate specificity in, primarily, GH26 enzymes.

Key words: β-mannanase, glycoside hydrolase, enzymology, structural biology, kinetics, HPLC, product profiles, structure-function

Classification system and/or index terms (if any)

Supplementary bibliographical information

Language
English

ISSN and key title

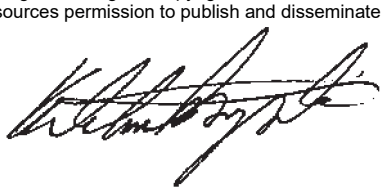
ISBN
978-91-7422-595-2

Recipient's notes

Number of pages 196

Price

Security classification

Signature 

Date 2018-09-07

The digestive machinery of a human gut bacterium

Structural enzymology of galactomannan utilisation

Viktoria Bågenholm

LUND
UNIVERSITY

Coverphoto: X-ray diffraction image of the GH26 β-mannanase *Bo*Man26A from *Bacteroides ovatus*, collected at the I911-3 beamline at the MAX-II synchotron, Lund. The image has been processed with the software Adxv.

Faculty of Science
Department of Chemistry
Division of Biochemistry and Structural Biology

ISBN 978-91-7422-595-2

Printed in Sweden by Media-Tryck, Lund University
Lund 2018

Table of Contents

List of papers

I. 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 let. 590, 2106-2118

II. 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: enzyme synergy and crystal structure of a β-mannanase. JBC 292, 229-243

III. Bågenholm, V., Wiemann, M., Reddy, S. K., Bhattacharya, A., Rosengren, A., Koropatkin, N. M., Logan, D. T. and Stålbrand, H. Manuscript: Dual GH26 β-mannanases in the galactomannan utilisation locus of *Bacteroides ovatus*: structure, role and phylogeny of BoMan26B

IV. Wernersson, S., Bågenholm, V., Persson, C., Upadhyay, S., Stålbrand, H. and Akke, M. Manuscript: Backbone 1H, 13C and 15N resonance assignments of BoMan26A, a β-mannanase of the glycoside hydrolase family 26 from the human gut bacterium Bacteroides ovatus

My contributions to the papers

I. I assisted with the SEC experiment and homology modelling. I took a major part in analysis of the homology model and preparing those parts of the manuscript, as well as participating in discussions and general manuscript revision.

II. I took part in planning the study and took a major part in experimental work. I performed basic enzyme characterisation, substrate profiling, and kinetics. I took a major part in crystallisation, data collection and structure refinement, as well as general analysis and drafting the manuscript.

III. I took a major part in designing the study. I performed expression and purification, metal stability, synergy and product length experiments. I assisted with preparation and analysis of the kinetics and took part in the analysis of the ^{18}O-labelling. I took a major part in crystallisation, data collection, structure refinement and bioinformatics analysis. I also performed a major part of data analysis and drafted the manuscript.

IV. I expressed and purified all protein used in the study, as well as drafting the sections of the manuscript related to that. I also took part in drafting the whole manuscript.

Svensk sammanfattning

Vi kryllar av bakterier. De bor framför allt i vår tjocktarm där de lever på de delar av vår mat som vi inte kan ta hand om själva. I utbyte omvandlar bakterierna en del av maten de äter till vitaminer och fettsyror som är viktiga för oss. Det bor flera olika arter av bakterier i vår tarm som äter lite olika saker, eftersom det är ont om mat för så många har de flesta specialiserat sig för att minska konkurrensen.

Våra tarmbakterier är viktiga för vår hälsa. Förutom att påverka relativt uppenbara saker som en irriterad tarm eller övervikt, har de också en inverkan på så skilda saker som immunförsvaret och blodtrycket. Eftersom vår tarmflora är så viktig och lever på det vi äter vill man förstå exakt vad olika arter äter för att kunna påverka vår tarmflora i en positiv riktning genom kosten. En del av det som finns i maten som bakterierna får tillgång till är kolhydrater vilket är olika sorters socker, som när de sitter ihop i långa kedjor kallas polysackarider. De mest välkända polysackariderna är nog cellulosa, som bland annat används i papperstillverkning, och stärkelse: en självklar del av vår kost. Stärkelse är den enda polysackarid i vår mat som vi kan bryta ner själva, men vi äter många olika sorters polysackarider dagligen. Det finns väldigt många andra typer av polysackarider, främst i växter, som består av ett stort antal olika socker, så kallad hemicellulosa. Hemicellulosa är en typ av kostfiber och är en av källorna till mat för våra tarmbakterier.

En sorts hemicellulosa heter galaktomannan och består av långa kedjor av sockret mannos med sockret galaktos som sticker ut från denna kedja. Galaktomannan finns i frön från vissa baljväxter, alltså i bönor och ärtor, och används som förtjockningsmedel i en del mat, till exempel glass. För att kunna bryta ner galaktomannan behövs enzymer. Enzymer är protein som tillverkas av alla levande celler. Enzymerna utför olika typer av reaktioner, till exempel att klippa sönder polysackarider eller bygga ihop DNA. För att kunna bryta ner galaktomannan behöver våra tarmbakterier framför allt två olika sorters enzymer: α-galaktosidas, som klipper bort galaktosen som sticker ut från mannoskedjan, samt β-mannanas som klipper sönder själva huvudkedjan.

En vanligt förekommande tarmbakterie som kan bryta ner galaktomannan heter *Bacteroides ovatus*. I denna bakterie, samt i många besläktade bakterier, finns det specifika delar av DNA som innehåller instruktionerna för alla proteiner som krävs för att bryta ner och använda en viss typ av polysackarid, som till exempel galaktomannan. Förutom enzymer finns i dessa DNA delar även proteiner som

håller fast polysackariden vid cellytan samt proteiner som transporterar den sönderklippta sockerkedjan in i cellen. Det är enzymerna i denna DNA del som denna avhandling har fokuserat på: ett α-galaktosidas och två β-mannanas.

α-Galaktosidaset, som framför allt studerades i **Artikel I**, kunde ta bort upp till 90% av alla galaktos från galaktomannan. Detta var förvånande eftersom alla liknande enzym som man känner till bara kan ta bort galaktos som sitter i änden på en sockerkedja, medan detta enzym kunde ta bort galaktoser som stack ut mitt i kedjan. För att förstå hur detta var möjligt gjordes en modell av hur enzymet ser ut. Andra liknande enzymer har en lång loop som ligger precis bredvid den del av enzymet klipper galaktos. Denna loop var mycket kortare i α-galaktosidaset från *Bacteroides ovatus*, vilket gjorde att en längre sockerkedja fick plats.

I DNA-delen finns två β-mannanaser. Den ena, mannanas A, studerades i **Artikel II** och var mycket bra på att klippa sönder korta mannoskedjor, men hade svårt för att göra det om det satt galaktos på kedjan. Om mannanas A fick jobba ihop med α-galaktosidaset blev den betydligt effektivare. Mannoskedjorna blev framför allt nerklippta till mannobios, en molekyl som består av två mannossocker. Strukturen av mannanas A, alltså hur den såg ut, visade att enzymet formar en klyfta som en mannoskedja kan lägga sig i, men där galaktoser har svårt att få plats. Ena änden av klyftan blockeras av en loop som gör att bara två mannossocker får plats. Detta är förbryllande eftersom enzymet kan klippa ner galaktomannan till längre sockerkedjor än mannobios, även om den inte föredrar det. **Artikel IV** förbereder för en studie som ska ta reda på hur mannanas A rör på sig för att se om den blockerande loopen kan flytta på sig när enzymet binder sockerkedjor.

Det andra β-mannanaset, mannanas B, studerades i **Artikel II** och **III** och, till skillnad från mannanas A, klippte mannoskedjan lika bra oavsett om det fanns galaktos på den eller inte. Mannanas B hjälpte α-galaktosidaset att bli mer effektivt, men blev inte själv effektivare av att de jobbade ihop. Dessutom visade det sig att detta enzym sitter på utsidan av cellen, vilket de andra två inte gör. I **Artikel III** visade strukturen av mannans B att det också hade en klyfta, men att den var öppnare, för att får plats med galaktos, och längre, vilket gör att enzymet föredrar längre sockerkedjor.

Detta ger en bild av hur *Bacteroides ovatus* bryter ner galaktomannan: β-mannanas B klipper galaktomannan till kortare kedjor på utsidan av cellen. Dessa kortare kedjor transporteras in i cellen där α-galaktosidaset tar bort galaktosen varefter β-mannanas A klipper sönder resten till mannobios som kan användas för energiproduktion. Detta ökar förståelsen för hur en tarmbakterie bryter ner en typ av kostfiber, vilket bidrar till helhetsbilden för hur vår tarmflora fungerar. Förhoppningen är att, genom att ha en detaljerad helhetsbild av vår tarmflora, så småningom kunna utveckla dieter och kosttillskott för att hjälpa dem vars tarmflora är ur balans, till exempel på grund av antibiotika.

Acknowledgements

There are many people I would like to thank for support and encouragement during these years.

Firstly I would like to thank my supervisor, Henrik Stålbrand, for letting me do a phD in his group. You have always taken a great interest in my work and provided a lot of valuable input over the years. I have really developed and learnt a lot during my time working with you.

To my co-supervisor, Derek Logan: I was glad to have you as a co-supervisor and for your help on the structural aspects of my work. Our meetings were generally short, but fruitful.

Thanks to all the current and former members of our group, for a fun work environment and interesting discussions about science and all kinds of other stuff. To the current group members, Anna, Samuel, Abhishek and Mathias: thanks for lots of support and help with lab work during my thesis writing, and Abhishek for the final time in the office together. Sumitha, thanks for good times sharing an office and all the help with phylogenetics for the current manuscript.

To the students I have supervised in our lab, Leila and Ioannis: I enjoyed working with you and you taught me a lot about supervision.

During my time as a phD I have collaborated with a number of people on different projects. Thanks to Eric Martens and Nicole Koropatkin for the work on several of the papers, such as the growth studies and cellular location in Paper II. To Oskar Aurelius, thanks for help with the *Bo*Man26A structure, and to Mikael Akke, Santosh Upadhyay, Angus J Robertson, Sven Wernersson and Cecilia Person for the collaboration on the NMR study. Polina Naidjonoka: thanks for the collaboration on the galactose release measurements. It was fun to get some insight into a more applied project.

CMPS has been a fun department to work at, with many great people who have come and gone. I would especially like to thank: Renzo Johansson, for many hours spent fussing over structures and talking about life. Andreas Kirscht, for the glorious adventures with Adorno and Lingneng. Sven, Olof, CJ and everyone else in the BC gang: you make Thursday a day to look forward to. A special thank you

of course also to Maryam, Adine and Magnus, without which this department would not function.

Apart from my colleauges I have had many friends and family that have supported me during my phD and made these years a lot more fun.

Tack till mamma och pappa, som alltid har stöttat och hjälpt till oavsett vad jag vill göra, och till mamma som fick mig att förstå att det var forskare jag ville bli. Till Anna och Isabell: jag är glad att vi hamnade i samma stad alla tre och det gjorde det mycket trevligare att bo här.

Till mina tjejer: Victoria, Wictoria, Bea, Augusta, Shqipe och Johanna. Vi har alla gått vidare i livet på olika håll, men håller ihop ändå. Det har alltid betytt mycket att vi fortsatt ses och fortfarande kan prata om precis allt.

Till Viktoria Tidqvist: jag är glad att vi fortfarande håller ihop och ses, även om det inte är lika ofta längre.

Thanks to the roleplaying gang, with Jonas, Martin, Steph and Sebastian, for all the crazy adventures and amazing characters. I hope we can spend many more evenings getting side-tracked from plot, and what we are doing in general, and frustrating our DM.

To Ben: we have had quite a journey together and I will always be happy you decided to come to Sweden with me. Soon we will hopefully both start new jobs and see where those take us. These last few months would have been a lot harder without you. Now I will finally have time for all the nerd things again. I love you.

Abbreviations

AgaA	GH36 α-galactosidase from *Geobacillus stearothermophilus*
B. theta	*Bacteroides thetaiotaomicron*
*B*750Man26A	GH26 β-mannanase A from *Bacillus* sp. strain JAMB-750
BACOVA_03432	GH30 glucuronoxylanase encoded by a xylan PUL from *Bacteroides ovatus*
*Bf*Man26	GH26 β-mannanase from *Bacteroides fragilis*
*Bo*Gal36A	GH36 α-galactosidase A from *Bacteroides ovatus*
*Bo*GH5A	GH5 endo-xyloglucanase encoded by a xyloglucan PUL from *Bacteroides ovatus*
*Bo*ManPUL	*Bacteroides ovatus* Galactomannan Polysaccharide Utilisation Locus *bacova_02087-97*
*Bo*Man26A	GH26 β-mannanase A from *Bacteroides ovatus*
*Bo*Man26B	GH26 β-mannanase B from *Bacteroides ovatus*
CAZymes	Carbohydrate Active Enzymes
CAZy	The Carbohydrate Active Enzymes database
CBM	Carbohydrate Binding Module
*Cj*Man26A	GH26 β-mannanase A from *Cellvibrio japonicus*
*Cj*Man26C	GH26 β-mannanase C from *Cellvibrio japonicus*
CUT	Carbohydrate Utilisation Locus containing TonB-dependent transporters
DNS	3,5-dinitrosalisylic acid
DP	Degree of Polymerisation
G1M4	6^3-α-D-Galactosyl-mannotetraose
G2M5	6^3,6^4-α-D-galactosyl-mannopentaose
GGM	Galactoglucomannan
GH	Glycoside Hydrolase
GM	α-1,6-galactosylmannose
GpPUL	Gram positive Polysaccharide Utilisation Locus
HPAEC-PAD	High Performance Anion Exchange Chromatography with Pulsed Amperometric Detection
HTCS	Hybrid Two Component System
INM	Ivory Nut Mannan
KGM	Konjac Glucomannan
LBG	Locus Bean Gum

MALDI-TOF	Matrix Assisted Laser Desorption-Ionisation Time Of Flight
M*n*	Mannooligosaccharide with degree of polymerisation *n*
MS	Mass Spectrometry
MST	Microscale Thermophoresis
NMR	Nuclear Magnetic Resonance
*Pa*Man26A	GH26 β-mannanase A from *Podospora anserina*
PUL	Polysaccharide Utilisation Locus
PULDB	Polysaccharide Utilisation Locus Database
RMSD	Root Mean Square Deviation
*Rs*Man26C	GH26 β-mannanase C from *Reticulitermes speratus*
SCFAs	Short Chain Fatty Acids
Sus	Starch Utilisation System
TROSY	Transverse Relaxation-Optimised Spectroscopy

Introduction

The living world around us is made up largely of plant material and in many organisms largely persists on various forms of plant material, from the grasses that cows eat to the fruit and vegetables that we humans consume. This is also true for many bacteria and other microbes, as one of the causes for a rotting log in the forest or as symbiotic scavengers in the intestinal tracts of larger animals. The human body is home to a large host of microbes, the majority of which are located in the intestinal tract: the gut microbiota. These microbes live off the things in our food that we can not digest ourselves, one such important group of compounds being various types of carbohydrates. Recently, understanding of the importance of gut microbiota for our health has increased. Today we know the organisms in our gut affects several health issues related to diet, like irritable bowel disease and obesity, but can also influence aspects less intuitively connected to our gut, like the immune system and blood pressure. Understanding exactly which organisms live in our gut, how they function and which of them are good for us has become increasingly important as we realise their impact on our wellbeing. We could then use this knowledge to develop ways of maintaining a healthy gut flora, especially in cases where it has been disrupted, such as during treatment with antibiotics.

A major source of nutrients for our gut microbiota is undigested parts of the food we eat. One way to help maintain a healthy gut flora could thus be through our diet. Compounds in food we eat that promote a healthy gut flora are called prebiotics. In the densely populated environment of the gut many bacteria have specialised to use certain types of nutrients to reduce competition for different food sources. One such food source is various types of carbohydrates. The main polymeric carbohydrate humans are capable of digesting is starch, but there are many other types of polysaccharides present in plants, and as such also in our diet but not digested by the human body. One example is hemicellulosic polysaccharides. But if we are to affect our gut microbiota through our diet, we must first understand what they eat.

We daily consume a large amount of hemicellulosic polysaccharides: they are part of the plant cell wall or are used as plant storage polysaccharides. One of the many types of polysaccharide present in our diet are the β-mannans, such as galactomannans, which are present in our diet as storage polysaccharides in

legume seeds and as food thickeners in, for example, ice cream. β-Mannan polysaccharides can be utilised as nutrients by some of our gut bacteria.

One of the common human gut bacteria capable of growing on galactomannan polysaccharides is *Bacteroides ovatus*. This thesis has aimed to increase the detailed understanding of galactomannan utilisation in *B. ovatus*, to further our knowledge about the human gut microbiota for potential development of new prebiotics. In addition, by detailed structural and biochemical analysis of the enzymes involved in the process of galactomannan utilisation an increased understanding of the fine-tuned specificity of galactomannan acting enzymes in *B. ovatus* and similar enzymes has been sought.

Background

Plant polysaccharides

Polysaccharides are an important part of our diet. They are either directly processed by our own enzymes and used as an energy source by us, as in the case of starch, or they are first processed and metabolised by our gut bacteria, in which case they are usually referred to as dietary fibres [1]. Polysaccharides consist primarily of sacchride units and these types of molecules are some of the most abundant in the living world with a wide variety of functions, from energy storage to structural support and signalling [2]. A prevalent function of polysaccharides is as a structural part of the plant cell wall [3].

The plant cell wall

Plant cells are surrounded by a cell wall: a barrier with multiple layers that provides rigidity, while some layers still being flexible enough to allow for communication and growth [3]. The primary components of the plant cell wall are the polysaccharides cellulose, hemicellulose and pectin, various types of proteins and, in woody material, the phenolic compound lignin [4]. The cell wall is made up of different layers, the details of which vary dependent on species [5]. During growth the cells are connected by the pectin rich middle lamella and the primary cell wall, which contains mostly pectin, cellulose, hemicellulose and some proteins [4]. After growth a secondary cell wall is developed, primarily present in woody plant material, that in itself has three layers with different structure [4]. These layers consist of cellulose, hemicellulose and lignin, with some protein [4]. For all layers cellulose fibres provide the framework for the wall, connected by a matrix of hemicellulose and pectin/lignin. The precise structure of this network is unknown and varies with different species, as reviewed by Shtein et al. [5]. A current model proposes a matrix with load bearing hotspots of tighter interaction between cellulose and hemicellulose [6].

Seed cell walls, which also contain polysaccharides, are different since their primary function is as a source of energy for the growing plant [3]. They are highly thickened and generally low on cellulose to make them easier to digest [7]. In many seeds and grains, such as wheat and rice, the primary energy storage

polysaccharide is starch [8,9], but there are also many examples of different heteroglycans being used as storage polysaccharides, as reviewed by Buckeridge [10].

Other sources of polysaccharides

Apart from their function in plant cell walls, polysaccharides are widely present across life [2,11,12]. A common post-translational modification of proteins is glycosylation: the attachment of a variety of saccharide structures. While mainly found in eukaryotes, protein glycosylation has also been found in bacteria and archaea [13]. An example of a group of glycosylated eukaryotic proteins is the mucins: a large component of the mucus layers in humans [11]. Many unicellular organisms produce cell walls or other structures surrounding the cell. Gram positive bacteria contain thick capsules composed of a mesh of peptides and polysaccharides [12] and many bacteria are capable of making carbohydrate rich biofilms for adhesion and protection. In some cases these have been implicated in pathogenicity and resistance to antibiotics [14-16]. Carbohydrate biofilms and cell walls are also present in other unicellular organisms, such as microalgae [17-19]. Gut bacteria produce a diverse range of extracellular polysaccharides, and genes encoding for proteins involved in their synthesis are more abundant in several gut bacteria compared to bacteria from environmental samples such as soil [20].

Cellulose, starch and pectin

Cellulose and starch are probably the most well known polysaccharides and are composed entirely of glucose [3,21]. Cellulose consists of straight β-1,4-linked glucose chains. These associate through hydrogen bonding to form microfibrils that are the backbone of the plant cell wall [22,23]. Cellulose has for a long time been used extensively for paper and textiles production [22], and is being developed into various renewable materials to replace fossil fuel-based products [24-26], and for use in medical applications [27,28].

Starch mainly functions as an energy storage polysaccharide in the plant kingdom. It is branched and has a backbone of α-1,4-linked glucose units, with the branches attaching through α-1,6-linkages [9]. Current and potential applications include production of biofules [29,30], bioplastics and other replacements for oil-based products [26,31], which causes problematic competition with food production [30].

The pectin carbohydrates are a group of heterogeneous polysaccharides that are composed of galacturonic acid units and are mainly found in the primary plant cell wall and middle lamella [32]. These components of the cell wall also have

established industrial applications, primarily in food industry to clarify fruit juices and as gelling agents in jams and similar [33,34]. They have also been investigated for a variety of pharmaceutical applications [34].

Hemicellulosic polysaccharides

Hemicelluloses are the other major family of heterogeneous polysaccharides in the plant cell wall, closely related in composition to hemicellulosic polysaccharides with different functions, such as seed storage polysaccharides [3]. They are overall structurally similar with β-1,4-linked backbones of xylose, glucose or mannose units with C1 and C4 in an equatorial conformation [3]. These hemicellulosic polysaccharides are divided into the subgroups xylan, glucan, xyloglucan and mannan, depending on the main constituents of their backbones [3]. Apart from the backbone of oligo- and polysaccharides they may contain substitutions: attachments to the backbone of other sugars or small molecules, such as acetyl groups [35,36].

Xylans are made up of a β-(1,4)-linked xylose unit backbone that can be substituted with several types of α-linked sugar units, such as arabinose and glucuronic acid, as well as other modifications such as acetylations [37,38]. Xylans are present in large amounts in some cereals (30-50%) [39] and is a major component of the cell wall of hardwoods (25-35%) [40]. β-(1,3)-linked xylans and those with a mixed β-(1,3)(1,4)-linked backbone have also been found in algae [41,42]. Several human gut bacteria have been shown to grow on xylan and xylo-oligosaccharides [43,44] and they are being investigated for applications in bioplastics, in drug delivery and for generation of xylitol as a sugar replacement, as reviewed by Naidu et al. [45].

β-Glucans have a β-(1,3)(1,4)-mixed linkage backbone and are mostly found in grasses, including grains like wheat and barely, both as part of the cell wall and as storage polysaccharides [46]. Xyloglucans, on the other hand, consist of a β-(1,4)-linked glucose unit backbone with α-(1,6)-xylose unit substitutions and are a part of the primary cell wall of all land plants [47]. Mixed linkage β-glucans are associated with a number of health benefits and has found application in immunological stimulation, as well as in anti-cancer and anti-inflammatory therapy, as reviewed by Bashir and Choi [48]. They promote growth of probiotic bacteria that are potentially beneficial for our health [49]. Xyloglucans, especially tamarind seed polysaccharide, are investigated for various drug delivery applications [50,51].

β-Mannans

The β-mannans, a group of hemicellulosic polysaccharides, have a backbone of β-1,4-linked mannose units and are grouped into linear mannans, galactomannans, glucomannans and galactoglucomannans (GGMs) (Figure 1) [3]. Linear mannans, galactomannans and glucomannans are mainly used by plants in seeds as energy storage, while GGMs are primarily part of the plant cell wall [3]. The linear mannans (Figure 1) are the simplest and can be found in the endosperm of the ivory palm nut (ivory nut mannan, INM) [52], African oil palm [53] and in coffee beans [10], as well as in some algae [54,55]. These polysaccharides are relatively short with a low degree of polymerisation (DP): INM has a DP of 15-20 [52,55]. Due to the linear nature and strong intermolecular interaction of these types of β-mannans they are insoluble in water, despite their relatively low DP [10]. There is currently little industrial use for linear mannan, it is mostly removed from food products such as coffee to reduce viscosity [56]. There has been some investigation into the production of oligosaccharides from African oil palm mannan for use in promoting a healthy gut flora, with initially positive results from *in vitro* growth studies [53].

Glucomannans have a backbone that contain a mix of mannose and β-1,4-glucose units [57] that can be partially acetylated (Figure 1) [58] and/or branched [59]. The acetylations and branches interfere with intermolecular interaction, making them viscous, but water soluble [57]. One such polysaccharide is konjac glucomannan (KKM, from *Amorphophallus konjac*), with a characterised DP of 4000-5000 [57]. It has been used as a flour in Japan for food stuffs like noodles for a long time [57] and is also used as an emulsifier and thickener in food [60]. Oligosaccharides derived from glucomannan have been investigated in relation to potential positive effects on human health and the growth of beneficial gut bacteria [59,61].

In galactomannan the mannan backbone is decorated with α-1,6-galactose side chains (Figure 1). Galactomannans are most commonly found as storage polysaccharides in legume seeds [3,62]. As they are water soluble, but highly viscous, they have seen wide application as various types of food additives such as thickeners and stabilisers [63,64], as well as for various pharmaceutical and industrial applications, as reviewed by Thombare et al. [65]. The two most commonly used galactomannans in industry are locust bean gum (LBG, from *Ceratonia siliqua*) [63] and guar gum (from *Cyamopsis tetragonoloba*) [64]. LBG has about 1 galactose unit for every 4 mannose units distributed along the chain in blocks of more or less galactose substitutions [63] and a DP of about 1500 [62]. Guar gum has a 1:2 galactose:mannose ratio with a more even galactosyl distribution [64] and a DP of about 900 [62].

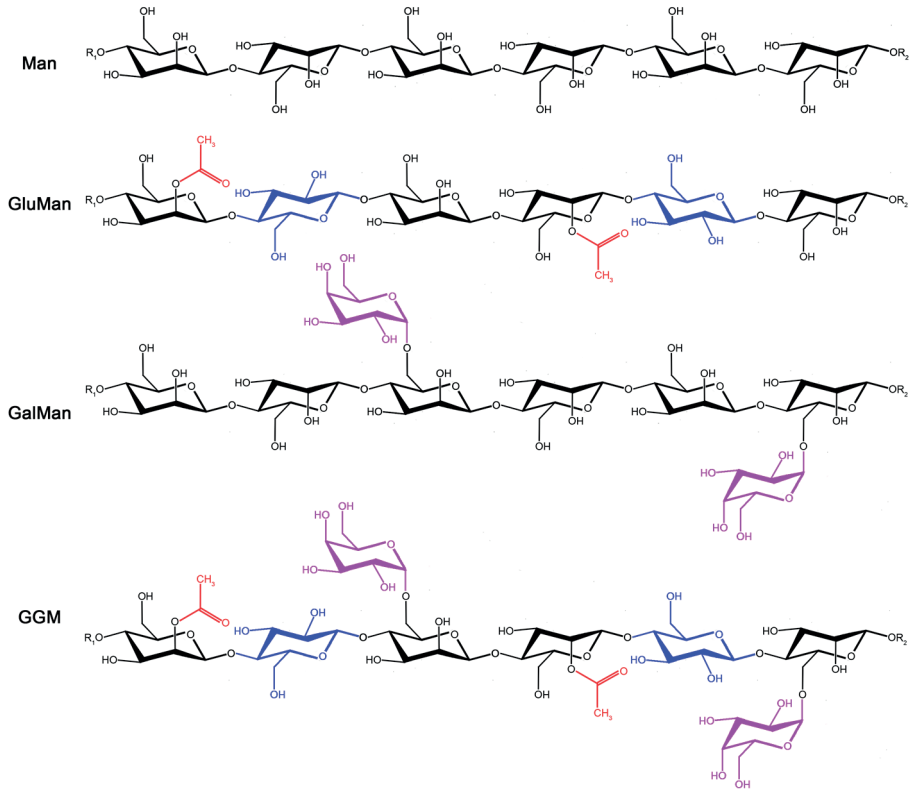


Figure 1 Schematic representation of different mannan polysaccharides, where R_1 and R_2 denote continuation of the sacchride chain, showing examples of linear mannan (Man), glucomannan (GluMan), galactomannan (GalMan) and galactoglucomannan (GGM). The galactomannan and GGM contain acetylations (red). The saccharides are coloured as follows: mannose is red, glucose is blue and galactose is violet. The saccharide backbone is connected through β-1,4-linkages, while the galactose substitutions are attached with α-1,6-linkages and the acetylations are attaced on the O2 of the mannosyl units.

GGM is the most complex of the characterised β-mannan polysaccharides as it consists of a β-(1,4)-linked backbone of mannose and glucose units with α-(1,6)-galactosyl substitutions and is often O-acetylated (Figure 1) [66]. It has a DP of 100-150 and is the main hemicellulose in softwoods, where it constitutes up to 25% of the dry weight [66,67]. GGM is also a part of cell walls in some fruits such as apple and kiwi, although the structures of these are less studied than their softwood counterparts [68,69]. GGM is being investigated for a range of applications, such as hydrogels [70], various coatings or composite films [71,72] and has been indicated to stimulate growth of human gut bacteria as judged by *in vitro* growth tests [73].

Human gut bacteria

Humans are estimated to have about as many bacterial cells as human cells in and on our bodies [74], the majority of which reside in our large intestine. This microbial community is called the gut microbiota [75]. The gut microbiota is a dense microbial environment dominated by bacteria, but to a lesser extent also includes fungi and other microorganisms [76]. Our bacteria are first introduced at birth, when we are exposed to the microbial flora of our mothers [77]. This shapes our gut flora early in life, the composition becoming increasingly affected by our diet [78,79]. In this symbiotic relationship our gut microbiota have a sheltered, anaerobic environment with access to nutrients, including carbohydrates [80,81]. In turn they supply us with many essential nutrients, such as short-chain fatty acids (SCFAs) that provide us with additional energy [80,81], as well as confer many other benefits, described in more detail below. The human body is incapable of using most polysaccharides for itself, starch being the major exception [80]. Other polysaccharides, such as hemicellulosic polysaccharides, are commonly termed dietary fibre and these are instead transported through our digestive system to be digested by our gut microbiota [82]. Some polysaccharides, such as cellulose, can not be extensively utilised by out gut microbiota and requires specialised organs, present in ruminants such as cows [83]. Apart from dietary fibre some organisms in the gut microbiota can live off the carbohydrate rich mucosal glycoproteins, such as mucins, that make up the mucus layer of the intestinal wall [84]. Utilisation of mucus saccharides appears to be especially prevalent during the suckling period when the amount and variety of dietary fibre is very low [85]. The microbial environment in the gut is very dense with an assumed fierce competition for different food sources [78]. This competition may be one reason why many of the microorganisms are specialised towards certain carbohydrates, with different organisms having partially overlapping carbohydrate degradation profiles [85-87]. While some bacteria are selfish [88] many are cooperative and improve the growth of each other [89,90].

Effects of the human microbiome relationship

In recent years it has become increasingly clear that the gut microbiota is important for many aspects of our health [91]. Apart from generating SCFAs [92] a healthy gut microbiota provides colonisation resistance against pathogens [93], thickens the intestinal mucosal layer [94] and increases vessel density in the small intestine [95]. Dysbiosis, an imbalance in the gut microbiota, has been implicated in a wide range of negative health effects such as inflammatory bowel disease [96], high blood pressure [97], obesity [98] and cancer [99,100]. A desire to treat dysbiosis and promote a healthy gut microbiota has given rise to development of

probiotics, defined as live microbial cultures that are consumed and confer health benefits on the host [101,102]. Several probiotic bacteria, for example *Lactobacillus acidophilus*, *Lactococcus lactis* and *Bifidobacterium* species, already see extensive use in the food industry for products like yoghurt [103,104]. Other potentially probiotic bacteria are currently under investigation such as *Roseburia* and *Faecalibacterium prausnitzii* [78]. Apart from probiotics there is also prebiotics, defined as non-digestible food components that have a positive effect on our gut microbiota and promote growth of beneficial bacteria [101]. Many of these are selective to certain bacteria, due to the specialisation of different species in the gut [91]. Prebiotic carbohydrates are generally different types of dietary fibre, including oligosaccharides, which can be a natural part of our diet or part of food supplements. The β-fructan inulin is a known prebiotic [105] and many oligo- and polysaccharides, such as resistant starch, galacto-oligosaccharides, xylo-oligosaccharides and β-manno-oligosaccharides have been shown to have prebiotic potential [43,44,106-108].

Common gut bacteria

The adult human gut is a dense but relatively non-diverse microbial community [75]. Though there are members of fungi, archea and several different families of bacteria found in the human gut, this environment is dominated by two bacterial phyla: Bacteroidetes and Firmicutes [109,110]. Bacteroidetes appears to vary more between individuals, while Firmicutes vary more over time in the same individual [111]. Together they make up about two thirds of the adult human gut microbiota [75]. Of these two phyla the Gram-negative Bacteroidetes, especially the *Bacteroides* genus, encodes the most enzymes for utilisation of complex carbohydrates: sugar chains consisting of at least two sugar types [1,112]. *Bacteroides* are one of the most abundant genera of the human gut microbiota and are generally found in animal gastrointestinal tracts [81]. Two prominent members of this genus are *Bacteroides thetaiotaomicron* (*B. theta*) and *Bacteroides ovatus*. Both bacteria are able to grow on several types of polysaccharides and together they can utilise most of the complex carbohydrates in our diet [82,113,114] (**Paper I**).

Polysaccharide Utilisation Loci

Degradation of polysaccharides is a complex process, with separate enzymes usually being required to hydrolyse different specific bonds between different specific sugar residues. In addition the polysaccharide and its hydrolysis products need to be bound to the cell surface, transported and often further processed for

energy production [115]. For this purpose many Bacteroidetes contain gene clusters termed polysaccharide utilisation loci (PULs). Each PUL encodes for all the enzymes and proteins required for the complete utilisation, binding and transport of one specific type of polysaccharide [110,113,115,116]. One of the most versatile polysaccharide degrading organisms in the human gut is *B. theta* with 88 PULs and 246 carbohydrate active enzymes (CAZymes) recorded so far. Apart from being capable of using several different types of dietary saccharides, it can also use mucins in the intestinal wall [112-114]. It is the organism where the first PUL was characterised: the *B. theta* Starch Utilisation System (Sus) [115,117-119], which allowed the identification of other PULs based on homology [115]. All identified PULs contain at least one homologue to the outer membrane proteins SusC, a transporter, and SusD, a sugar binding protein, from Sus in *B. theta* [115]. PULs that target highly branched or complex glycans can have two SusCD pairs [120].

Predicted and characterised PULs have been assembled in the PUL database (PULDB, http://www.cazy.org/PULDB/) [121] in the carbohydrate active enzymes database (CAZy) (www.cazy.org) [122]. Some evidence of PUL-like structures in non-Bacteroidetes has been described, with the proposed name Carbohydrate Utilisation locus containing TonB-dependent transporters (CUT) [116]. In Firmicutes, gene clusters encoding carbohydrate utilisation proteins have been identified that lack SusC and SusD homologues, termed Gram-positive PULs (GpPULs) [123]. However, generally the genes for polysaccharide processing in other bacteria tend to be more spread out [124,125].

In contrast to *B. theta*, *B. ovatus* is capable of growing on hemicellulose [113]. It has 112 putative PULs, where six of these were shown to be upregulated when grown on hemicelluloses: xylan, xyloglucan, galacto- and glucomannan and β-glucan [113]. Out of these, PULs for xyloglucan [126], xylan [127-129] and β-glucan [130] utilisation have previously been characterised. However, while it had been shown that *B. ovatus* can grow on galactomannan and had cell associated enzymes with β-mannanase activity [131], these have only recently been characterised (**Papers I-III**).

The following sections will shortly summarise the proteins typically seen in PULs, with the exception of the CAZymes, which will be detailed in a later section.

Sugar binding and transport: the SusCD complex

The two universal proteins for all PULs are SusC and SusD. They are both typically located at the extracellular membrane where SusD binds carbohydrates to the surface of the cell [118], while SusC transports cleaved products into the periplasm [117]. SusD is a membrane anchored, mostly α-helical protein with one

variable β-sheet region responsible for substrate binding, with the remaining majority of the structure forming a conserved superhelix [120]. SusC is a TonB-dependent transporter with a 22-stranded β-barrel and a plug domain in the periplasmic space [132]. The plug domain becomes partially disordered upon substrate binding, associating with the C-terminal domain of TonB, an inner membrane protein [132]. TonB initiates a conformational change that opens the SusC channel, allowing diffusion of the substrate into the periplasmic space [132]. SusC and SusD form a complex (SusCD) with a pedal bin mechanism: SusD is located on top of the β-barrel of SusC in the extracellular space, like a lid, primarily interacting with one loop from SusC, like a hinge [133]. The lid (SusD) opens to expose the carbohydrate binding site and closes upon binding, placing the sugar chain at SusC for transport across the membrane [133].

Transcriptional regulation and common PUL proteins

PULs encode for a transcriptional regulator commonly located in the periplasmic space, spanning the inner membrane [134]. In most cases this is a hybrid two component system (HTCS) or extracytoplasmic function σ/anti-σ factor type regulator [113]. The HTCS type is found in plant polysaccharide and non-mucin O-linked glucosaminoglycan PULs [120]. It binds carbohydrate products from enzymatic degradation in the periplasm, causing a conformational change that transmits to the other side of the membrane for transcriptional activation [120].

In addition to a transcriptional regulator and the SusC and SusD homologues there are usually several other proteins present in a PUL. The number and type of additional proteins for each PUL varies greatly dependent on the type of polysaccharide it is aimed at and often share little sequence homology with the original Sus system [115]. These proteins include additional extracellularly anchored sugar binding proteins [115], inner membrane transporters [126,129] and enzymes involved in glycan utilisation such as esterases [135] and phosphatases [136]. One of the main enzyme groups present in PULs are the glycoside hydrolases.

Glycoside hydrolases

CAZymes required for hydrolysis of the glycosidic bond are known as glycoside hydrolases (GHs, EC 3.2.1.-) [137]. This is a large group of enzymes and most of them are very specialised, cleaving a specific type of linkage at a specific type of sugar residue [138]. They are present across life for many different purposes: utilisation of energy reserves, remodelling of the plant cell wall, for lysosomal

degradation and for signalling and metabolism [2,139-141]. Humans are predicted to express relatively few GHs (97 identified so far), with only 17 of these proposed to be involved in nutrient breakdown of starch, sucrose and lactose, a number that pales in comparison to gut bacteria like *B. theta* or *B. ovatus* [142]. The human GHs not involved in nutrient breakdown are often part of internal carbohydrate processing. Deficiency in certain GHs is, for example, involved in lysosomal storage diseases [141,143].

The GHs are divided into families based on their sequence similarity and can be accessed through the CAZy database, which currently harbours 153 GH families (www.cazy.org) [122]. These families vary greatly in size and it is not uncommon for one family to harbour several types of activity (www.cazy.org) [122]. Some families have been grouped into 17 clans, previously called superfamilies, based on conserved protein fold, catalytic residues and catalytic mechanism [144-146]. Clan GH-A is the largest, currently with 21 families, and enzymes in this clan share a $(\beta/\alpha)_8$-barrel fold [122,146]. Some of the larger families, like GH5, have been further classified into subfamilies [147]. The nomenclature for naming GHs is as follows: the first two letters, in italics, designate the origin species, followed by three letters referring to the specificity of the catalytic domain, then the family number and a letter signifying the order in which the enzymes of that family have been reported [148]. For example, *Bo*Man26A, from **Papers II-IV**, is the first GH26 β-mannanase characterised in *B. ovatus*.

Most GHs have their active site in a cleft, tunnel or pocket, to facilitate substrate binding [138]. In the cases where the enzyme is oligomeric this cleft or pocket may be the combination of two or more monomers [124,149,150] (**Papers I-III**). GHs may be exo- or endo-acting. An exo-acting enzyme cleaves at the end of a sugar chain, and will often cleave off monosugars [151], but exo-enzymes that generate di- or trisaccharides are also known [124,152]. Endo-acting enzymes instead cleave internally in a saccharide chain and can generate a range of products [150,153]. Enzymes that are endo-acting or cleave crystalline substrates, such as endoglucanases active on cellulose, often bind their substrate in a groove or cleft [150,153] (**Paper II and III**), while strict exo-acting GHs often have a pocket shaped active site cleft [124,154]. Some GHs may be processive: upon substrate binding they perform several consecutive catalytic events before dissociating, moving along the polysaccharide chain [155]. The number of consecutive catalytic events before dissociation, the degree of processivity, varies between different enzymes [155] and may be initiated by an endo- or exo-type hydrolysis [156]. Processive enzymes tend to have their active site in a groove or tunnel [150]. Processivity was first studied in α-amylases [155] and has also been studied for enzymes like cellobiohydrolases and chitobiohydrolases, which act on insoluble substrates [157].

Reaction mechanism

GHs cleave the glycosidic bond with an inverting or retaining mechanism, reflecting the final configuration of the anomeric carbon [137]. Both mechanisms commonly use two carboxylic catalytic residues (glutamates or aspartates) [138]. These function as an acid and a base for the inverting mechanism, while for retaining GHs one residue serves as a nucleophile, while the other functions as a general acid/base [137,138]. Some GHs use a substrate-assisted mechanism and as such lacks one of the carboxylates [158-161]. In some cases a cofactor such as NAD+ or an ascorbic acid can be used [162-164]. Retaining GHs with an acetamido nucleophile [165] and a retaining GH with a tyrosine as the nucleophile [166,167] have been observed. The inverting mechanism requires only one step [137], while the retaining mechanism uses a double displacement mechanism, first discovered by Koshland [168], that requires two steps (Figure 2). As the GHs studied in this thesis (**Papers I-IV**) all use a retaining mechanism it will be described here. In the first step a nucleophilic attack is performed on the anomeric carbon, inverting the bond and generating a glycosyl ester intermediate (Figure 2) [169]. This is facilitated by the acid/base acting as a proton donor that breaks the glycosidic bond [169]. The leaving group dissociates. A water molecule is deprotonated by the acid/base and makes a nucleophilic attack on the enzyme-substrate intermediate, breaking the bond and inverting the anomeric configuration back to the original one before the carbohydrate dissociates (Figure 2) [137]. If a nucleophile other than water is used for the final attack, such as a sugar or an alcohol, it is a transglycosylation reaction [137,170-172]. Transglycosylation in general has great potential for making new materials and value added products [173].

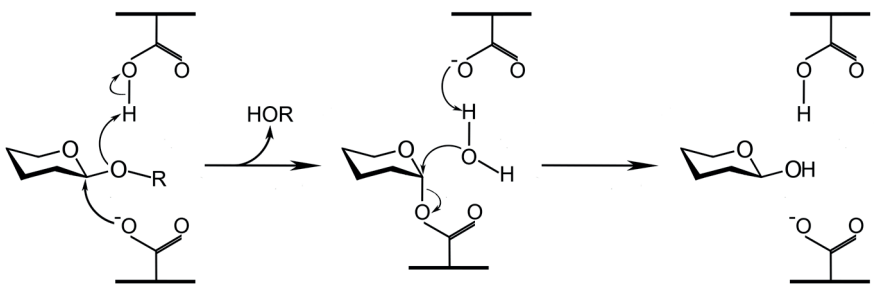


Figure 2 Schematic representation of the retaining catalytic mechanism of GHs. R represents additional sugar residues. In transglycosylation the water molecule is replaced with another nucleophile.

Sugar binding: subsites and CBMs

To confer increased binding and fine-tuned specificity most GHs bind several sugar units on each side of the catalytic residues [174] (**Paper II** and **III**). These additional binding sites are referred to as subsites, with those on the aglycone side (the leaving group) denoted with positive numbers and those on the glycone side (participates in the covalent enzyme-substrate intermediate) instead having negative numbering (Figure 3) [174]. The point of bond cleavage is between the monomer units bound in subsite -1 and +1, with numbering increasing as you move away from it (-2, -1, +1, +2 and so on, Figure 3) [174]. The number and conformation of these subsites varies greatly between enzymes, but as the cleavage occurs at the bond of the anomeric carbon of the monosaccharide in -1, this subsite is generally the most conserved [125,138,175]. Exo-acting enzymes tend to have only one or two negatively numbered subsites [124,154] (**Paper I**), while endo-acting enzymes have a more varied number of subsites [150,153] (**Paper II** and **III**). Endo-enzymes often generate a range of different products due to binding of the substrate along different positions on the sugar chain [176] (**Paper II** and **III**). Mutagenesis of subsite residues may alter substrate length preference and product formation [171]. The loops around the active site can have a large effect on fine-tuned substrate specificity or endo- vs exo-action of the enzyme [124,177-179] (**Papers II** and **IV**).

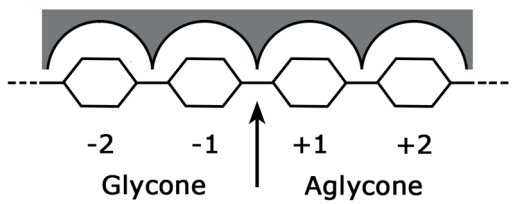


Figure 3 An overview of subsite organisation, each subsite in the enzyme represented as half circles. Cleavage, indicated by the arrow, occurs between the -1 and +1 subsites. A representation of a saccharide chain, containing at least 4 sugar units, is shown as hexagons situated in different subsites. The glycone and aglycone sides have been labelled.

Many GHs are modular: apart from the catalytic module, they may contain modules such as carbohydrate binding modules (CBMs) [180-182]. CBMs are among the best characterised modules associated with GHs. They assist in binding and may thereby increase activity of the attached GH [183]. The enzymes studied in this thesis do not have CBMs, but these modules can be found in similar enzymes [125,180]. CBMs are currently listed in 33 families in the CAZy database (www.cazy.org) [122]. They can be a part of GHs with many different fine-tuned

specificities, both for soluble oligo- or polysaccharides [180,184,185] and for enzymes that hydrolyse crystalline substrates [182,186]. In the latter case the CBM does not always bind the same saccharide chain as the catalytic module, something that could assist in targeting cell walls or other mixed structures of polysaccharide [187,188]. For soluble substrates the CBM tends to have the same specificity as the appended catalytic module [183].

Mannan active enzymes

There are several different types of β-mannan polysaccharides, see above, that require several different enzymes for complete hydrolysis [189]. For GGM, the most complex type of mannan polysaccharide, 5 different enzymes are required for complete degradation: the 4 GHs endo-β-1,4-mannanases (EC 3.2.1.78), exo-β-1,4-mannosidases (EC 3.2.1.25), exo-α-1,6-galactosidases (EC 3.2.1.22) and exo-β-1,4-glucosidases (EC 3.2.1.21), as well as acetyl esterases (EC 3.1.1.6) [35] (Figure 4). β-Mannanases hydrolyse internal β-mannosidic bonds, while β-mannosidases are exo-acting enzymes that cleave the ends of β-mannan chains (Figure 4) [189]. Any galactose substitutions may be removed by α-galactosidases (Figure 4). There are variations among α-galactosidases in the capacity to cleave substitutions carried by internal mannosyls, versus cleavage of galactosyls carried by a terminal mannosyl [189]. For glucomannan and GGM, glucose residues are removed by exo-acting β-glucosidases, in some cases after initial cleavage of the polysaccharide [35,190]. β-Mannan-acting acetyl esterases remove O-acetylations from the β-mannan chain (Figure 4) [189,191,192]. α-Galactosidases (**Paper I**) and β-mannanases (**Papers II-IV**) will be covered in greater detail below.

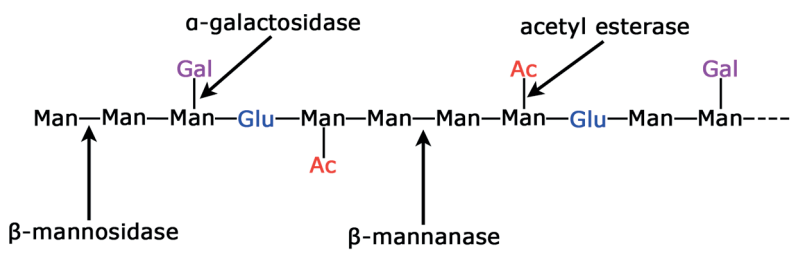


Figure 4 Schematic overview of the primary enzymes required for GGM hydrolysis. The β-1,4-linked backbone is composed of mannosyl (black) and glucosyl (blue) units and is cleaved by endo-β-1,4-Mannanases and exo-β-1,4-mannosidases. Saccharide chains with a terminal glucosyl unit are cleaved by exo-β-1,4-glucsodiases (not shown). The substitutions, galactosyl units (violet) and acetylations (red), are removed by exo-α-1,6-galactosidases and acetyl esterases, respectively.

β-Mannanases

β-Mannanases, studied in **papers II-IV**, generally hydrolyse β-1,4-linkages in mannan based oligo- and polysaccharides in a random endo-wise fashion [193], though a few exceptions have been found with mannobiohydrolase activity, generating mannobiose (M2) [124,194,195]. Precise specificity varies between enzymes, with differences in tolerance to substitutions and preferred number of bound sugar units [125,176,196] (**Paper II**). β-Mannanases are found in both prokaryotes and eukaryotes and are important for polysaccharide degradation in a number of environments such as soil and the gut [124,125,181,197]. The β-mannanases belong to GH families 5, 26, 113 and 134 (www.cazy.org) [122], out of which all except 134 are a part of clan GH-A [144,198]. GH134 enzymes are inverting, with a lysozyme like fold and the catalytic residues are an aspartate and a glutamate [199]. GH-A, the largest of the GH clans, has a $(\beta/\alpha)_8$-barrel fold [144,198] and uses a retaining catalytic mechanism with two glutamate catalytic residues that, in most known cases, are about 5 Å apart at the end of strands β4 (acid/base) and β7 (nucleophile) [144,146,189]. In addition a Tyr/His at the end of strand β6 and a Trp/Phe at the end of strand β8 flanking the nucleophile in GH-A [200] are conserved in GH26 as Tyr258 and Trp307, respectively, numbered according to GH26 β-mannanase C from a symbiotic protist of *Reticulitermes speratus* (*Rs*Man26C) [175] (**Papers II** and **III**). Apart from β-mannanases, clan GH-A harbours a large number of activities such as β-xyloglucanase [201], β-galactosidase [200] and β-glucosaminidase [202]. β-Mannanases are currently being used for reducing the viscosity of coffee extract [56] and are also under investigation for a number of different applications such as bioethanol production [203], production of prebiotics and clarification of fruit juices [59,204] and for addition in detergents [205]. In addition, transglycosylating β-mannanases have great potential for generating value added products [173].

One of the largest GH families is GH5, which currently has 54 subfamilies with a wide variety of enzymes such as β-glucanases, licheninases, endo-glyceroamidases, and β-mannanases (www.cazy.org) [122]. β-Mannanases have so far been found in subfamilies 7, 8 10, 17, 36 and 55 (www.cazy.org) [122]. Subfamilies 7 and 10 primarily contain β-mannanases of eukaryotic origin, for example from fungi [125,171] and molluscs [206,207], while bacterial enzymes dominate in subfamilies 8, 17, 36 and 55 [122,179,208] (www.cazy.org). The GH5 β-mannanases are usually monomeric and in some cases have additional modules, for example CBMs [187,188,209]. A number of β-mannanase structures have been determined for this family [125,179,206,207,209-217] which has seven conserved residues: Arg54, Asn168, Glu169 (acid/base), His241, Tyr243, Glu276 (nucleophile) and Trp306, numbered according to GH5 β-mannanase A from *Trichoderma reesei* [216]. Transglycosylation capacity has been detected for several GH5 β-mannanases [171,206,218-220].

GH26 mainly contains β-mannanases, but also contains a few characterised β-xylanases and a β-glucanase (www.cazy.org) [122]. GH26 β-mannanases are mainly found in prokaryotes [124,181,195] (**Papers II** and **III**), but there are also some enzymes from this family found in fungi [125,196]. No transglycosylation has so far been detected for a wild type enzyme in this family [125]. GH26 enzymes may contain several modules, such as CBMs, and are generally monomeric [125,180,181,221]. Precise specificity of products generated and preferred substrate length varies between enzymes in this family. So far, most enzymes in GH26 seem restricted by galactose substitutions [176,196,208,222-224], with any exceptions primarily being of fungal origin [225]. Generally, the activity of GH26 β-mannanases decreases on substrates with higher galactose content, such as guar gum compared to LBG [176,196,208,222-224]. There have been a number of structures determined experimentally in this family [124,125,153,175,177,212,221,226] (**Papers II** and **III**). GH26 β-mannanases generally have cleft-shaped active sites with seven conserved residues, primarily located around the -1 subsite [153,175]: His124, His190, Glu191 (acid/base), Trp196, Tyr258, Glu288 (nucleophile) and Trp307 (*Rs*Man26C numbering) [175]. The main difference compared to the conserved residues in GH5 is that the residue preceding the acid/base is a histidine instead of an asparagine. Of the conserved residues in GH26 His124, His190 and Trp307 interact with the -1 mannosyl, His190 interacting with the axial O2 of the sugar unit, helping to confer mannosyl specificity, as studied by Tsukagoshi et al. [175]. Tyr258 stabilises the nucleophile, while Trp196 provides stacking interactions with the +1 sugar [175]. In some GH26 β-mannanases the -1 mannose unit has been shown to adopt a $B_{2,5}$ boat conformation [227].

Several GH26 mannanases with known structures contain metal binding sites [124,153,175,212,226] (**Paper III**). However, most of these are remote from the active site with a purely structural role and have perhaps therefore not been widely studied [124,153,175,212]. A small number of thermostable β-mannanases, two from GH26 and two from GH5, have so far been shown to have metal binding sites with zinc (GH26) or calcium (GH5) that significantly impact thermal stability [226,228-230]. In addition, two β-mannanases that show increased activity in the presence of various divalent metal ions have been described [231,232]. In **Paper III** we add to this picture by characterising a GH26 β-mannanase that shows increased stability in the presence of calcium ions and with a determined calcium binding site.

α-Galactosidases

The α-galactosidases cleave α-linked galactosyl units. Examples include α-1,6-galactosyl linkages in β-mannans [189] (**Paper I**) and raffinose family oligosaccharides [233] (**Paper I**), as well as α-1,2-, α-1,3- and α-1,4-linked galactosyls in various mammalian glycoproteins [234] and human glycolipids [235]. They have a wide range of applications, including treatment of Fabrys disease [141], in processing of food products [236,237] and in the pulp and paper industry [238]. α-Galactosidases belong to GH families 4, 27, 31, 36, 57, 97 and 110 in the CAZy database (www.cazy.org) [122]. Of these, families GH27, 31 and 36 constitute the three members of clan GH-D (www.cazy.org) [122]. Clan GH-D enzymes, similarly to GH-A, also have a $(\beta/\alpha)_8$-barrel fold and use a retaining mechanism [189], but the catalytic residues are aspartates located at the end of strand β4 (nucleophile) and in the loop after strand β6 (acid/base) [239,240]. The majority of characterised enzymes present in GH27 and GH36 are α-1,6-galactosidases (www.cazy.org) [122], while the enzymes present in GH31 have more varied activities and include α-glucosidase [239] and α-xylosidase [241]. α-1,6-Galactosidases can hydrolyse internal galactose substitutions in β-mannans, an activity mainly found in GH27, or cleave terminal α-1,6-galactosyl linkages, which is the primary specificity of GH36 enzymes [242,243]. Several α-galactosidases have been shown to act synergistically with β-mannanases and β-mannosidases in the breakdown of different mannan type substrates [222,244-246] (**Paper II** and **III**). They are found in both eukaryotic and prokaryotic organisms from a variety of environments, such as soil [243,247] and gut bacteria [233,248,249] (**Paper I**). Several α-galactosidases have been shown to be capable of transglycosylation [172,250-252].

An enzyme from family GH36 was characterised in **Paper I**. The GH36 family primarily acts on terminal galactose linkages on oligosaccharide substrates [233,243,253,254], with one recently characterised gut bacterial enzyme active on terminal arabinose units [255]. The GH36 α-galactosidase from **Paper I** is rare among the characterised GH36 enzymes, since it shows significant activity on polysaccharide substrates. This family has been divided into four subgroups based on conserved sequence motifs and phylogenetic analysis [149,233] (**Paper I**). GH36 family enzymes from subgroup I are generally tetrameric with two conserved motifs: CxxGxxR involved in galactose recognition and GxxLxxxG related to tetramer formation [149,256,257] (**Paper I**). The conserved cysteine of the CxxGxxR motif interacts with the bound galactose in the active site [233]. Members of other subgroups may be monomers or other oligomeric states [240,254,258]. GH36 enzymes generally contain three domains per monomer: one N-terminal and one C-terminal domain, both containing mostly β-sheets, and a central catalytic domain [233]. The tetrameric form is primarily facilitated through the N- and C-terminal domains, meaning that the largest structural differences

between enzymes of different oligomeric state are found here [233]. While only one monomer interacts with the galactose and performs hydrolysis, adjacent monomers position the substrate and the acid/base [233] (**Paper I**). It has been found that a serine and an arginine (S548 and R556, numbered according to a GH36 α-galactosidase from *Flavobacterium* sp. TN17) are important for gut living α-galactosidases [259].

A predicted galactomannan PUL in *B. ovatus*

B. ovatus was shown to be able to grow on galactomannans already in 1977 by Salyers et al. [82], as one of few gut bacteria in that study able to do so. Subsequently both α-galactosidase and β-mannanase activity was found in this organism, but the sequences or identity of these enzymes were never determined [131,248,260,261]. Much later, in 2011, a potential galactomannan PUL in *B. ovatus* strain ATCC 8483 (in this thesis termed *Bo*ManPUL, gene locus *bacova_02087-97*) was identified by Martens et al. [113]. *Bo*ManPUL contained three putative GHs, predicted by sequence homology: one GH36 α-galactosidase (*Bo*Gal36A) and two GH26 β-mannanases (*Bo*Man26A and *Bo*Man26B) (**Paper I**), further studied in **Papers II-IV**. *Bo*ManPUL also contained a HTCS regulator that was activated by manno-oligosaccharides [113] and was predicted to encode a SusC- and a SusD-like as well as a SusE-positioned protein, a symporter, an isomerase and a mannoglucosyl phosphorylase (**Paper I**).

GH enzymes with the same specificity present in the same organism may have differing fine-tuned substrate specificities [124,125,135], something which has also been seen for some GH enzymes with the same specificity present in the same PUL [115,129,262]. This raises questions regarding differences between the two putative β-mannanases, *Bo*Man26A and *Bo*Man26B, of the *Bo*ManPUL. In addition, GH36 α-galactosidases tend to be less active on internal galactosyl substitutions, raising questions as to the role of *Bo*Gal36A in the galactomannan acting *Bo*ManPUL. **Papers I-IV** focus on the GH enzymes of *Bo*ManPUL, to delineate the differences between the two GH26 β-mannanases and the role of the GH36 α-galactosidase, resulting in a model of PUL function. The differences between the two GH26 β-mannanases are analysed in context of similar *Bacteroides* PULs and are related to the biochemistry and structure of other characterised GH26 enzymes.

Methods

The structure of an enzyme is intimately tied to its function. The orientation and position of amino acid residues in an active site or binding pocket can shed light on the nature of ligand interaction, or the cause for a certain fine-tuned specificity. To understand this structure function relationship both biochemical data on the nature of the fine-tuned specificity of the enzyme, as well as its three dimensional structure need to be studied. This section aims to give an overview of some of the main biochemical and structural methods used in this thesis for this purpose.

Biochemical methods

GHs often have very fine-tuned substrate specificities and the activity often varies based on length of saccharide, restrictions due to substrate substitutions, different sugar units in the backbone of the substrate or other glycan modifications [124,196] (**Paper II**). The number of subsites and the binding strength of each one is specific for each enzyme, leading to variations in preferred substrate binding mode, minimum substrate length and generated products [124,174,175]. This fine-tuned specificity, products generated and preferred binding mode are all studied with a range of biochemical methods, some of which are briefly described here.

Measuring activity by determining produced reducing ends

The reducing end of a sugar chain or monosaccharide is the sugar unit with an unbound anomeric C1 carbon [263]. The unbound anomeric C1 carbon changes between two cyclic conformations (α and β) through an equilibrium with an open conformation. While in the open conformation the anomeric C1 carbon can be oxidised, making it a reducing agent. For most hemicellulosic oligo- or polysaccharides all but one of the sugar units are blocked in this capacity since C1 is involved in the glycosidic bond [263]. One method to assess the activity and kinetics of GH hydrolysis on a polysaccharide is to measure the increase in reducing ends using the 3,5-dinitrosalisylic acid (DNS) assay [264], as was done in **Papers II-IV**. The basis for the assay is the reduction of DNS to 3-amino-5-nitrosalisylic acid, which absorbs light at 540 nm, by the sugar unit reducing ends

present in solution [265]. The enzyme and a polysaccharide substrate are incubated together at the desired reaction conditions, in **Papers II-IV** most commonly for 15 min at 37ºC. A DNS solution containing a strong alkali is added that, due to its high pH, inactivates the enzyme and stops the reaction. Subsequent boiling increases the reaction speed of DNS reduction in the presence of strong alkali by the reducing ends present. The measured absorbance is subtracted by a blank containing enzyme and substrate before the released reducing ends can be determined by comparison with a standard curve with DNS, polysaccharide and known mannose concentrations. This method is difficult to use for oligosaccharides, as the number of reducing ends present in solution before hydrolysis is very high, and it is mostly used to analyse GH activity on polysaccharides.

Product profiles and oligosaccharide kinetics

Studying the product profile from GH hydrolysis of different oligo- and polysaccharides can give valuable information such as the types of substrates that can be productively bound in the active site cleft and preferred mode of binding. Through kinetic analysis of oligosaccharide hydrolysis it is also possible to deduce the number of subsites required for efficient hydrolysis. In **Papers I-III** product profiles from a number of β-mannan substrates were analysed and quantified using high performance anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD) [266-268]. The identification of products analysed with this method relies on comparisons with standards of equivalent carbohydrates. This is problematic when analysing hydrolysis products from endo-acting enzymes on galactomannans as there are few characterised galactomanno-oligosaccharides available for use as standards. In **paper III** this issue was partially solved by hydrolysing the generated galactomanno-oligosaccharide products with an α-galactosidase before HPAEC-PAD analysis.

As the DNS assay is not suitable for measuring oligosaccharide hydrolysis the determination of oligosaccharide kinetics is instead carried out by HPAEC-PAD analysis. The measurement of standard Michaelis-Menten kinetics using this method is cumbersome when several different substrates need to be analysed, with many separate incubations and several optimisation steps usually required. In addition, there is the risk of secondary hydrolysis of the initial products disturbing the measurements of kinetic constants. A method was previously developed for determining catalytic efficiency (k_{cat}/K_M) based on Michaelis-Menten kinetics, but with a more limited experimental setup and lower risk of interference from secondary hydrolysis [269,270]. k_{cat}/K_M is useful for analysing substrate preferences of an enzyme and for comparing different enzymes. k_{cat} is the turnover number: the number of substrate molecules converted per unit time and the

Michaelis constant K_M is the substrate concentration when the initial reaction rate is $V_{max}/2$. Thus k_{cat}/K_M takes into account both the rate of the reaction and the strength of enzyme-substrate binding [263]. To determine k_{cat}/K_M using the method described by Matsui et al. [269] hydrolysis must be performed at low enzyme and substrate concentrations such that: $[E]<<[S]<<K_M$, quantifying the decrease in substrate concentration over time. It can then be regarded as a first order reaction with the rate constant $k=(k_{cat}/K_M)[E]$, resulting in equation 1:

$$kt = \ln\left(\frac{[S_0]}{[S_t]}\right) \quad (1)$$

where k is the rate constant, t is time and $[S_0]$ and $[S_t]$ are substrate concentrations at time 0 and t, respectively [269]. This method was used to perform oligosaccharide kinetics in **Papers I** and **II**.

Analysing preferred mode of binding

An endo-acting enzyme hydrolysing a sufficiently long oligosaccharide usually generates a range of products, meaning that the sugar chain can bind available subsites in different ways [176]. A variation in product profiles between endo-acting enzymes [153,176,221] indicates a preferred mode of binding where the occupation of some subsites is more favourable. The hydrolysis of mannopentaose (M5) into M2 and mannotriose (M3), for example, could be due to substrate binding in subsites -2 to +3 or between subsites -3 and +2. This can not be differentiated by HPAEC-PAD analysis alone. To determine preferred binding mode a method utilising hydrolysis in the presence of $H_2{}^{18}O$ [171,176] was used in **Papers II** and **III**. An oligosaccharide is hydrolysed by the studied enzyme in the presence of $H_2{}^{18}O$, an ^{18}O is incorporated in the newly generated reducing end and the leaving group remains unlabelled. The products are analysed with matrix assisted laser desorption-ionisation time of flight mass spectrometry (MALDI-TOF MS), a highly sensitive method for precise determination of the mass of carbohydrates where the relative amounts of labelled and unlabelled saccharide can be compared [271,272]. The hydrolysate mixture is cocrystallised with a matrix of 2,5-dihydrobenzoic acid [273] and the ratio of non-labelled and labelled peaks for the same DP oligosaccharide is determined. As MALDI-TOF MS does not allow absolute quantification it is coupled with HPAEC-PAD analysis for determination of preferred mode of binding.

Structural studies

The preferred substrate binding mode and specificity of a GH is related to its structure. The overall shape around the active site helps determine fine-tuned specificity [138] (**Paper I-III**). In GH5 and GH26 β-mannanases the active site is usually located in a cleft with at least four subsites for oligo- and polysaccharide binding [189]. The -1 subsite of β-mannanases is highly conserved within families, providing specific saccharide interactions that directly relate to the type of enzyme activity [175,189]. Subsites with a stronger binding preference usually interact with the sugar unit through aromatic stacking interactions or hydrogen bonding [124,125,189], which in turn affects the product profile and binding mode (**Paper I-III**). The loops surrounding the active site cleft are often crucial for shaping fine-tuned specificity [124,179,210] (**Paper I-III**) and loop flexibility may affect binding and activity [274,275]. Methods used for studying structure and dynamics in **Papers II-IV** are presented here.

Structure determination by X-ray crystallography

A common method for high resolution protein structure determination is X-ray crystallography. Protein crystals are often generated by vapour diffusion [124,175,276]: a drop of protein solution equilibrates by evaporation against a reservoir of lower water content that causes an increase in protein concentration in the drop [277]. This leads to supersaturation of the protein, resulting in either precipitation or crystallisation [277]. What differentiates conditions that cause precipitation from those that crystallise varies for each protein and is difficult to predict, so various screens are utilised to determine optimal crystallisation conditions. Upon data collection the crystal is mounted in a nylon loop and exposed to an intense, narrow beam of X-ray radiation, a small portion of which scatter due to the electrons in the crystal, forming a diffraction pattern [263]. To reduce radiation damage to the protein crystal caused by the ionising nature of X-rays [278,279] it is commonly kept in a stream of nitrogen at 100 K [280]. In order to prevent the formation of ice the protein crystal is typically transferred to a drop including a cryoprotectant such as polyethylene glycol or glycerol, before mounting in the loop [281] (**Papers II** and **III**).

X-ray diffraction patterns are collected from the protein crystal over a wide range of angles. The square root of the observed spot intensity can be converted into an electron density map using a Fourier transform [263]. In **Papers II** and **III** initial processing of the diffraction patterns was performed with the XDS program suite [282] and CAD from CCP4 [283]. To obtain a complete electron density map the phase of the diffracted X-rays must be known, something that is not possible to measure directly. Several methods have been developed to solve this phase

problem [284-287]. In **Papers II** and **III** molecular replacement was used, a relatively quick and easy method which commonly requires the known structure of a protein with at least 30% identity (the search model) to the protein of interest (the target) [287]. The search model is fitted in to the asymmetric unit of the target, in **Papers II** and **III** using the software Phaser [288] in Phenix [289], allowing the phase information of the search model to be transferred to the target [287,290]. The initial target phases are estimates based on the search model [291]. Manual editing of the initial target model, generated based on the search model, to improve its fit to the electron density map increases the accuracy of the estimated phases upon automated refinement of the model to the data [291]. Several consecutive rounds of manual editing and automated refinement are usually required for a high quality structure. In **Papers II** and **III** refinement was performed with Phenix [289] and manual editing with Coot [292].

For many proteins it is often desirable to obtain structures containing a bound ligand. Generating protein-ligand crystals is mainly achieved in one of two ways: soaking or co-crystallisation [293]. In co-crystallisation the ligand is included in the crystallisation conditions [124], which may be necessary if there is not enough space in the original crystal packing to accommodate it [293]. Soaking involves adding ligand to an already generated protein crystal, either directly in the crystallisation drop [227] or by moving the crystal to a new drop containing the ligand [175,221] (**Paper III**). The new drop with added ligand can also contain cryoprotectant and allow different conditions than those used for crystallisation, which may not be favourable for ligand binding [293]. In **Paper III** soaking was used to obtain a ligand bound crystal structure of *Bo*Man26B.

Isotopic labelling and NMR assignment

Enzyme loop flexibility can have a significant impact on GH enzyme activity, as well as substrate binding and release [275,294,295]. X-ray crystal structures provide static pictures and the underlying data, including the B-factor, provide limited information on flexibility in the protein structure. This makes enzyme flexibility and any impact it may have on function difficult to assess: very flexible loops may not be visible and regions that are flexible in solution could be stabilised by crystal contacts. Solution state nuclear magnetic resonance spectroscopy (NMR) was performed on *Bo*Man26A in **Paper IV**, yielding a nearly complete (about 95%) backbone peak assignment. This will form the basis for future studies of substrate interactions and protein dynamics.

NMR measures the spinning frequency of nuclear magnetic moments around a static magnetic field, which arise from the nuclear spin, a quantum mechanical property described by the spin quantum number I. Isotopes with $I = 0$, like ^{12}C, are NMR inactive, whereas those with $I = ½$, like ^{1}H, ^{13}C and ^{15}N, have favourable

properties and those with $I > ½$, like ^{14}N, are more challenging to study [263]. For this reason protein NMR requires that samples are isotope labelled with ^{13}C and ^{15}N. In addition, replacing all ^{1}H with ^{2}H, except at the backbone amides, improves the signal by increasing the relaxation time of the adjacent atoms [296,297]. To produce $^{2}H^{13}C^{15}N$-labelled protein in **Paper IV** the expression strain was grown and *Bo*Man26A was expressed with $^{2}H_2O$, $^{15}NH_4Cl$ and ^{13}C-glucose as the sole sources of these elements. The bacteria initially need to be adapted to growing in $^{2}H_2O$, which is done through a series of small overnight cultures successively replacing H_2O with $^{2}H_2O$ [298] (**Paper IV**, Angus J Robertson, personal communication). The backbone amides readily exchange their deuterons to protons upon exposure to $^{1}H_2O$ [299]. To ensure hydrogen exchange at sites buried in the interior of the protein it can be unfolded using urea, followed by refolding through slow dialysis with the target storage buffer (**Paper IV**). The isotopically labelled protein can then be characterised using multidimensional NMR experiments, as was done in **Paper IV**.

The collection of 3D NMR assignment spectra with transverse relaxation-optimised spectroscopy (TROSY) [300] enables studies of proteins larger than about 25 kDa [297]. By assigning each peak in the NMR spectra to its corresponding atom in the protein chain one essentially achieves atomic resolution [297]. The position of a peak will shift upon interaction of the corresponding atom, enabling identification of ligand binding sites. Furthermore, NMR assignments provide a starting point for atomic-resolution studies of protein structure, dynamics and flexibility [297]. The assignment of *Bo*Man26A with TROSY NMR (**Paper IV**) will be used in further studies to elucidate the residues involved in substrate binding and to characterise loop flexibility around the active site.

Results and discussion

A limited number of *B. ovatus* strains are capable of growing on galactomannans [82]. The previously identified *Bo*ManPUL from strain ATCC 8483 [113] was shown to encode the three GHs, *Bo*Gal36A, *Bo*Man26A and *Bo*Man26B, in **Papers I** and **II**. Appended **Papers I-IV** has focused on understanding the structure and function of these GHs, their role in *B. ovatus* galactomannan utilisation and the relation of *Bo*Man26A and *Bo*Man26B to other GH26 enzymes. A summary of this work will be presented here.

The GH enzymes of the *Bo*ManPUL

Characterisation of β-mannan active GHs from human gut bacteria has been limited and not previously carried out on enzymes expressed from a PUL [181,195,208,301,302]. Early studies of *B. ovatus* galactomannan utilisation revealed outer membrane associated β-mannanase activity [260], as well as soluble β-mannanase activity located inside the cell [131] and three assumed soluble α-galactosidases [248,261]. The identity of the enzymes responsible for the *B. ovatus* β-mannanase and α-galactosidase activities was not established. However, based on cellular location and endo-action on guar gum, the previously detected outer membrane associated β-mannanase from *B. ovatus* cultures [260] could perhaps be *Bo*Man26B from the *Bo*ManPUL, characterised in **Papers II** and **III**. Potential identity of the previously detected α-galactosidases is more difficult to suggest: *Bo*Gal36A (**Paper I**) is the only characterised α-galactosidase in *B. ovatus*, but the CAZy database contains an additional 18 sequences putatively assigned as α-galactosidases from the same strain (www.cazy.org) [122]. In **Paper I** several strains of *B. ovatus* and *Bacteroides xylanisolvens*, all containing homologous PULs to *Bo*ManPUL, were shown to be able to grow on galactomannan (**Paper I**, Figure 1). The gene organisation of these PULs was overall similar, but one striking difference was the presence or absence of a gene coding for a putative GH36 α-galactosidase, resulting in two types of homologous PULs. Type I contained the putative GH36 α-galactosidase gene (*Bo*ManPUL type I), while type II did not (*Bo*ManPUL type II, **Paper I**, Figure S1). Studies on these putative mannan PUL variants in **Paper II** showed that *B. ovatus* strain ATCC

8483, containing the type I *Bo*ManPUL, was able to grow on both galactomannan and glucomannan. Strains containing the type II PUL grew more slowly on galactomannan. In addition, a separate *B. ovatus* ATCC 8483 PUL (gene locus *bacova_03386-406*) was found containing a putative GH26 encoding gene. However, knockout of this second PUL did not alter the growth of *B. ovatus* on the previously tested β-mannans, while a deletion of the *Bo*ManPUL abolished or greatly reduced growth on these polysaccharides (**Paper II**, Figure 1). Thus the *Bo*ManPUL seems to be the main PUL responsible for the ability of *B. ovatus* to grow on β-mannans.

PULs typically encode several GH enzymes, some of which are attached to the outer membrane and some which are periplasmic, for different stages of polysaccharide processing [115,116,126,303]. The three GH enzymes encoded by *Bo*ManPUL also have differences in their cellular localisation: *Bo*Man26B is anchored to the extracellular membrane, while *Bo*Man26A and *Bo*Gal36A are periplasmic (**Paper II**, Figure 2). The differences in cellular location hint towards an order of processing β-mannan substrates of the GH enzymes encoded by the *Bo*ManPUL where the outer membrane *Bo*Man26B would make the initial substrate attack, creating products further processed by *Bo*Man26A and *Bo*Gal36A. Further biochemical characterisation of the enzymes made it possible to suggest this model, as explained in the following sections.

An odd GH36 α-galactosidase: *Bo*Gal36A

The presence of an α-galactosidase appears to be important for growth on galactomannan. Strains of *B. ovatus* containing the type II PUL, that lacks an α-galactosidase encoding gene, grew more poorly on galactomannan than those strains with type I PULs (**Paper I**, Figure 1). *Bacteroides fragilis* contains a known gene cluster encoding enzymes active on mannans [195,304] and a putative PUL encoding for a GH26 β-mannanase (predicted in PULDB, http://www.cazy.org/PULDB/ [121]) but *B. fragilis* did not grow at all in the presence of galactomannan (**Paper I**, Figure 1). The type I PULs contain a putative α-galactosidase gene, exemplified by the presence of the *Bo*Gal36A gene in the *Bo*ManPUL. However, *Bo*Gal36A is part of the GH36 family, enzymes which in several cases have been shown to hydrolyse internal galactose substitutions poorly [233,242,253,254] (**Paper I**). By contrast, several GH27 α-galactosidases have previously been shown to efficiently hydrolyse internal galactose substitutions and to act synergistically with β-mannanases [222,243].

Biochemical characterisation

*Bo*Gal36A belongs to GH36 subgroup I. The enzyme is a tetramer in solution, as determined by SEC analysis, and contains the two sequence motifs typical for this

subgroup (**Paper I**, Figure 2). However, *Bo*Gal36A efficiently hydrolyses internal galactose substitutions from galactomanno-oligo- and polysaccharides (**Paper I**). Upon hydrolysis of guar gum galactomannan, 90% of galactose substitutions were released, causing aggregation of the remaining polysaccharide backbone (**Paper I**, Figure 4). While low activity on galactomannan polysaccharides has previously been reported for some GH36 α-galactosidases [242,301,305], *Bo*Gal36A is the first enzyme in this family where hydrolysis of internal galactosyl side-groups is efficient and which causes guar gum aggregation. The specific activity of *Bo*Gal36A is higher for galactomanno-oligosaccharides than polysaccharides (**Paper I**) and the enzyme is located in the periplasm (**Paper II**). This indicates that *Bo*Gal36A is unlikely to perform the initial attack on galactomannans.

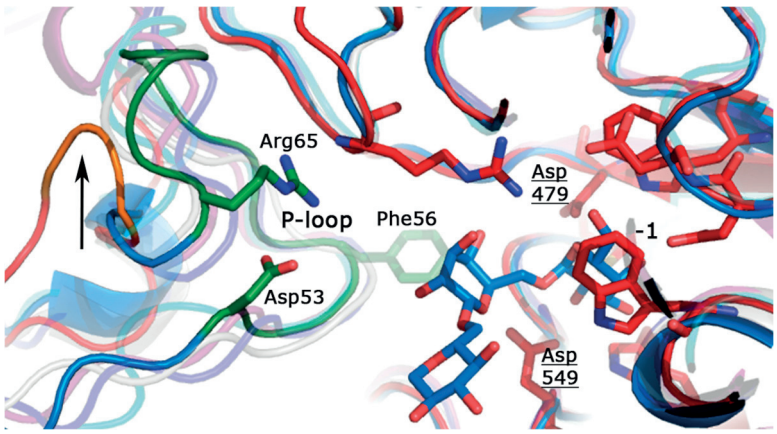


Figure 5 The homology model of *Bo*Gal36A (red) superposed with AgaA (blue), as well as the other determined GH36 subgroup I structures (semi-transparent), PDB IDs: 3MI6 (grey), 2XN2 (pink), 4FNQ (light blue) and 2YFO (purple). Stachyose (blue) present in the AgaA active site cleft is shown in stick representation. The P-loop is labelled and coloured green, while the equivalent loop in *Bo*Gal36A is marked with an arrow and coloured orange, for clarity. Conserved residues in the -1 subsite are shown in stick representation, the catalytic residues are labelled (underlined, *Bo*Gal36A numbering). Residues in the P-loop relevant for substrate interaction are shown in stick represenation and labelled (AgaA numbering).

Homology model

In previously determined structures of GH36 subgroup I enzymes the active site is located in a deep pocket (Figure 5) [149,301,306]. The -1 subsite is located entirely in one monomer, while the positive subsites also involve two adjacent monomers [149,301,306]. A homology model of *Bo*Gal36A was made in **Paper I** with a GH36 α-galactosidase from *Geobacillus stearothermophilus* (AgaA, PDB ID 4FNU) [306] and overlaid with all determined GH36 subgroup I structures

[149,301,306] using PyMOL [307] (Figure 5). The -1 subsite binds the galactose unit and contains the catalytic residues. It is conserved for all compared structures, in line with the outcome of the sequence alignment in **Paper I**, Figure 2. In AgaA a loop adjacent to the active site (the P-loop), present on a different monomer than the active site, provides hydrophobic stacking to the sugar unit in +1 through Phe56 and hydrogen bonds to the sugar unit in +2 through Asp53 and Arg65 [306]. This P-loop restricts the space around the active site, likely reducing the ability to accommodate internal galactose residues. In all currently determined GH36 subgroup I crystal structures the backbone conformation of the P-loop is conserved [149,301,306] (Figure 5). The three residues involved in binding positive subsites are also conserved in all structures except the subgroup I GH36 from *Lactobacillus brevis* (PDB ID 3MI6, determined by the Northeast Structural Genomics Consortium). It would be interesting to see what effect the lack of these conserved P-loop residues has on substrate binding, however biochemical data for this enzyme is currently not available. Uniquely for subgroup I, *Bo*Gal36A lacks the P-loop, causing the area around the -1 subsite to be more open. This lack of the P-loop probably enables the accommodation of internal galactose units seen in the biochemical characterisation (Figure 5) (**Paper I**, Figure 6).

Differences in fine-tuned substrate specificity of *Bo*Man26A and *Bo*Man26B

None of the three previously characterised gut bacterial β-mannanases are known to be expressed from a PUL [181,195,208]. For one of them the gene BF0771 codes for the GH26 β-mannanase from *Bacteroides fragilis* (*Bf*Man26) and is part of a small gene cluster suggested to be involved in β-mannan utilisation [195,304]. This gene cluster flanks a putative PUL that also contains a gene putatively encoding for a GH26 enzyme (**Paper I**). The two other characterised gut bacterial β-mannanases, from *Bifidobacterium adolescentis* and *Bifidobacterium animalis* subspecies lactis, are not part of known gene clusters for β-mannan utilisation and contain mannan-binding CBMs [181,208]. Neither *Bo*Man26A nor *Bo*Man26B contain CBMs, something that is generally the case in GHs expressed from PULs, which may be because they are co-expressed with polysaccharide binding proteins [115,116,126] (**Paper II**). *Bo*Man26A and *Bo*Man26B generally hydrolyse β-mannan substrates longer than DP2, but were not active on 6^3,6^4-α-D-galactosyl-mannopentaose (G2M5) (**Papers II** and **III**). Based on manno-oligosaccharide kinetics *Bo*Man26A requires at least 4 subsites for efficient hydrolysis, while *Bo*Man26B requires at least 6 subsites (**Paper II**, Table 2). k_{cat}/K_m for *Bo*Man26A on oligosaccharides was similar to what has been seen previously for other GH26 β-mannanases [124,153,221], but was in comparison very low for *Bo*Man26B on mannohexaose (M6) (**Paper II**, Table 2). No activity could be detected for

*Bo*Man26B on M5 and lower manno-oligosaccharides, indicating that *Bo*Man26B prefers longer substrates (**Paper II**, Table 2).

The specific activity of *Bo*Man26A on LBG galactomannan was about three times that of *Bo*Man26B (**Papers II** and **III**), both of which were in the same range as other GH26 enzymes (Table 1). Neither enzyme hydrolysed non-mannose based polysaccharides and had very low activity on the insoluble β-mannan INM (**Paper II**, Table 1). One of the most drastic differences between *Bo*Man26A and *Bo*Man26B was a major decrease in specific activity on guar gum compared to LBG for *Bo*Man26A, while it remained relatively unchanged for *Bo*Man26B (**Paper II**, Table 1). The decrease in activity for *Bo*Man26A on guar gum compared to LBG, where guar gum has a higher level of galactose substitution, indicates a restriction by the galactosyl side-groups carried by mannans. Despite the higher level of galactose substitution in guar gum compared to LBG the specific activity for *Bo*Man26B remained very similar, indicating little or no restriction by galactose substitutions for this enzyme (**Paper II**, Table 1).

Table 1 Specific activity (kat/mol) and enzyme kinetic parameters on LBG and guar gum for GH26 β-mannanases.

Enzyme	References	LBG kat/mol	LBG K_M gL^{-1}	LBG k_{cat} s^{-1}	LBG k_{cat}/K_M $s^{-1}gL^{-1}$	Guar kat/mol	Guar K_M gL^{-1}	Guar k_{cat} s^{-1}	Guar k_{cat}/K_M $s^{-1}gL^{-1}$
*Bo*Man26A	**Paper II**	301				0.8			
*Bo*Man26B	**Papers II and III**	100	11	250	23	86	13	122	9.5
*Cj*Man26A	[153]	2167	8.5	2904	342				
*Cj*Man26C	[124]	233							
*Cf*Man26A-50K	[176,221]	376	1.8	137	76	84	16	137	9
*Pa*Man26A	[308]		5.5	83	15				
*Rs*Man26C	[223]		5.2	833	160		28	490	17
*Bs*Man26A	[212]		4.5	5.5	1.2				
BCman	[226]		10	61	6.1				
*Ba*Man26A-53K	[181]	176	21	444	20.8	14			
*Tr*Man26A	[218]		0.6	240	400				
*An*Man26A	[222]	264				149			

The k_{cat}/K_M for *Bo*Man26B on LBG was similar to some characterised GH26 β-mannanases (Table 1) [181,226,308], yet is in the lower range of all enzymes compared in Table 1. On guar gum the k_{cat}/K_M for *Bo*Man26B was halved, compared to LBG (**Paper III**, Table 1). There have been a limited number of kinetics studies on other GH26 β-mannanases with LBG and guar gum [176,223] and in these cases the decrease in k_{cat}/K_M was instead about 90% for guar gum compared to LBG (Table 1). Thus, out of the enzymes with available kinetic data on LBG and guar gum, *Bo*Man26B is the least affected by the higher level of galactose substitution in guar gum, as judged by the halved k_{cat}/K_M on guar gum compared to LBG (Table 1).

In addition, the 90% decrease in k_{cat}/K_M for other GH26 β-mannanases on guar gum compared to LBG is primarily due to an increase in K_M for guar gum, while in the case of *Bo*Man26B the K_M remains relatively similar (Table 1). This means that, while the K_M of *Bo*Man26B for LBG is the second highest out of the GH26 β-mannanases compared in Table 1, the K_M on guar gum is the lowest. The above comparison implies that *Bo*Man26B is one of the more adept β-mannanases at accommodating galactose substitutions within the GH26 family.

Thus, *Bo*Man26A appears restricted by galactose substitutions, similar to several other GH26 β-mannanases (Table 1 and references therein), with a preference for hydrolysing manno-oligosaccharides (**Paper II**, Tables 1 and 2). *Bo*Man26B instead hydrolyses polysaccharides, with poor oligosaccharide hydrolysis, and is more unusual in the GH26 family as the level of restriction from galactose substitutions appears to be low (**Paper II**, Table 2, and **Paper III**, Table 1). In context of galactomannan utilisation encoded by the *Bo*ManPUL this naturally suggests a role for *Bo*Man26B as the enzyme of initial attack on polysaccharide galactomannan substrates, while *Bo*Man26A would hydrolyse subsequent mannooligosaccharides after hydrolysis by *Bo*Man26B and a removal of galactose substitutions by *Bo*Gal36A (**Paper II**, Figure 8).

Product profiles of *Bo*Man26A and *Bo*Man26B

Upon β-mannan oligo- and polysaccharide hydrolysis, *Bo*Man26A produced M2 as a major product, with smaller amounts of longer oligosaccharides produced for LBG and guar gum (Figure 6) (**Paper II**, Figure 3). *Bo*Man26B produced more varied hydrolysis products compared to *Bo*Man26A, especially for guar gum, where M2 was no longer a major product of end point hydrolysis (Figure 6) (**Paper II**, Figure 3). LBG and guar gum hydrolysis by *Bo*Man26B produced products primarily of DP2-5, as revealed by guar α-galactosidase processing of the *Bo*Man26B LBG and guar gum degradation products (**Paper III**, Figure 2). The generation of random length products identifies *Bo*Man26B as an endo-acting β-mannanase. *Bo*Man26A primarily generates M2 making it a mannobiohydrolase

(**Paper II**), several of which have previously been identified in GH26 [124,194,195]. The longer products seen for *Bo*Man26A upon LBG and guar gum hydrolysis are indicative of endo-action (**Paper II**, Figure 3), likely caused by restrictions due to galactose substitutions. The dominance of M2 products may indicate that *Bo*Man26A is a processive enzyme, but that any galactose substitutions would restrict this activity (**Paper II**). In GH26 one of the previously characterised mannobiohydrolases has also been suggested to be processive [194].

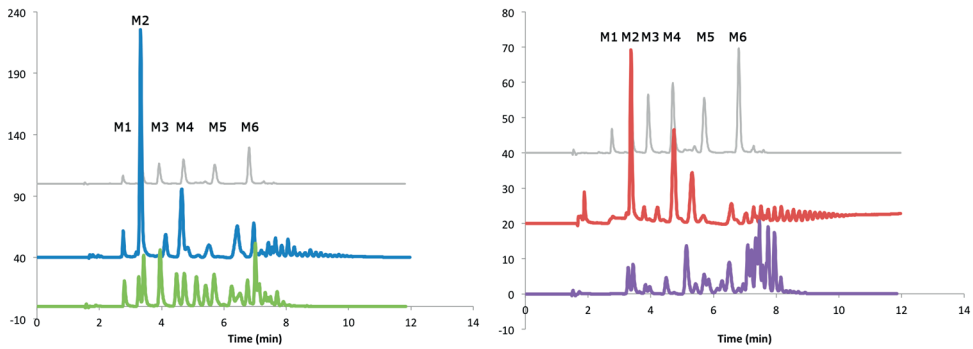


Figure 6 HPAEC-PAD analysis of poduct profiles after 24 h hydrolysis of LBG (left) and guar gum (right) by *Bo*Man26A (blue and red, respectively, **Paper II**, Figure 3) and *Bo*Man26B (green and purple, respectively, **Paper III**, Figure 2), shown with a 2.5 µM M1-M6 standard (grey). The M1-M6 peaks have been labelled.

To summarise the biochemical characterisation of *Bo*Man26A and *Bo*Man26B: *Bo*Man26B prefers longer substrates, which it hydrolyses to oligosaccharides of DP2-5. Since *Bo*Man26B is anchored to the outer membrane and is not sensitive to galactose substitutions it is likely the first GH in galactomannan utilisation encoded by the *Bo*ManPUL. *Bo*Man26A instead primarily degrades β-mannan substrates into M2 and hydrolyses manno-oligosaccharides more efficiently than *Bo*Man26B does. This endo-capable mannobiohydrolase is periplasmic and restricted by galactose substitutions, making it likely to have a place downstream in galactomannan utilisation.

Synergy: enzymes helping each other

The characterisation of *Bo*Man26A, *Bo*Man26B and *Bo*Gal36A described above indicates an order of GH attack on galactomannans by these GHs expressed form the *Bo*ManPUL. According to this view, the galactomannan attack is initiated by *Bo*Man26B, followed by *Bo*Gal36A and finally *Bo*Man26A. To assess the synergy between *Bo*Man26A, *Bo*Man26B and *Bo*Gal36A co-incubations of either β-

mannanase with *Bo*Gal36A were carried out in **Papers II** (Table 3) and **III** (Table 1). There have been several studies on pathways for microbial polysaccharide utilisation based on differences in substrate specificity [88,116,124,126,129,135,195,309,310] and synergy between GH enzymes on β-mannan substrates [222,245,246,253,311]. However, GH synergy within PUL-related enzymes is relatively unexplored, having previously been shown for GHs from xylan PULs by coincubations of different substrates with various enzyme mixtures [312,313].

In the presence of *Bo*Gal36A the release of M2 increased for both *Bo*ManPUL β-mannanases (**Paper II**, Table 3 and **Paper III**, Table 1). Conversely, an increase in galactose release was only seen when *Bo*Gal36A was co-incubated with *Bo*Man26B (**Paper III**, Table 1) but not when co-incubated with *Bo*Man26A (**Paper II**, Table 3). The increase in M2 release by *Bo*Man26B was partly due to *Bo*Gal36A hydrolysis of 6^1-α-D-galactosyl-mannobiose produced by *Bo*Man26B, meaning it is not clear that *Bo*Gal36A assists *Bo*Man26B galactomannan hydrolysis. For *Bo*Man26A, however, M2 release increases roughly 3-fold for LBG and hydrolysis of G2M5 became possible in the presence of *Bo*Gal36A. G2M5 was likely de-galactosylated into M5, a substrate of *Bo*Man26A (**Paper II**, Table 3). The increased release of M2 from *Bo*Man26A in the presence of *Bo*Gal36A clearly indicates a synergistic relationship.

The increase in galactose release by *Bo*Gal36A when co-incubated with *Bo*Man26B was significantly higher than the increase in M2 release by *Bo*Man26B (**Paper III**, Table 1). Upon *Bo*Gal36A co-incubation with *Bo*Man26A there was instead a decrease in galactose produced (**Paper II**, Table 3), which may be due to competition over similar substrate binding sites between the two enzymes. The synergy fits the biochemical data: *Bo*Man26B synergises with *Bo*Gal36A, which hydrolyses oligosaccharides more efficiently than polysaccharides (**Paper I**). *Bo*Gal36A then synergises with *Bo*Man26A by removing galactose substitutions, which *Bo*Man26A is restricted by (**Paper II**). This provides support for the view described above relating to the order in which the GH26 enzymes encoded by the *Bo*ManPUL degrade galactomannan.

Potential PUL interaction: the SusD-like protein

The binding of the SusD-like protein from the *Bo*ManPUL to various manno-oligosaccharides was briefly studied using microscale thermophoresis (MST) in **Paper III**. No reliable binding could be seen to M6 and low viscosity LBG, but a K_D of 3.9 ± 0.7 mM was determined for G2M5 (**Paper III**, Figure 3). SusD-like protein can bind specific saccharides with high affinity, such as in *B. fragilis* where a SusD-like protein binds sialic acid with a K_D of 397 nM [314], but may in other cases be more specific toward polysaccharide binding [80]. That SusD-like

protein binds to G2M5 fits well with the longer galactomanno-oligosaccharide products generated by *Bo*Man26B (**Paper III**, Figure 2). SusD forms a complex with the transporter SusC [119,120,314,315] and the only available crystal structures of the complex show that binding and transport are facilitated by a pedal bin mechanism [133]. In Sus the binding of SusD is reduced in absence of SusC [119], but SusD and similar proteins are capable of binding without SusC [80,314]. Thus, the *Bo*ManPUL SusD-like protein may bind more efficiently in the presence of SusC. In addition the outer membrane proteins of the Sus system, SusC, SusD, the α-amylase SusG and the sugar binding proteins SusE and SusF have been shown to interact [118,315], possibly through substrate induced complex formation [316]. It would be interesting to continue the binding studies of SusD to determine binding specificity and potential protein interaction partners within the *Bo*ManPUL.

A model of galactomannan utilisation by the *Bo*ManPUL

While the GHs encoded by several PULs from human gut bacteria have been characterised [88,115,126,129,135,184,309,310,313,317-320], *Bo*Gal36A, *Bo*Man26A and *Bo*Man26B are the first encoded by a β-mannan PUL. The cellular location (**Paper II**, Figure 2) and biochemical data from **papers I-III** reveal a model of galactomannan utilisation by the *Bo*ManPUL (Figure 7) (**Paper II**, Figure 8). Polysaccharide binding to the outer membrane surface is possibly facilitated by the putative sugar binding SusE-positioned protein. *Bo*Man26B, outer membrane anchored and insensitive to galactosyl side-groups (**Paper II**), cleaves galactomannan into shorter oligosaccharides primarily of DP2-5 (**Paper III**, Figure 2). These are transported into the periplasm by a putative SusCD-like complex, in which the SusD-like protein is known to bind G2M5 (Figure 7) (**Paper III**, Figure 3). *Bo*Gal36A is periplasmic and preferentially hydrolyses internal galactose substitutions (**Paper I**), generating manno-oligosaccharides. As previously studied for *Bo*ManPUL, the HTCS regulator, which binds manno-oligosaccharides above DP3, activates transcription (Figure 7) [113]. *Bo*Man26A is restricted by galactose substitutions (**Paper II**) and generates M2 from the degalactosylated manno-oligosaccharides that are likely transported into the cytoplasm by a putative symporter. There, a predicted isomerase converts M2 into mannosyl glucose that is cleaved by a confirmed mannoglucosyl phosphorylase (unpublished data) into mannose and glucose.

In the case of other characterised GHs encoded by PULs from human gut bacteria the periplasmic enzymes are often exo-acting [126,309,310], similar to the mannobiohydrolase activity of *Bo*Man26A (**Paper II**, Table 1). In most of these cases endo-capable activity on polysaccharides, as seen for *Bo*Man26A (**Paper II**, Table 1 and Figure 3), has not been shown. The ability to hydrolyse

polysaccharides is generally not necessary for periplasmic PUL proteins, as the polysaccharides are commonly hydrolysed into oligosaccharides before transport into the periplasm [126,309,310].

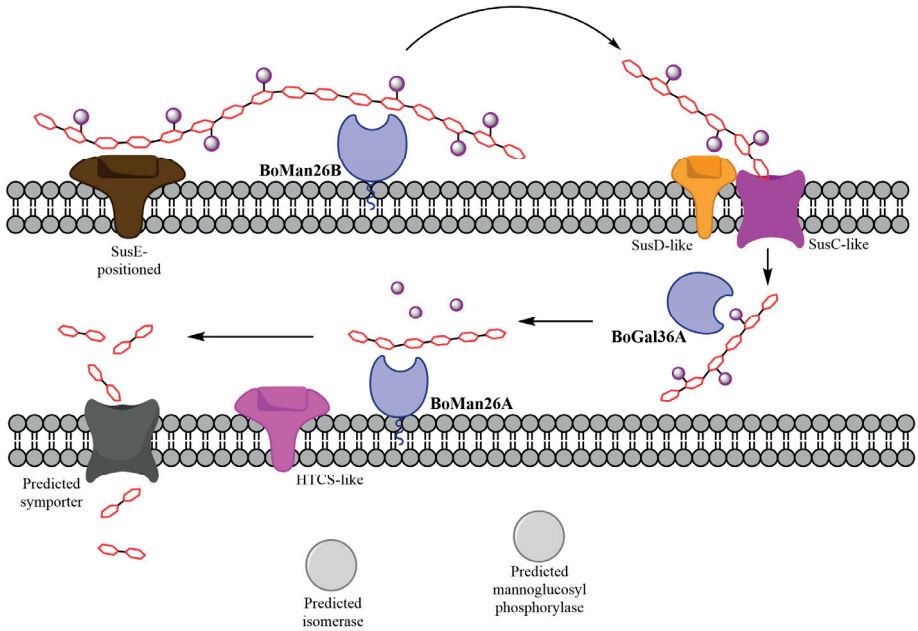


Figure 7 Proposed action of the proteins encoded by the *Bo*ManPUL (as shown in **Paper II**, Figure 8) on galactomannan (mannosyl backbone is red, with purple galactosyl spheres), showing the outer membrane (top), perplasmic space and inner membrane (bottom). All proteins have been labelled with their suggested role in galactomannan utilisation or confirmed enzyme name. GH enzymes characterised in this thesis are coloured blue and labelled in bold. Other coloured proteins are the SusE-positioned protein (brown), SusD-like protein (yellow), SusC-like protein (pink) and the HTCS-like regulator (pink). The proteins are coloured according to **Paper II**, Figure 1. The predicted function of the symporter and the isomerase is based on a BLAST search and on Senoura et al. [304].

Structural studies of *Bo*Man26A and *Bo*Man26B

*Bo*Man26A and *Bo*Man26B are both unusual among currently characterised GH26 β-mannanases. *Bo*Man26A is an endo-capable mannobiohydrolase (**Paper II**), one of few characterised mannobiohydrolases in GH26 [124,194,195]. *Bo*Man26B appears significantly less restricted by galactose substitutions (**Papers II** and **III**) than other GH26 β-mannanases [176,196,222,223]. Comparing *Bo*Man26A and *Bo*Man26B reveals a difference in preferred substrate length: *Bo*Man26A

efficiently hydrolyses manno-oligosaccharides, while *Bo*Man26B prefers longer substrates (**Paper II**). The differences in fine-tuned substrate specificity between *Bo*Man26A and *Bo*Man26B are likely correlated with differences in the structure of the active site cleft. In **Papers II** and **III** the crystal structures were determined for *Bo*Man26A and *Bo*Man26B, respectively. Both are $(\beta/\alpha)_8$-barrel structures typical for GH-A [144] and contain the seven residues conserved in GH26, including the nucleophile and acid/base [175], seen in all current GH26 β-mannanase structures [124,125,153,175,212,221,226].

Active site cleft substrate restrictions in *Bo*Man26A

The active site cleft of *Bo*Man26A is narrow (**Paper II**, Figure 7) and shows a large degree of conservation with the closest structural homologue: GH26 exo-β-mannanase C from *Cellvibrio japonicus* (*Cj*Man26C, PDB ID 2VX6) [124]. Like *Bo*Man26A, *Cj*Man26C is a mannobiohydrolase that requires at least four subsites for efficient hydrolysis [124]. The determined *Cj*Man26C structure includes a galactosyl-mannotetraose (G1M4) bound from subsite -2 to +2 with the galactosyl substitution in subsite -1 [124], which was used for subsite determination and analysis of substrate interaction in *Bo*Man26A via structural overlay (**Paper II**, Figure 6).

The positive subsites

In subsites -2 to +2 only three residues proposed to be involved in substrate binding are not conserved between *Cj*Man26C and *Bo*Man26A: Gln385, Asp266 and Arg269 (*Cj*Man26C numbering, Figure 8, **Paper II**, Figure 6). Asp266 and Arg269 (Gly232 and Glu234 in *Bo*Man26A, respectively) are located in the +2 subsite, where the greatest overall differences in active site cleft structure between *Bo*Man26A and *Cj*Man26C are seen (Figure 8) (**Paper II**, Figure 6). The shorter side chains of Gly232 and Glu234 in *Bo*Man26A, compared to Asp266 and Arg269 in *Cj*Man26C, makes the +2 subsite more open in *Bo*Man26A and may suggest weaker substrate interaction (Figure 8). However, both *Bo*Man26A and *Cj*Man26C require 4 subsites for efficient hydrolysis, based on differences in catalytic efficiency between M3 and M4, and generate M2 hydrolysis products [124] (**Paper II)**. Their similarities in oligosaccharide hydrolysis indicate that the +2 subsite still plays a role in both *Bo*Man26A and *Cj*Man26C, despite their structural differences in this subsite. Potential interactions in the +2 subsite by *Bo*Man26A may involve hydrophobic stacking by Phe261 or polar interactions with Glu234 (Figure 8). However, for the sugar residue in the +2 subsite to interact with either Phe261 or Glu234 a shift in position of the saccharide chain would be required compared to what is seen for the *Cj*Man26C structure. The residue Glu382 in *Cj*Man26C corresponds to His332 in *Bo*Man26A and both are at

a distance suitable for potential polar interaction with the mannosyl unit (**Paper II**, Figure 6). Thus, the main differences in the active site clefts of *Cj*Man26C and *Bo*Man26A are found in the positive subsites beyond subsite +1 (Figure 8).

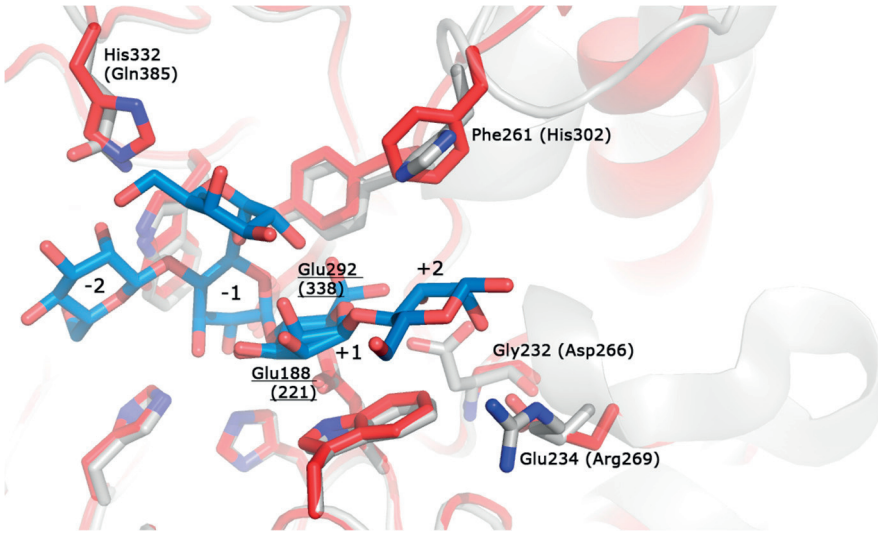


Figure 8 An overlay of *Bo*Man26A (red) and *Cj*Man26C (grey) with the G1M4 from *Cj*Man26C shown in stick representation (blue). The seven residues conserved in all GH26 β-mannanases, as well as residues around the active site cleft that differ between *Bo*Man26A and *Cj*Man26C are shown in stick representation. The catalytic residues (underlined) and differing residues are labelled, *Bo*Man26A numbering (corresponding *Cj*Man26C residue in brackets). A semi-transparent cartoon of the backbone is shown for both residues.

The negative subsites: restricted by loops

While *Bo*Man26A is capable of endo-action (**Paper II**), *Cj*Man26C is an exo-acting enzyme [124], suggesting a structural difference in the negative subsites. However, the negative subsites are largely conserved between *Bo*Man26A and *Cj*Man26C (**Paper II**, Figure 6). ^{18}O-labelling of M5 hydrolysis by *Bo*Man26A showed that, consistent with a preferential production of M2, *Bo*Man26A was about 5 times more likely to accommodate substrate in subsites -2 to +3 than in -3 to +2, yet binding in the -3 to +2 subsites was clearly possible (**Paper II**, Figure 4). In both *Bo*Man26A and *Cj*Man26C two loops, loop 2 and loop 8, restrict substrate access beyond subsite -2 [124] (**Paper II**, Figure 6). Loop 8 is situated to one side of the -3 subsite and hydrogen bonds with the backbone of loop 2, the tip of which is located in the -3 subsite, blocking it (**Paper II**, Figure 6). Removal of loop 2 through mutation in *Cj*Man26C causes it to become endo-acting, heavily implying that loop 2 is responsible for exo-activity in this enzyme [124]. While

loop 2 likely restricts the region beyond subsite -2 in *Bo*Man26A it apparently does not confer the same strict exo-action that it does in *Cj*Man26C.

Due to steric clashes with loop 2, it is currently not clear how a longer saccharide chain would be accommodated in the negative subsites beyond subsite -2 in *Bo*Man26A. One possible explanation for the endo-capability observed for *Bo*Man26A is loop flexibility (**Paper II**). The flexibility of loops surrounding the active site cleft in β-mannanases has previously been shown to affect substrate specificity and binding in GH5 enzymes [275,321]. Crystallographic structures are static images composed of the average of a very large number of protein molecules in an unnatural environment that may perturb the natural dynamic features of the protein. NMR peak assignment of *Bo*Man26A was performed in **Paper IV** to assess possible loop flexibility in future dynamics studies. In the NMR assignment of *Bo*Man26A, loop 8 was not possible to assign due to peak broadening beyond detection, which most likely reflects dynamic exchange between alternative conformations. Further NMR-based dynamics studies with a focus on loops 2 and 8 will hopefully shed light on potential loop flexibility and substrate accommodation in a -3 subsite by *Bo*Man26A.

Accommodation of galactose substitutions

Saccharide binding to *Bo*Man26A is restricted by galactose substitutions, the structural basis of which can be assessed using an overlay of the G1M4 substrate from *Cj*Man26C [124] (**Paper II**, Figure 6). In subsite +2 the area around the mannosyl O6 is very open, the closest atom is CE3 of Trp193 5.4 Å away, enabling the presence of a substitution (Figure 8) (**Paper II**, Figure 6). In *Cj*Man26C a very weak density was seen in the +2 subsite, indicating partial occupancy of a galactose substitution [124]. The +1 subsite mannosyl O6 is situated in a narrow part of the active site cleft and any galactosyl unit would likely clash with the aromatic residue Phe261 in *Bo*Man26B (His302 in *Cj*Man26, would possibly cause the same clash) (Figure 8) (**Paper II**, Figure 6). In addition, substitutions in both the -1 and +1 subsites simultaneously would be at risk of severe clashes with each other.

Galactosyl side-group accommodation in the -1 subsite is shown for *Cj*Man26C [124] and is likely similar in *Bo*Man26A due to the large conservation in this subsite. In the -2 subsite the mannosyl O6 faces the enzyme and interacts with Asp99 (**Paper II**, Figure 6), blocking the ability to fit a galactosyl side-group. Any galactosyl accommodation in a potential -3 subsite is speculative, but if loops 2 and 8 undergo conformational change to make space for the substrate, this may restrict the space available for a substitution around a possible -3 subsite. Of the known subsites in *Bo*Man26A, only the -2 and +1 subsites are unlikely to accommodate galactose substitutions, a restriction that correlates with the

observed reduction in activity of *Bo*Man26A on guar gum compared to LBG and their different levels of galactose substitution.

A wide active site cleft in *Bo*Man26B

*Bo*Man26B requires at least 6 subsites for efficient hydrolysis (**Paper II**, Table 2), which is reflected by a long active site cleft with 5 determined negatively numbered subsites (**Paper III**, Figure 8). A structure could be determined with a G1M4 saccharide bound in subsites -5 to -2 (**Paper III**, Figure 6) and a bound calcium ion in a metal binding site (**Paper III**, Figure 5). The presence of calcium in solution improves the stability of *Bo*Man26B, but does not affect activity (**Paper III**), indicating that it is not involved in catalysis or substrate binding. Metal binding with an impact on thermal stability has previously been seen in thermostable β-mannanases [226,228-230], however these metal binding sites are not conserved in *Bo*Man26B.

G1M4 binding

The G1M4 bound in subsites -5 to -2 of *Bo*Man26B includes a galactose substitution at the -4 subsite (**Paper III**, Figure 6). In the -5 subsite Trp112 forms a clear aromatic stacking platform for the mannosyl unit and in subsite -2 the mannosyl unit O6 interacts with Tyr315 and the O2 interacts with Trp314. No clear interaction was seen with the mannosyl unit in the -4 subsite, while the galactosyl unit interacts with Tyr148 and Lys149, which may imply a preference for galactose substitutions in this subsite. In subsite -3 the main interaction with the sugar residue is hydrophobic stacking with Tyr317. Thus, the strongest interactions with mannosyl units are present in subsites -5 and -2. The possible galactosyl side-group preference in the -4 subsite may be one reason for the low levels of restriction by substitutions carried by galactomannan substrates displayed by *Bo*Man26B. (**Paper III**, Figure 6)

Accommodation of galactose side-groups in BoMan26B and its closest structural homologues

The two closest structural homologues to *Bo*Man26B are GH26 β-mannanase A from *Podospora anserina* (*Pa*Man26A, PDB ID 3ZM8) [125] and *Rs*Man26C (PDB ID 3WDR) [175]. While *Rs*Man26C displays a level of restriction by galactose substitution typical for GH26, *Bo*Man26B and *Pa*Man26A are two of the least restricted enzymes in the family so far (Table 1). To understand the differences in the ability of *Bo*Man26B, *Pa*Man26A and *Rs*Man26C to accommodate galactosyl side-groups the crystal structures of these three enzymes were compared using an overlay with the *Bo*Man26B G1M4 complex structure (Figure 9) (**Paper III**, Figure 8).

Similar to *Bo*Man26B, *Pa*Man26A and *Rs*Man26C also have long active site clefts with up to 4 and 5 significant negatively numbered subsites. Both *Bo*Man26B and *Rs*Man26C have determined structures with bound substrate in subsites -5 to -2 (G1M4 and M4, respectively), while no substrate bound structure has been determined for *Pa*Man26A. For the purposes of this comparison the G1M4 oligosaccharide from *Bo*Man26B was superposed with *Pa*Man26A to assess galactosyl side-group accommodation.

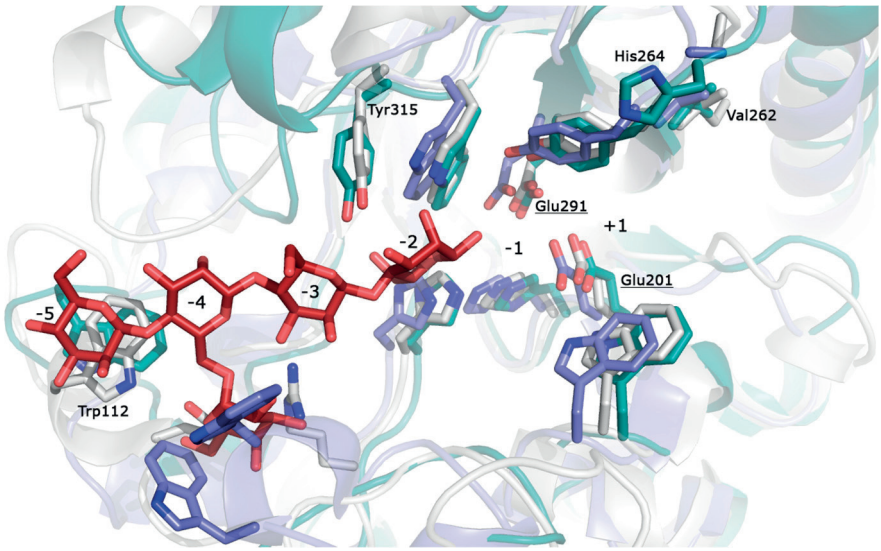


Figure 9 Overlay of *Bo*Man26B (green), *Rs*Man26C (grey) and *Pa*Man26A (blue) with the G1M4 from the *Bo*Man26B structure shown in stick representation (red). The side chains of the seven residues conserved in all GH26 β-mannanases, Trp112 in the -5 subsite and residues relating to accommodation of galactose substitutions are shown in stick representation, while a semi-transparent cartoon is shown for the backbone. The subsites, Trp112 in the -5 subsite, the catalytic residues (underlined) and residues potentially affecting the accommodation of galactose substitutions are labelled according to *Bo*Man26B numbering.

Regarding the negative subsites, *Rs*Man26C and *Bo*Man26B are so far the only GH26 β-mannanase structures with an identified -5 subsite, for both involving stacking with a tryptophan and in both cases open enough to accommodate a galactosyl side-group (Figure 9) [175] (**Paper III**, Figure 7). Accommodation of a galactose substitution in the -4 subsite is only possible for *Bo*Man26B and would cause clashes with Glu92 and Arg126 in *Rs*Man26C and Trp244 and Trp245 in *Pa*Man26A, where they stack with the -4 subsite mannosyl (Figure 9) (**Paper III**, Figure 7). The mannosyl O6 in the -3 subsite points out of the active site cleft in all three cases, making accommodation of a galactose substitution likely (**Paper**

III, Figure 7) (Figure 9). Thus, *Bo*Man26B is the only one out of the three compared enzymes that can accommodate a galactosyl side-group in all three of the -3 to -5 subsites, as *Rs*Man26C and *Pa*Man26A appear to be blocked in this capacity at subsite -4.

In the -2 subsite the mannosyl O6 in *Bo*Man26B and *Rs*Man26C is partially responsible for mannosyl interaction with the enzyme, blocking a galactose substitution (**Paper III**, Figure 7). *Pa*Man26A has instead been shown to be able to harbour a galactosyl side-group in subsite -2 [196]. In *Bo*Man26B and *Rs*Man26C a galactosyl side-group in the -2 subsite would be blocked primarily by Tyr315 (Figure 9). As has been shown previously, the -1 subsite is highly conserved and also capable of accommodating a galactose substitution [124,175,196]. This means that only *Pa*Man26A can fit a galactosyl side-group in both subsites -2 and -1, while *Bo*Man26A and *Rs*Man26C are restricted at subsite -2.

*Pa*Man26C has been indicated to be able to accommodate a galactose substitution in the +1 subsite [196], the primary residues that seem to enable this are Val361 (Val262 in *Bo*Man26B) and Ala363 (Ala259 in *Rs*Man26C) [196] (Figure 9). In *Rs*Man26C Val361 corresponds to Ile257, while in *Bo*Man26B Ala363 corresponds to His264 (Figure 9). Isoleucine and histidine are larger residues than valine and alanine, respectively, yet in both cases the side chains point away from the active site cleft making it unlikely that they would hinder a galactose substitution in any of the three enzymes (Figure 9).

Thus, *Rs*Man26C is unlikely to harbour galactose substitutions in subsite -2 and -4, while *Bo*Man26B and *Pa*Man26A appear to be restricted only in the -2 or -4 subsites, respectively. Based on ligand docking with *Pa*Man26A [196] and conservation with *Bo*Man26B and *Rs*Man26C all three enzymes are proposed to be able to simultaneously accommodate a galactosyl side-group in subsites -1 and +1. The possible restriction of galactose substitutions in subsites -4 and -2 in *Rs*Man26C is reflected in a roughly 90% reduction in catalytic efficiency on guar gum compared to LBG [223]. For *Bo*Man26B this reduction in catalytic efficiency is instead only about 50% (**Paper III**), similar to the reduction in initial rate seen for *Pa*Man26A [196], which relates to them only being restricted in their ability to accommodate galactose substitutions in one subsite.

The preference for binding in subsite -5

*Bo*Man26B requires at least 6 subsites for efficient substrate binding (**Paper II**, Table 2). The mannosyl interactions in the -5 subsite primarily consist of the hydrophobic stacking provided by Trp112 (**Paper III**, Figure 6). *Rs*Man26C and two close homologues from the same source (*Rs*Man26A and *Rs*Man26B) also have a similarly positioned tryptophan in the -5 subsite which, when mutated,

negatively affects binding, but not catalytic rate (k_{cat}) [223]. To assess the importance of Trp112 in *Bo*Man26B ^{18}O-labelling analysis of M6 hydrolysis by *Bo*Man26B and the two enzyme variants Trp112Phe and Trp112Ala confirmed the importance of the -5 subsite (**Paper III**, Figure 10). In *Bo*Man26B binding in the -5 to +1 and -4 to +2 subsites were preferred, while the binding mode preference was shifted away from the -5 subsite in the two enzyme variants (**Paper III**, Figure 10). While wild type (WT) *Bo*Man26B had a 33% binding preference for the -5 to +1 subsites, in the variants Trp112Phe and Trp112Ala the preference was only 12% and 4%, respectively (**Paper III**, Figure 10). The most dramatic decrease was seen between *Bo*Man26B and Trp112Phe, rather than between Trp112Phe and Trp112Ala, even though phenylalanine is another aromatic residue. Carbohydrates generally have a preference for interaction with aromatic residues, but phenylalanine is less favoured than tryptophan due to reduced electronegativity and hydrophobic area [322,323]. The shift in binding preference in the enzyme variants Trp112Phe and Trp112Ala away from the -5 subsites clearly demonstrates the importance of Trp112 for substrate binding in the -5 subsite.

Bioinformatic comparisons of *Bo*Man26A and *Bo*Man26B

In several ways *Bo*Man26A and *Bo*Man26B exemplify the spectrum of fine-tuned specificity and structure of characterised GH26 β-mannanases. β-Mannanases in GH26 range from endo-activity [153,175] to mannobiohydrolases [124,195], from severe [175,176] to limited [196,225] restriction by galactose substitutions and from two [124,153] to five [175,223] significant negatively numbered subsites (Table 2). The *Bo*Man26A and *Bo*Man26A pair encoded by the *Bo*ManPUL was mirrored by two types of homologous PULs identified in **Paper I**, which all encoded for a pair of putative GH26 β-mannanases (**Paper I**, Figure 1). In these homologous GH26 pairs one of the sequences was more similar to *Bo*Man26A, while the other was more similar to *Bo*Man26B (**Paper III**). To relate these GH26 pairs to other GH26 enzymes, phylogenetic analysis was carried out in **Paper III**. Based on the generated phylogenetic tree (**Paper III**, Figure 12) a multiple sequence alignment assessed the active site cleft conservation of *Bo*Man26B (**Paper III**, Figure S4).

Phylogenetic analysis of *Bo*Man26A and *Bo*Man26B

A phylogenetic tree in **Paper III**, Figure 12, contains five major branches. All characterised β-mannanases were found in branches A and B (**Paper III**, Figure 12). *Bo*Man26A and similar GH26 pair enzymes encoded by type I and type II PULs (**Paper I**) clustered in branch A, while *Bo*Man26B and similar GH26 pair enzymes clustered in branch B (**Paper III**, Figure 12). While details in fine-tuned specificity are difficult to predict within branch A and B there appears to be a difference in the length of the active site cleft, where those enzymes in branch A seem to have shorter active site clefts based on available crystal structures (Table 2) (**Paper III**). All GH26 pair enzymes similar to *Bo*Man26B clustered together with four additional sequences, including *Rs*Man26C [175], in a region of branch B termed "subbranch *Bo*Man26B" (**Paper III**, Figure 12). Out of the GH26 β-mannanses with a determined structure *Rs*Man26C, *Pa*Man26A and two enzymes from *Bacillus* clustered in branch B, while the others clustered in branch A (Table 2). A phylogenetic tree with a limited number of GH26 sequences previously indicated that *Bo*Man26A and *Bo*Man26B cluster in different branches [324], yet due to the significantly smaller scale than the phylogenetic tree in **Paper III** the evolutionary distance was not as clear. That *Bo*Man26A and *Bo*Man26B, together with their respective GH26 pair enzymes, cluster in two different major branches indicates a significant evolutionary distance between them.

Active site cleft conservation of *Bo*Man26B

The level of conservation in the active site cleft of *Bo*Man26B was assessed using a multiple sequence alignment of all GH26 pair enzymes, as well as the additional sequences from subbranch *Bo*Man26B (**Paper III**, Figure S4). Apart from seven active site cleft residues conserved in GH26 β-mannanases [153,175], the only residue conserved across all compared sequences was Asp101 (*Bo*Man26B numbering) (**Paper III**, Figure S4). Enzymes in "subbranch *Bo*Man26B" are likely to contain five negatively numbered subsites and be restricted in their ability to accommodate galactose substitutions in the -2 subsite as the only additional conserved active site cleft residues in this subbranch were Trp112 and Tyr315 (**Paper III**, Figures 8 and S4). The level of conservation at the -4 subsite was low, as exemplified by *Rs*Man26C [175], meaning that the ability to accommodate galactosyl side-groups in this subsite is difficult to predict. The relatively low levels of conservation in the active site cleft, even between enzymes with a close evolutionary relationship, indicate that fine-tuned specificity may be very highly specific for different enzymes.

Comparing *Bo*Man26A and *Bo*Man26B with other GH26 β-mannanase structures

The fine-tuned specificities of *Bo*Man26A and *Bo*Man26B differ significantly: *Bo*Man26A preferentially generates M2 from unsubstituted oligosaccharides, while *Bo*Man26B produces a range of products from galactomannan polysaccharides (**Papers II** and **III**). These biochemical differences between *Bo*Man26A and *Bo*Man26B are reflected in the structure of their active site clefts: *Bo*Man26B contains 5 significant negative subsites with a more open active site cleft compared to *Bo*Man26A with only 2 confirmed negative subsites (**Paper III**, Figure 8). Other currently experimentally determined GH26 β-mannanase structures also contain between 2 and 5 significant negatively numbered subsites, with corresponding differences in preferred substrate length and product profile (Table 2) [124,125,153,175,212,221,226]. For a comparison of GH26 β-mannanases with a determined structure [124,125,175,212,221,226,227] (**Papers II** and **III**) the catalytic module of these enzymes were superposed with *Bo*Man26A using the software PyMOL [307], the statistics for which can be found in Table 2. In all cases the seven residues conserved throughout GH26 were aligned, indicating that the overlay could be used for comparison.

Apart from the seven active site cleft residues conserved throughout GH26 [153], all situated round the -1 and +1 subsites, the degree of conservation in the active site cleft between the determined GH26 β-mannanase structures is low [124,125,175,212,221,226,227] (**Papers II** and **III**). When comparing the determined structures of GH26 β-mannanases, only one residue other than those conserved in GH26 β-mannanases is conserved in all of them (except *Cj*Man26C): Asp86 (*Bo*Man26A numbering), situated by the -2 subsite (Figure 10) [125,153,175,212,221,226]. This is the same residue that was conserved in the multiple sequence alignment with *Bo*Man26B (Asp101, **Paper III**, Figure S4). In *Cj*Man26C the corresponding residue is the conservative substitution Glu117 [124]. Moreover, the side chain carboxylate group is similarly positioned in all structures, including *Cj*Man26C, indicating a similar role for all compared enzymes (Figure 10). Out of all available GH β-mannanase structures containing a bound oligosaccharide [124,175,221,227] (**Paper III**) Asp86 is only reported to be involved in substrate binding for *Bo*Man26B (**Paper III**, Figure 6) and *Rs*Man26C [175]. This could indicate that Asp86 generally is involved in substrate binding, or the conserved nature of Asp86 and its position deep in the active site cleft may indicate an important structural role.

Table 2 Overview of determined subsites, preferred product lengths and product profiles for the GH26 β-mannanases with a determined structure. The branch that each enzyme belongs to in the phylogenetic tree from **Paper III** is also shown.

Enzyme	Branch	Subsites	Preferred minimun substrate length	Product profile[1]	PDB ID	RMSD (nr of Cα)[2]	References
*Bo*Man26A	A	-2 to +2	M4	Mainly M2	4ZXO		**Paper II**
*Bo*Man26B	B	-5 to +1	M6	Endo	6HF2	2.30 (151)	**Paper III**
*Cj*Man26A	A	-2 to +1	M3	Endo on polymannans, M2 from M4	1GVY	0.92 (222)	[153,227]
*Cj*Man26C	A	-2 to +2	M4	M2	2VX6	0.70 (243)	[124]
*Cf*Man26A	A	-3 to +2	M4	Endo on M3-M6	2BVT	1.17 (263)	[221]
*Pa*Man26A	B	-4 to +1	M5	Mainly M4 from M5 and M6	3ZM8	2.28 (161)	[125]
*Rs*Man26C	B	-5 to +1	M4	Endo	3WDR	1.88 (156)	[175]
*Bs*Man26A	B	-3 to +2	M4	M2 from M4 and M2+M3 from M5	2WHK	3.99 (225)	[212]
BCman	B	-4 to +2	M6	M4 and M2 from INM	2QHA	3.72 (227)	[226]
*B*750Man26A	A				4YN5	1.01 (237)	To be published

[1] In the case the products profile states endo a range of products is generally generated. In the cases where a limited set of substrates have been used for the product profile the substates tested are stated, if nothing is stated both oligomeric and polymeric β-mannan substrates have been tested.

[2] Root mean square deviation (RMSD) for an overlay with *Bo*Man26A with the number of aligned Cα atoms shown in brackets.

Loop structures in the glycone region of the active site cleft

The loop structures surrounding the active site cleft appear to be crucial for shaping fine-tuned substrate specificity. In *Bo*Man26A loops 2 and 8 likely cause the preference for producing M2 (**Paper II**, Figure 6). While loops 2 and 8 are conserved between *Bo*Man26A and *Cj*Man26C [124], in the other GH26 β-mannanase structures loop 2 is either shorter [221,227] or has an entirely different conformation (Figure 10) [125,175,212,226] (**Paper III**). Some GH26 β-mannanases have a shortened loop 2 in a similar conformation to *Bo*Man26A: these are the endo-acting enzymes with the shortest active site clefts (Table 2) [221,227]. In *Bo*Man26B and *Rs*Man26C loop 2 harbours Trp112 responsible for

preferential binding in the -5 subsite and is thus situated considerably further away from the active site than in *Bo*Man26A [175] (**Paper III**, Figure 9). Thus, the position of loop 2 could possibly play a major role in the fine-tuned specificity of several of the GH26 β-mannanases.

The structure of loop 8 is very variable between the different GH26 β-mannanase structures (Figure 10) [124,125,175,212,221,226,227] (**Papers II** and **III**). In some cases a shorter loop opens up the active site cleft around the -2 and -3 subsites [125,212,226], while in others a large residue (Arg, Tyr or Phe) occupy the same space as Arg324 in *Bo*Man26A and can provide potential interaction with a -2 mannosyl (Figure 10) [175,221,227] (**Paper III**). Interactions between loops 2 and 8 that potentially restrict the active site cleft are only seen in the mannobiohydrolases *Cj*Man26C and *Bo*Man26A [124] (**Paper II**, Figure 6), in contrast to the other determined structures of endo-acting GH26 β-mannanases [125,175,212,221,226,227] (**Papers III**).

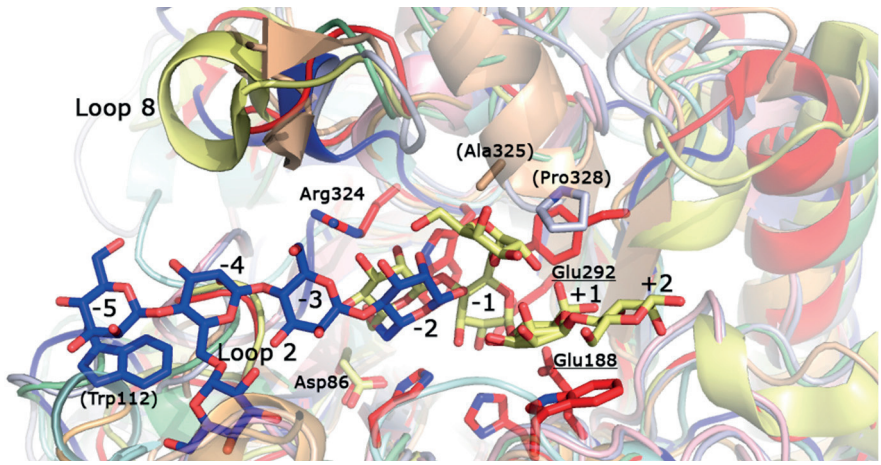


Figure 10 Overlay of all GH26 β-mannanases with a determined structure. *Bo*Man26A is shown in red and *Bo*Man26B in blue. Other GH26 β-mannanases are coloured according to the following: 1GVY beige, 2BVT green, 2QHA light blue, 2VX6 yellow, 2WHK pink, 3WDR cyan, 3ZM8 orange and 4YN5 grey. The G1M4 structures from *Bo*Man26B (**Paper III**) and *Cj*Man26C (2VX6) [124] are shown in stick representation. Also shown in stick representation are conserved residues in GH26 (shown residues are from *Bo*Man26A, catalytic residues are labelled and underlined) and residues relevant for discussions about GH26 conservation, loop conservation and galactose accommodation (labelled, non-*Bo*Man26A numbering in brackets). A transparent cartoon of the backbone is shown. Loops 2 and 8 are labelled and not rendered transparent.

Accommodation of galactose substitutions in GH26 β-mannanases

The GH26 β-mannanases with a determined structure vary in terms of their ability to accommodate galactose substitutions, as judged primarily from specific activity analysis or enzyme kinetics using mannans with different degrees of substitution (Table 1). Through an overlay of determined GH26 β-mannanase structures, one aspect of this restriction by galactosyl side-groups may be initially assessed. Two of the determined GH26 β-mannanase structures contain a galactose-substitution bound in the active site cleft: *Bo*Man26B, in the -4 subsite (**Paper III**, Figure 6), and *Cj*Man26C, in subsite -1 [124].

The -1 subsite is conserved in determined GH26 β-mannanase structures [124,125,175,212,221,226,227] (**Papers II** and **III**). However, the region where a galactose substitution would be situated differs. Based the G1M4 substrate from *Cj*Man26C in the current overlay of GH26 β-mannanase structures (Table 2, Figure 10), two out of these enzymes potentially clash with the galactose substitution in the -1 subsite: GH26 β-mannanases A from *Cellvibrio japonicus* (*Cj*Man26A, PDB ID 1GVY [227]) and *Bacillus* sp. strain JAMB-750 (*B*750Man26A, PDB ID 4YN5, to be published). However, in both cases the clash is minor and the closest distances are 1.9 Å and 1 Å for *Cj*Man26A and *B*750Man26A, respectively, making it likely a slight shift in position of saccharide or amino acid residues would allow galactosyl side-group accommodation (Figure 10). As such, the ability to accommodate a galactose substitution in the -1 subsite generally seems possible in GH26 β-mannanases. This has previously been suggested for some of the compared structures [124,175,196] (**Papers II** and **III**) and has also been indicated for GH26 β-mannanases without determined crystal structures [196,224,325].

The suggested ability to harbour a galactose substitution in the -2 subsite in *Pa*Man26A [196] seems unlikely for all other determined GH26 β-mannanase structures. In these structures a tyrosine, phenylalanine or an arginine (Arg324 in *Bo*Man26A, Figure 10) restricts the space in which a galactosyl side-group would fit. For all GH26 β-mannanase structures with a mannosyl unit in the -3 subsite [125,175,212,221,226] (**Paper III**) a galactosyl side-group would likely fit, as the mannosyl O6 points out of the active site cleft.

The galactosyl side-group in the conformation found in the -4 subsite of *Bo*Man26B (**Paper III**, Figure 6) would clash with all determined GH26 β-mannanase structures except *Cj*Man26A [227] and GH26 β-mannanase A from *Cellulomonas fimi* (*Cf*Man26A, PDB ID 2BVT [221]) (Figure 10). In some cases, such as stated above for *Pa*Man26A, amino acid side chains or whole loops fill the space of the galactosyl side-group in subsite -4 of *Bo*Man26B [125,212,226]. In the case of *Cj*Man26C, *Rs*Man26C and *Bo*Man26A [124,175] (**Paper II**) there is little direct overlap and the clashes are mainly due to proximity (Figure 10).

Interesting to note is that the GH26 β-mannanases with little or no clashes with the superposed -4 subsite galactosyl side group in *Bo*Man26B are not reported to have -4 subsites [124,221,227] (**Paper II**), with the exception of *Rs*Man26C [175]. This means that in all relevant structures, except *Bo*Man26B, the -4 subsite seems restricted in its ability to accommodate a galactosyl side-group.

Similar to the -2 subsite, *Pa*Man26A may also be able to accommodate a galactose substitution in the +1 subsite [196]. This is less likely in all other currently determined GH26 β-mannanase structures [124,175,212,221,226,227] (**Paper II**), where this area is more restricted in the current overlay, either by loop backbone or side chains, with the possible exception of *Bo*Man26B (**Paper III**) and *Rs*Man26C [175].

Summarising potential galactosyl side-group accommodation for the determined GH26 β-mannanases, *Pa*Man26A and *Bo*Man26B are the only enzymes shown to be restricted in only one subsite (-4 and -2, respectively) [196] (**Paper III**). All other currently determined GH26 β-mannanase structures appear to be restricted in at least subsites -2 and, where relevant, -4. With the exception of *Rs*Man26C, the above statement is also true for the +1 subsite. In accordance with the low structural restriction of galactosyl side-groups in *Pa*Man26A and *Bo*Man26B, these two enzymes also show the lowest reduction in catalytic ability with increasing galactose content [196] (**Paper III**) out of available results from other GH26 β-mannanases with a determined structure [176,223] (**Paper II**).

General discussion

Through biochemical and structural characterisation of *Bo*Man26A, *Bo*Man26B and *Bo*Gal36A in **Papers I-IV**, a model for the previously proposed galactomannan utilisation encoded by the *Bo*ManPUL [113] could be presented. In the process, a structural basis for the differences in fine-tuned specificity between *Bo*Man26A and *Bo*Man26B was revealed. For example, loop 2 in *Bo*Man26A restricts a potential -3 subsite, resulting in mannobiohydrolase activity (**Paper II**), while in *Bo*Man26B, loop 2 harbours the tryptophan mainly responsible for preferential binding in the -5 subsite (**Paper III**). In the following sections the structure-function relationship of fine-tuned specificity will be discussed based on three aspects relevant for *Bo*Man26A and *Bo*Man26B: 1) the effect that the number of different subsites has on the products produced; 2) the structural basis for the preference for, or restriction by, substitutions to the polysaccharide backbone; 3) the impact of loops surrounding the active site and their potential flexibility. These aspects will be related to other GHs, especially those encoded by other PULs from human gut bacteria. As none of the GHs characterised in this thesis contain CBMs, the focus has been on comparing the catalytic subunits of different enzymes, regardless of CBMs.

The effect of the number of subsites on product profile

The number of subsites and available space around the active site can be assumed to influence the preferred substrate length of a given GH. In β-mannanases, with the active site located in a cleft, the number of subsites usually relates to a minimum preferred substrate length [153,198,218]. For instance, *Bo*Man26A has 4 significant subsites, while *Bo*Man26B has 6 subsites as judged from their preferred β-mannan substrates of at least DP4 or DP6, respectively, and confirmed by their crystal structures (**Papers II** and **III**). For exo-acting enzymes with the active site located in a pocket, such as GH36 α-galactosidases, the presence of additional positively numbered subsites may restrict the ability to bind certain types of longer substrates due to space limitations. This was seen in **Paper I** when comparing *Bo*Gal36A with other GH36 subgroup I enzymes [149,301,306].

Similarly, the number of subsites can affect the hydrolysis products produced. Examples include *Pa*Man26A and *Cj*Man26A, which have 4 and 2 determined negatively numbered subsites and preferentially generate DP4 and DP2 oligosaccharides, respectively [125,153]. However, the product profiles of endo-acting enzymes such as *Cf*Man26A [221] and *Rs*Man26C [175] do not always correlate with their number of glycone subsites. These enzymes have 3 or 5 identified negative subsites, respectively, but generate a range of products of varying DP (Table 2). In the case of *Rs*Man26C the major initial product from M6 hydrolysis is M5, reflecting the number of significant glycone subsites, while the DPs of products from later stages of hydrolysis are more varied [175]. *Cf*Man26C instead in general produces varied DP products as short as DP1 even at initial time points, despite the presence of a -3 subsite [221].

For characterised endo-acting GHs, encoded by PULs from human gut bacteria (Table 3), the effects that the number of subsites has on the product profile vary [126,135,184,309,313,318] (**Paper III**). In some of the characterised cases, the GH enzymes primarily produce oligosaccharides of a certain DP, dependant on the enzyme, as final products [126,135,309,313]. The DP produced corresponds to the number of significant glycone subsites, when known [126,309]. The other three characterised endo-acting GH enzymes encoded by PULs from human gut bacteria all produce products with a range of DP [184,318] (**Paper III**, Figures 2 and 11) and, with the exception of *Bo*Man26B, the number of glycone subsites is unknown.

Thus, the number of subsites in endo-acting enzymes may govern the products they produce, but this is not necessarily the case. Enzymes with several glycone subsites, such as *Bo*Man26B, may still generate a range of shorter oligosaccharides in their product profiles. The differences in mode of binding and product profiles may in part be due to differences in subsite affinities, as has previously been discussed for various GH enzymes such as amylases and glucanases [326,327]. Some endo-acting enzymes have longer active site clefts, such as *Bo*Man26B, yet still produce a range of products (**Paper III**, Figure 2). In these cases a few stronger subsites, in *Bo*Man26B suggested to be subsites -5 and -2, may affect the range of products produced, but the ability to accommodate sugar units in other subsites may still result in products of a varying DP.

Table 3 Overview of some currently characterised GHs encoded by PULs from *Bacteroides* residing in the human gut

Organism	PUL target polysaccharide	Characterised GHs	GHs with determined structure	References
B. ovatus	β-mannan	2 GH26 and a GH36	2 GH26	**Papers I-IV**
B. ovatus	Xyloglucan	GH5, GH9, 2 GH3, 2 GH43, GH31 and a GH2	GH5, GH3, 2 GH43 and a GH31	[126,328]
B. ovatus	Xylan	GH3, 5 GH43, GH67, GH31, GH97, GH95, 2 GH10, GH30, GH115 and aGH98	GH95, GH115	[129]
B. ovatus	Mixed linkage glucan	GH16 and a GH3	GH16	[309]
B. cellulolyticus	Undetermined	GH39	GH39	[320]
B. intestinalis	Xylan	2 GH8		[135]
B. intestinalis	Xylan	GH10		[184]
B. intestinalis	Xylan	2 GH10, GH10/GH43, GH8, GH5, GH43 and a GH67		[313]
B. plebeius	Seaweed glycans	2 GH116 and a GH86	GH86	[318]
B. theta	α-mannan	GH99, 2 GH76, 2 GH92, GH38, 3 GH76 and 2 GH125	GH99, GH76 and GH125	[88]
B. theta	Starch	2 GH13 and a GH97	GH97	[115,329]
B. theta	Pectic glycans	GH2, GH51, GH146, GH28, GH27, GH147		[310]
B. theta	N-glycans	GH130		[317]
B. theta	Fructan	3 GH32		[319]

The preference for, or restriction by, substitutions

Substitutions of the backbone are a defining feature for several types of polysaccharides. These can be monosaccharide units, as for xyloglucans [47] and galactomannans [63], or other modifications, like acetylation of xylans [37] and GGMs [66]. As exemplified by *Bo*Man26A and *Bo*Man26B, the degree to which substitutions affect their activity varies (**Papers II** and **III**). *Bo*Man26A, which is restricted by substitutions, is part of the later stages of galactomannan processing conferred by *Bo*ManPUL (**Paper II**, Figure 8). Such restriction by substitutions is common for several other GH enzymes involved in later stages of PUL-mediated polysaccharide backbone degradation in human gut bacteria [126,129,135,313,328] (**Paper II**). For some large PULs encoding several GHs involved in the later stages of polysaccharide processing, a few of these enzymes

may be more limited in their restriction by substitutions (Table 3) [126,129]. A significant factor causing restriction by substitutions is lack of space due to clashes with the protein structure [313,328] (**Paper II**). Based on currently characterised GHs encoded by PULs from human gut bacteria (Table 3) [126,129,135,313,328] (**Paper II**), a common feature appears to be that enzymes involved in the final stages of backbone hydrolysis are often exo-acting and sensitive to substitutions on the saccharide chain.

GHs encoded by PULs from human gut bacteria (Table 3) that perform the initial attack on the target polysaccharide are, unsurprisingly, generally not restricted by substitutions [126,129,135,313] (**Paper III**). While this lack of restriction may be due to simply being able to accommodate a substitution, the enzyme may also have a preference for binding to them. To accommodate a substitution there only needs to be enough space for it to fit into the protein structure. This is suggested for a possible galactosyl side-group positioned in the -3 subsite of several GH26 endo-β-mannanases (**Paper III**) and the +1 subsite of *Pa*Man26A [196], where any specific interactions with a potential substitution are unclear.

The preference of an enzyme to bind a substituted saccharide is correlated with some sort of interaction with the enzyme, such as indicated for *Bo*Man26B (**Paper III**) and has previously been shown of a GH5 arabionxylan-specific xylanase [330]. Two GHs encoded by PULs from human gut bacteria have previously been confirmed to prefer substituted substrates: a GH5 endo-xyloglucanase encoded by a xyloglucan PUL (*Bo*GH5A) [126] and a GH30 glucuronoxylanase encoded by a xylan PUL (BACOVA_03432) [129], both from *B. ovatus*. *Bo*GH5A hydrolyses xyloglucan, a glucose-based polysaccharide that is highly substituted by xylose residues. A crystal structure of *Bo*GH5A (PDB ID 3ZMR) [126] shows that in subsites -2 and -3, out of four reported glycone subsites, the xylose substitutions interact via aromatic stacking to a tryptophan and a tyrosine, respectively. In subsite -4 the xylosyl side-group interacts with the enzyme backbone [126]. BACOVA_03432, which hydrolyses a xylan backbone with methylated glucuronic acid substitution, has no determined crystal structure. However, as reported in Rogowski et al. [129], the reason for the substitution preference is likely similar to what has been seen for a homologous enzyme: weak interactions with the backbone, but extensive hydrogen bonding with the substitution [331]. The preference for substitutions for both *Bo*GH5A and BACOVA_03432 is also evident from their activity, which is reduced or absent in the absence of substitutions. The increased activity on substituted substrates of *Bo*GH5A and BACOVA_03432 differs from *Bo*Man26B, which shows a reduction in activity of more highly substituted substrates (**Paper III**). Thus, while the -4 subsite of *Bo*Man26B may have a preference for a galactose substitution, the potential interaction through hydrogen bonding does not appear sufficient to have an overall impact on fine-tuned specificity, possibly due to the restrictions in the -2 subsite.

In summary, PUL encoded GHs from human gut bacteria that initiate utilisation of substituted polysaccharides are often not restricted by substitutions [126,129,135,313] (**Paper III**). In some cases endo-acting GHs even prefer a substituted substrate for efficient hydrolysis [126,129]. By contrast, GHs involved in the later stages of backbone hydrolysis of substituted polysaccharides are often exo-acting and restricted by substitutions [126,129,135,313,328] (**Paper II**). This restriction does not usually hinder polysaccharide utilisation, as the substitutions are removed during processing of the polysaccharide.

How loops affect fine-tuned specificity

Loops adjacent to the active site of GHs are often essential for shaping the surrounding region into a pocket [306] (**Paper I**), cleft [196] (**Papers II** and **III**) or tunnel [332] and as such they have a great impact on fine-tuned specificity. *Bo*Man26A and *Bo*Man26B are excellent examples showing the importance of loop structures in an active site cleft, as the differences in fine-tuned specificity between the two enzymes is in part due to differences in loop conformation (**Papers II** and **III**). The following sections will briefly review two major aspects of loop impact on fine-tuned specificity: how loop conformation shapes endo- and exo-activity and the potential impact of loop flexibility.

Loop determinants of endo- versus exo-activity

Whether a GH enzyme is endo- or exo-acting is mainly defined by the structure of loops around the active site cleft. The effect of these loops on endo- and exo-activity has been well studied for cellulose-acting enzymes from GH6 and GH7 [295,332-334]. GH6 and GH7 both contain endo-glucanases, for which the loops form a cleft, and processive, exo-acting cellobiohydrolases, where the loops form a tunnel [295,332-334]. In several other types of exo-acting GHs, the active site is located in a cleft, which is blocked at one end by one or several loops. Examples include a GH43 arabinanase from *Cellvibrio japonicus* [335], a GH5 mannosidase from *Cellvibrio mixtus* [214] and *Cj*Man26C [124]. Deletion of loops blocking the active site cleft or creating the tunnel in cellobiohydrolases increases the endo-capability of these normally exo-acting enzymes [124,332,335].

In PUL encoded polysaccharide processing systems from human gut bacteria (Table 3), the GHs that initially attack the target polysaccharide are commonly endo-acting, while backbone-cleaving GHs involved in processing at later stages are usually exo-acting [115,126,129,135,184,309,313,318] (**Papers II** and **III**). Endo-acting PUL-encoded GHs with a determined structure have their active site

located in a cleft [126,129,309,318] (**Paper III**, Figure 7). Out of the few available structures of PUL-encoded exo-acting enzymes that cleave the saccharide backbone, the endo-capable *Bo*Man26A has an active site cleft (**Paper II**, Figure 6), blocked by a loop, while a GH3 exo-β-glucosidase encoded by a xylan PUL in *B. ovatus* has the active site located in a pocket [328]. The importance of loops around the active site for endo- or exo-action is in this thesis exemplified by *Bo*Man26A and *Bo*Man26B, in which loop 2 either restricts the active site or is part of substrate binding in the -5 subsite (**Paper II** and **III**).

In addition, the loops surrounding the active site play an important part in conferring fine-tuned specificity to exo-enzymes, such as those GHs that remove substitutions. This is exemplified by *Bo*Gal36A, where a heavily truncated loop adjacent to the active site enables the removal of internal substitutions from a saccharide chain (Figure 5) (**Paper I**, Figure 6). For other enzymes of this family the equivalent loop restricts the area around the active site cleft, limiting activity to bind terminal galactosyl units [149,301,306].

Loop flexibility

The flexibility of loops around the active site may influence fine-tuned specificity and action of different GHs. For cellobiohydrolases loop flexibility has been implicated in processivity [334] and facilitates product release in a GH6 cellobiohydrolase from *Thermobifida fusca* [295]. In two endo-acting GHs, a GH19 chitinase from *Hordeum vulgare* [336] and a GH11 xylanase from *Bacillus subtilis* [294], a flexible loop is involved in substrate binding and fine-tuned specificity. In both cases, mutated variants of these enzymes that alter loop flexibility also affect specificity and activity [294,336].

The effect of flexibility and protein dynamics surrounding the active site in GH enzymes encoded by PULs from human gut bacteria (Table 3) has not been well studied. An observation among these enzymes relating to flexibility in the active site cleft is an altered conformation of two tryptophan residues upon substrate binding in a GH16 endo-glucanase encoded by a PUL involved in mixed-linkage β-glucan utilisation from *B. ovatus* [309]. Loops 2 and 8 in *Bo*Man26A have been suggested to possibly be flexible in response to substrate binding (**Paper II**) and part of loop 8 was not possible to assign during NMR analysis (**Paper IV)**. While there are examples of GH enzymes with flexible loops involved in substrate binding [294,336] none of these are similar to *Bo*Man26A and so further studies will be required to shed light on this matter.

The above sections on loops around the active site exemplify the importance of these structures. Apart from determining the overall shape of the area, into a cleft, pocket or tunnel, loops also determine the endo- or exo-mode of action of their

respective enzymes [124,332,335]. In addition, loop flexibility can impact binding, product release and fine-tuned specificity [294,295,336].

Conclusions and future perspectives

PUL-encoded polysaccharide processing systems break down polysaccharides using a range of GH enzymes that each fill specific roles in the pathway. While it is not uncommon for one PUL to harbour several GHs from the same family, these GHs may differ in their fine-tuned specificity. Differences in fine-tuned specificity can be related to the enzyme structure surrounding the active site.

*Bo*ManPUL is the main PUL in *B. ovatus* responsible for galactomannan utilisation (**Paper I**). Based on growth studies of a number of *Bacteroides* strains containing homologs to *Bo*ManPUL, the presence of an α-galactosidase appears important for growth on galactomannan (**Paper I**).

*Bo*Man26B is extracellularly located, not restricted by galactose substitutions and prefers longer saccharides, making it the enzyme that performs the initial attack, in a model of *B. ovatus* galactomannan utilisation (**Papers II** and **III**). The levels of restriction by substitutions in *Bo*Man26B are some of the lowest of characterised enzymes from the GH26 family (**Paper III**). One product that is likely produced by *Bo*Man26B, G2M5, is bound by the SusD-like protein (**Paper III)**, which is in accordance with a potential role in galactomannan utilisation (**Paper II**). Studies of interaction between SusC and SusD are limited and any interaction of all outer membrane PUL proteins has mainly been studied for the Sus of *B. theta*. It would be interesting to further expand the knowledge of PUL protein interactions and synergy through similar studies in *Bo*ManPUL.

*Bo*Man26B displays sequential synergy with the periplasmic *Bo*Gal36A (**Paper III**), which preferentially cleaves internal galactose substitutions from oligosaccharides, an unusual activity in this GH family (**Paper I**). *Bo*Gal36A in turn shows sequential synergy with the periplasmic mannobiohydrolase *Bo*Man26A, which is restricted by galactose substitutions, and efficiently hydrolyses mannooligosaccharides (**Paper II**). *Bo*Man26A produces longer oligosaccharides from substituted substrates, likely due to restriction by the galactosyl side-groups (**Paper II**).

In a model of PUL-encoded galactomannan utilisation, *Bo*Man26B cleaves the polysaccharide that is then transported into the periplasm, where it is hydrolysed by *Bo*Gal36A and subsequently by *Bo*Man26A (**Paper II** and **III**). This organisation is seen in several PULs: one or several outer membrane GHs that are

insensitive to substitutions cleave the target polysaccharide into oligosaccharides that are then transported into the periplasm. There a number of, usually exo-acting, GHs hydrolyse the resulting oligosaccharides. Further characterisation of the remaining proteins encoded by *Bo*ManPUL, together with more extensive synergy studies, would be interesting to shed additional light on the complete function of this galactomannan utilisation system.

Loop structures surrounding the active site of GHs define their fine-tuned specificity. In *Bo*Gal36A the absence of a P-loop conserved in other GH36 subgroup I α-galactosidases enables hydrolysis of internal galactose substitutions (**Paper I)**. The absence of this loop is currently based on a homology model and sequence alignments (**Paper I**) so it would be interesting to confirm this absence in an experimentally determined structure.

The evolutionary distance between *Bo*Man26A and *Bo*Man26B is relatively large (**Paper III**) and they have distinct active site clefts where loop 2 plays a pivotal role. In *Bo*Man26A loop 2 blocks the active site cleft (**Paper II**) and in *Bo*Man26B it is part of forming the -5 subsite (**Paper III**). For *Bo*Man26B loop 2 harbours Trp112, an aromatic stacking platform important for substrate binding in the -5 subsite (**Paper III**). In *Bo*Man26A loops 2 and 8 are indicated to be responsible for exo-action, despite the enzyme being endo-capable (**Paper II**). Further analysis of binding dynamics with NMR, initiated in **Paper IV**, aim to shed light on the role of loops 2 and 8 in substrate binding, especially as loop flexibility can be an important factor in substrate binding.

Together with *Pa*Man26A, *Bo*Man26B is capable of accommodating the largest number of galactose residues of known GH26 β-mannanase structures, conversely showing low levels of restriction by galactosyl side-groups (**Paper III**). The active site cleft of GH26 enzymes generally shows low levels of conservation (except in subsites -1 and +1), even when comparing enzymes that are relatively closely related, meaning similar restrictions in other GH26 enzymes is difficult to predict (**Paper III**).

The work in this thesis contributes to an increased understanding of the detailed polysaccharide utilisation by the human gut microbiota, providing further insight into the synergistic action of PUL-encoded proteins. The structure-function relationship of fine-tuned specificity, primarily in GH26, with specific attention to the ability to accommodate substitutions, is highlighted. In addition it points to several areas of continued interest for research.

References

1. El Kaoutari, A., Armougom, F., Gordon, J. I., Raoult, D., and Henrissat, B. (2013) The abundance and variety of carbohydrate-active enzymes in the human gut microbiota. *Nat Rev Microbiol* **11**, 497-504
2. Minic, Z. (2008) Physiological roles of plant glycoside hydrolases. *Planta* **227**, 723-740
3. Scheller, H. V., and Ulvskov, P. (2010) Hemicelluloses. *Annu Rev Plant Biol* **61**, 263-289
4. Gibson, L. J. (2012) The hierarchical structure and mechanics of plant materials. *J R Soc Interface* **9**, 2749-2766
5. Shtein, I., Bar-On, B., and Popper, Z. A. (2018) Plant and algal structure: from cell walls to biomechanical function. *Physiologia plantarum*
6. Park, Y. B., and Cosgrove, D. J. (2012) A revised architecture of primary cell walls based on biomechanical changes induced by substrate-specific endoglucanases. *Plant Physiol* **158**, 1933-1943
7. Burton, R. A., and Fincher, G. B. (2014) Evolution and development of cell walls in cereal grains. *Front Plant Sci* **5**, 456
8. Housley, T. L., Kirleis, A. W., Ohm, H. W., and Patterson, F. L. (1981) An Evaluation of Seed Growth in Soft Red Winter-Wheat. *Can J Plant Sci* **61**, 525-534
9. Vandeputte, G. E., and Delcour, J. A. (2004) From sucrose to starch granule to starch physical behaviour: a focus on rice starch. *Carbohyd Polym* **58**, 245-266
10. Buckeridge, M. S. (2010) Seed cell wall storage polysaccharides: models to understand cell wall biosynthesis and degradation. *Plant Physiol* **154**, 1017-1023
11. Dhanisha, S. S., Guruvayoorappan, C., Drishya, S., and Abeesh, P. (2018) Mucins: Structural diversity, biosynthesis, its role in pathogenesis and as possible therapeutic targets. *Crit Rev Oncol Hematol* **122**, 98-122
12. Delcour, J., Ferain, T., Deghorain, M., Palumbo, E., and Hols, P. (1999) The biosynthesis and functionality of the cell-wall of lactic acid bacteria. *Antonie Van Leeuwenhoek* **76**, 159-184
13. Spiro, R. G. (2002) Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology* **12**, 43R-56R
14. Casillo, A., Lanzetta, R., Parrilli, M., and Corsaro, M. M. (2018) Exopolysaccharides from Marine and Marine Extremophilic Bacteria: Structures, Properties, Ecological Roles and Applications. *Mar Drugs* **16**
15. Maunders, E., and Welch, M. (2017) Matrix exopolysaccharides; the sticky side of biofilm formation. *FEMS Microbiol Lett* **364**

16. Lam, J., Chan, R., Lam, K., and Costerton, J. W. (1980) Production of mucoid microcolonies by Pseudomonas aeruginosa within infected lungs in cystic fibrosis. *Infect Immun* **28**, 546-556
17. Sabater, S., and Romani, A. M. (1996) Metabolic changes associated with biofilm formation in an undisturbed Mediterranean stream. *Hydrobiologia* **335**, 107-113
18. Cheng, Y. S., Labavitch, J. M., and VanderGheynst, J. S. (2015) Elevated CO2 concentration impacts cell wall polysaccharide composition of green microalgae of the genus Chlorella. *Lett Appl Microbiol* **60**, 1-7
19. Liberman, G. N., Ochbaum, G., Arad, S., and Bitton, R. (2016) The sulfated polysaccharide from a marine red microalga as a platform for the incorporation of zinc ions. *Carbohyd Polym* **152**, 658-664
20. Porter, N. T., and Martens, E. C. (2017) The Critical Roles of Polysaccharides in Gut Microbial Ecology and Physiology. *Annu Rev Microbiol* **71**, 349-369
21. Fasahat, P., Rahman, S., and Ratnam, W. (2014) Genetic controls on starch amylose content in wheat and rice grains. *J Genet* **93**, 279-292
22. Klemm, D., Heublein, B., Fink, H. P., and Bohn, A. (2005) Cellulose: Fascinating biopolymer and sustainable raw material. *Angew Chem Int Edit* **44**, 3358-3393
23. Somerville, C. (2006) Cellulose synthesis in higher plants. *Annu Rev Cell Dev Bi* **22**, 53-78
24. Siracusa, V., Rocculi, P., Romani, S., and Dalla Rosa, M. (2008) Biodegradable polymers for food packaging: a review. *Trends Food Sci Tech* **19**, 634-643
25. Yadav, C., Saini, A., and Maji, P. K. (2018) Cellulose nanofibres as biomaterial for nano-reinforcement of poly[styrene-(ethylene-co-butylene)-styrene] triblock copolymer. *Cellulose* **25**, 449-461
26. Guimaraes, M., Teixeira, F. G., and Tonoli, G. H. D. (2018) Effect of the nano-fibrillation of bamboo pulp on the thermal, structural, mechanical and physical properties of nanocomposites based on starch/poly(vinyl alcohol) blend. *Cellulose* **25**, 1823-1849
27. Rao, Z. Q., Ge, H. Y., Liu, L. L., Zhu, C., Min, L., Liu, M., Fan, L. H., and Li, D. (2018) Carboxymethyl cellulose modified graphene oxide as pH-sensitive drug delivery system. *International Journal of Biological Macromolecules* **107**, 1184-1192
28. Courtenay, J. C., Sharma, R. I., and Scott, J. L. (2018) Recent advances in modified cellulose for tissue culture applications. *Molecules* **23**, 654
29. Bothast, R. J., and Schlicher, M. A. (2005) Biotechnological processes for conversion of corn into ethanol. *Appl Microbiol Biot* **67**, 19-25
30. Naik, S. N., Goud, V. V., Rout, P. K., and Dalai, A. K. (2010) Production of first and second generation biofuels: A comprehensive review. *Renew Sust Energ Rev* **14**, 578-597
31. Lu, D. R., Xiao, C. M., and Xu, S. J. (2009) Starch-based completely biodegradable polymer materials. *Express Polym Lett* **3**, 366-375
32. Atmodjo, M. A., Hao, Z. Y., and Mohnen, D. (2013) Evolving Views of Pectin Biosynthesis. *Annual Review of Plant Biology, Vol 64* **64**, 747-+

33. May, C. D. (1990) Industrial Pectins - Sources, Production and Applications. *Carbohyd Polym* **12**, 79-99
34. Naqash, F., Masoodi, F. A., Rather, S. A., Wani, S. M., and Gani, A. (2017) Emerging concepts in the nutraceutical and functional properties of pectin-A Review. *Carbohyd Polym* **168**, 227-239
35. Moreira, L. R., and Filho, E. X. (2008) An overview of mannan structure and mannan-degrading enzyme systems. *Appl Microbiol Biotechnol* **79**, 165-178
36. Sun, R. C., Sun, X. F., and Tomkinson, I. (2004) Hemicelluloses and their derivatives. *Acs Sym Ser* **864**, 2-22
37. Thorsheim, K., Siegbahn, A., Johnsson, R. E., Stalbrand, H., Manner, S., Widmalm, G., and Ellervik, U. (2015) Chemistry of xylopyranosides. *Carbohyd Res* **418**, 65-88
38. Saha, B. C. (2003) Hemicellulose bioconversion. *J Ind Microbiol Biot* **30**, 279-291
39. Izydorczyk, M. S., and Biliaderis, C. G. (1995) Cereal arabinoxylans: Advances in structure and physicochemical properties. *Carbohyd Polym* **28**, 33-48
40. Ebringerova, A., and Heinze, T. (2000) Xylan and xylan derivatives - biopolymers with valuable properties, 1 - Naturally occurring xylans structures, procedures and properties. *Macromol Rapid Comm* **21**, 542-556
41. Atkins, E. D. T., Parker, K. D., and Preston, R. D. (1969) Helical Structure of Beta-1,3-Linked Xylan in Some Siphoneous Green Algae. *Proc R Soc Ser B-Bio* **173**, 209-+
42. Viana, A. G., Noseda, M. D., Goncalves, A. G., Duarte, M. E. R., Yokoya, N., Matulewicz, M. C., and Cerezo, A. S. (2011) beta-D-(1 -> 4), beta-D-(1 -> 3) 'mixed linkage' xylans from red seaweeds of the order Nemaliales and Palmariales. *Carbohyd Res* **346**, 1023-1028
43. Immerzeel, P., Falck, P., Galbe, M., Adlercreutz, P., Karlsson, E. N., and Stalbrand, H. (2014) Extraction of water-soluble xylan from wheat bran and utilization of enzymatically produced xylooligosaccharides by Lactobacillus, Bifidobacterium and Weissella spp. *Lwt-Food Sci Technol* **56**, 321-327
44. 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. *Crit Rev Food Sci* **51**, 178-194
45. Naidu, D. S., Hlangothi, S. P., and John, M. J. (2018) Bio-based products from xylan: A review. *Carbohyd Polym* **179**, 28-41
46. Sorensen, I., Pettolino, F. A., Wilson, S. M., Doblin, M. S., Johansen, B., Bacic, A., and Willats, W. G. (2008) Mixed-linkage (1-->3),(1-->4)-beta-D-glucan is not unique to the Poales and is an abundant component of Equisetum arvense cell walls. *Plant J* **54**, 510-521
47. Fry, S. C., Nesselrode, B. H. W. A., Miller, J. G., and Mewburn, B. R. (2008) Mixed-linkage (1 -> 3,1 -> 4)-beta-D-glucan is a major hemicellulose of Equisetum (horsetail) cell walls. *New Phytol* **179**, 104-115
48. Bashir, K. M. I., and Choi, J. S. (2017) Clinical and Physiological Perspectives of beta-Glucans: The Past, Present, and Future. *Int J Mol Sci* **18**

49. Charalampopoulos, D., Wang, R., Pandiella, S. S., and Webb, C. (2002) Application of cereals and cereal components in functional foods: a review. *Int J Food Microbiol* **79**, 131-141
50. Kulkarni, A. D., Joshi, A. A., Patil, C. L., Amale, P. D., Patel, H. M., Surana, S. J., Belgamwar, V. S., Chaudhari, K. S., and Pardeshi, C. V. (2017) Xyloglucan: A functional biomacromolecule for drug delivery applications. *International Journal of Biological Macromolecules* **104**, 799-812
51. Sumathi, S., and Ray, A. R. (2002) Release behaviour of drugs from tamarind seed polysaccharide tablets. *J Pharm Pharm Sci* **5**, 12-18
52. Meier, H. (1958) On the structure of cell walls and cell wall mannans from ivory nuts and from dates. *Biochim Biophys Acta* **28**, 229-240
53. Kalidas, N. R., Saminathan, M., Ismail, I. S., Abas, F., Maity, P., Islam, S. S., Manshoor, N., and Shaari, K. (2017) Structural characterization and evaluation of prebiotic activity of oil palm kernel cake mannanoligosaccharides. *Food Chem* **234**, 348-355
54. Dunn, E. K., Shoue, D. A., Huang, X. M., Kline, R. E., MacKay, A. L., Carpita, N. C., Taylor, I. E. P., and Mandoli, D. F. (2007) Spectroscopic and biochemical analysis of regions of the cell wall of the unicellular 'mannan weed', Acetabularia acetabulum. *Plant Cell Physiol* **48**, 122-133
55. Frei, E., and Preston, R. D. (1968) Non-Cellulosic Structural Polysaccharides in Algal Cell Walls .3. Mannan in Siphoneous Green Algae. *Proc R Soc Ser B-Bio* **169**, 127-+
56. Chauhan, P. S., Sharma, P., Puri, N., and Gupta, N. (2014) A process for reduction in viscosity of coffee extract by enzymatic hydrolysis of mannan. *Bioproc Biosyst Eng* **37**, 1459-1467
57. Ratcliffe, I., Williams, P. A., Viebke, C., and Meadows, J. (2005) Physicochemical characterization of konjac glucomannan. *Biomacromolecules* **6**, 1977-1986
58. Maekaji, K. (1978) Determination of Acidic Component of Konjac Mannan. *Agr Biol Chem Tokyo* **42**, 177-178
59. Albrecht, S., van Muiswinkel, G. C., Xu, J., Schols, H. A., Voragen, A. G., and Gruppen, H. (2011) Enzymatic production and characterization of konjac glucomannan oligosaccharides. *J Agric Food Chem* **59**, 12658-12666
60. Behera, S. S., and Ray, R. C. (2017) Nutritional and potential health benefits of konjac glucomannan, a promising polysaccharide of elephant foot yam, Amorphophallus konjac K. Koch: A review. *Food Rev Int* **33**, 22-43
61. Tester, R. F., and Al-Ghazzewi, F. H. (2016) Beneficial health characteristics of native and hydrolysed konjac (Amorphophallus konjac) glucomannan. *J Sci Food Agric* **96**, 3283-3291
62. Dea, I. C. M., and Morrison, A. (1975) Chemistry and Interactions of Seed Galactomannans. in *Advances in Carbohydrate Chemistry and Biochemistry* (Tipson, R. S., and Horton, D. eds.), Academic Press. pp 241-312
63. Barak, S., and Mudgil, D. (2014) Locust bean gum: processing, properties and food applications--a review. *Int J Biol Macromol* **66**, 74-80

64. Mudgil, D., Barak, S., and Khatkar, B. S. (2014) Guar gum: processing, properties and food applications-A Review. *J Food Sci Technol* **51**, 409-418
65. Thombare, N., Jha, U., Mishra, S., and Siddiqui, M. Z. (2016) Guar gum as a promising starting material for diverse applications: A review. *International Journal of Biological Macromolecules* **88**, 361-372
66. Timell, T. E. (1967) Recent progress in the chemistry of wood hemicelluloses. *Wood Science and Technology* **1**, 45-70
67. Lundqvist, J., Jacobs, A., Palm, M., Zacchi, G., Dahlman, O., and Stalbrand, H. (2003) Characterization of galactoglucomannan extracted from spruce (Picea abies) by heat-fractionation at different conditions. *Carbohyd Polym* **51**, 203-211
68. Dheilly, E., Gall, S. L., Guillou, M. C., Renou, J. P., Bonnin, E., Orsel, M., and Lahaye, M. (2016) Cell wall dynamics during apple development and storage involves hemicellulose modifications and related expressed genes. *BMC Plant Biol* **16**, 201
69. Schroder, R., Nicolas, P., Vincent, S. J. F., Fischer, M., Reymond, S., and Redgwell, R. J. (2001) Purification and characterisation of a galactoglucomannan from kiwifruit (Actinidia deliciosa). *Carbohyd Res* **331**, 291-306
70. Elgueta, E., Sanchez, J., Dax, D., Xu, C. L., Willfor, S., Rivas, B. L., and Gonzalez, M. (2016) Functionalized galactoglucomannan-based hydrogels for the removal of metal cations from aqueous solutions. *J Appl Polym Sci* **133**
71. Oinonen, P., Krawczyk, H., Ek, M., Henriksson, G., and Moriana, R. (2016) Bioinspired composites from cross-linked galactoglucomannan and microfibrillated cellulose: Thermal, mechanical and oxygen barrier properties. *Carbohyd Polym* **136**, 146-153
72. Willfor, S., Sundberg, K., Tenkanen, M., and Holmbom, B. (2008) Spruce-derived mannans - A potential raw material for hydrocolloids and novel advanced natural materials. *Carbohyd Polym* **72**, 197-210
73. Polari, L., Ojansivu, P., Makela, S., Eckerman, C., Holmbom, B., and Salminen, S. (2012) Galactoglucomannan Extracted from Spruce (Picea abies) as a Carbohydrate Source for Probiotic Bacteria. *J Agr Food Chem* **60**, 11037-11043
74. Sender, R., Fuchs, S., and Milo, R. (2016) Revised Estimates for the Number of Human and Bacteria Cells in the Body. *PLoS biology* **14**
75. D'Argenio, V., and Salvatore, F. (2015) The role of the gut microbiome in the healthy adult status. *Clin Chim Acta* **451**, 97-102
76. Lloyd-Price, J., Abu-Ali, G., and Huttenhower, C. (2016) The healthy human microbiome. *Genome Med* **8**, 51
77. Mandar, R., and Mikelsaar, M. (1996) Transmission of mother's microflora to the newborn at birth. *Biol Neonate* **69**, 30-35
78. Koropatkin, N. M., Cameron, E. A., and Martens, E. C. (2012) How glycan metabolism shapes the human gut microbiota. *Nat Rev Microbiol* **10**, 323-335

79. Rothschild, D., Weissbrod, O., Barkan, E., Kurilshikov, A., Korem, T., Zeevi, D., Costea, P. I., Godneva, A., Kalka, I. N., Bar, N., Shilo, S., Lador, D., Vila, A. V., Zmora, N., Pevsner-Fischer, M., Israeli, D., Kosower, N., Malka, G., Wolf, B. C., Avnit-Sagi, T., Lotan-Pompan, M., Weinberger, A., Halpern, Z., Carmi, S., Fu, J., Wijmenga, C., Zhernakova, A., Elinav, E., and Segal, E. (2018) Environment dominates over host genetics in shaping human gut microbiota. *Nature* **555**, 210-215

80. Koropatkin, N. M., Martens, E. C., Gordon, J. I., and Smith, T. J. (2008) Starch catabolism by a prominent human gut symbiont is directed by the recognition of amylose helices. *Structure* **16**, 1105-1115

81. Backhed, F., Ley, R. E., Sonnenburg, J. L., Peterson, D. A., and Gordon, J. I. (2005) Host-bacterial mutualism in the human intestine. *Science* **307**, 1915-1920

82. Salyers, A. A., Vercellotti, J. R., West, S. E., and Wilkins, T. D. (1977) Fermentation of mucin and plant polysaccharides by strains of Bacteroides from the human colon. *Applied and environmental microbiology* **33**, 319-322

83. Hooper, L. V., Midtvedt, T., and Gordon, J. I. (2002) How host-microbial interactions shape the nutrient environment of the mammalian intestine. *Annu Rev Nutr* **22**, 283-307

84. Corfield, A. P. (2015) Mucins: a biologically relevant glycan barrier in mucosal protection. *Biochim Biophys Acta* **1850**, 236-252

85. Bjursell, M. K., Martens, E. C., and Gordon, J. I. (2006) Functional genomic and metabolic studies of the adaptations of a prominent adult human gut symbiont, Bacteroides thetaiotaomicron, to the suckling period. *The Journal of biological chemistry* **281**, 36269-36279

86. Mahowald, M. A., Rey, F. E., Seedorf, H., Turnbaugh, P. J., Fulton, R. S., Wollam, A., Shah, N., Wang, C., Magrini, V., Wilson, R. K., Cantarel, B. L., Coutinho, P. M., Henrissat, B., Crock, L. W., Russell, A., Verberkmoes, N. C., Hettich, R. L., and Gordon, J. I. (2009) Characterizing a model human gut microbiota composed of members of its two dominant bacterial phyla. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 5859-5864

87. McNulty, N. P., Wu, M., Erickson, A. R., Pan, C., Erickson, B. K., Martens, E. C., Pudlo, N. A., Muegge, B. D., Henrissat, B., Hettich, R. L., and Gordon, J. I. (2013) Effects of diet on resource utilization by a model human gut microbiota containing Bacteroides cellulosilyticus WH2, a symbiont with an extensive glycobiome. *PLoS biology* **11**, e1001637

88. Cuskin, F., Lowe, E. C., Temple, M. J., Zhu, Y., Cameron, E. A., Pudlo, N. A., Porter, N. T., Urs, K., Thompson, A. J., Cartmell, A., Rogowski, A., Hamilton, B. S., Chen, R., Tolbert, T. J., Piens, K., Bracke, D., Vervecken, W., Hakki, Z., Speciale, G., Munoz-Munoz, J. L., Day, A., Pena, M. J., McLean, R., Suits, M. D., Boraston, A. B., Atherly, T., Ziemer, C. J., Williams, S. J., Davies, G. J., Abbott, D. W., Martens, E. C., and Gilbert, H. J. (2015) Human gut Bacteroidetes can utilize yeast mannan through a selfish mechanism. *Nature* **517**, 165-169

89. Egan, M., Motherway, M. O., Kilcoyne, M., Kane, M., Joshi, L., Ventura, M., and van Sinderen, D. (2014) Cross-feeding by Bifidobacterium breve UCC2003 during co-cultivation with Bifidobacterium bifidum PRL2010 in a mucin-based medium. *BMC Microbiol* **14**, 282

90. Rakoff-Nahoum, S., Foster, K. R., and Comstock, L. E. (2016) The evolution of cooperation within the gut microbiota. *Nature* **533**, 255-259
91. Appanna, V. D. (2018) Dysbiosis, Probiotics, and Prebiotics: In Diseases and Health. in *Human Microbes - The Power Within*, 1 Ed., Springer Singapore. pp 81-119
92. Adam, R. D. (2001) Biology of Giardia lamblia. *Clin Microbiol Rev* **14**, 447-475
93. Lawley, T. D., and Walker, A. W. (2013) Intestinal colonization resistance. *Immunology* **138**, 1-11
94. Sharma, R., Schumacher, U., Ronaasen, V., and Coates, M. (1995) Rat intestinal mucosal responses to a microbial flora and different diets. *Gut* **36**, 209-214
95. Reinhardt, C., Bergentall, M., Greiner, T. U., Schaffner, F., Ostergren-Lunden, G., Petersen, L. C., Ruf, W., and Backhed, F. (2012) Tissue factor and PAR1 promote microbiota-induced intestinal vascular remodelling. *Nature* **483**, 627-631
96. Petersen, C., and Round, J. L. (2014) Defining dysbiosis and its influence on host immunity and disease. *Cell Microbiol* **16**, 1024-1033
97. Kim, S., Goel, R., Kumar, A., Qi, Y., Lobaton, G., Hosaka, K., Mohammed, M., Handberg, E. M., Richards, E. M., Pepine, C. J., and Raizada, M. K. (2018) Imbalance of gut microbiome and intestinal epithelial barrier dysfunction in patients with high blood pressure. *Clin Sci (Lond)*
98. Kasselman, L. J., Vernice, N. A., DeLeon, J., and Reiss, A. B. (2018) The gut microbiome and elevated cardiovascular risk in obesity and autoimmunity. *Atherosclerosis* **271**, 203-213
99. Garrett, W. S. (2015) Cancer and the microbiota. *Science* **348**, 80-86
100. O'Keefe, S. J., Ou, J., Aufreiter, S., O'Connor, D., Sharma, S., Sepulveda, J., Fukuwatari, T., Shibata, K., and Mawhinney, T. (2009) Products of the colonic microbiota mediate the effects of diet on colon cancer risk. *J Nutr* **139**, 2044-2048
101. Markowiak, P., and Slizewska, K. (2017) Effects of Probiotics, Prebiotics, and Synbiotics on Human Health. *Nutrients* **9**
102. FAO. (30 April-1 May 2002) Food and Agriculture Organization (FAO). *Guidelines for the Evaluation of Probiotics in Food*; Report of a Joint FAO/WHO Working Group on Drafting Guidelines for the Evaluation of Probiotics in Food. FAO, London, ON, Canada
103. Kailasapathy, K., and Chin, J. (2000) Survival and therapeutic potential of probiotic organisms with reference to Lactobacillus acidophilus and Bifidobacterium spp. *Immunol Cell Biol* **78**, 80-88
104. Kimoto-Nira, H. (2018) New lactic acid bacteria for skin health via oral intake of heat-killed or live cells. *Anim Sci J*
105. Vandeputte, D., Falony, G., Vieira-Silva, S., Wang, J., Sailer, M., Theis, S., Verbeke, K., and Raes, J. (2017) Prebiotic inulin-type fructans induce specific changes in the human gut microbiota. *Gut* **66**, 1968-1974
106. Cheng, W., Lu, J., Lin, W., Wei, X., Li, H., Zhao, X., Jiang, A., and Yuan, J. (2018) Effects of a galacto-oligosaccharide-rich diet on fecal microbiota and metabolite profiles in mice. *Food Funct* **9**, 1612-1620

107. Zhang, Y., Wang, Y., Zheng, B., Lu, X., and Zhuang, W. (2013) The in vitro effects of retrograded starch (resistant starch type 3) from lotus seed starch on the proliferation of Bifidobacterium adolescentis. *Food Funct* **4**, 1609-1616

108. Brufau, M. T., Campo-Sabariz, J., Carne, S., Ferrer, R., and Martin-Venegas, R. (2017) Salmosan, a beta-galactomannan-rich product, in combination with Lactobacillus plantarum contributes to restore intestinal epithelial barrier function by modulation of cytokine production. *J Nutr Biochem* **41**, 20-24

109. Arumugam, M., Raes, J., Pelletier, E., Le Paslier, D., Yamada, T., Mende, D. R., Fernandes, G. R., Tap, J., Bruls, T., Batto, J. M., Bertalan, M., Borruel, N., Casellas, F., Fernandez, L., Gautier, L., Hansen, T., Hattori, M., Hayashi, T., Kleerebezem, M., Kurokawa, K., Leclerc, M., Levenez, F., Manichanh, C., Nielsen, H. B., Nielsen, T., Pons, N., Poulain, J., Qin, J., Sicheritz-Ponten, T., Tims, S., Torrents, D., Ugarte, E., Zoetendal, E. G., Wang, J., Guarner, F., Pedersen, O., de Vos, W. M., Brunak, S., Dore, J., Meta, H. I. T. C., Antolin, M., Artiguenave, F., Blottiere, H. M., Almeida, M., Brechot, C., Cara, C., Chervaux, C., Cultrone, A., Delorme, C., Denariaz, G., Dervyn, R., Foerstner, K. U., Friss, C., van de Guchte, M., Guedon, E., Haimet, F., Huber, W., van Hylckama-Vlieg, J., Jamet, A., Juste, C., Kaci, G., Knol, J., Lakhdari, O., Layec, S., Le Roux, K., Maguin, E., Merieux, A., Melo Minardi, R., M'Rini, C., Muller, J., Oozeer, R., Parkhill, J., Renault, P., Rescigno, M., Sanchez, N., Sunagawa, S., Torrejon, A., Turner, K., Vandemeulebrouck, G., Varela, E., Winogradsky, Y., Zeller, G., Weissenbach, J., Ehrlich, S. D., and Bork, P. (2011) Enterotypes of the human gut microbiome. *Nature* **473**, 174-180

110. Ndeh, D., and Gilbert, H. J. (2018) Biochemistry of complex glycan depolymerisation by the human gut microbiota. *FEMS Microbiol Rev* **42**, 146-164

111. Lloyd-Price, J., Mahurkar, A., Rahnavard, G., Crabtree, J., Orvis, J., Hall, A. B., Brady, A., Creasy, H. H., McCracken, C., Giglio, M. G., McDonald, D., Franzosa, E. A., Knight, R., White, O., and Huttenhower, C. (2017) Strains, functions and dynamics in the expanded Human Microbiome Project. *Nature* **550**, 61-66

112. Martens, E. C., Chiang, H. C., and Gordon, J. I. (2008) Mucosal glycan foraging enhances fitness and transmission of a saccharolytic human gut bacterial symbiont. *Cell Host Microbe* **4**, 447-457

113. Martens, E. C., Lowe, E. C., Chiang, H., Pudlo, N. A., Wu, M., McNulty, N. P., Abbott, D. W., Henrissat, B., Gilbert, H. J., Bolam, D. N., and Gordon, J. I. (2011) Recognition and degradation of plant cell wall polysaccharides by two human gut symbionts. *PLoS biology* **9**, e1001221

114. Xu, J., Bjursell, M. K., Himrod, J., Deng, S., Carmichael, L. K., Chiang, H. C., Hooper, L. V., and Gordon, J. I. (2003) A genomic view of the human-Bacteroides thetaiotaomicron symbiosis. *Science* **299**, 2074-2076

115. Martens, E. C., Koropatkin, N. M., Smith, T. J., and Gordon, J. I. (2009) Complex glycan catabolism by the human gut microbiota: the Bacteroidetes Sus-like paradigm. *The Journal of biological chemistry* **284**, 24673-24677

116. Grondin, J. M., Tamura, K., Dejean, G., Abbott, D. W., and Brumer, H. (2017) Polysaccharide Utilization Loci: Fueling Microbial Communities. *Journal of bacteriology* **199**

117. Reeves, A. R., DElia, J. N., Frias, J., and Salyers, A. A. (1996) A Bacteroides thetaiotaomicron outer membrane protein that is essential for utilization of maltooligosaccharides and starch. *Journal of bacteriology* **178**, 823-830

118. Reeves, A. R., Wang, G. R., and Salyers, A. A. (1997) Characterization of four outer membrane proteins that play a role in utilization of starch by Bacteroides thetaiotaomicron. *Journal of bacteriology* **179**, 643-649

119. Shipman, J. A., Berleman, J. E., and Salyers, A. A. (2000) Characterization of four outer membrane proteins involved in binding starch to the cell surface of Bacteroides thetaiotaomicron. *Journal of bacteriology* **182**, 5365-5372

120. Bolam, D. N., and Koropatkin, N. M. (2012) Glycan recognition by the Bacteroidetes Sus-like systems. *Curr Opin Struct Biol* **22**, 563-569

121. Terrapon, N., Lombard, V., Drula, E., Lapebie, P., Al-Masaudi, S., Gilbert, H. J., and Henrissat, B. (2018) PULDB: the expanded database of Polysaccharide Utilization Loci. *Nucleic Acids Res* **46**, D677-D683

122. Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., and Henrissat, B. (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res* **42**, D490-495

123. Sheridan, P. O., Martin, J. C., Lawley, T. D., Browne, H. P., Harris, H. M., Bernalier-Donadille, A., Duncan, S. H., O'Toole, P. W., Scott, K. P., and Flint, H. J. (2016) Polysaccharide utilization loci and nutritional specialization in a dominant group of butyrate-producing human colonic Firmicutes. *Microb Genom* **2**, e000043

124. Cartmell, A., Topakas, E., Ducros, V. M. A., Suits, M. D. L., Davies, G. J., and Gilbert, H. J. (2008) 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. *Journal of Biological Chemistry* **283**, 34403-34413

125. Couturier, M., Roussel, A., Rosengren, A., Leone, P., Stalbrand, H., and Berrin, J. G. (2013) Structural and Biochemical Analyses of Glycoside Hydrolase Families 5 and 26 beta-(1,4)-Mannanases from Podospora anserina Reveal Differences upon Manno-oligosaccharide Catalysis. *Journal of Biological Chemistry* **288**, 14624-14635

126. Larsbrink, J., Rogers, T. E., Hemsworth, G. R., McKee, L. S., Tauzin, A. S., Spadiut, O., Klinter, S., Pudlo, N. A., Urs, K., Koropatkin, N. M., Creagh, A. L., Haynes, C. A., Kelly, A. G., Cederholm, S. N., Davies, G. J., Martens, E. C., and Brumer, H. (2014) A discrete genetic locus confers xyloglucan metabolism in select human gut Bacteroidetes. *Nature* **506**, 498-502

127. Whitehead, T. R., and Hespell, R. B. (1990) The genes for three xylan-degrading activities from Bacteroides ovatus are clustered in a 3.8-kilobase region. *Journal of bacteriology* **172**, 2408-2412

128. Weaver, J., Whitehead, T. R., Cotta, M. A., Valentine, P. C., and Salyers, A. A. (1992) Genetic analysis of a locus on the Bacteroides ovatus chromosome which contains xylan utilization genes. *Applied and environmental microbiology* **58**, 2764-2770

129. Rogowski, A., Briggs, J. A., Mortimer, J. C., Tryfona, T., Terrapon, N., Lowe, E. C., Basle, A., Morland, C., Day, A. M., Zheng, H., Rogers, T. E., Thompson, P., Hawkins, A. R., Yadav, M. P., Henrissat, B., Martens, E. C., Dupree, P., Gilbert, H. J., and Bolam, D. N. (2015) Glycan complexity dictates microbial resource allocation in the large intestine. *Nat Commun* **6**, 7481
130. Temple, M. J., Cuskin, F., Basle, A., Hickey, N., Speciale, G., Williams, S. J., Gilbert, H. J., and Lowe, E. C. (2017) A Bacteroidetes locus dedicated to fungal 1,6-beta-glucan degradation: Unique substrate conformation drives specificity of the key endo-1,6-beta-glucanase. *The Journal of biological chemistry* **292**, 10639-10650
131. Gherardini, F. C., and Salyers, A. A. (1987) Purification and characterization of a cell-associated, soluble mannanase from Bacteroides ovatus. *Journal of bacteriology* **169**, 2038-2043
132. Bolam, D. N., and van den Berg, B. (2018) TonB-dependent transport by the gut microbiota: novel aspects of an old problem. *Curr Opin Struct Biol* **51**, 35-43
133. Glenwright, A. J., Pothula, K. R., Bhamidimarri, S. P., Chorev, D. S., Basle, A., Firbank, S. J., Zheng, H., Robinson, C. V., Winterhalter, M., Kleinekathofer, U., Bolam, D. N., and van den Berg, B. (2017) Structural basis for nutrient acquisition by dominant members of the human gut microbiota. *Nature* **541**, 407-411
134. D'Elia, J. N., and Salyers, A. A. (1996) Effect of regulatory protein levels on utilization of starch by Bacteroides thetaiotaomicron. *Journal of bacteriology* **178**, 7180-7186
135. Hong, P. Y., Iakiviak, M., Dodd, D., Zhang, M., Mackie, R. I., and Cann, I. (2014) Two new xylanases with different substrate specificities from the human gut bacterium Bacteroides intestinalis DSM 17393. *Applied and environmental microbiology* **80**, 2084-2093
136. Abbott, D. W., Martens, E. C., Gilbert, H. J., Cuskin, F., and Lowe, E. C. (2015) Coevolution of yeast mannan digestion: Convergence of the civilized human diet, distal gut microbiome, and host immunity. *Gut Microbes* **6**, 334-339
137. Sinnott, M. L. (1990) Catalytic Mechanisms of Enzymatic Glycosyl Transfer. *Chem Rev* **90**, 1171-1202
138. Davies, G., and Henrissat, B. (1995) Structures and mechanisms of glycosyl hydrolases. *Structure* **3**, 853-859
139. Jamet, E., Canut, H., Boudart, G., and Pont-Lezica, R. F. (2006) Cell wall proteins: a new insight through proteomics. *Trends Plant Sci* **11**, 33-39
140. Zeeman, S. C., Kossmann, J., and Smith, A. M. (2010) Starch: its metabolism, evolution, and biotechnological modification in plants. *Annu Rev Plant Biol* **61**, 209-234
141. Ferreira, C. R., and Gahl, W. A. (2017) Lysosomal storage diseases. *Transl Sci Rare Dis* **2**, 1-71
142. Cantarel, B. L., Lombard, V., and Henrissat, B. (2012) Complex carbohydrate utilization by the healthy human microbiome. *PLoS One* **7**, e28742

143. Uchino, Y., Fukushige, T., Yotsumoto, S., Hashiguchi, T., Taguchi, H., Suzuki, N., Konohana, I., and Kanzaki, T. (2003) Morphological and biochemical studies of human beta-mannosidosis: identification of a novel beta-mannosidase gene mutation. *Br J Dermatol* **149**, 23-29

144. Henrissat, B., Callebaut, I., Fabrega, S., Lehn, P., Mornon, J. P., 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 United States of America* **92**, 7090-7094

145. Henrissat, B., and Bairoch, A. (1996) Updating the sequence-based classification of glycosyl hydrolases. *The Biochemical journal* **316 (Pt 2)**, 695-696

146. Jenkins, J., Leggio, L. L., Harris, G., and Pickersgill, R. (1995) Beta-Glucosidase, Beta-Galactosidase, Family-a Cellulases, Family-F Xylanases and 2 Barley Glycanases Form a Superfamily of Enzymes with 8-Fold Beta/Alpha-Architecture and with 2 Conserved Glutamates near the Carboxy-Terminal Ends of Beta-Strand-4 and Beta-Strand-7. *Febs Lett* **362**, 281-285

147. Aspeborg, H., Coutinho, P. M., Wang, Y., Brumer, H., 3rd, and Henrissat, B. (2012) Evolution, substrate specificity and subfamily classification of glycoside hydrolase family 5 (GH5). *BMC Evol Biol* **12**, 186

148. Henrissat, B., Teeri, T. T., and Warren, R. A. (1998) A scheme for designating enzymes that hydrolyse the polysaccharides in the cell walls of plants. *Febs Lett* **425**, 352-354

149. Fredslund, F., Hachem, M. A., Larsen, R. J., Sorensen, P. G., Coutinho, P. M., Lo Leggio, L., and Svensson, B. (2011) Crystal structure of alpha-galactosidase from Lactobacillus acidophilus NCFM: insight into tetramer formation and substrate binding. *Journal of molecular biology* **412**, 466-480

150. Breyer, W. A., and Matthews, B. W. (2001) A structural basis for processivity. *Protein Sci* **10**, 1699-1711

151. Saburi, W., Rachi-Otsuka, H., Hondoh, H., Okuyama, M., Mori, H., and Kimura, A. (2015) Structural elements responsible for the glucosidic linkage-selectivity of a glycoside hydrolase family 13 exo-glucosidase. *Febs Lett* **589**, 865-869

152. Chen, Y. C., Chen, W. T., Liu, J. C., Tsai, L. C., and Cheng, H. L. (2014) A highly active beta-glucanase from a new strain of rumen fungus Orpinomyces sp.Y102 exhibits cellobiohydrolase and cellotriohydrolase activities. *Bioresour Technol* **170**, 513-521

153. 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. *The Journal of biological chemistry* **276**, 31186-31192

154. Honda, Y., Arai, S., Suzuki, K., Kitaoka, M., and Fushinobu, S. (2016) The crystal structure of an inverting glycoside hydrolase family 9 exo-beta-D-glucosaminidase and the design of glycosynthase. *The Biochemical journal* **473**, 463-472

155. Robyt, J. F., and French, D. (1967) Multiple attach hypothesis of alpha-amylase action: action of porcine pancreatic, human salivary, and Aspergillus oryzae alpha-amylases. *Arch Biochem Biophys* **122**, 8-16

156. Kurasin, M., and Valjamae, P. (2011) Processivity of cellobiohydrolases is limited by the substrate. *The Journal of biological chemistry* **286**, 169-177

157. Beckham, G. T., Stahlberg, J., Knott, B. C., Himmel, M. E., Crowley, M. F., Sandgren, M., Sorlie, M., and Payne, C. M. (2014) Towards a molecular-level theory of carbohydrate processivity in glycoside hydrolases. *Curr Opin Biotechnol* **27**, 96-106

158. Tiels, P., Baranova, E., Piens, K., De Visscher, C., Pynaert, G., Nerinckx, W., Stout, J., Fudalej, F., Hulpiau, P., Tannler, S., Geysens, S., Van Hecke, A., Valevska, A., Vervecken, W., Remaut, H., and Callewaert, N. (2012) A bacterial glycosidase enables mannose-6-phosphate modification and improved cellular uptake of yeast-produced recombinant human lysosomal enzymes. *Nat Biotechnol* **30**, 1225-1231

159. van Aalten, D. M., Komander, D., Synstad, B., Gaseidnes, S., Peter, M. G., and Eijsink, V. G. (2001) Structural insights into the catalytic mechanism of a family 18 exo-chitinase. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 8979-8984

160. Terwisscha van Scheltinga, A. C., Armand, S., Kalk, K. H., Isogai, A., Henrissat, B., and Dijkstra, B. W. (1995) Stereochemistry of chitin hydrolysis by a plant chitinase/lysozyme and X-ray structure of a complex with allosamidin: evidence for substrate assisted catalysis. *Biochemistry* **34**, 15619-15623

161. Brameld, K. A., Shrader, W. D., Imperiali, B., and Goddard, W. A. (1998) Substrate assistance in the mechanism of family 18 chitinases: Theoretical studies of potential intermediates and inhibitors. *Journal of molecular biology* **280**, 913-923

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

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

164. Burmeister, W. P., Cottaz, S., Rollin, P., Vasella, A., and Henrissat, B. (2000) High resolution x-ray crystallography shows that ascorbate is a cofactor for myrosinase and substitutes for the function of the catalytic base. *Journal of Biological Chemistry* **275**, 39385-39393

165. Gloster, T. M., and Vocadlo, D. J. (2010) Mechanism, Structure, and Inhibition of O-GlcNAc Processing Enzymes. *Curr Signal Transduct Ther* **5**, 74-91

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

167. Ndeh, D., Rogowski, A., Cartmell, A., Luis, A. S., Basle, A., Gray, J., Venditto, I., Briggs, J., Zhang, X., Labourel, A., Terrapon, N., Buffetto, F., Nepogodiev, S., Xiao, Y., Field, R. A., Zhu, Y., O'Neil, M. A., Urbanowicz, B. R., York, W. S., Davies, G. J., Abbott, D. W., Ralet, M. C., Martens, E. C., Henrissat, B., and Gilbert, H. J. (2017) Complex pectin metabolism by gut bacteria reveals novel catalytic functions. *Nature* **544**, 65-70

168. Koshland, D. E. (1953) Stereochemistry and the Mechanism of Enzymatic Reactions. *Biol Rev* **28**, 416-436

169. Zechel, D. L., and Withers, S. G. (2001) Dissection of nucleophilic and acid-base catalysis in glycosidases. *Curr Opin Chem Biol* **5**, 643-649

170. Mangas-Sanchez, J., and Adlercreutz, P. (2015) Enzymatic preparation of oligosaccharides by transglycosylation: A comparative study of glucosidases. *J Mol Catal B-Enzym* **122**, 51-55

171. Rosengren, A., Hagglund, P., Anderson, L., Pavon-Orozco, P., Peterson-Wulff, R., Nerinckx, W., and Stalbrand, H. (2012) The role of subsite +2 of the Trichoderma reesei beta-mannanase TrMan5A in hydrolysis and transglycosylation. *Biocatal Biotransfor* **30**, 338-352

172. Nakai, H., Baumann, M. J., Petersen, B. O., Westphal, Y., Hachem, M. A., Dilokpimol, A., Duus, J. O., Schols, H. A., and Svensson, B. (2010) Aspergillus nidulans alpha-galactosidase of glycoside hydrolase family 36 catalyses the formation of alpha-galacto-oligosaccharides by transglycosylation. *Febs J* **277**, 3538-3551

173. Morrill, J., Rosengren, A., Butler, S., and Stalbrand, H. (2017) Biocatalytic functionalization of softwood galactoglucomannan through transglycosylation and enzyme engineering. *Abstr Pap Am Chem S* **253**

174. Davies, G. J., Wilson, K. S., and Henrissat, B. (1997) Nomenclature for sugar-binding subsites in glycosyl hydrolases. *The Biochemical journal* **321 (Pt 2)**, 557-559

175. 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 beta-Mannanase from a Symbiotic Protist of the Termite Reticulitermes speratus. *Journal of Biological Chemistry* **289**, 10843-10852

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

177. Taylor, E. J., Goyal, A., Guerreiro, C. I. P. D., Prates, J. A. M., Money, V. A., Ferry, N., Morland, C., Planas, A., Macdonald, J. A., Stick, R. V., Gilbert, H. J., Fontes, C. M. G. A., and Davies, G. J. (2005) How family 26 glycoside hydrolases orchestrate catalysis on different polysaccharides - Structure and activity of a clostridium thermocellum lichenase, CtLic26A. *Journal of Biological Chemistry* **280**, 32761-32767

178. Ye, Y., Saburi, W., Odaka, R., Kato, K., Sakurai, N., Komoda, K., Nishimoto, M., Kitaoka, M., Mori, H., and Yao, M. (2016) Structural insights into the difference in substrate recognition of two mannoside phosphorylases from two GH130 subfamilies. *Febs Lett* **590**, 828-837
179. Kumagai, Y., Yamashita, K., Tagami, T., Uraji, M., Wan, K., Okuyama, M., Yao, M., Kimura, A., and Hatanaka, T. (2015) The loop structure of Actinomycete glycoside hydrolase family 5 mannanases governs substrate recognition. *Febs J* **282**, 4001-4014
180. Stoll, D., Boraston, A., Stalbrand, H., McLean, B. W., Kilburn, D. G., and Warren, R. A. J. (2000) Mannanase Man26A from Cellulomonas fimi has a mannan-binding module. *FEMS Microbiology Letters* **183**
181. Kulcinskaja, E., Rosengren, A., Ibrahim, R., Kolenova, K., and Stalbrand, H. (2013) Expression and characterization of a Bifidobacterium adolescentis beta-mannanase carrying mannan-binding and cell association motifs. *Applied and environmental microbiology* **79**, 133-140
182. Nakamura, A., Tasaki, T., Ishiwata, D., Yamamoto, M., Okuni, Y., Visootsat, A., Maximilien, M., Noji, H., Uchiyama, T., Samejima, M., Igarashi, K., and Iino, R. (2016) Single-molecule Imaging Analysis of Binding, Processive Movement, and Dissociation of Cellobiohydrolase Trichoderma reesei Cel6A and Its Domains on Crystalline Cellulose. *The Journal of biological chemistry* **291**, 22404-22413
183. Boraston, A. B., Bolam, D. N., Gilbert, H. J., and Davies, G. J. (2004) Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *The Biochemical journal* **382**, 769-781
184. Zhang, M., Chekan, J. R., Dodd, D., Hong, P. Y., Radlinski, L., Revindran, V., Nair, S. K., Mackie, R. I., and Cann, I. (2014) Xylan utilization in human gut commensal bacteria is orchestrated by unique modular organization of polysaccharide-degrading enzymes. *Proceedings of the National Academy of Sciences of the United States of America* **111**, E3708-3717
185. Ye, L., Su, X., Schmitz, G. E., Moon, Y. H., Zhang, J., Mackie, R. I., and Cann, I. K. (2012) Molecular and biochemical analyses of the GH44 module of CbMan5B/Cel44A, a bifunctional enzyme from the hyperthermophilic bacterium Caldicellulosiruptor bescii. *Applied and environmental microbiology* **78**, 7048-7059
186. Zheng, F., and Ding, S. (2013) Processivity and enzymatic mode of a glycoside hydrolase family 5 endoglucanase from Volvariella volvacea. *Applied and environmental microbiology* **79**, 989-996
187. Hagglund, P., Eriksson, T., Collen, A., Nerinckx, W., Claeyssens, M., and Stalbrand, H. (2003) A cellulose-binding module of the Trichoderma reesei beta-mannanase Man5A increases the mannan-hydrolysis of complex substrates. *J Biotechnol* **101**, 37-48
188. Hogg, D., Pell, G., Dupree, P., Goubet, F., Martin-Orue, 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. *The Biochemical journal* **371**, 1027-1043

189. Gilbert, H. J., Stalbrand, H., and Brumer, H. (2008) How the walls come crumbling down: recent structural biochemistry of plant polysaccharide degradation. *Curr Opin Plant Biol* **11**, 338-348

190. Cairns, J. R. K., and Esen, A. (2010) beta-Glucosidases. *Cell Mol Life Sci* **67**, 3389-3405

191. Mai-Gisondi, G., Maaheimo, H., Chong, S. L., Hinz, S., Tenkanen, M., and Master, E. (2017) Functional comparison of versatile carbohydrate esterases from families CE1, CE6 and CE16 on acetyl-4-O-methylglucuronoxylan and acetyl-galactoglucomannan. *Bba-Gen Subjects* **1861**, 2398-2405

192. Acosta, P. B., and Gross, K. C. (1995) Hidden sources of galactose in the environment. *Eur J Pediatr* **154**, S87-92

193. Mccleary, B. V. (1988) Beta-D-Mannanase. *Method Enzymol* **160**, 596-610

194. Tsukagoshi, H., Nakamura, A., Ishida, T., Otagiri, M., Moriya, S., Samejima, M., Igarashi, K., Kitamoto, K., and Arioka, M. (2014) The GH26 beta-mannanase RsMan26H from a symbiotic protist of the termite Reticulitermes speratus is an endo-processive mannobiohydrolase: heterologous expression and characterization. *Biochem Biophys Res Commun* **452**, 520-525

195. Kawaguchi, K., Senoura, T., Ito, S., Taira, T., Ito, H., Wasaki, J., and Ito, S. (2014) The mannobiose-forming exo-mannanase involved in a new mannan catabolic pathway in Bacteroides fragilis. *Archives of microbiology* **196**, 17-23

196. von Freiesleben, P., Spodsberg, N., Blicher, T. H., Anderson, L., Jorgensen, H., Stalbrand, H., Meyer, A. S., and Krogh, K. B. (2016) An Aspergillus nidulans GH26 endo-beta-mannanase with a novel degradation pattern on highly substituted galactomannans. *Enzyme Microb Technol* **83**, 68-77

197. Stoll, D., Stalbrand, H., and Warren, R. A. (1999) Mannan-degrading enzymes from Cellulomonas fimi. *Applied and environmental microbiology* **65**, 2598-2605

198. 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 AaManA of Alicyclobacillus acidocaldarius reveals a novel glycoside hydrolase family belonging to clan GH-A. *The Journal of biological chemistry* **283**, 31551-31558

199. Jin, Y., Petricevic, M., John, A., Raich, L., Jenkins, H., Portela De Souza, L., Cuskin, F., Gilbert, H. J., Rovira, C., Goddard-Borger, E. D., Williams, S. J., and Davies, G. J. (2016) A beta-Mannanase with a Lysozyme-like Fold and a Novel Molecular Catalytic Mechanism. *ACS Cent Sci* **2**, 896-903

200. Durand, P., Lehn, P., Callebaut, I., Fabrega, S., Henrissat, B., and Mornon, J. P. (1997) Active-site motifs of lysosomal acid hydrolases: Invariant features of clan GH-A glycosyl hydrolases deduced from hydrophobic cluster analysis. *Glycobiology* **7**, 277-284

201. Gloster, T. M., Ibatullin, F. M., Macauley, K., Eklof, J. M., Roberts, S., Turkenburg, J. P., Bjornvad, M. E., Jorgensen, P. L., Danielsen, S., Johansen, K. S., Borchert, T. V., Wilson, K. S., Brumer, H., and Davies, G. J. (2007) Characterization and three-dimensional structures of two distinct bacterial xyloglucanases from families GH5 and GH12. *Journal of Biological Chemistry* **282**, 19177-19189

202. Cote, N., Fleury, A., Dumont-Blanchette, E., Fukamizo, T., Mitsutomi, M., and Brzezinski, R. (2006) Two exo-beta-D-glucosaminidases/exochitosanases from actinomycetes define a new subfamily within family 2 of glycoside hydrolases. *Biochemical Journal* **394**, 675-686

203. Ishii, J., Okazaki, F., Djohan, A. C., Hara, K. Y., Asai-Nakashima, N., Teramura, H., Andriani, A., Tominaga, M., Wakai, S., Kahar, P., Yopi, Prasetya, B., Ogino, C., and Kondo, A. (2016) From mannan to bioethanol: cell surface co-display of beta-mannanase and beta-mannosidase on yeast Saccharomyces cerevisiae. *Biotechnol Biofuels* **9**, 188

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

205. David, A., Singh Chauhan, P., Kumar, A., Angural, S., Kumar, D., Puri, N., and Gupta, N. (2018) Coproduction of protease and mannanase from Bacillus nealsonii PN-11 in solid state fermentation and their combined application as detergent additives. *Int J Biol Macromol* **108**, 1176-1184

206. Larsson, A. M., Anderson, L., Xu, B., Munoz, I. G., Uson, I., Janson, J. C., Stalbrand, H., and Stahlberg, J. (2006) Three-dimensional crystal structure and enzymic characterization of beta-mannanase Man5A from blue mussel Mytilus edulis. *Journal of molecular biology* **357**, 1500-1510

207. Mizutani, K., Tsuchiya, S., Toyoda, M., Nanbu, Y., Tominaga, K., Yuasa, K., Takahashi, N., Tsuji, A., and Mikami, B. (2012) Structure of beta-1,4-mannanase from the common sea hare Aplysia kurodai at 1.05 angstrom resolution. *Acta Crystallogr F* **68**, 1164-1168

208. Morrill, J., Kulcinskaja, E., Sulewska, A. M., Lahtinen, S., Stalbrand, H., Svensson, B., and Abou Hachem, M. (2015) The GH5 1,4-beta-mannanase from Bifidobacterium animalis subsp. lactis Bl-04 possesses a low-affinity mannan-binding module and highlights the diversity of mannanolytic enzymes. *BMC Biochem* **16**, 26

209. Takasuka, T. E., Acheson, J. F., Bianchetti, C. M., Prom, B. M., Bergeman, L. F., Book, A. J., Currie, C. R., and Fox, B. G. (2014) Biochemical Properties and Atomic Resolution Structure of a Proteolytically Processed beta-Mannanase from Cellulolytic Streptomyces sp SirexAA- E. *Plos One* **9**

210. 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-beta-1,4-D-mannanase from the Antarctic springtail, Cryptopygus antarcticus. *Proteins* **82**, 3217-3223

211. 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 beta-mannanase from Thermomonospora fusca - substrate specificity in glycosyl hydrolase family 5. *Struct Fold Des* **6**, 1433-1444

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

213. 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 Crystallogr D* **60**, 1490-1492

214. Dias, F. M., Vincent, F., Pell, G., Prates, J. A., Centeno, M. S., Tailford, L. E., Ferreira, L. M., Fontes, C. M., 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**, 25517-25526

215. Zhou, P., Liu, Y., Yan, Q. J., Chen, Z. Z., Qin, Z., and Jiang, Z. Q. (2014) Structural insights into the substrate specificity and transglycosylation activity of a fungal glycoside hydrolase family 5 beta-mannosidase. *Acta Crystallogr D* **70**, 2970-2982

216. Sabini, E., Schubert, H., Murshudov, G., Wilson, K. S., Siika-Aho, M., and Penttila, M. (2000) The three-dimensional structure of a Trichoderma reesei beta-mannanase from glycoside hydrolase family 5. *Acta crystallographica. Section D, Biological crystallography* **56**, 3-13

217. Bourgault, R., Oakley, A. J., Bewley, J. D., and Wilce, M. C. J. (2005) Three-dimensional structure of (1,4)-beta-D-mannan mannanohydrolase from tomato fruit. *Protein Sci.* **14**, 1233-1241

218. Rosengren, A., Reddy, S. K., Sjoberg, J. S., Aurelius, O., Logan, D. T., Kolenova, K., and Stalbrand, H. (2014) An Aspergillus nidulans beta-mannanase with high transglycosylation capacity revealed through comparative studies within glycosidase family 5. *Appl Microbiol Biotechnol* **98**, 10091-10104

219. Mizutani, K., Fernandes, V. O., Karita, S., Luis, A. S., Sakka, M., Kimura, T., Jackson, A., Zhang, X., Fontes, C. M., Gilbert, H. J., and Sakka, K. (2012) Influence of a mannan binding family 32 carbohydrate binding module on the activity of the appended mannanase. *Applied and environmental microbiology* **78**, 4781-4787

220. Schroder, R., Wegrzyn, T. F., Sharma, N. N., and Atkinson, R. G. (2006) LeMAN4 endo-beta-mannanase from ripe tomato fruit can act as a mannan transglycosylase or hydrolase. *Planta* **224**, 1091-1102

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

222. Malgas, S., van Dyk, S. J., and Pletschke, B. I. (2015) beta-mannanase (Man26A) and alpha-galactosidase (Aga27A) synergism - a key factor for the hydrolysis of galactomannan substrates. *Enzyme Microb Technol* **70**, 1-8

223. Hsu, Y., Koizumi, H., Otagiri, M., Moriya, S., and Arioka, M. (2018) Trp residue at subsite - 5 plays a critical role in the substrate binding of two protistan GH26 beta-mannanases from a termite hindgut. *Appl Microbiol Biotechnol*

224. Katsimpouras, C., Dimarogona, M., Petropoulos, P., Christakopoulos, P., and Topakas, E. (2016) A thermostable GH26 endo-beta-mannanase from Myceliophthora thermophila capable of enhancing lignocellulose degradation. *Appl Microbiol Biot* **100**, 8385-8397

225. von Freiesleben, P., Spodsberg, N., Stenbæk, A., Stålbrand, H., Krogh, K. B. R. M., and Meyer, A. S. (2018) Boosting of enzymatic softwood saccharification by fungal GH5 and GH26 endomannanases. *Biotechnol Biofuels* **11**, 194

226. Yan, X. X., An, X. M., Gui, L. L., and Liang, D. C. (2008) From structure to function: insights into the catalytic substrate specificity and thermostability displayed by Bacillus subtilis mannanase BCman. *Journal of molecular biology* **379**, 535-544

227. Ducros, V. M., Zechel, D. L., Murshudov, G. N., Gilbert, H. J., Szabo, L., Stoll, D., Withers, S. G., and Davies, G. J. (2002) Substrate distortion by a beta-mannanase: snapshots of the Michaelis and covalent-intermediate complexes suggest a B(2,5) conformation for the transition state. *Angewandte Chemie* **41**, 2824-2827

228. Kumagai, Y., Usuki, H., Yamamoto, Y., Yamasato, A., Arima, J., Mukaihara, T., and Hatanaka, T. (2011) Characterization of calcium ion sensitive region for beta-mannanase from Streptomyces thermolilacinus. *Biochim Biophys Acta* **1814**, 1127-1133

229. Kumagai, Y., Kawakami, K., Mukaihara, T., Kimura, M., and Hatanaka, T. (2012) The structural analysis and the role of calcium binding site for thermal stability in mannanase. *Biochimie* **94**, 2783-2790

230. Srivastava, P. K., Appu Rao, G. A., and Kapoor, M. (2016) Metal-dependent thermal stability of recombinant endo-mannanase (ManB-1601) belonging to family GH 26 from Bacillus sp. CFR1601. *Enzyme Microb Technol* **84**, 41-49

231. Yin, L. J., Tai, H. M., and Jiang, S. T. (2012) Characterization of mannanase from a novel mannanase-producing bacterium. *J Agric Food Chem* **60**, 6425-6431

232. Nadaroglu, H., Adiguzel, G., Adiguzel, A., and Sonmez, Z. (2017) A thermostable-endo-beta-(1,4)-mannanase from Pediococcus acidilactici (M17): purification, characterization and its application in fruit juice clarification. *Eur Food Res Technol* **243**, 193-201

233. Abou Hachem, M., Fredslund, F., Andersen, J. M., Larsen, R. J., Majumder, A., Ejby, M., Van Zanten, G., Lahtinen, S. J., Barrangou, R., Klaenhammer, T., Jacobsen, S., Coutinho, P. M., Lo Leggio, L., and Svensson, B. (2012) Raffinose family oligosaccharide utilisation by probiotic bacteria: insight into substrate recognition, molecular architecture and diversity of GH36 alpha-galactosidases. *Biocatalysis and Biotransformation* **30**, 316-325

234. Thall, A., and Galili, U. (1990) Distribution of Gal alpha 1----3Gal beta 1----4GlcNAc residues on secreted mammalian glycoproteins (thyroglobulin, fibrinogen, and immunoglobulin G) as measured by a sensitive solid-phase radioimmunoassay. *Biochemistry* **29**, 3959-3965

235. Wherrett, J. R., and Hakomori, S. I. (1973) Characterization of a blood group B glycolipid, accumulating in the pancreas of a patient with Fabry's disease. *The Journal of biological chemistry* **248**, 3046-3051

236. Liu, C. Q., and He, G. Q. (2012) Multiple alpha-galactosidases from Aspergillus foetidus ZU-G1: purification, characterization and application in soybean milk hydrolysis. *Eur Food Res Technol* **234**, 743-751

237. Pessela, B. C. C., Fernandez-Lafuente, R., Torres, R., Mateo, C., Fuentes, M., Filho, M., Vian, A., Garcia, J. L., Guisan, J. M., and Carrascosa, A. V. (2007) Production of a thermoresistant alpha-galactosidase from Thermus sp strain T2 for food processing. *Food Biotechnol* **21**, 91-103

238. Clarke, J. H., Davidson, K., Rixon, J. E., Halstead, J. R., Fransen, M. P., Gilbert, H. J., and Hazlewood, G. P. (2000) A comparison of enzyme-aided bleaching of softwood paper pulp using combinations of xylanase, mannanase and alpha-galactosidase. *Appl Microbiol Biotechnol* **53**, 661-667

239. Ernst, H. A., Lo Leggio, L., Willemoes, M., Leonard, G., Blum, P., and Larsen, S. (2006) Structure of the Sulfolobus solfataricus alpha-glucosidase: Implications for domain conservation and substrate recognition in GH31. *Journal of molecular biology* **358**, 1106-1124

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

241. Larsbrink, J., Izumi, A., Ibatullin, F. M., Nakhai, A., Gilbert, H. J., Davies, G. J., and Brumer, H. (2011) Structural and enzymatic characterization of a glycoside hydrolase family 31 alpha-xylosidase from Cellvibrio japonicus involved in xyloglucan saccharification. *The Biochemical journal* **436**, 567-580

242. Luonteri, E., Tenkanen, M., and Viikari, L. (1998) Substrate specificities of Penicillium simplicissimum alpha-galactosidases. *Enzyme Microb Technol* **22**, 192-198

243. Ademark, P., Larsson, M., Tjerneld, F., and Stalbrand, H. (2001) Multiple alpha-galactosidases from Aspergillus niger: purification, characterization and substrate specificities. *Enzyme Microb Tech* **29**, 441-448

244. Margolles-Clark, E., Tenkanen, M., Luonteri, E., and Penttila, M. (1996) Three alpha-galactosidase genes of Trichoderma reesei cloned by expression in yeast. *Eur J Biochem* **240**, 104-111

245. Ademark, P., de Vries, R. P., Hagglund, P., Stalbrand, H., and Visser, J. (2001) Cloning and characterization of Aspergillus niger genes encoding an alpha-galactosidase and a beta-mannosidase involved in galactomannan degradation. *Eur J Biochem* **268**, 2982-2990

246. Wang, H., Luo, H., Li, J., Bai, Y., Huang, H., Shi, P., Fan, Y., and Yao, B. (2010) An alpha-galactosidase from an acidophilic Bispora sp. MEY-1 strain acts synergistically with beta-mannanase. *Bioresour Technol* **101**, 8376-8382

247. Shankar, S. K., Dhananjay, S. K., and Mulimani, V. H. (2009) Purification and characterization of thermostable alpha-galactosidase from Aspergillus terreus (GR). *Appl Biochem Biotechnol* **152**, 275-285

248. Gherardini, F., Babcock, M., and Salyers, A. A. (1985) Purification and characterization of two alpha-galactosidases associated with catabolism of guar gum and other alpha-galactosides by Bacteroides ovatus. *Journal of bacteriology* **161**, 500-506

249. Aguilera, M., Rakotoarivonina, H., Brutus, A., Giardina, T., Simon, G., and Fons, M. (2012) Aga1, the first alpha-Galactosidase from the human bacteria Ruminococcus gnavus E1, efficiently transcribed in gut conditions. *Res Microbiol* **163**, 14-21

250. Wang, C., Wang, H., Ma, R., Shi, P., Niu, C., Luo, H., Yang, P., and Yao, B. (2016) Biochemical characterization of a novel thermophilic alpha-galactosidase from Talaromyces leycettanus JCM12802 with significant transglycosylation activity. *J Biosci Bioeng* **121**, 7-12

251. Zhao, H., Lu, L., Xiao, M., Wang, Q., Lu, Y., Liu, C., Wang, P., Kumagai, H., and Yamamoto, K. (2008) Cloning and characterization of a novel alpha-galactosidase from Bifidobacterium breve 203 capable of synthesizing Gal-alpha-1,4 linkage. *FEMS Microbiol Lett* **285**, 278-283

252. Zhou, J., Lu, Q., Zhang, R., Wang, Y., Wu, Q., Li, J., Tang, X., Xu, B., Ding, J., and Huang, Z. (2016) Characterization of two glycoside hydrolase family 36 alpha-galactosidases: novel transglycosylation activity, lead-zinc tolerance, alkaline and multiple pH optima, and low-temperature activity. *Food Chem* **194**, 156-166

253. Malgas, S., van Dyk, J. S., and Pletschke, B. I. (2015) A review of the enzymatic hydrolysis of mannans and synergistic interactions between beta-mannanase, beta-mannosidase and alpha-galactosidase. *World J Microbiol Biotechnol* **31**, 1167-1175

254. Huang, Y., Zhang, H., Ben, P., Duan, Y., Lu, M., Li, Z., and Cui, Z. (2018) Characterization of a novel GH36 alpha-galactosidase from Bacillus megaterium and its application in degradation of raffinose family oligosaccharides. *Int J Biol Macromol* **108**, 98-104

255. Sasaki, Y., Togo, N., Kitahara, K., and Fujita, K. (2018) Characterization of a GH36 β-L-Arabinopyranosidase in Bifidobacterium adolescentis. *Journal of Applied Glycoscience* **65**, 23-30

256. Carrera-Silva, E. A., Silvestroni, A., LeBlanc, J. G., Piard, J. C., Savoy de Giori, G., and Sesma, F. (2006) A thermostable alpha-galactosidase from Lactobacillus fermentum CRL722: genetic characterization and main properties. *Curr Microbiol* **53**, 374-378

257. Bakunina, I. Y., Balabanova, L. A., Pennacchio, A., and Trincone, A. (2016) Hooked on alpha-d-galactosidases: from biomedicine to enzymatic synthesis. *Crit Rev Biotechnol* **36**, 233-245

258. Adamson, C., Pengelly, R. J., Shamsi Kazem Abadi, S., Chakladar, S., Draper, J., Britton, R., Gloster, T. M., and Bennet, A. J. (2016) Structural Snapshots for Mechanism-Based Inactivation of a Glycoside Hydrolase by Cyclopropyl Carbasugars. *Angewandte Chemie* **55**, 14978-14982

259. Zhou, J., Shi, P., Huang, H., Cao, Y., Meng, K., Yang, P., Zhang, R., Chen, X., and Yao, B. (2010) A new alpha-galactosidase from symbiotic Flavobacterium sp. TN17 reveals four residues essential for alpha-galactosidase activity of gastrointestinal bacteria. *Appl Microbiol Biotechnol* **88**, 1297-1309

260. Gherardini, F. C., and Salyers, A. A. (1987) Characterization of an outer membrane mannanase from Bacteroides ovatus. *Journal of bacteriology* **169**, 2031-2037

261. Valentine, P. J., Gherardini, F. C., and Salyers, A. A. (1991) A Bacteroides ovatus chromosomal locus which contains an alpha-galactosidase gene may be important for colonization of the gastrointestinal tract. *Applied and environmental microbiology* **57**, 1615-1623
262. Martens, E. C., Kelly, A. G., Tauzin, A. S., and Brumer, H. (2014) The devil lies in the details: how variations in polysaccharide fine-structure impact the physiology and evolution of gut microbes. *Journal of molecular biology* **426**, 3851-3865
263. Berg, J. M., Tymoczko, J. L., Gatto Jr., G. J., and Stryer, L. (2015) *Biochemistry*, 8th ed., W. H. Freeman and Company, New York
264. Stalbrand, H., Siikaaho, M., Tenkanen, M., and Viikari, L. (1993) Purification and Characterization of 2 Beta-Mannanases from Trichoderma-Reesei. *J Biotechnol* **29**, 229-242
265. Miller, G. L. (1959) Use of Dinitrosalicylic Acid Reagent for Determination of Reducing Sugar. *Anal Chem* **31**, 426-428
266. Rocklin, R. D., and Pohl, C. A. (1983) Determination of Carbohydrates by Anion-Exchange Chromatography with Pulsed Amperometric Detection. *J Liq Chromatogr* **6**, 1577-1590
267. Grey, C., Edebrink, P., Krook, M., and Jacobsson, S. P. (2009) Development of a high performance anion exchange chromatography analysis for mapping of oligosaccharides. *J Chromatogr B Analyt Technol Biomed Life Sci* **877**, 1827-1832
268. Westereng, B., Agger, J. W., Horn, S. J., Vaaje-Kolstad, G., Aachmann, F. L., Stenstrom, Y. H., and Eijsink, V. G. (2013) Efficient separation of oxidized cello-oligosaccharides generated by cellulose degrading lytic polysaccharide monooxygenases. *J Chromatogr A* **1271**, 144-152
269. Matsui, I., Ishikawa, K., Matsui, E., Miyairi, S., Fukui, S., and Honda, K. (1991) Subsite structure of Saccharomycopsis alpha-amylase secreted from Saccharomyces cerevisiae. *Journal of biochemistry* **109**, 566-569
270. Suganuma, T., Matsuno, R., Ohnishi, M., and Hiromi, K. (1978) A study of the mechanism of action of Taka-amylase A1 on linear oligosaccharides by product analysis and computer simulation. *Journal of biochemistry* **84**, 293-316
271. Corradini, C., Lantano, C., and Cavazza, A. (2013) Innovative analytical tools to characterize prebiotic carbohydrates of functional food interest. *Anal Bioanal Chem* **405**, 4591-4605
272. Harvey, D. J. (2012) Analysis of carbohydrates and glycoconjugates by matrix-assisted laser desorption/ionization mass spectrometry: an update for 2007-2008. *Mass Spectrom Rev* **31**, 183-311
273. Harvey, D. J. (1993) Quantitative aspects of the matrix-assisted laser desorption mass spectrometry of complex oligosaccharides. *Rapid Commun Mass Spectrom* **7**, 614-619
274. Bu, L., Crowley, M. F., Himmel, M. E., and Beckham, G. T. (2013) Computational investigation of the pH dependence of loop flexibility and catalytic function in glycoside hydrolases. *The Journal of biological chemistry* **288**, 12175-12186

275. de Oliveira, L. C., da Silva, V. M., Colussi, F., Cabral, A. D., de Oliveira Neto, M., Squina, F. M., and Garcia, W. (2015) Conformational changes in a hyperthermostable glycoside hydrolase: enzymatic activity is a consequence of the loop dynamics and protonation balance. *PLoS One* **10**, e0118225
276. Chayen, N. E., and Saridakis, E. (2008) Protein crystallization: from purified protein to diffraction-quality crystal. *Nature Methods* **5**, 147-153
277. Chayen, N. E. (2004) Turning protein crystallisation from an art into a science. *Curr Opin Struc Biol* **14**, 577-583
278. Blake, C. C. F., and Phillips, D. C. (1962) Effects of {X}-Irradiation on Single Crystals of Myoglobin. in *Biological Effects of Ionizing Radiation At the Molecular Level, {IAEA}, Vienna, 1962 {STI/PUB/60}*, Int. Atomic Energy Agency
279. Nave, C. (1995) Radiation-Damage in Protein Crystallography. *Radiat Phys Chem* **45**, 483-490
280. Garman, E. F. (2010) Radiation damage in macromolecular crystallography: what is it and why should we care? *Acta Crystallogr D* **66**, 339-351
281. Garman, E. F., and Owen, R. L. (2006) Cryocooling and radiation damage in macromolecular crystallography. *Acta Crystallogr D* **62**, 32-47
282. Kabsch, W. (2010) Xds. *Acta Crystallogr D* **66**, 125-132
283. 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 Crystallogr D* **67**, 235-242
284. Hendrickson, W. A., Horton, J. R., and Lemaster, D. M. (1990) Selenomethionyl Proteins Produced for Analysis by Multiwavelength Anomalous Diffraction (Mad) - a Vehicle for Direct Determination of 3-Dimensional Structure. *Embo J* **9**, 1665-1672
285. Uson, I., and Sheldrick, G. M. (1999) Advances in direct methods for protein crystallography. *Curr Opin Struc Biol* **9**, 643-648
286. Rose, J. P., Wang, B. C., and Weiss, M. S. (2015) Native SAD is maturing. *Iucrj* **2**, 431-440
287. Scapin, G. (2013) Molecular replacement then and now. *Acta Crystallogr D* **69**, 2266-2275
288. Mccoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *Journal of Applied Crystallography* **40**, 658-674
289. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta crystallographica. Section D, Biological crystallography* **66**, 213-221
290. Rossmann, M. G., and Blow, D. M. (1962) Detection of Sub-Units within Crystallographic Asymmetric Unit. *Acta Crystallogr* **15**, 24-&

291. Dodson, E. (2008) The befores and afters of molecular replacement. *Acta crystallographica. Section D, Biological crystallography* **64**, 17-24
292. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta crystallographica. Section D, Biological crystallography* **66**, 486-501
293. Mueller, I. (2017) Guidelines for the successful generation of protein-ligand complex crystals. *Acta Crystallographica Section D-Structural Biology* **73**, 79-92
294. Pollet, A., Vandermarliere, E., Lammertyn, J., Strelkov, S. V., Delcour, J. A., and Courtin, C. M. (2009) Crystallographic and activity-based evidence for thumb flexibility and its relevance in glycoside hydrolase family 11 xylanases. *Proteins-Structure Function and Bioinformatics* **77**, 395-403
295. Wu, M., Bu, L. T., Vuong, T. V., Wilson, D. B., Crowley, M. F., Sandgren, M., Stahlberg, J., Beckham, G. T., and Hansson, H. (2013) Loop Motions Important to Product Expulsion in the Thermobifida fusca Glycoside Hydrolase Family 6 Cellobiohydrolase from Structural and Computational Studies. *Journal of Biological Chemistry* **288**, 33107-33117
296. Torchia, D. A., Sparks, S. W., and Bax, A. (1988) Delineation of Alpha-Helical Domains in Deuteriated Staphylococcal Nuclease by 2d Noe Nmr-Spectroscopy. *Journal of the American Chemical Society* **110**, 2320-2321
297. Fernandez, C., and Wider, G. (2003) TROSY in NMR studies of the structure and function of large biological macromolecules. *Curr Opin Struc Biol* **13**, 570-580
298. Reed, M. A. C., Hounslow, A. M., Sze, K. H., Barsukov, I. G., Hosszu, L. L. P., Clarke, A. R., Craven, C. J., and Waltho, J. P. (2003) Effects of domain dissection on the folding and stability of the 43 kDa protein PGK probed by NMR. *Journal of molecular biology* **330**, 1189-1201
299. Krishna, N. R., Huang, D. H., Glickson, J. D., Rowan, R., and Walter, R. (1979) Amide Hydrogen-Exchange Rates of Peptides in H2o Solution by H-1 Nuclear Magnetic-Resonance Transfer of Solvent Saturation Method - Conformations of Oxytocin and Lysine Vasopressin in Aqueous-Solution. *Biophys J* **26**, 345-366
300. Pervushin, K., Riek, R., Wider, G., and Wuthrich, K. (1997) Attenuated T-2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 12366-12371
301. Bruel, L., Sulzenbacher, G., Cervera Tison, M., Pujol, A., Nicoletti, C., Perrier, J., Galinier, A., Ropartz, D., Fons, M., Pompeo, F., and Giardina, T. (2011) alpha-Galactosidase/sucrose kinase (AgaSK), a novel bifunctional enzyme from the human microbiome coupling galactosidase and kinase activities. *The Journal of biological chemistry* **286**, 40814-40823
302. Tailford, L. E., Money, V. A., Smith, N. L., Dumon, C., Davies, G. J., and Gilbert, H. J. (2007) Mannose foraging by Bacteroides thetaiotaomicron: structure and specificity of the beta-mannosidase, BtMan2A. *The Journal of biological chemistry* **282**, 11291-11299

303. Mackenzie, A. K., Naas, A. E., Kracun, S. K., Schuckel, J., Fangel, J. U., Agger, J. W., Willats, W. G., Eijsink, V. G., and Pope, P. B. (2015) A polysaccharide utilization locus from an uncultured bacteroidetes phylotype suggests ecological adaptation and substrate versatility. *Applied and environmental microbiology* **81**, 187-195

304. Senoura, T., Ito, S., Taguchi, H., Higa, M., Hamada, S., Matsui, H., Ozawa, T., Jin, S., Watanabe, J., Wasaki, J., and Ito, S. (2011) New microbial mannan catabolic pathway that involves a novel mannosylglucose phosphorylase. *Biochem Biophys Res Commun* **408**, 701-706

305. Schroder, C., Janzer, V. A., Schirrmacher, G., Claren, J., and Antranikian, G. (2017) Characterization of two novel heat-active alpha-galactosidases from thermophilic bacteria. *Extremophiles : life under extreme conditions* **21**, 85-94

306. Merceron, R., Foucault, M., Haser, R., Mattes, R., Watzlawick, H., and Gouet, P. (2012) The molecular mechanism of thermostable alpha-galactosidases AgaA and AgaB explained by x-ray crystallography and mutational studies. *The Journal of biological chemistry* **287**, 39642-39652

307. Schrödinger, LLC. (2015) The PyMOL Molecular Graphics System, Version 1.7.

308. Couturier, M., Haon, M., Coutinho, P. M., Henrissat, B., Lesage-Meessen, L., and Berrin, J. G. (2011) Podospora anserina hemicellulases potentiate the Trichoderma reesei secretome for saccharification of lignocellulosic biomass. *Applied and environmental microbiology* **77**, 237-246

309. Tamura, K., Hemsworth, G. R., DeJean, G., Rogers, T. E., Pudlo, N. A., Urs, K., Jain, N., Davies, G. J., Martens, E. C., and Brumer, H. (2017) Molecular Mechanism by which Prominent Human Gut Bacteroidetes Utilize Mixed-Linkage Beta-Glucans, Major Health-Promoting Cereal Polysaccharides. *Cell Rep* **21**, 417-430

310. Luis, A. S., Briggs, J., Zhang, X. Y., Farnell, B., Ndeh, D., Labourel, A., Basle, A., Cartmell, A., Terrapon, N., Stott, K., Lowe, E. C., McLean, R., Shearer, K., Schuckel, J., Venditto, I., Ralet, M. C., Henrissat, B., Martens, E. C., Mosimann, S. C., Abbott, D. W., and Gilbert, H. J. (2018) Dietary pectic glycans are degraded by coordinated enzyme pathways in human colonic Bacteroides. *Nature Microbiology* **3**, 210-219

311. 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 beta-1,4-Mannanase Belonging to a New Glycoside Hydrolase Family in Aspergillus nidulans. *The Journal of biological chemistry* **290**, 27914-27927

312. Armstrong, Z., Mewis, K., Liu, F., Morgan-Lang, C., Scofield, M., Durno, E., Chen, H. M., Mehr, K., Withers, S. G., and Hallam, S. J. (2018) Metagenomics reveals functional synergy and novel polysaccharide utilization loci in the Castor canadensis fecal microbiome. *ISME J*

313. Wang, K., Pereira, G. V., Cavalcante, J. J. V., Zhang, M. L., Mackie, R., and Cann, I. (2016) Bacteroides intestinalis DSM 17393, a member of the human colonic microbiome, upregulates multiple endoxylanases during growth on xylan. *Sci Rep-Uk* **6**

314. Phansopa, C., Roy, S., Rafferty, J. B., Douglas, C. W., Pandhal, J., Wright, P. C., Kelly, D. J., and Stafford, G. P. (2014) Structural and functional characterization of NanU, a novel high-affinity sialic acid-inducible binding protein of oral and gut-dwelling Bacteroidetes species. *The Biochemical journal* **458**, 499-511

315. Cho, K. H., and Salyers, A. A. (2001) Biochemical analysis of interactions between outer membrane proteins that contribute to starch utilization by Bacteroides thetaiotaomicron. *Journal of bacteriology* **183**, 7224-7230

316. Tuson, H. H., Foley, M. H., Koropatkin, N. M., and Biteen, J. S. (2018) The Starch Utilization System Assembles around Stationary Starch-Binding Proteins. *Biophys J*

317. Nihira, T., Suzuki, E., Kitaoka, M., Nishimoto, M., Ohtsubo, K., and Nakai, H. (2013) Discovery of beta-1,4-D-mannosyl-N-acetyl-D-glucosamine phosphorylase involved in the metabolism of N-glycans. *The Journal of biological chemistry* **288**, 27366-27374

318. Hehemann, J. H., Kelly, A. G., Pudlo, N. A., Martens, E. C., and Boraston, A. B. (2012) Bacteria of the human gut microbiome catabolize red seaweed glycans with carbohydrate-active enzyme updates from extrinsic microbes. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 19786-19791

319. Sonnenburg, E. D., Zheng, H., Joglekar, P., Higginbottom, S. K., Firbank, S. J., Bolam, D. N., and Sonnenburg, J. L. (2010) Specificity of polysaccharide use in intestinal bacteroides species determines diet-induced microbiota alterations. *Cell* **141**, 1241-1252

320. Ali-Ahmad, A., Garron, M. L., Zamboni, V., Lenfant, N., Nurizzo, D., Henrissat, B., Berrin, J. G., Bourne, Y., and Vincent, F. (2017) Structural insights into a family 39 glycoside hydrolase from the gut symbiont Bacteroides cellulosilyticus WH2. *J Struct Biol* **197**, 227-235

321. Liang, P. H., Lin, W. L., Hsieh, H. Y., Lin, T. Y., Chen, C. H., Tewary, S. K., Lee, H. L., Yuan, S. F., Yang, B., Yao, J. Y., and Ho, M. C. (2018) A flexible loop for mannan recognition and activity enhancement in a bifunctional glycoside hydrolase family 5. *Bba-Gen Subjects* **1862**, 513-521

322. Hudson, K. L., Bartlett, G. J., Diehl, R. C., Agirre, J., Gallagher, T., Kiessling, L. L., and Woolfson, D. N. (2015) Carbohydrate-Aromatic Interactions in Proteins. *J Am Chem Soc* **137**, 15152-15160

323. Asensio, J. L., Arda, A., Canada, F. J., and Jimenez-Barbero, J. (2013) Carbohydrate-aromatic interactions. *Acc Chem Res* **46**, 946-954

324. Song, J., Kim, S.-Y., Kim, D.-H., Lee, Y.-S., Sim, J.-S., Hahn, B.-S., and Lee, C.-M. (2018) Characterization of an inhibitor-resistant endo-1,4-β-mannanase from the gut microflora metagenome of Hermetia illucens. *Biotechnol Lett*

325. Srivastava, P. K., Panwar, D., Prashanth, K. V., and Kapoor, M. (2017) Structural Characterization and in Vitro Fermentation of beta-Mannooligosaccharides Produced from Locust Bean Gum by GH-26 endo-beta-1,4-Mannanase (ManB-1601). *J Agric Food Chem* **65**, 2827-2838

326. Robyt, J. F., and French, D. (1970) Multiple attack and polarity of action of porcine pancreatic alpha-amylase. *Arch Biochem Biophys* **138**, 662-670

327. Hrmova, M., and Fincher, G. B. (2001) Structure-function relationships of beta-D-glucan endo- and exohydrolases from higher plants. *Plant Mol Biol* **47**, 73-91
328. Hemsworth, G. R., Thompson, A. J., Stepper, J., Sobala, L. F., Coyle, T., Larsbrink, J., Spadiut, O., Goddard-Borger, E. D., Stubbs, K. A., Brumer, H., and Davies, G. J. (2016) Structural dissection of a complex Bacteroides ovatus gene locus conferring xyloglucan metabolism in the human gut. *Open Biol* **6**
329. Kitamura, M., Okuyama, M., Tanzawa, F., Mori, H., Kitago, Y., Watanabe, N., Kimura, A., Tanaka, I., and Yao, M. (2008) Structural and Functional Analysis of a Glycoside Hydrolase Family 97 Enzyme from Bacteroides thetaiotaomicron. *Journal of Biological Chemistry* **283**, 36328-36337
330. Correia, M. A. S., Mazumder, K., Bras, J. L. A., Firbank, S. J., Zhu, Y. P., Lewis, R. J., York, W. S., Fontes, C. M. G. A., and Gilbert, H. J. (2011) Structure and Function of an Arabinoxylan-specific Xylanase. *Journal of Biological Chemistry* **286**, 22510-22520
331. St John, F. J., Hurlbert, J. C., Rice, J. D., Preston, J. F., and Pozharski, E. (2011) Ligand Bound Structures of a Glycosyl Hydrolase Family 30 Glucuronoxylan Xylanohydrolase. *Journal of molecular biology* **407**, 92-109
332. Meinke, A., Damude, H. G., Tomme, P., Kwan, E., Kilburn, D. G., Miller, R. C., Warren, R. A. J., and Gilkes, N. R. (1995) Enhancement of the Endo-Beta-1,4-Glucanase Activity of an Exocellobiohydrolase by Deletion of a Surface Loop. *Journal of Biological Chemistry* **270**, 4383-4386
333. Payne, C. M., Knott, B. C., Mayes, H. B., Hansson, H., Himmel, M. E., Sandgren, M., Stahlberg, J., and Beckham, G. T. (2015) Fungal Cellulases. *Chem Rev* **115**, 1308-1448
334. Momeni, M. H., Payne, C. M., Hansson, H., Mikkelsen, N. E., Svedberg, J., Engstrom, A., Sandgren, M., Beckham, G. T., and Stahlberg, J. (2013) Structural, Biochemical, and Computational Characterization of the Glycoside Hydrolase Family 7 Cellobiohydrolase of the Tree-killing Fungus Heterobasidion irregulare. *Journal of Biological Chemistry* **288**, 5861-5872
335. Proctor, M. R., Taylor, E. J., Nurizzo, D., Turkenburg, J. P., Lloyd, R. M., Vardakou, M., Davies, G. J., and Gilbert, H. J. (2005) Tailored catalysts for plant cell-wall degradation: Redesigning the exo/endo preference of Cellvibrio japonicus arabinanase 43A. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 2697-2702
336. Fukamizo, T., Miyake, R., Tamura, A., Ohnuma, T., Skriver, K., Pursiainen, N. V., and Juffer, A. H. (2009) A flexible loop controlling the enzymatic activity and specificity in a glycosyl hydrolase family 19 endochitinase from barley seeds (Hordeum vulgare L.). *Bba-Proteins Proteom* **1794**, 1159-1167