Biochemical and Chemoenzymatic Strategies for the Characterisation of Carbohydrate Active Enzymes

A thesis submitted to The University of Manchester for the degree of

Doctor of Philosophy in the Faculty of Science and Engineering

2021

Gregory S. Bulmer

Department of Chemistry, School of Natural Sciences

Table of Contents

Word Count: 45,754

Abbreviations

Araf	L-Arabinofuranose
CAZy	Carbohydrate Active Enzyme
CBP	Cellobiose Phosphorylase
CDP	Cellodextrin Phosphorylase
CNC	Cellulose Nanocrystals
CNF	Cellulose Nanofibers
CSC	Cellulose Synthase Complex
DP	Degree of Polymerisation
FA	Ferulic Acid
FAE	Feruloyl esterase/ ferulic acid esterase
Gal	D-Galactose
Galf	D-Galactofuranose
GalNAc	N-Acetyl-D-galactosamine
GH	Glycoside Hydrolase
Glc	D-Glucose
GlcNAc	N-Acetyl-D-glucosamine
GP	Glycoside phosphorylase
GS	Glycosynthase
GT	Glycosyltransferase
HPLC	High-performance liquid chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LPMO	Lytic Polysaccharide Monooxgenase
MALDI	Matrix-assisted laser desorption ionization
MS	Mass Spectrometry
PCW	Primary Cell Wall
PNP	p-nitrophenol
SCW	Secondary Cell Wall
SuSy	Sucrose Synthase
UDP	Uridine diphosphate

Abstract

The growing ramifications of climate change, coupled with a finite pool of natural resources, highlight a pertinent need for a rapid transition to a renewables-based bio-economy. One essential aspect of sustainability is the valorisation of plant carbohydrates into renewable biofuels, value-added chemicals and bio-polymers through the use of Carbohydrate Active Enzymes (CAZymes). Involved in the synthesis and deconstruction of carbohydrates, CAZymes offer excellent regio- and stereoselectivity that can be exploited for the production of desired structures. Although the utility of CAZymes has been demonstrated extensively, technologies for their large-scale characterisation are lacking. Consequently, the vast majority of predicted CAZymes are not biochemically characterised leaving the true scope of enzyme activities unknown.

The overall aim of this thesis was to advance the knowledge on, and assess biotechnological application of, microbial enzymes for the processing of lignocellulose and its components. In this project we identified a glycosyltransferase capable of polymerising glucose into β-1,4-linked oligosaccharides. Polymerisation of glucose onto imidazolium-based tags, coupled with chemoenzymatic derivatisation, enabled a sensitive, mass spectrometry-based assay that could characterise the substrate scope and specificity of glycoside hydrolases and LPMOs. Furthermore, following phylogenetic analysis of the GH43 subfamily 34 and molecular docking studies, a β-galactofuranosidase (XynD) from *Aspergillus niger* was biochemically characterised. Incubations against a variety of oligo- and polysaccharides revealed activity against pNP-Gal*f,* suggesting the natural substrate is a disaccharide or a glycan with a different linkage type from β-(1,5), β-(1,6) or α-(1,2)-linked Gal*f.* Finally, a thermotolerant feruoyl esterase from *Thermobacillus xylanilyticus* (Tx-Est1) was further characterised, with activities against ferulated oligosaccharides and plant polysaccharides investigated. The enzyme revealed an important role in the processing of nutritionally important, pre-biotic xylooligosaccharides. This thesis demonstrates the versatile nature and utility offered by CAZymes.

Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright statement

i. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the "Copyright") and s/he has given the University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

iii. The ownership of certain Copyright, patents, designs, trademarks and other intellectual property (the "Intellectual Property") and any reproductions of copyright works in the thesis, for example graphs and tables ("Reproductions"), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=2442 0), in any relevant Thesis restriction declarations deposited in the University Library, the University Library's regulations (see http://www.library.manchester.ac.uk/ about/regulations/) and in the University's policy on Presentation of Theses.

Dedicated to my grandma

Shelia Joyce Robinson

1925-2015

Acknowledgements

Science is never an individual effort and, for that reason, I owe so much, to so many, who have helped me get to where I am now. First, I would like to thank my supervisors Dr. Jolanda van Munster and Prof. Sabine Flitsch for giving me this opportunity — their support, patience and creativity throughout this process has enabled me to grow as a scientist and learn so much across a varied range of fields. Thanks also to the EPSRC for funding my PhD and conference travel.

I would like to thank Dr. Fabio Parmeggiani, Dr. Antonio Angelastro, Dr. Peter Both, Dr. Kun Huang and Dr. Andrea Marchesi for their glycobiology help throughout my time in the MIB. Especially, I would like to thank Dr. Ashley Mattey for his help, chemistry knowledge and friendship. Working in the Turner-Flitsch group has been an excellent environment to learn in, for this I would particularly like to thank my friends Maximillian Lubberink, Grayson Ford, Joe Sharratt, Ryan Palmer, James Marshall, Laura Rodríguez Pérez, Matthew Willmott, Bethan Jones and Tilly Allen. Furthermore, my thanks go to the Glyco gals (Gareth Surman, Matthew Treadall and David Evans), the next generation of MALDIers, who have been a great team to work with.

Support has not only existed within the workplace, I want to thank my friends Louis Cane, Dr. Josh Whitehurst, Daniel Holford and Alyx Adams for their support and humour in often-difficult times. Additionally thanks to Dr. Holly Allen, Ellie Welch and Tasha Spoor, who I initially met through science and now have the privilege of calling my friends. I must not forget my gratitude to Dr. Simon Snowden and Mathew Chatterton, for all the laughs we share on our cactus-based trips together.

I thank my parents, Theresa and Trevor, for their unconditional support throughout my education and for all the sacrifices they have made to ensure I was able to achieve my full potential. I am so grateful for all that you do for me. Finally, Amy, my sister for always looking out for me and, of course, my niece Rose — who can always cheer me up on the most fruitless of lab days.

Thesis structure

This thesis is prepared in 'Journal Format'. Chapter 1 introduces the structure and enzymatic deconstruction of the plant cell wall, the structure of the fungal cell wall and the use of Carbohydrate Active Enzymes (CAZymes) for the production of high-value chemicals. Chapter 2 provides the objectives of this project. Chapters 3, 4, 5 and 6 are prepared in manuscript format. Chapter 3 has been published as a review in Carbohydrate Research. The article is available at https://doi.org/10.1016/j.carres.2021.108411. Chapter 4 has been published in Organic and Biomolecular Chemistry and is available at https://doi.org/10.1039/D1OB00971K. Chapter 5 has been prepared in paper format with submission to the Journal of Biological Chemistry anticipated. Chapter 6 has been prepared in paper format with submission to Food Chemistry anticipated after additional experiments are undertaken. Chapter 7 summarises the scope of this project and discusses the potential for future work.

Chapter 1: Introduction

The biological importance of carbohydrates

Carbohydrates are compounds formed of carbon, hydrogen and oxygen, which play important structural and functional roles throughout nature. Glycosidic bonds can link individual monosaccharides to form disaccharides (2 units), oligosaccharides (3-~10 units) or polysaccharides (polymers > 10 units) [1]. Polysaccharides can offer substantial mechanical strength; the glucose-based polysaccharide cellulose holds an important structural role in plants [2] whilst chitin, formed of *N*-acetylglucosamine, is found in arthropod exoskeletons and fungal cell walls [3]. Additionally polysaccharides can hold energy storage capabilities, with starch and glycogen acting as storage molecules for glucose in plants and animals respectively [4,5].

Carbohydrates are ubiquitous on cell surfaces with those present in glycoproteins (glycans) holding key roles in molecular recognition and enable cell-to-cell communication within multicellular organisms or the obtaining of nutrients [6]. These external glycans can enable the immune system to recognise which cells belong to the host organism or which are non-self, however some pathogens employ molecular mimicry of host glycans to avoid detection [7]. Two of the most common types of glycosylation are N-linked and O-linked glycosylation (Fig. 1), which are predominantly found in eukaryotes and archaea but examples have also been identified in bacteria [8]. In addition to important natural roles, carbohydrates are of great importance in the production of fuels, materials and pharmaceuticals. Carbohydrates from plant waste can be used in the production of bioethanol [9] or bioplastics [10]. In the production of medicines carbohydrates can improve or modify the solubility of therapeutic biomolecules in addition to improving their stability [11].

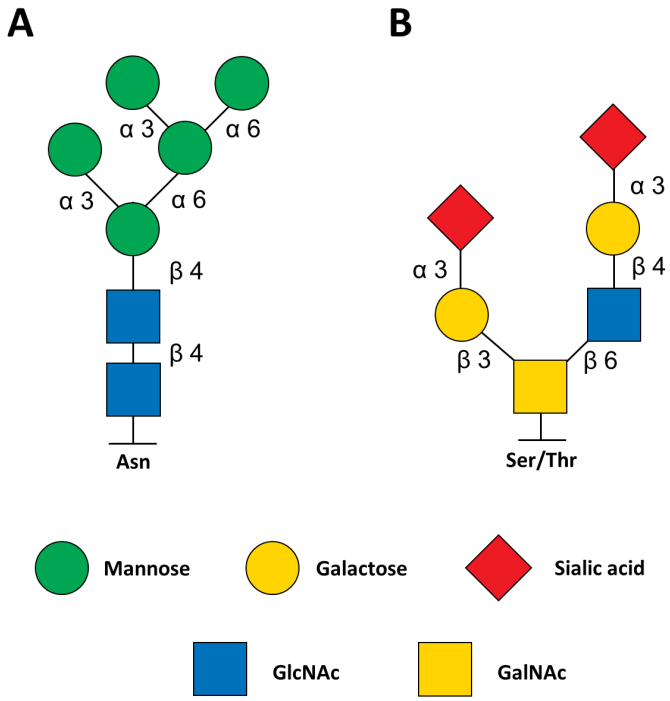


Figure 1: Representative structures of N-linked (A) and O-linked (B) glycans. N-linked glycosylation (**A**) consists of the linkage of N-acetylglucosamine (GlcNAc) to the nitrogen atom of an asparagine (Asn) side chain. These Asn-linked glycoconjugates contain a pentasaccharide core of GlcNAc units linked to three mannoses, to which further glycosylation (commonly galactosylation, GlcNAclyation, sialylation and fucosylation) can occur. O-linked glycosylation (**B**) consists of the linkage of (predominantly) N-acetylglucosamine (GlcNAc) or N-acetylgalactosamine (GalNAc) to the oxygen atom in the hydroxyl groups of serine and threonine, followed by further decoration as seen in N-linked glycosylation [12].

Bio-based economy

Plant polysaccharides comprise 75% of carbon produced by the biosphere, of this vast source, carbohydrates and lignin from lignocellulosic biomass waste constitute 2 x 10^{11} tonnes of material on an annual basis [13]. Consequently, this waste represents a promising renewable source of carbon for both the production of biomaterials and biofuels and shall play a fundamental role in the development of a future bio-based economy. This concept of a renewable, sustainable and financially viable system includes concepts such as microbial production of enzymes, enzyme technology, green chemistry and advanced physical/ chemical processing [14].

The valorisation of plant polysaccharides by using them as a biofuel feedstock offers a route to carbon neutral sources of energy. Initially, first generation biofuels were produced from crops such as sugarcane or wheat grain. Starch from these crops was converted into simple sugars a subsequently fermented into ethanol, propanol and butanol [15]. However, such fuels are controversial due to the occupying of agricultural land for fuel production rather than food production. Second generation biofuels circumvent this dilemma by utilising sugars within agricultural waste or by-products such as wheat straw lignocellulose. For example, due to its cheap and abundant nature, coupled with its by-product status, within European agriculture wheat straw is a key candidate for an economically viable, renewable source of second generation biofuel [16]. Other such by-products eg bagasse (the residue from milled sugar cane after juice extraction) constitute a vast feed stock potential in many Asian and South American economies [17].

The enzymatic functionalisation of lignocellulose can enable the production of value added materials such as biopolymers, nanomaterials or substrates for further chemical synthesis. For example, via pre-treatment with NaOH, followed by acid hydrolysis and further purification, lignocellulose is converted into cellulose nanocrystals (CNC). Due to a high proportion of free hydroxyl groups CNCs can be highly functionalised with a wide range of uses including printed electronics, aerogels and biosensors [18]. Enzymatic alternatives for the synthesis of such CNCs are reviewed in more detail in chapter 3. Wood can be treated with mild conditions toform cellulose scaffolds onto which CBM(Carbohydrate Binding Module)-tagged proteins can be immobilised for the use of continuous flow biocatalysis [19].

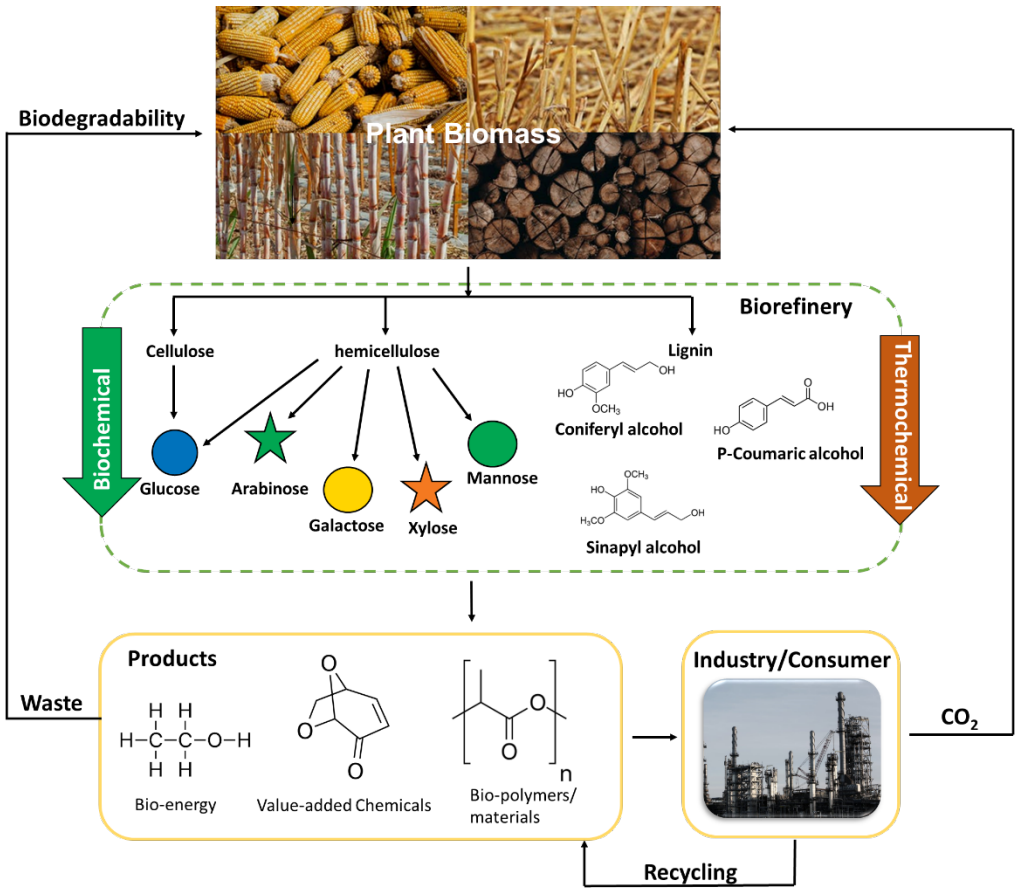


Figure 2: Biorefinery concept involving the utilisation of plant biomass to produce products with life cycles suitable for a circular bio-economy.

(Hemi)cellulose oligosaccharides can be modified enzymatically and functionalised through simple click chemistry. In one example, GlcA-containing xylooligoaccharides were functionalised at the carboxylic acid groups with click functionalities. Xylooligosaccharides were then oxidised at the terminal xylose position for further click functionalisation, allowing the addition of different groups at either end of the oligosaccharide. These bifunctional xylooligosaccharide monomers can then undergo Copper(I)-catalyzed azide-alkyne cycloaddition (CuAAC) "click" polymerisation to form a gel with unique properties, such as transition metal coordination [20]. Such approaches show how CAZymes should be thought of as tools in the degradation, modification or synthesis of lignocellulose.

However, despite the advent of novel technologies mentioned prior, the recalcitrant nature of lignocellulosic biomass and its diverse, complex structure means that its degradation to simple sugars or purification of individual polysaccharides is both difficult and often expensive, and therefore negatively

impacts industries ability to synthesise carbohydrate-derived materials [21]. To combat these complications the “biorefinery” concept has been developed, which combines a mix of degradation and conversion processes including chemical, mechanical and thermal pre-treatment combined with biotechnological steps in order to generate biofuels, high value biochemical and biomaterials (Fig. 2). The processes required to refine the starting material vary depending on the initial feedstock (agricultural waste, food waste or specifically designed crops) and the end product aim [22].

In order to improve technologies associated with the valorisation of lignocellulosic biomass and the production of value-added chemicals for industry a comprehensive understanding of several biological and chemical factors must be understood. By understanding the structure and regulation of the plant cell wall, genetic modification can enable the production of plants that are more amenable to enzymatic deconstruction or saccharification. The development of tools to characterise CAZymes involved in the plant biomass modification or degradation will enable increased awareness of the degradative capabilities of microbes, both in reaction mechanism and substrate scope. Chapter 4 describes the development of such a tool, a mass spectrometry-based assay for cellulolytic enzymes. In tandem, through characterising enzymes involved in the synthesis or degradation of polysaccharides, more efficient routes can be found to deconstruct or modify properties of lignocellulose and it’s components. Chapters 4, 5 and 6 describe the characterisation and exploitation of such microbial carbohydrate active enzymes.

Structure of the plant cell wall

Understanding the structure of plant cell walls (Fig. 3) and their formation is fundamental to understanding how to exploit plants for the uses mentioned prior. A complex, supramolecular structure, the plant wall is formed of two discrete regions, the primary cell wall (PCW) and secondary cell wall (SCW) [24]. The formation of these are highly regulated both spatially and temporally; the PCW allows the cell to grow and elongate and after the final cell size has been reached the SCW then strengthens the cell wall by forming between the PCW and plasma membrane [25].

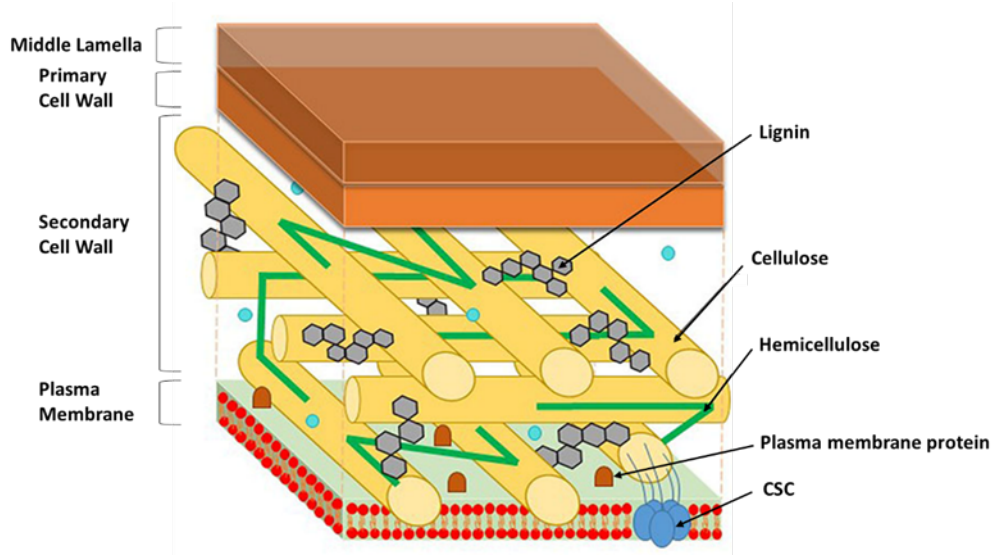


Figure 3: Structure of plant Secondary Cell Wall. SCW forms between the PCW and the plasma membrane and consists predominantly of cellulose microfibrils, hemicelluloses (such as xylan and xyloglucan). The phenolic polymer lignin is intercalated within these structures. Adapted from [23].

The genetics regulating the formation of the SCW are an area of plentiful research as wood is predominately formed from the SCW thickened xylem vessels thus making the subject economically important for increased xylem yields [26]. The PCW and SCW are a complex, supramolecular structure and thus the plant cell wall is held together through multiple mechanisms. Cellulose is cross-linked by hemicellulose and this structure in the primary cell wall is embedded in pectin - a highly hydrated network of polysaccharides containing large quantities of galacturonic acid. In the secondary cell wall lignin performs this structural role whilst a combination ofcovalent and noncovalent bonding hold all of these structural components together [27].

Cellulose is a linear, crystalline polymer of β-1,4-linked glucose and constitutes the most abundant molecule in the plant cell wall by dry weight (~30%) [28,29]. Cellulose chains consist of a reducing end and a non-reducing end (Fig. 4): at the reducing end lies a hemiacetal with a reducing aldehyde, whilst at the non-reducing end the C1 carbon is involved in a β-1,4 glycosidic bond [30]. While cellulose is naturally produced by a complex, membrane-associated biosynthetic machinery, considerable progress has been made in developing biocatalytic approaches to synthesize cellulose and its derivatives, as reviewed in chapter 3 and explored experimentally in chapter 4.

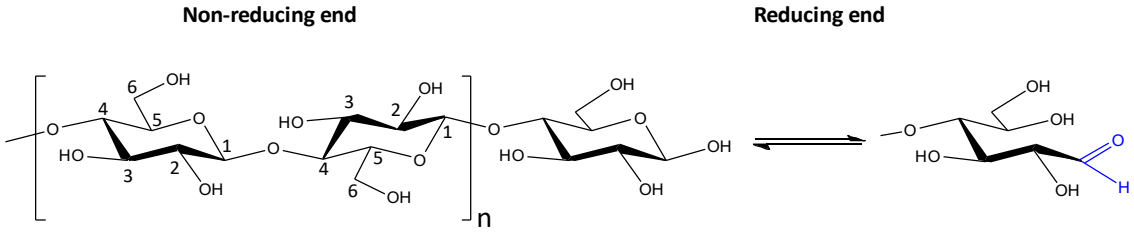


Figure 4: Structure of cellulose highlighting non-reducing and reducing ends. Free aldehyde denoted in the blue on open-ring terminal glucose. Numbering of carbon atoms starts from the anomeric carbon (the carbon participating in CHO aldehyde or C=O ketone in the open-chain form of the carbohydrate molecule) going down the chain.

Structure of hemicelluloses

Hemicelluloses contain β-1,4-linked backbones of the sugars glucose, mannose, or xylose and show structural similarity in part due to the equatorial configuration at C1 and C4. The hemicellulose group includes the polymers xylan, glucuronoxylan, arabinoxylan, glucomannan, and xyloglucan [31](Fig. 5). Xylans are the most abundant hemicelluloses within the cell wall and form between 20-35% of all terrestrial biomass [32]. This abundance makes them economically relevant targets in a future bio-economy.

Xylan structure (both in molecule size and the presence of functional groups) varies widely and is not only species dependent but also shows variation depending on the cell type the xylan is found in, as well as the developmental stage of the plant [33]. Although showing this great structural diversity the vast majority of xylans contain a β-1,4-linked xylose backbone which can be enzymatically modified to

contain different functional groups. β-1,3/β-1,4-D-xylans have been described in the red algae *Palmariapalmata*, however these are the only non-β-1,4-linked exception [34]. In β-1,4-linked xylan, the xylose units within the chains can be substituted with glucuronic acid (GlcA), 4-O-methylglucuronic acid, acetyl groups and arabinose [31]. GlcA and D-glucuronic acid 4-O-methyl are α-1,2 linked to the xylose backbone. In dicots, such as *Arabidopsis*, GlcA substitutions are common with around 1 in 8

residues substituted, however in monocots such as the *Poaceae* these GlcA substitutions are uncommon [35]. In contrast arabinosyl substitutions are far more frequent in the *Poaceae* yet have not been reported for *Arabidopsis*. These arabinose groups can be α-1,2 or α-1,3-linked to the xylan backbone and may be further substituted with arabinose, coumaric acid, or ferulic acid groups (Fig.5)[31]. Here both coumaric and ferulic acid are linked to the O-5 of the arabinose side chains of arabinoxylan [36,37].

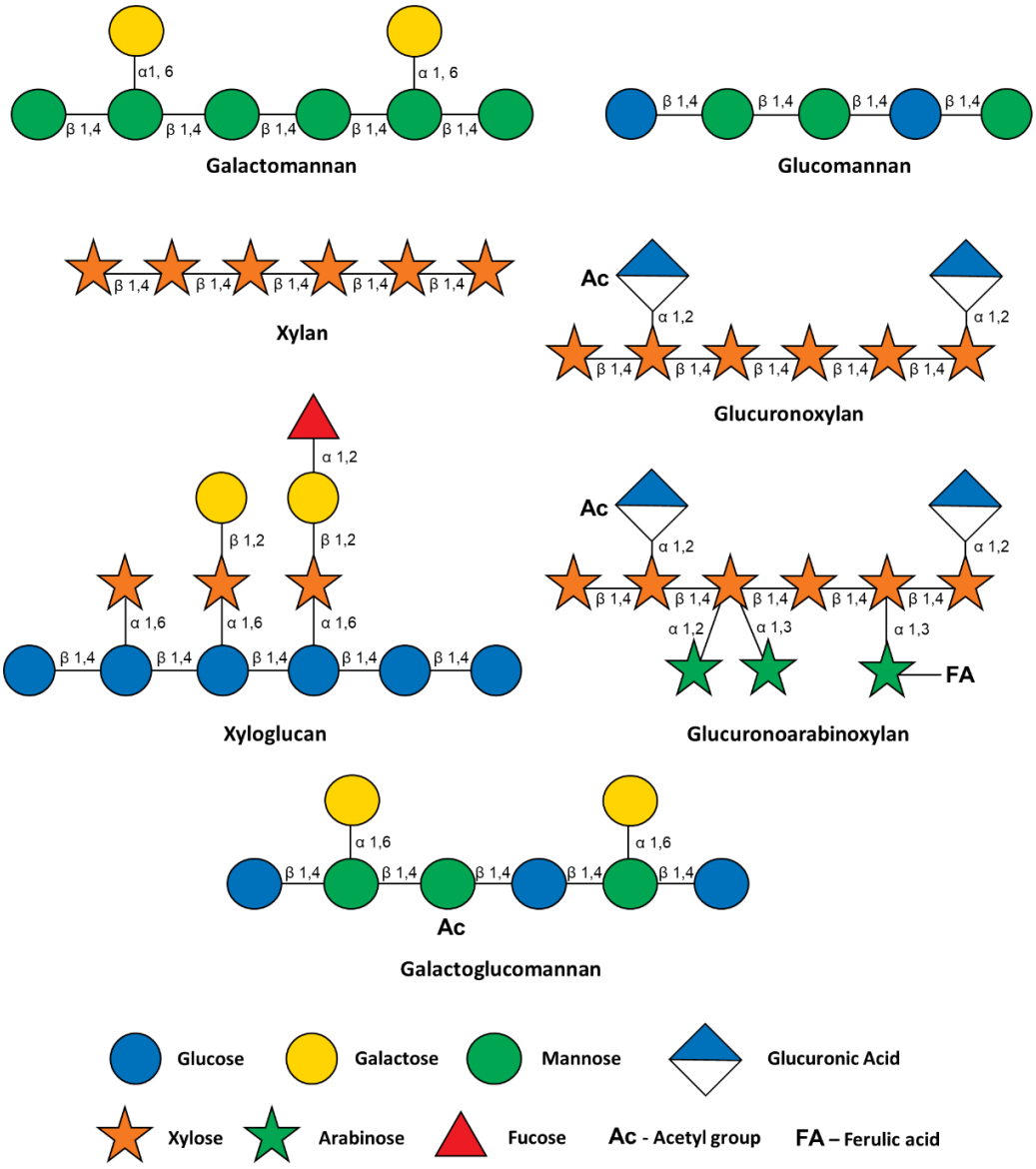


Figure 5: Hemicelluloses found in the plant cell wall

The other predominant hemicelluloses are xyloglucan, galactomannan, glucomannan, galactoglucomannan. Xyloglucan is a formed of β-1,4–linked glucose backbone onto which the majority of glucosyl residues are substituted with α-1,6 xylosyl residues. Multiple, further substitutions can be present, which are often (β-1,2) galactose or (α-1,2) fucose – the nature of such substitutions vary between both tissue type or plant species. Acetylation is found within xyloglucans, with acetyl groups present on non-xylose-linked backbone glucosyl residues or on side chain galactosyl and arabinofuranosyl residues [38]. Galactomannan has a more simple structure, consisting of a β-1,4-linked mannose backbone with ~ 25–50 %

(depending on species source) of residues substituted with a α-1,6-linked galactose [39]. Glucomannan is formed of a linear, β-1,4-linked mannose and glucose (1.6:1) backbone and has a low frequency of β-1,6- glucose side chains [40]. Glucogalactomannan comprises a very similar structure with the exception of α-1,6 galactosyl side chains [41]. Hemicelluloses are synthesised within the golgi then transported to the plasma membrane for integration into the secondary cell wall [27].

Structure of plant cell wall phenolic compounds

Plants are also the source of the second most common biopolymer – lignin. The organic polymer is key to strengthening the cell wall and was fundamental in the evolution of vascular plants 400 million years ago [42]. The cross linked phenolic polymer lignin is comprised of 3 monomeric precursors called monolignols: p-hydroxylphenyl (H), guaiacyl (G) and syringyl (S) which are formed in the cytoplasm and are initially derived from phenylalanine [43]. The structure of lignin varies widely across the Tracheophyta (vascular plants), showing variation not only between species but also within different tissues of the same plant. This is partially due to the different composition of the 3 monolignols. G- lignin (35-49% of total lignin) and S- lignin (40-61% of total lignin) are the predominant monolignols in both monocots and dicots however H-lignin is particularly prevalent in the *Poaceae* (the grasses) often forming up to 15% of the total lignin content [44].

Unlike cellulose, which is synthesised at the plasma membrane then added to the cell wall, lignin biosynthesis predominantly takes place within the plant cell wall. Here the activity of wall bound laccases and peroxidases (either singly or in tandem) results in the production of free radicals which catalyse lignin polymerisation [45]. The exact route precursors take from the cytoplasm and subsequent translocation to the cell wall is currently unclear, however with the great variety seen within and amongst plant lignin the probability of multiple biosynthesis pathways is likely [46].

Biotechnological potential of lignocellulose-degrading microorganisms

Plant cell walls can be synthesized, modified and degraded via carbohydrate active enzymes from plant or microbial sources. Fungi such as *Aspergillus* have been utilised in biotechnological settings for over a century, for example *A. niger* has

been known as a producer of citric acid since 1917 [47], with the fungus first employed on an industrial scale in 1919 [48]. The further potential of *Aspergillus* as a producer of lignocellulose degrading enzymes was realised with the discovery of multiple oxidases, hydrolases, cellulases and pectinases. *A. niger* glucoamylase (responsible for turning starch to glucose) has been generally regarded as safe (GRAS) for several decades and plays a major role in the fermentation and food industries [48]. Similarly, *Trichodermareesei* has become a frequently used tool for lignocellulose degradation, its degradative potential was realised early on and subsequently the first cloning of a eukaryotic cellulase was CBHI from *T. reesei* [49].The advent of mutagenesis led to the production of numerous mutants with increased cellulase production, many of which have been adopted for use in the textiles, paper and food industries [50]. As microbial cell factory, *T. reesei* was one of the first expression hosts for heterologous production of mammalian proteins [51]. Other microbes, such as the bacterium *Thermobacillus xylanilyticus,* offer desirable traits required by industry, such as enzymes with high thermotolerance for reactions requiring higher temperatures [52]. These examples of the varied range of lignocellulose-degrading capabilities by microorganisms and their enzymes highlight their potential utility in biocatalysis — the synthesis of desired organic products by enzymatic means.

Glycoside hydrolases

Glycoside hydrolases (GHs), catalyse the breaking of glycosidic bonds in oligo-/polysaccharides and glycoconjugates [53]. Conversely, under certain conditions such enzyme can also form glycosidic bonds. Despite much research, large quantities of GHs only have predicted functions leaving great, untapped potential for biotechnology. The vast variety of GHs is currently organised into 168 separate families that are categorised on the basis of sequence conservation and the similarity of structural motifs, the Carbohydrate Active Enzyme database (CAZy) has provided an essential service in grouping these enzymes into a relevant families in a systematic manner [54].

GHs catalyse glycosidic bond hydrolysis, which occurs by action of general acid catalysis. Such catalysis is dependent on two essential residues within a GH active

site, that of a proton donor and a nucleophile/base [55]. GHs are grouped according to the whether the anomeric configuration of the hydrolysed glycoside is retained or inverted. Retaining GHs demonstrate a double displacement mechanism whereby nucleophilic attack of the anomeric centre by an amino acid residue, supported by another residue acting as an acid/base creates a glycosyl-enzyme-intermediate. This is followed by hydrolysis of the glycosyl-enzyme-intermediate by a nucleophilic water (assisted by the acidic carboxylate from the previous step) resulting in retention of the anomeric configuration (Fig. 6a). Inverting GHs, however, employ a single displacement mechanism in which two amino acid residues function separately as acid and base (Fig. 6b) resulting in the inversion of the anomeric configuration.

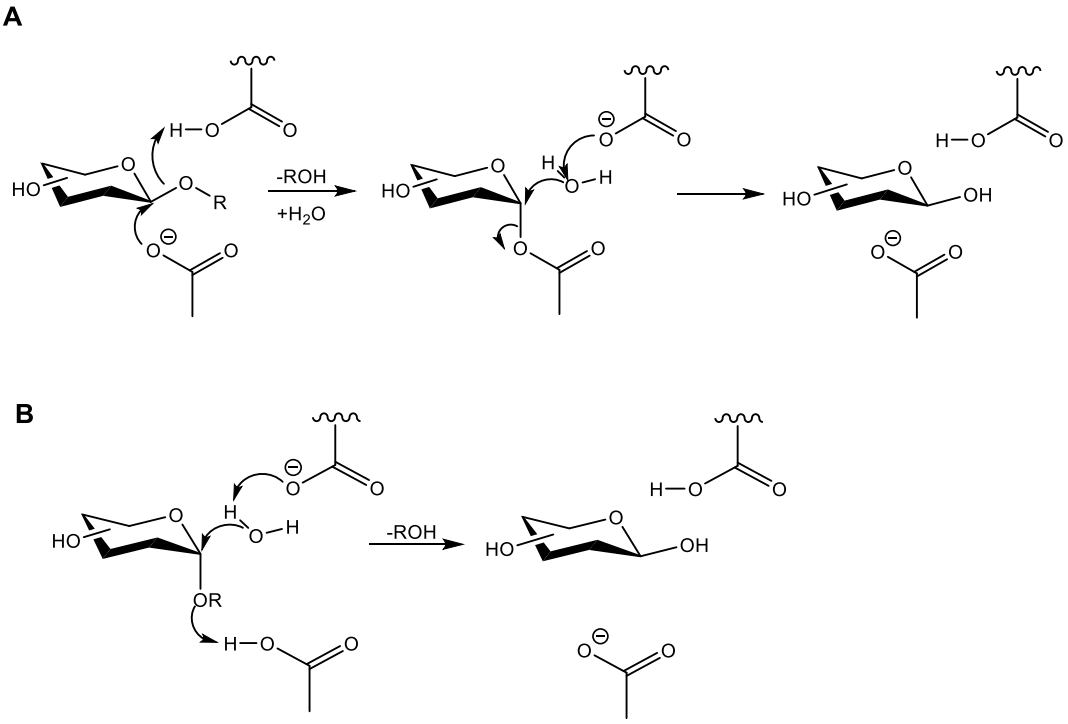


Figure 6: Mechanisms of retaining (A) and inverting (B) glycosylhydrolases.

Enzymatic degradation of cellulose

As the structures of plant polysaccharides are elucidated the use of microorganisms and their enzymes to degrade complex plant carbohydrates (Table 1) in an economically favourable manner is increasingly becoming a viable option. Glycohydrolase enzymes that degrade cellulose consist of endocellulases (through cleaving within the cellulose chain) and exocellulases (those that cleave at the end

of the cellulose chain or degrade cellobiose). The degradation of cellulose requires the action of 3 types of GHs: β-1,4-endoglucanases (EGL), cellobiohydrolases (CBH) and β- glucosidase (BGL). EGLs cleave cellulose into glucooligosaccharides,CBHs cleave cellulose in a processive manner, releasing mainly cellobiose from the ends of the cellulose chain, whilst BGLsare responsible for the final step of degrading cellobiose into glucose.

Lytic polysaccharide monooxygenases (LPMO) are enzymes capable of breaking the glycosidic bonds within cellulose via oxidative cleavage [56]thus differing from the usual GH-mediated hydrolytic cleavage by using molecular oxygen. Although initially described in microorganisms, LPMOs encompass a variety of sources including crustaceans, molluscs and algae [57]. The breakage of the β-1,4-linkage in an oxidative manner makes the cellulose substrate more susceptible to hydrolysis by GHs [58]. As such, LPMOs allow substantial increases in the efficiency of commercial enzymatic cocktails for the degradation of lignocellulose and have become an area of intensive research since their discovery in the early 2010s [59].

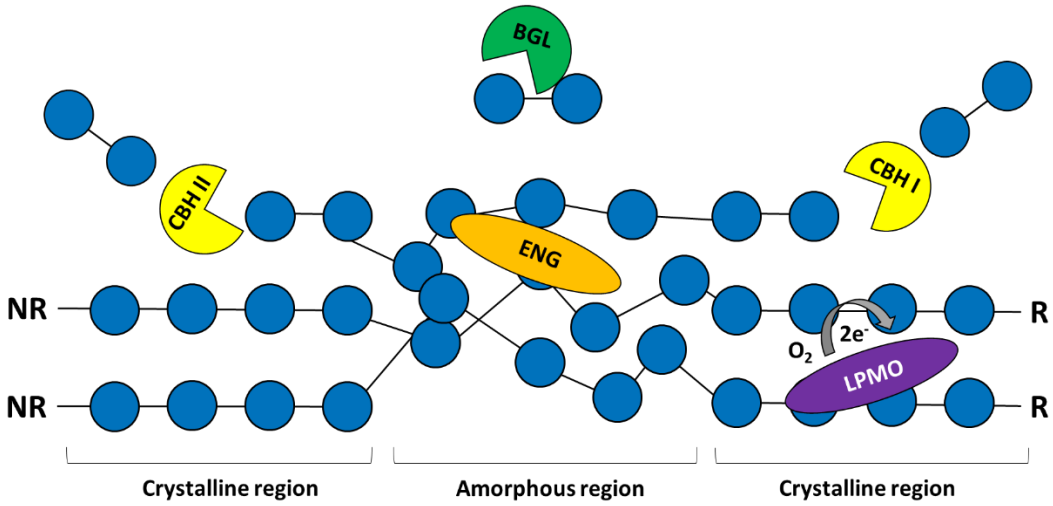


Figure 7: Cellulose degrading enzymes. Endoglucanase (ENG) cleaves in the amorphous regions of the cellulose chain. Cellobiohydrolase I (CBH I) cleaves cellobiose from the reducing end (R) whilst Cellobiohydrolase II (CBH II) cleaves cellobiose from the non-reducing end (NR). Cellobiose is then cleaved by β-glucosidases (BGL). LPMOs cleave the cellulose chain by oxidative cleavage that harnesses molecular oxygen and electrons.

Enzymatic degradation of hemicelluloses

For the degradation of linear xylan the presence of 2 GH enzymes are required: β-1,4-endoxylanase and β-1,4-xylosidase. Here endoxylanases cleave within the xylan backbone thus producing a variety of xylooligosaccharides (van den Brink and de Vries, 2011). Although endo-β-1,4 xylanases are found in multiple GH families, most reside in GH families 10 and 11 [60]. GH 10 and GH 11 family members differ by the structure of the catalytic domain, with the former containing a TIM-barrel fold and the latter a β-jelly roll structure [61,62]. This structural difference is variable in substrate specificity between the 2 families, with GH10 demonstrating broader substrate specificity whilst GH11 endoxylanses demonstrate preference for unsubstituted xylan. GH 10 family endoxylanases can break glycosidic bonds of the xylose backbone at sites closer to substituted side groups, whereas GH11 cannot cleave this close to decorations meaning the products are often longer [63]. After xyloogliosaccharides and xylobiose are released by endoxylanses they can then be further degraded by the action of β-xylosidases. β-xylosidases hydrolyse xylobiose and xylooligosaccharides at the non-reducing end to produce xylose [64].

Previously LMPOs had been shown to degrade substrates such as the glucose based cellulose but not xylans [65]. However, recently an LMPO from the basidiomycete *Pycnoporus coccineus* was shown to target xylan covering cellulose fibres thus offering an expanded variety of enzymes available to combat the recalcitrance of lignocellulostic material [66].

α-arabinofuranosidases remove L-arabinose from arabinoxylan and arabinoxylan oligosaccharides. Arabinofuranosidases belong to multiple GH families and are capable of removing arabinose residues from xylose oligosaccharides, although only members of GH 43 and 51 are able to remove both mono- and di-substituted arabinose units [67]. Chapter 5 describes the characterisation of a GH43 enzyme that is expressed by *A. niger* during its exposure to plant biomass. α-glucuronidases hydrolyse α-1,2-linked GlcA/4-O-MeGlcA from glucuronoxylan and arabinoglucuronoxylan. GH families show different substrate preference, with GH 67 active against short GlcA-substitued xylooligosaccharides and GH 155 active against larger, polymeric xylans decorated with GlcA [69]

Table 1: (Hemi)cellulases active against plant biomass, adapted from [68]

Enzyme	EC Family	CAZy family	Target	Product(s)
Cellulases				
Endoglucanase	EC 3.2.1.4	GH7	Randomly on the cellulose chain	Cello–oligomers
Cellobiohydrolase I	EC 3.2.1.91	GH5, GH6, GH7, GH8, GH9, GH10, GH12, GH44, GH45, GH48, GH51, GH74, GH124	Reducing end of cellulose	Cellobiose
Cellobiohydrolase II	EC 3.2.1.176	GH7, GH48	Non-reducing end of cellulose	Cellobiose
β–glucosidase	EC 3.2.1.21	GH1, GH2, GH3, GH5, GH30, GH39, GH116	Cellobiose	Glucose
LPMO	EC1.14.99.54, EC 1.14.99.56	AA9, AA10, AA15, AA16	C1 or C4 of glucose	Aldonic acid
Depolymerisation Hemicellulases				
Endo–1,4–β–xylanase	EC 3.2.1.8	GH5, GH8, GH10, GH11, GH30, GH43, GH51, GH98, GH141	β–1,4 xylan	Xylo–oligomers
β–xylosidase	EC 3.2.1.37	GH1, GH2, GH3, GH30, GH39, GH43, GH51, GH52, GH54, GH116, GH120	Exo β–1,4 xylo–oligomers	Xylose
Endo–1,4–mannanase	EC 3.2.1.78	GH5, GH26, GH45, GH113, GH134	β–1,4 mannan	Mannan oligomers
β–Mannosidase	EC 3.2.1.25	GH1, GH2, GH5, GH164	β–1,4 mannans oligomers	Mannose

Accessory Hemicellulases				
α–L–Arabinofuranosidase	EC 3.2.1.55	GH2 GH3 GH5 GH39 GH43 GH51 GH54 GH62	Positions O-2 and O-3 of α–L–arabinofuranosides	L–arabinose
α–L–Arabinanase	EC 3.2.1.99	GH43	α–1,5–arabinan	Arabinose
α–Galactosidases	EC 3.2.1.22	GH4, GH27, GH31, GH36, GH57, GH97, GH110	Oligosaccharides with α–galactosyl groups	Galactose and sucrose
α–D–Glucuronidase	EC 3.2.1.139	GH4, GH67	α–1,2–glycoside between 4–O–methyl–glucuronic acid and xylo–oligomers	Glucuronic acid
Acetyl xylan esterase	EC 3.1.1.72	CE1, CE2, CE3, CE4, CE5, CE6, CE7, CE12	Positions 2– or 3 of –O–acetyl xylan	Acetic acid
Feruloyl xylan esterase	E. C.3.1.1.73	CE1	Ferulic acid substitutions	Ferulic acid
Glucuronyl esterase	EC 3.2.1.55	GH2, GH3, GH5, GH39, GH43, GH51, GH54, GH62	Positions O-2 and O-3 of α–L–arabinofuranosides	L–arabinose

Enzymatic degradation of phenolic containing compounds

The crosslinking of lignin with hemicelluloses is a major barrier to the efficient degradation of the plant cell wall. Carbohydrate esterases (CEs) catalyze the de-O-acylation/de-N-acylation by removing the ester-linked decorations from carbohydrates and assist in this degradative process [70]. This varied group includes acetylxylan esterases, feruloyl esterases and glucuronoyl esterases [71]. Glucuronoyl esterases belong to CE family 15 and are responsible for the cleavage of ester bonds between GlcA/4-O-MeGlcAresidues in glucuronoxylan and the alcohol functional groups of lignin [72]. Acetylxylanesterases remove acetyl residues from the O-2 or O-3 positions of xylose in the xylan backbone and are often essential for the degradation of xylan [73]. Feruloyl esterases (FAE) catalyse

the cleavage of the ester bonds between hydroxycinnamates (e.g. ferulic acid or *p*-coumaric acid) and arabinoxylan [74]. Chapter 6 describes the biochemical characterisation of a feruloyl esterase.

Chemo-enzymatic synthesis of glycans

Humans have exploited the activities of enzymes for 1000s of years in processes such as the brewing of beer, baking of bread, cheese making and the tanning of leather [75]. In the 20th century, enzymes were employed more extensively in chemical production, in processes such as the synthesis of acrylamide [76]. Biocatalysis is the use of enzymes for the production of organic compounds and offers a greener route than a multitude of traditional processes. More recently, enzymes have been used for the synthesis of more complex molecules, such as pharmaceuticals. Advances in the high-throughput screening of enzyme activities and their subsequent directed evolution has caused rapid expansion in the scope in the role of enzymes in organic synthesis [77]. The regioselectivity, chemoselectivity and stereoselectivity of enzymes, combined with their ability to undertake reactions in water and mild conditions enables them to fulfil many of the principles of green chemistry [78,79]. Many reactions are amenable to one-pot enzymatic cascades in which multiple enzymes can undertake reactions simultaneously within the same reaction, greatly reducing the time and financial constraints present with step-wise synthesis [80]. The abundance of GHs in combination with a wide scope of substrate specificities and affordable substrates makes them desirable candidates for biocatalysts. By understanding the activity and specificity of enzymes active against plant lignocellulose, biocatalytic routes to desired carbohydrate-derived products can be envisaged. In addition to their normal activities as degradative enzyme, many GHs are also capable of functioning in the reverse direction through either the displacement of equilibrium (thermodynamic control) or through the use of activated sugar donors that intercept the reactive glycosyl-enzyme intermediate via kinetically controlled transglycosylation[81]. These methods offer opportunities to leverage the large natural repertoire of GH specificities, and enable routes for the enzymatic synthesis of designer glycans.

References

[1] A. Anadón, M.R. Martínez-Larrañaga, I. Ares, M.A. Martínez, Prebiotics: Safety and Toxicity Considerations, Nutraceuticals Effic. Saf. Toxic. (2016) 757–775.

[2] J. Liu, S. Willför, C. Xu, A review of bioactive plant polysaccharides: Biological activities, functionalization, and biomedical applications, Bioact. Carbohydrates Diet. Fibre. 5 (2015) 31–61.

[3] V. Zargar, M. Asghari, A. Dashti, A Review on Chitin and Chitosan Polymers: Structure, Chemistry, Solubility, Derivatives, and Applications, ChemBioEng Rev. 2 (2015) 204–226.

[4] E. Bertoft, Understanding starch structure: Recent progress, Agronomy. 7 (2017).

[5] C. Prats, T.E. Graham, J. Shearer, The dynamic life of the glycogen granule, J. Biol. Chem. 293 (2018) 7089–7098.

[6] Y.Y. Zhao, M. Takahashi, J.G. Gu, E. Miyoshi, A. Matsumoto, S. Kitazume, N. Taniguchi, Functional roles of N-glycans in cell signaling and cell adhesion in cancer, Cancer Sci. 99 (2008) 1304–1310.

[7] A. Varki, Biological roles of glycans, Glycobiology. 27 (2017) 3–49.

[8] A. Dell, A. Galadari, F. Sastre, P. Hitchen, Similarities and differences in the glycosylation mechanisms in prokaryotes and eukaryotes, Int. J. Microbiol. (2010)148178.

[9] H.B. Aditiya, T.M.I. Mahlia, W.T. Chong, H. Nur, A.H. Sebayang, Second generation bioethanol production: A critical review, Renew. Sustain. Energy Rev. 66 (2016) 631–653.

[10] A. Di Bartolo, G. Infurna, N.T. Dintcheva, A review of bioplastics and their adoption in the circular economy, Polymers (Basel). 13 (2021) 1229.

[11] R. Jefferis, Glycosylation as a strategy to improve antibody-based therapeutics, Nat. Rev. Drug Discov. 8 (2009) 226–234.

[12] C. Reily, T.J. Stewart, M.B. Renfrow, J. Novak, Glycosylation in health and disease, Nat. Rev. Nephrol. 2019 156. 15 (2019) 346–366.

[13] C.O. Tuck, E. Pérez, I.T. Horváth, R.A. Sheldon, M. Poliakoff, Valorization of biomass: Deriving more value from waste, Science. 337 (2012) 695–699.

[14] L. Lange, K.O. Connor, S. Arason, U. Bundgård-Jørgensen, A. Canalis, D. Carrez, J. Gallagher, N. Gøtke, C. Huyghe, B. Jarry, P. Llorente, M. Marinova, L.O. Martins, P. Mengal, P. Paiano, C. Panoutsou, L. Rodrigues, D.B. Stengel, Y. van der Meer, H. Vieira, Developing a Sustainable and Circular Bio-Based Economy in EU: By Partnering Across Sectors, Upscaling and Using New Knowledge Faster, and For the Benefit of Climate, Environment & Biodiversity, and People & Business, Front. Bioeng. Biotechnol. (2021) 8:619066.

[15] P. Ning, G. Yang, L. Hu, J. Sun, L. Shi, Y. Zhou, Z. Wang, J. Yang, Recent advances in the valorization of plant biomass, Biotechnol. Biofuels 2021 141. 14 (2021) 1–22.

[16] A. Thorenz, L. Wietschel, D. Stindt, A. Tuma, Assessment of agroforestry residue potentials for the bioeconomy in the European Union, J. Clean. Prod. 176 (2018) 348–359.

[17] T.L. Bezerra, A.J. Ragauskas, A review of sugarcane bagasse for second-generation bioethanol and biopower production, Biofuels, Bioprod. Biorefining. 10 (2016) 634–647.

[18] D. Trache, A.F. Tarchoun, M. Derradji, T.S. Hamidon, N. Masruchin, N. Brosse, M.H. Hussin, Nanocellulose: From Fundamentals to Advanced Applications, Front. Chem. 8(2020) 392.

[19] A.D. Roberts, K.A.P. Payne, S. Cosgrove, V. Tilakaratna, I. Penafiel, W. Finnigan, N.J. Turner, N.S. Scrutton, Enzyme immobilisation on wood-derived cellulose scaffolds via carbohydrate-binding module fusion constructs, Green. Chem. 23 (2021) 4716–4732.

[20] B. MacCormick, T. V. Vuong, E.R. Master, Chemo-enzymatic Synthesis of

Clickable Xylo-oligosaccharide Monomers from Hardwood 4-O-Methylglucuronoxylan, Biomacromolecules. 19 (2018) 521–530.

[21] A. Martínez-Abad, A.C. Ruthes, F. Vilaplana, Enzymatic-assisted extraction and modification of lignocellulosic plant polysaccharides for packaging applications, J. Appl. Polym. Sci. 133 (2016) 1-15.

[22] M. Valdivia, J.L. Galan, J. Laffarga, J.L. Ramos, Biofuels 2020: Biorefineries based on lignocellulosic materials, Microb. Biotechnol. 9 (2016) 585–594.

[23] C. Loix, M. Huybrechts, J. Vangronsveld, M. Gielen, E. Keunen, A. Cuypers, Reciprocal Interactions between Cadmium-Induced Cell Wall Responses and Oxidative Stress in Plants, Front. Plant Sci. 8 (2017) 1867.

[24] B. Zhang, Y. Gao, L. Zhang, Y. Zhou, The plant cell wall: Biosynthesis, construction, and functions, J. Integr. Plant Biol. 63 (2021) 251–272.

[25] S. Wolf, S. Greiner, Growth control by cell wall pectins, Protoplasma 2012 2492. 249 (2012) 169–175.

[26] S. Escamez, M.L. Gandla, M. Derba-Maceluch, S.-O. Lundqvist, E.J. Mellerowicz, L.J. Jönsson, H. Tuominen, A collection of genetically engineered Populus trees reveals wood biomass traits that predict glucose yield from enzymatic hydrolysis, Sci. Reports 2017 71. 7 (2017) 1–11.

[27] M. Kumar, L. Campbell, S. Turner, Secondary cell walls: Biosynthesis and manipulation, J. Exp. Bot. 67 (2016) 515–531.

[28] H. Allen, D. Wei, Y. Gu, S. Li, A historical perspective on the regulation of cellulose biosynthesis, Carbohydr. Polym. 252 (2021) 117022.

[29] M. Ochoa-Villarreal, E. Aispuro-Hernndez, I. Vargas-Arispuro, M. Angel, Plant Cell Wall Polymers: Function, Structure and Biological Activity of Their Derivatives, Polymerization. (2012).

[30] P.K. Gupta, S.S. Raghunath, D.V. Prasanna, P. Venkat, V. Shree, C. Chithananthan, S. Choudhary, K. Surender, K. Geetha, An Update on Overview of Cellulose, Its Structure and Applications, Cellulose. (2019).

[31] E.A. Rennie, H.V. Scheller, Xylan biosynthesis, Curr. Opin. Biotechnol. 26 (2014) 100–107.

[32] P.J. Smith, H.-T. Wang, W.S. York, M.J. Peña, B.R. Urbanowicz, Designer biomass for next-generation biorefineries: leveraging recent insights into xylan structure and biosynthesis, Biotechnol. Biofuels 2017 101. 10 (2017) 1–14.

[33] A. Ebringerová, T. Heinze, Xylan and xylan derivatives - Biopolymers with valuable properties, 1: Naturally occurring xylans structures, isolation procedures and properties, Macromol. Rapid Commun. 21 (2000) 542–556.

[34] E. Deniaud, B. Quemener, J. Fleurence, M. Lahaye, Structural studies of the mix-linked-(1 → 3)/-(1 → 4)-d-xylans from the cell wall of Palmaria palmata (Rhodophyta), Int. J. Biol. Macromol. 33 (2003) 9–18.

[35] J.R. Bromley, M. Busse-Wicher, T. Tryfona, J.C. Mortimer, Z. Zhang, D.M. Brown, P. Dupree, GUX1 and GUX2 glucuronyltransferases decorate distinct domains of glucuronoxylan with different substitution patterns, Plant J. 74 (2013) 423–434.

[36] S. Mathew, T.E. Abraham, Ferulic Acid: An Antioxidant Found Naturally in Plant Cell Walls and Feruloyl Esterases Involved in its Release and Their Applications, Crit. Rev. Biotechnol. 24 (2008) 59–83.

[37] M.M. Marcia, Feruloylation in Grasses: Current and Future Perspectives, Mol. Plant. 2 (2009) 861–872.

[38] M. Pauly, K. Keegstra, Biosynthesis of the Plant Cell Wall Matrix Polysaccharide Xyloglucan∗, Annu. Rev. Plant Biol. 67 (2016) 235–259.

[39] V.D. Prajapati, G.K. Jani, N.G. Moradiya, N.P. Randeria, B.J. Nagar, N.N. Naikwadi, B.C. Variya, Galactomannan: A versatile biodegradable seed polysaccharide, Int. J. Biol. Macromol. 60 (2013) 83–92.

[40] K. Katsuraya, K. Okuyama, K. Hatanaka, R. Oshima, T. Sato, K. Matsuzaki, Constitution of konjac glucomannan: Chemical analysis and 13C NMR

spectroscopy, Carbohydr. Polym. 53 (2003) 183–189.

[41] H.V. Scheller, P. Ulvskov, Hemicelluloses, Annu. Rev. Plant Biol. 61 (2010) 263–289.

[42] W. Boerjan, J. Ralph, M. Baucher, Lignin Biosynthesis, Annu. Rev. Plant Biol. 54 (2003) 519–565.

[43] N.D. Bonawitz, C. Chapple, The Genetics of Lignin Biosynthesis: Connecting Genotype to Phenotype, Annu. Rev. Genet. 44 (2010) 337–363.

[44] J. Vogel, Unique aspects of the grass cell wall, Curr. Opin. Plant Biol. 11 (2008) 301–307.

[45] J. Ralph, K. Lundquist, G. Brunow, F. Lu, H. Kim, P.F. Schatz, J.M. Marita, R.D. Hatfield, S.A. Ralph, J.H. Christensen, W. Boerjan, Lignins: Natural polymers from oxidative coupling of 4-hydroxyphenyl-propanoids, Phytochem. Rev. 3 (2004) 29–60.

[46] C.J. Liu, Y.C. Miao, K.W. Zhang, Sequestration and transport of lignin monomeric precursors, Molecules. 16 (2011) 710–727.

[47] J.N. Currie, THE CITRIC ACID FERMENTATION OF ASPERGILLUS NIGER, J. Biol. Chem. 31 (1917) 15–37.

[48] T.C. Cairns, C. Nai, V. Meyer, How a fungus shapes biotechnology: 100 years of Aspergillus niger research, Fungal Biol. Biotechnol. 2018 51. 5 (2018) 1–14.

[49] S. Shoemaker, V. Schweickart, M. Ladner, D. Gelfand, S. Kwok, K. Myambo, M. Innis, Molecular Cloning of Exo–Cellobiohydrolase I Derived from Trichoderma Reesei Strain L27, Bio/Technology 1983 18. 1 (1983) 691–696.

[50] R.H. Bischof, J. Ramoni, B. Seiboth, Cellulases and beyond: the first 70 years of the enzyme producer Trichoderma reesei, Microb. Cell Factories 2016 151. 15 (2016) 1–13.

[51] A. Harkki, J. Uusitalo, M. Bailey, M. Penttilä, J.K.C. Knowles, A Novel Fungal Expression System: Secretion of Active Calf Chymosin from the Filamentous Fungus Trichoderma Reesei, Bio/Technology 1989 76. 7 (1989) 596–603.

[52] H. Rakotoarivonina, B. Hermant, N. Monthe, C. Rémond, The hemicellulolytic enzyme arsenal of Thermobacillus xylanilyticus depends on the composition of biomass used for growth, Microb. Cell Factories 2012. 11 (2012) 159.

[53] P.M. Danby, S.G. Withers, Advances in Enzymatic Glycoside Synthesis, ACS Chem. Biol. 11 (2016) 1784–1794.

[54] B.I. Cantarel, P.M. Coutinho, C. Rancurel, T. Bernard, V. Lombard, B. Henrissat, The Carbohydrate-Active EnZymes database (CAZy): An expert resource for glycogenomics, Nucleic Acids Res. 37 (2009).

[55] D.E. KOSHLAND, Stereochemistry and the Mechanism of Enzymatic Reactions, Biol. Rev. 28 (1953) 416–436.

[56] V.G.H. Eijsink, D. Petrovic, Z. Forsberg, S. Mekasha, Å.K. Røhr, A. Várnai, B. Bissaro, G. Vaaje-Kolstad, On the functional characterization of lytic polysaccharide monooxygenases (LPMOs), Biotechnol. Biofuels 2019 121. 12 (2019) 1–16.

[57] F. Sabbadin, G.R. Hemsworth, L. Ciano, B. Henrissat, P. Dupree, T. Tryfona, R.D.S. Marques, S.T. Sweeney, K. Besser, L. Elias, G. Pesante, Y. Li, A.A. Dowle, R. Bates, L.D. Gomez, R. Simister, G.J. Davies, P.H. Walton, N.C. Bruce, S.J. McQueen-Mason, An ancient family of lytic polysaccharide monooxygenases with roles in arthropod development and biomass digestion, Nat. Commun. 2018 91. 9 (2018) 1–12.

[58] L. Lo Leggio, T.J. Simmons, J.-C.N. Poulsen, K.E.H. Frandsen, G.R. Hemsworth, M.A. Stringer, P. von Freiesleben, M. Tovborg, K.S. Johansen, L. De Maria, P. V. Harris, C.-L. Soong, P. Dupree, T. Tryfona, N. Lenfant, B. Henrissat, G.J. Davies, P.H. Walton, Structure and boosting activity of a starch-degrading lytic polysaccharide monooxygenase, Nat. Commun. 2015 61. 6 (2015) 1–9.

[59] G. Müller, P. Chylenski, B. Bissaro, V.G.H. Eijsink, S.J. Horn, The impact of hydrogen peroxide supply on LPMO activity and overall saccharification efficiency of a commercial cellulase cocktail, Biotechnol. Biofuels. 11 (2018) 1–17.

[60] V. Puchart, K. Šuchová, P. Biely, Xylanases of glycoside hydrolase family 30 – An overview, Biotechnol. Adv. 47 (2021) 107704.

[61] L. Liu, X. Sun, P. Yan, L. Wang, H. Chen, Non-Structured Amino-Acid Impact on GH11 Differs from GH10 Xylanase, PLoS One. 7 (2012)e45762.

[62] A. Pollet, J.A. Delcour, C.M. Courtin, Structural determinants of the substrate specificities of xylanases from different glycoside hydrolase families, Crit. Rev. Biotechnol. 30 (2010) 176–191.

[63] P. Biely, M. Vršanská, M. Tenkanen, D. Kluepfel, Endo-β-1,4-xylanase families: Differences in catalytic properties, J. Biotechnol. 57 (1997) 151–166.

[64] A. Rohman, B.W. Dijkstra, N.N.T. Puspaningsih, β-xylosidases: Structural diversity, catalytic mechanism, and inhibition by monosaccharides, Int. J. Mol. Sci. 20 (2019) 7–11.

[65] Y. Kojima, A. Várnai, T. Ishida, N. Sunagawa, D.M. Petrovic, K. Igarashi, J. Jellison, B. Goodell, G. Alfredsen, B. Westereng, V.G.H. Eijsink, M. Yoshida, A lytic polysaccharide monooxygenase with broad xyloglucan specificity from the brown-rot fungus Gloeophyllum trabeum and its action on cellulose-xyloglucan complexes, Appl. Environ. Microbiol. 82 (2016) 6557–6572.

[66] M. Couturier, S. Ladevèze, G. Sulzenbacher, L. Ciano, M. Fanuel, C. Moreau, A. Villares, B. Cathala, F. Chaspoul, K.E. Frandsen, A. Labourel, I. Herpoël-Gimbert, S. Grisel, M. Haon, N. Lenfant, H. Rogniaux, D. Ropartz, G.J. Davies, M.N. Rosso, P.H. Walton, B. Henrissat, J.G. Berrin, Lytic xylan oxidases from wood-decay fungi unlock biomass degradation, Nat. Chem. Biol. 14 (2018) 306–310.

[67] S. Lagaert, A. Pollet, C.M. Courtin, G. Volckaert, β-Xylosidases and α-L-arabinofuranosidases: Accessory enzymes for arabinoxylan degradation, Biotechnol. Adv. 32 (2013) 316-332.

[68] A.A. Houfani, N. Anders, A.C. Spiess, P. Baldrian, S. Benallaoua, Insights from enzymatic degradation of cellulose and hemicellulose to fermentable sugars–a review, Biomass and Bioenergy. 134 (2020) 105481.

[69] S. Malgas, M.S. Mafa, L. Mkabayi, B.I. Pletschke, A mini review of xylanolytic enzymes with regards to their synergistic interactions during hetero-xylan degradation, World J. Microbiol. Biotechnol. 35 (2019) 187.

[70] A.M. Nakamura, A.S. Nascimento, I. Polikarpov, Structural diversity of carbohydrate esterases, Biotechnol. Res. Innov. 1 (2017) 35–51.

[71] B. P, Microbial carbohydrate esterases deacetylating plant polysaccharides, Biotechnol. Adv. 30 (2012) 1575–1588.

[72] R.N. Monrad, J. Eklöf, K.B.R.M. Krogh, P. Biely, Glucuronoyl esterases: diversity, properties and biotechnological potential. A review, Crit. Rev. Biotechnol. 38 (2018) 1121–1136.

[73] M.A. Kabel, C.J. Yeoman, Y. Han, D. Dodd, C.A. Abbas, J.A.M. de Bont, M. Morrison, I.K.O. Cann, R.I. Mackie, Biochemical Characterization and Relative Expression Levels of Multiple Carbohydrate Esterases of the Xylanolytic Rumen Bacterium Prevotella ruminicola 23 Grown on an Ester-Enriched Substrate, Appl. Environ. Microbiol. 77 (2011) 5671–5681.

[74] E.N. Underlin, M. Frommhagen, A. Dilokpimol, G. van Erven, R.P. de Vries, M.A. Kabel, Feruloyl Esterases for Biorefineries: Subfamily Classified Specificity for Natural Substrates, Front. Bioeng. Biotechnol. 8 (2020) 332.

[75] C.M. Heckmann, F. Paradisi, Looking Back: A Short History of the Discovery of Enzymes and How They Became Powerful Chemical Tools, ChemCatChem. 12 (2020) 6082–6102.

[76] O. Kirk, T.V. Borchert, C.C. Fuglsang, Industrial enzyme applications, Curr. Opin. Biotechnol. 13 (2002) 345–351.

[77] E.L. Bell, W. Finnigan, S.P. France, A.P. Green, M.A. Hayes, L.J. Hepworth, S.L. Lovelock, H. Niikura, S. Osuna, E. Romero, K.S. Ryan, N.J. Turner, S.L. Flitsch, Biocatalysis, Nat. Rev. Methods Prim. 2021 11. 1 (2021) 1–21.

[78] R.A. Sheldon, J.M. Woodley, Role of Biocatalysis in Sustainable Chemistry, Chem. Rev. 118 (2018) 801–838.

[79] R.A. Sheldon, D. Brady, M.L. Bode, The Hitchhiker's guide to biocatalysis:

recent advances in the use of enzymes in organic synthesis, Chem. Sci. 11 (2020) 2587–2605.

[80] N. Losada-Garcia, Z. Cabrera, P. Urrutia, C. Garcia-Sanz, A. Andreu, J.M. Palomo, Cascade Processes, Catalysts. 10 (2020) 1–19.

[81] S.R. Marsden, L. Mestrom, D.G.G. McMillan, U. Hanefeld, Thermodynamically and Kinetically Controlled Reactions in Biocatalysis – from Concepts to Perspectives, ChemCatChem. 12 (2020) 426–437.

Chapter 2: Project objectives

Carbohydrate Active Enzymes (CAZymes) are responsible for the synthesis and degradation of carbohydrates and are routinely exploited for the production of specific structures or for the valorisation of plant biomass. However, due to a lack of high-throughput methodologies, the vast majority of predicted CAZymes are not biochemically characterised and, as such, a great amount of biotechnological potential remains untapped. The overall aim of this thesis is to advance the knowledge on, and assess biotechnological application of, microbial enzymes for processing lignocellulose and its components.

The first aim of this thesis was to exploit a novel biocatalyst capable of producing cellulose oligosaccharides in the development of an analytical mass spectrometry-based technique for the (semi)high-throughput characterisation of cellulolytic enzymes. Via the screening of a panel of glycosyltransferases an enzyme was discovered that was capable of polymerising glucose onto an imidazolium-based tag. The subsequent product of this reaction and their derivatives were employed to characterise the substrate scope, specificity and mode of action of a variety of enzymes, as described in chapter 4. Chapter 3 reviews recent progress in biocatalytic synthesis of beta-glucans and cellulose.

The second aim was the characterisation of microbial CAZymes produced by their host during exposure to lignocellulose. A gene upregulated during lignocellulose degradation by *A. niger* was hypothesised as a arabinoxylan active enzyme. Subsequent phylogenetic analysis and molecular docking suggested this enzyme was potentially a β-galactofuranosidase, which was supported experimentally after cloning, expression and characterisation. Additionally, the use of an enzyme in the processing of nutritionally relevant oligosaccharides from wheat bran was assessed. For this, a thermotolerant feruoyl esterase from *Thermobacillus xylanilyticus* was cloned, expressed and biochemically characterised.

Chapter 3: Recent advances in enzymatic synthesis of β-glucan and cellulose

This chapter is composed of one published review article.

G.S. Bulmer, P. de Andrade, R.A. Field, J.M. van Munster, Recent advances in enzymatic synthesis of β-glucan and cellulose, Carbohydrate Research, https://doi.org/10.1016/ j.carres.2021.108411.

Foreword

β-glucans hold key roles in the nutrition, biomaterials and the renewables-based biotechnology sector. As such, the enzymatic synthesis of these products are highly relevant to both industry and academia, with this chapter discussing the various routes to their production. The review focusses on the biocatalysts available, enzyme promiscuity and the engineering and modification of reaction conditions to drive both functionalised and non-functionalised β-glucan synthesis.

Contribution

As joint first authors, G. S. Bulmer and P. de Andrade were responsible for the research and writing of the original draft. R. A. Field and J. M. van Munster reviewed and edited the manuscript.

Recent advances in enzymatic synthesis of β-glucan and cellulose

Gregory S. Bulmer[1]*, Peterson de Andrade[1]*, Robert A. Field[1] and Jolanda M. van Munster[1,2#]

[1]Manchester Institute for Biotechnology, School of Natural Sciences, University of Manchester, Manchester, UK
[2]Scotland's Rural College, Edinburgh, UK

* equal contribution # corresponding author; Jolanda.van-munster@sruc.ac.uk, jolanda.vanmunster@manchester.ac.uk

Abstract
Bottom-up synthesis of β-glucans such as callose, fungal β-(1,3)(1,6)-glucan and cellulose, can create the defined compounds that are needed to perform fundamental studies on glucan properties and develop applications. With the importance of β-glucans and cellulose in high-profile fields such as nutrition, renewables-based biotechnology and materials science, the enzymatic synthesis of such relevant carbohydrates and their derivatives has attracted much attention. Here we review recent developments in enzymatic synthesis of β-glucans and cellulose, with a focus on progress made over the last five years. We cover different types of biocatalysts employed, their incorporation in cascades, the exploitation of enzyme promiscuity and their engineering, and reaction conditions affecting the production as well as in situ self-assembly of (non)functionalised glucans. The recent achievements in application of glycosyl transferases and β-1,4- and β-1,3-glucan phosphorylases demonstrate the high potential and versatility of these biocatalysts in glucan synthesis in both industrial and academic contexts.

Keywords: glucans, cellulose, glycosyltransferases, glycoside hydrolases, glycoside phosphorylases, nanostructures

1. Introduction

1.1 β-glucan structure

β-glucans (Figure 1a) are major structural components of plants and fungi that are of significant importance to a large range of fields such as nutrition, microbial pathogenicity and the renewable fuel sector. β-glucans, including the major biopolymer cellulose, circumscribe a range of large, natural glucose oligosaccharides linked by β-glycosidic bonds of varying linkage types and as a group comprise some of the most abundant carbon sources on Earth. The type and proportion of linkage types within β-glucans has an impact on both structure and behavior, with different forms of β-glucans occurring depending on their source [1]. The enzymatic production of these glucans by glycosyl transferases, glycosynthases and glycoside phosphorylases is a rapidly changing field, of which the recent advances will be discussed in detail.

1.1.1 β-1,4 glucan

The β-1,4 glucan cellulose is the most abundant biopolymer on Earth with 2 x 10^{11} tons produced annually [2]. Consequently, this renewable source is of great interest in the production of bioethanol, other high-value products and the overall success of a future bio-based economy [3]. Cellulose is held together by van der Waals forces and an extensive network of intra/inter chain hydrogen bonds that yield microfibrils with highly ordered regions (crystalline), thus resulting in structural stability and robustness [4]. Wood and cotton fibres are the most common sources of cellulose [5], and the polysaccharide is also produced by bacteria [6]. The production of nanocellulose (cellulose nanocrystals - CNC and cellulose nanofibers - CNF), based on its isolation from cellulosic biomass, is considered a top-down method due to high energy consumption and harsh chemical conditions (strongly acidic and basic) [7]. Nanocellulose has become a valuable class of material due its inherent characteristics and potential for a wide range of advanced industrial applications (composites, electronics, sensors, cosmetics, pharmaceutical, biomedical, etc) [8–10].

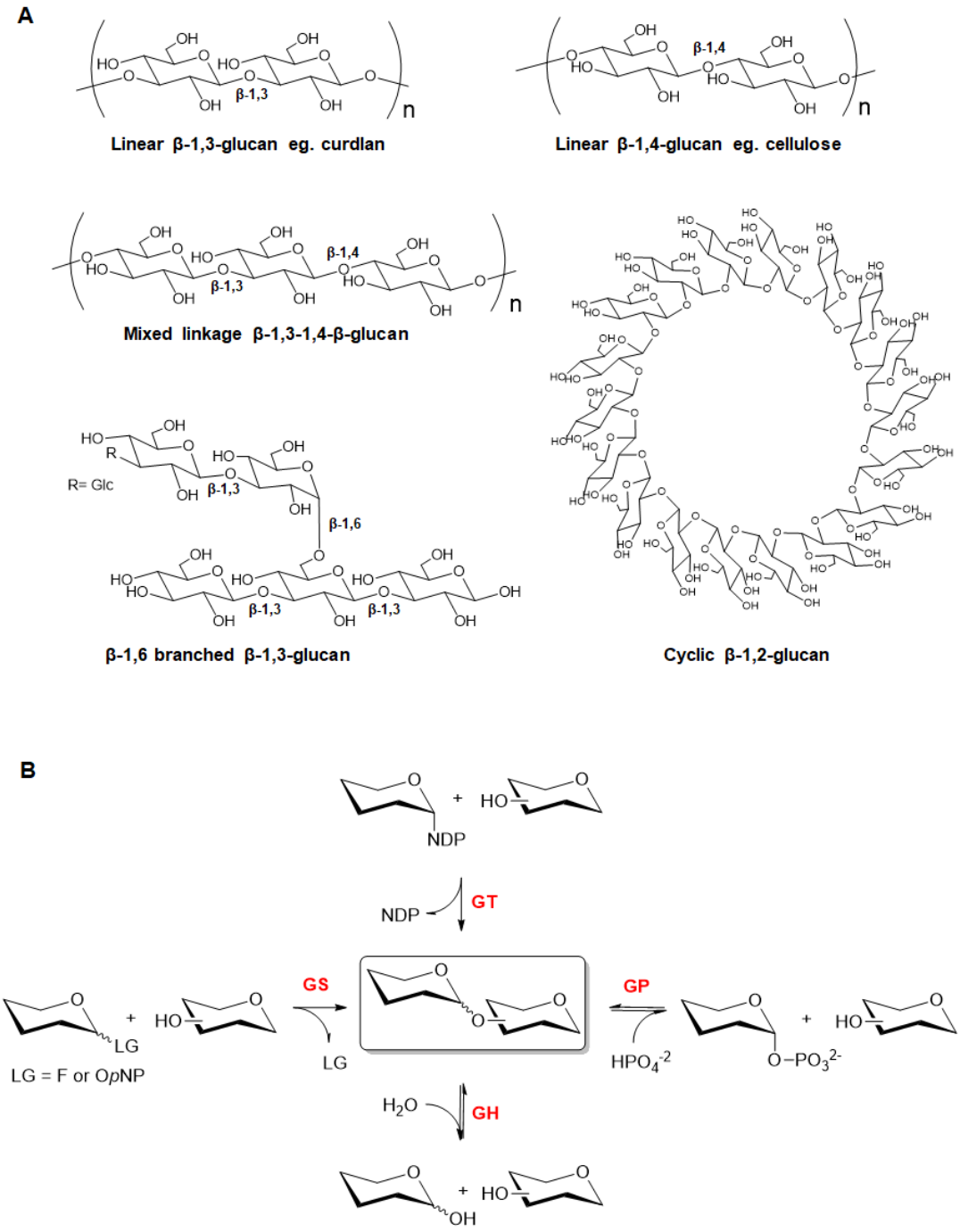


Figure 1. a structure of natural glucans. **b**Glycosidic bond formation and/or cleavage catalysed by different CAZy families. **GS** = glycosynthases (Leaving Group: Fluorine or O-*para*-nitrophenyl); **GT** = glycosyltransferases (NDP: nucleotide diphosphate); **GP** = glycoside phosphorylases and **GH** = glycoside hydrolases. (Part **B** was adapted from [23]).

1.1.2 β-1,2 glucan

Cyclic β-1,2-glucans (CβG) are found in a limited number of symbiotic and pathogenic gram negative bacteria, including the genera *Agrobacterium*, *Rhizobium* and *Brucella* [11]. Within these genera CβGs enable the invasion of, and survival within, host cells. CβGs are non-cytotoxic, can activate mammalian dendritic cells and enhance antigen-specific $CD4^+$/$CD8^+$ T cell responses, thus offering potential as an adjuvant in antimicrobial therapies [12]. In bacteria CβGs are known to sequester iron in order to protect against iron-induced toxicity [13]. With a backbone of around 17-25 glucose units, these glucans reside within the periplasmic space and are responsible for osmotic homeostasis on the cell wall [12].

1.1.3 β-1,3 glucan

Cereal β-glucans found in oats, barley and rye have a combination of β-1,4 and β-1,3 linkages and form long-chain linear polysaccharides of high molecular weight, with the β-1,3 linkages causing an irregular coiled structure. Fungal β-glucans are β-1,3 linked polysaccharides with side chains of varying length attached to the backbone through β-1,6-linkages [14]. They are a major constituent of the fungal cell wall that can form up to 40% of the dry wall mass [15]. As seen in the plantae, fungal glucan composition varies by species, with mushrooms normally containing short β-1,6-linked branches whilst the yeasts display longer β-1,6 side branches [16]. Such glucans are a major component of dietary fibre and have food stabilizing properties. Furthermore β-1,3-glucans have been shown to hold anti-tumor activities [17] in addition to their use as β-glucan micro-particles for the delivery of therapeutics [18]. Functioning as key immunomodulatory agents, they stimulate the immune system and induce cellular immunity [19]. Linear β-1,3 glucan polysaccharides such as curdlan (bacterial), callose (plant) and paramylon (microalgal) havehigh-molecular weight and are water-insoluble [20]. These polysaccharides are stabilised by a broad network of intra/inter chain hydrogen bonds that can exist in helical or random coil structures [21,22]. Their potential application in several industrial sectors ranges from farming, nutrition, cosmetics to therapeutics. For instance, curdlan has been used as gelling agent in foods, and

there has been also interest in using β-1,3 glucans as functional dietary fibres and immunomodulatory agents [24,25]. In fact, the regulatory effects of dietary β-glucans (food supplements or part of daily diets) have been recently reviewed to highlight a potential role in cancer control due to their ability to modulate a variety of biological responses [26].

1.2 Current need for enzymatic synthesis of β-glucans

Oligo- and polysaccharides have conventionally been produced with low yields from natural sources by hydrolysis and used for biological assessment mainly as heterogeneous mixtures. Inefficient processes for extraction and purification have contributed to high-cost production that limits the access to pure and structurally well-defined short/long linearβ-glucans [21,24]. Chemical synthesis of glycans, despite all the progress in the field [27], remains challenging due to the need of specific building blocks as reactants, solubility challenges and poor regio-/stereoselective reactions that yield heterogenous materials. In this respect, enzymatic synthesis has naturally become an attractive approach to enable the bottom-up preparation of oligo- and polysaccharides in a regio- and stereocontrolled manner, thus resulting in high synthetic precision and site-specific modifications [23,28,29].

Enzymes offer a variety of benefits in comparison to chemical catalysts, specifically their ability to operate under mild conditions thus allowing for greater process efficiency and enabling more sustainable routes to chemicals [30,31]. Briefly, ideal biocatalysts constitute the following properties in respect to catalysis: high activity; high regio-/stereo-selectivity; lack of unwanted side activities such as donor or product hydrolysis; stability both when stored and within reactions. Additionally, simple, high-yield heterologous expression within popular expression hosts such as *Escherichia coli* is desirable, as ispromiscuity towards a variety of modified substrates to produce a diverse repertoire of natural and non-natural products [32]. Attainment of this suite of traits can be enabled by enzyme discovery or engineering to improve activity, scope and yield. Examples of such progress in the

production of β-glucans and cellulose have been demonstrated across a variety of glycoenzymes, recent examples of which will be covered in this review.

2. Glycosyltransferases

2.1 Glycosyltransferase activity

Glycosyltransferases (GTs) (Figure 1b) are found widely across nature and are responsible for the synthesis of glycosidic linkages between a non-activated acceptor carbohydrate and (usually) an activated sugar donor [33]. This activity can occur onto a variety of different biomolecules such as mono-, di- and oligosaccharides [34,35], LPS [36] and peptidoglycan [37]. GTs that utilize activated nucleotide sugars (carbohydrates linked to anucleoside mono- or diphosphate e.g. CMP, UDP, GDP or TDP) are termed Leloir glycosyltransferases. Non-Leloir GTs utilise non-nucleotide donors such as activated phosphorylated sugars (phosphorylases, discussed in more detail in section 3) or non-activated donors in the form of sucrose or starch-derived oligosaccharides (transglycosidases) [35,38]. Leloir GTs are responsible for the synthesis of the majority of carbohydrates in nature and as such are of great interest (and one of the first enzymatic tools considered) for those wanting to enzymatically synthesize commercially or medically important glycans. Glycosyltransferases can be promiscuous in both acceptor and donor substrate acceptance, and thus provide biocatalysts with potentially exploitable side reactions with rates that in many cases have been demonstrated to be sufficient for exploitation in glycoside synthesis [32]. Generally, GTs display high glycosyl donor and acceptor selectivity resulting in products with high regio- and stereo-selectivity and have frequently been used in preparative scale, one-pot single/multi-enzyme glycosylation cascades to synthesise a variety of natural products as recently reviewed in [33,38].

2.2 Glycosyl donor recycling

Although commercially available, the activated glycosyl donor monosaccharides required by Leloir GTs are relatively expensive, thus presenting a problem in chemoenzymatic reactions whereby larger quantities of product are required. To rectify this issue there has been a concerted effort in the in situ recycling of

nucleotide sugar donors. UDP-glucose can be produced via activity of sucrose synthase (SuSy), whereby sucrose is hydrolysed into fructose and UPD-glucose when in the presence of UDP. SuSy has been utilised in a variety of biocatalytic processes as reviewed here [39]. Mutants of bacterial SuSys have enabled the development of more efficient variants allowing improved UDP-Glc generation. Furthermore, analysis of the plant SuSys, such as *Glycine max* (soybean), has revealed members more active than their bacterial counterparts [40]. The lower activities of bacterial SuSys have been attributed to residues in the nucleoside binding site having higher affinities for ADP, rather than UDP. Introduction of plant residues at these sites resulted in up to 60-fold decreases in K_m for UDP, producing enzymes more suitable for industrial applications [41]. In GT-dependent biocatalysis, achieving the typical efficiency targets required for industrial processes is often problematic. Therefore, the ~100g scale production of UDP sugars such as UDP-Glc is vital for commercially viable biotransformations. Bioreactor cultivation of *E. coli* expressing *Acidithiobacillus caldus* SuSy was shown to produce 100 $g_{product}$/L and operate a space-time yield of 10 g/L/h. UDP-Glc could then be purified using chromatography-free downstream processing with yield of 86% [42].

2.3 Glycosyltransferase mediated glucan synthesis

2.3.1 β-1,4 glucan synthesis

The in vitro biosynthetic synthesis of (1,4)-β-D-glucan has proved a challenging task that for many years was unattainable [43]. Within plants, the predominant producers of cellulose, the natural biosynthetic machinery consists of a complex membrane-embedded multi component system called the cellulose synthase complex (CSC) (Figure 2a, b) [44]. This machinery has proven to be extremely difficult to express and maintain in a functional form. While CSC activity from a variety of plant cell extracts had initially been demonstrated, [47–49], only recently a CSC's has been functionally reconstituted in vitro [50]. Heterologous expression of *Populus tremula x tremuloides* CesA8 (*Ptt*CesA8) enabled the in vitro production of cellulose and its subsequent processing into cellulose microfibrils. This polymerisation was dependent on a lipid bilayer environment (in addition to Mn^{2+}

ion presence), as solubilisation of these lipid vesicles by detergent removed this activity. Whether activity is abolished due to enzymatic sensitivity to the detergents or alteration of the enzyme quaternary structure is unknown [50].

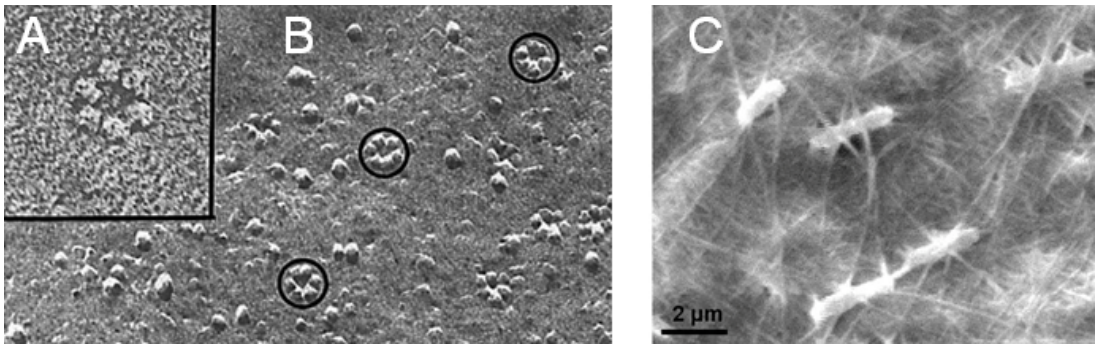


Figure 2: A & **B** Freeze fracture replica of CSC (inset) within plasma membrane of a *Zinnia elegans*, adapted from [45]. **C** SEM of *K. rhaeticus* surrounded by bacterial cellulose, adapted from [46].

Reconstitution of such bacterial cellulose synthases (Bcs), which bridge the periplasm through the use of subunits [51], has also been achieved. Purified cellulose synthase BcsA and membrane-anchored, periplasmic subunit BcsB from *Rhodobacter sphaeroides* were utilised in vitro for the production of cellulose of DP 200-300, whereby BcsB inclusion was essential for the catalytic activity of BscA. Unlike *Ptt*CesA8, BcsA remains active in a variety of non-denaturing detergents and is independent of lipid-linked reactants, but requires the allosteric activator cyclic-di-GMP [52]. The tailoring of reaction conditions for bacterial cellulose synthases can drive cellulose morphology in vitro. Detergent extracted CSC's from *Komagataeibacter xylinus* demonstrated how varying temperature and centrifuging reactions, to mimic an anisotropic environment, can alter the product profile. Optimal production of cellulose, with a II crystal structure and average DP of 60-80, occurred at 35 °C without centrifugation. Lower temperatures (e.g. 13 °C) resulted in lower yield, likely due to reduced particle movement. At higher temperatures (42 °C) protein denaturation and aggregation affected cellulose morphology. Increasing rates of centrifugation resulted in decreased cellulose production and a reduction in reducing end detection by BCA assay [53].

The application of synthetic biology has enabled the effective tuning of bacterial cellulose production, as demonstrated in the genetic manipulation of

Komagataeibacter rhaeticus (Figure 2c). Development of such a genetic toolkit allowed engineering of inducible cellulose production when required, thus improving yields by avoiding the emergence of cellulose-nonproducing mutants that can arise when cellulose is synthesised constitutively. Gene expression of *K. rhaeticus* can be induced within the cellulose pellicle and offers a route to the controlled production of new cellulose-based spatially and temporally patterned biomaterials. Post hoc addition of *E. coli* expressed proteins with a fused CBD (Cellulose Binding Domain) to cellulose produced functionalised materials and was implemented in the production of cellulose-based garments containing GFP demonstrating feasibility at macroscale [46].

The synthesis of Glc-β-1,4-Glc linkages from nucleotide sugars by glycosyltransferases is predominantly limited to the use of plant and bacterial cellulose synthases. However, promiscuity in donor and acceptor specificity of other glycosyl transferases has been exploited. Bovine β4GalT1 has been demonstrated to transfer Glc from UDP-Glc onto GlcNAc and Glc-terminated acceptors at low efficiencies, producing β-1,4 linkages [32]. More recently it has been demonstrated that *Neisseria meningitidis* glycosyltransferase, LgtB, can be used for the *in vitro* polymerization of glucose from UDP-glucose via the generation of β-1,4-glycosidic linkages, producing oligosaccharides from DP 2-10. Furthermore LgtB is permissible to biocatalytic cascades and the production of glucose oligosaccharides with tailored functionalities that can be utilized to inform on the activity of various plant cell wall degrading enzymes [54].

2.3.2 β-1,3 glucan synthesis

The study of β-1,3 glucan synthesis in vivo has identified the responsible glycosyltransferases, which are also commonly referred to as glucan synthases (GLS) [14]. Various groups have demonstrated how membrane bound GLS' could be extracted from fungal and plant cells for use in in vitro assays, e.g. via the preparation of microsomal membranes and subsequent detergent fractionation [55–57]. Recently, through inhibition of cellulose synthesis, the stress response formation of β-1,3callose can be induced with cultured protoplasts of *Nicotiana*

tabacum (tobacco) and *Arabidopsis thaliana.* Fibres of different morphologies being produced depending on the plant species, e.g. hollow with average width of 12 µm or densely packed with a width of 8 µm respectively [58]. Detergent-isolated *Betula* protoplasts containing callose synthase were shown to secrete β-1,3callose bundles in vitro when incubated with a high concentration of Ca^{2+} ions. The lengths of these fibres were considerably longer as those from the two herbaceous plants [59]. These GTs require membrane protein solubilisation (through the use of detergents such as CHAPS) for use in in vitro reactions which, in tandem with their large nature and poor stability in solution, means there are few examples of their successful use in in vitro reactions.

3 Glycoside hydrolases

3.1 Glycoside hydrolase

Glycoside hydrolases (GHs) catalyse the breaking of glycosidic bonds but many GHs are also capable of functioning in the reverse direction through either the displacement of equilibrium (thermodynamic control) or through the use of activated sugar donors that intercept the reactive glycosyl-enzyme intermediate via kinetically controlled transglycosylation [60–62]. These methods offer opportunities to leverage the large natural repertoire of GH specificities, to enable routes for the enzymatic synthesis of designer glycans.

A variety of enzymes with natural transglycosylation abilities are reported in the literature and therefore of interest when making building blocks for the production of derivatised oligo- and polysaccharides. One such example is a β-glucosidase from *Penicillium funiculosum*. The recombinantly expressed rBgl6 demonstrated the transglycosylation of glucose to form the β-1,6 linked glucose dimer gentiobiose with a yield of 23 % [63]. The bacterial endo-1,3-β-d-glucanase from *Formosa algae* is able to perform transglycosylation of the acceptors methyl-β-d-xylopyranoside, methyl-α-d-glucopyranoside, 4-methylumbelliferyl-β-d-glucoside and glycerol, and utilizes laminarin as the donor. Of these acceptors glycerol was the most efficient whereby the most predominant species in the reaction mix were the

transglycosylation products with a DP of 2-12 [64]. Other exoglucanases, such as Exg-D, carry out transglycosylation using various primary alcohols to produce alkyl-glycosides. Using methanol, ethanol and propanol as acceptors and *p*-nitrophenyl cellobioside as the glycosyl donor, methyl-, ethyl- and propyl-cellobiosides were synthesised with yields of 87 %, 55% and 30% respectively – offering enzymatic routes to non-ionic surfactants normally produced by acid catalysis [65].

3.2 Glycosynthases

3.2.1 Glycosynthase activity

The basis for glycosynthase research began in 1991 with the first non-natural, biosynthetic, *in vitro* route to cellulose that was performed using a wild-type retaining *Trichoderma viride* cellulase whereby β-D-cellobiosyl fluoride was utilised as the substrate in a acetonitrile/acetate buffer mix (required to promote transglycosylation) [66]. However, the use of wild type glycosidases has significant drawbacks, as the polymerisation products can act as a substrate for the enzyme, resulting in hydrolysis of the product – vastly reducing potential yields [67]. To alleviate this issue, in 1998, Mackenzie et al. developed the first glycosynthase (Figure 1b); a genetically engineered exo-glycosidase capable of synthesising oligosaccharides through utilisation of an activated glycosyl donor in an anomeric configuration opposite to the natural substrate. The glycosynthase is unable to hydrolyze reaction products due to mutation of the catalytic residue, resulting in an inability to form the requisite R-glycosyl-enzyme intermediate. Mutation of endo-glycosidase, to enable the use of oligosaccharides of different degree of polymerizations to act as glycosyl donors, has also been reported[68]. Besides glucosynthases, a broad range of synthases accepting other glycans has been developed, as recently reviewed elsewhere [69].

3.2.3 β-1,2 glycosynthases

Various phenolic compounds are known to have positive impacts on human health, however such compounds often have poor bioavailability which can be improved

with glycosylation. Protein engineering enabled the production of the β-1,2 transglycosylating BGL-1-E521G that was able to utilise phenolic acceptors such as epigallocatechin gallate (EGCC) in addition to pNP-sugars, whilst using α-glucosyl-fluoride as the donor. Consequently a variety of β-1,2-glucosylated phenolic compounds were produced in a regioselective manner [70].

3.2.4 β-1,3 glycosynthases

Glycosynthases have been successfully employed to generate β-1,3-glucan for example as demonstrated for *Hordeum vulgare* E231G mediated self-condensation of α-laminaribiosyl fluoride and 3-thio-α-laminaribiosyl fluoride to polymers of > DP 20 and DP 6-8 respectively [71]. Synthesis of mixed-linked 1,3-1,4 β-glucans from di-, tri- and tetrasaccharide donors has also been achieved, whereby tuning of the β-1,3 and β-1,4 linkage ratio produced glucans that do not occur in nature, thus demonstrating the potential to modify glucan structures through enzymatic synthesis [72]. Recently progress has been made towards the development of β-1,3-glucan synthases that dispense with the need for fluorinated donors.

The glycosyl fluoride donors commonly required by glycosynthases are often unstable at high temperatures, meaning reactions are limited to temperatures of 30-37 °C which is suboptimal for many thermophilic enzymes of interest. Consequently the development of *in situ*-generated α-glycosyl formate donors, in an example of chemical rescue, allowed the use of both the fluoride donor or a exogeneous formate nucleophile to produce a β-1,3 disaccharide [73]. Such an approach offers an attractive alternative to the use of expensive activated donors and was further utilised in the development of a thermotolerant β-glycosynthase (*Tn*Bgl3B) from a *Thermotoganeapolitana* β-glucosidase. In addition to using a fluoride donor, rational design of the glycosynthase also allowed the use of the formate “non-classical’’ system (without a glycosyl fluoride donor) and enabled reaction temperatures of 80 °C permissible for the thermotolerant enzyme. Using pNp-Glc as both donor and acceptor the *Tn*Bgl3B variant consequently produced pNP-laminaribioside (pNP-Glc-1,3-Glc), after longer incubations pNP-cellobioside was also produced [74]. The recent application of a glycosynthase immobilised and used in flow for the synthesis of speciality glycans suggests an emerging field for the

design and implementation of glycosynthases in flow chemistry and could be key to improving their commercial viability [75].

3.2.5 β-1,4 glycosynthases

Mutation of an endoglucanase from *Humicola insolens* (producing HiCel7B E197A) enabled the polymerisation of cellobiosyl fluorides for the production of regioselectively modified oligo- and polysaccharides. Both unsubstituted and modified mono- and disaccharide acceptors were utilised to create a variety of natural and novel β-1,4 linked compounds. HiCel7B was able to accept C-6 position functionalised α-cellobiosyl fluorides containing various groups including bromine and thioglucosyl groups and represented the first self-condensation of a cellobiosyl fluoride donor to produce cellulose [76]. HiCel7B E197A was further exploited to produce modified cellulose derivatives through polymerization of 6'-azido-α-cellobiosyl fluoride to produce alternating 6-Azido-6-deoxycellulose. Such alternating 6-Azido-6-deoxycelluloses can be acetylated then subsequently reduced to amines. Conversely, alternating 6-Azido-6-deoxycelluloses can undergo copper(I)-catalysed azide-alkyne cycloaddition (CuAAc - click chemistry) with Alexa Fluor 488 Alkyne to form conjugated alternating 6-Azido-6-deoxycellulose [77].

The need for high-throughput technologies in enzyme discovery and development has driven the creation of assays. The development of colorimetric assays for glycosynthase screening allows the improved discovery and selection in a high-throughput manner [69] whilst the screening of diverse, synthetic gene libraries allows a systematic approach to glycosynthase discovery [78].

4. Glycoside Phosphorylases

Glycoside phosphorylases (GPs) naturally catalyse the cleavage of glycosidic bonds in the presence of inorganic phosphate, releasing sugar 1-phosphate and shorter glycans (Figure 1b). Since these phosphorolysis reactions are reversible, due to equivalent free energy released by the anomeric bond cleavage in both directions, these enzymes are effectively used to catalyse the synthesis of a wide range of glycosides in a regio- and stereospecific manner. These carbohydrate-active

enzymes (CAZy) are capable of producing disaccharides or oligosaccharides with a broad range of glycosidic linkages (except β-1,1 and α/β-1,6) from several acceptors (D-Glc, D-GlcNAc, maltose, trehalose, sucrose, nigerose, etc) and mostly α-D-glucose 1-phosphate (α-D-Glc-1P) as donor substrate. GPs have been classified into glycoside hydrolase (GH) and glycosyltransferase (GT) families according to their sequence similarity (CAZy database; http://www.cazy.org/). Most GPs (EC 2.4.1.x) belong to GH families and are grouped into retaining and inverting phosphorylases, [23]. Some good reviews have covered in more detail phosphorylases classification [23], structure and mechanism [79,80], and use as catalysts for glycoside synthesis [23,29,81]. Herein, we provide a brief overview on recent use of GPs to synthesise β-1,3 and β-1,4 glucans focussing on the following enzymes: cellodextrin and cellobiose phosphorylases, as well as Pro_7066/laminaridextrin and laminaribiose phosphorylases.

4.1 β-1,4-D-Glucan phosphorylases (cellodextrin and cellobiose phosphorylases)

4.1.1 Cellulose-like materials (long-chain insoluble oligosaccharides)

Characterised from various anaerobic bacteria and obtained in recombinant form in *E. coli* under standard conditions, cellodextrin phosphorylase (CDP, EC 2.4.1.49) is a homodimer that belongs to the GH94 family and is the most studied phosphorylase polymerising β-1,4-glucan. CDP is an inverting enzyme that produces cellodextrins (cello-oligosaccharides, oligocelluloses or cellulose oligomers) from α-D-Glc-1P and D-Glc/short β-1,4-glucans as natural donor and acceptor substrates, respectively [82]. Different degree of polymerisation (DP) towards long-chain insoluble (DP ≥ 7) or short-chain soluble (DP ≤ 6) oligosaccharides can be achieved depending on the reaction conditions. Although optimum pH and temperature are dependent on the enzyme source, CDPs are very stable under a wide range of pH (5.0 - 8.6) and temperature (37 - 60 °C), as well as in mixture of aqueous buffer and organic solvents. Equally important, the nature of the substrates has a significant impact upon structural features, such as crystallinity and morphology. CDP's great potential as biocatalyst for synthetic applications relies on the fact that non-natural

substrates can be reasonably tolerated, thus giving rise to tailor-made cellulose-like materials with different nanostructures.

A comparison of the acceptor substrate specificity of CDP from different sources has shown its permissiveness towards a broad range of natural and functionalised D-glucosyl at the reducing end, whereas the same promiscuity is not observed to its donor substrate. For instance, D-cellobiose and longer cello-oligosaccharides are considerably better acceptors than plain D-Glc, which yields very poor activity (0.5 to < 8%) [28]. The ability of CDP to better accommodate longer (DP ≥ 2) and modified acceptor substrates is consistent with the structural data (PDB code 5NZ8). CDP has an open active site with a more accessible acceptor binding pocket due to the disposition of the two subunits within the dimer interface and the significant rearrangement of the catalytic and opposing loops that delineate the active site [82]. In order to harness its poor substrate specificity, most recent works have been focussed on studying self-assembly, crystallinity, morphology and potential properties of nanostructures produced from iterative glycosylation of β-D-glucose acceptors (primers) functionalised at the anomeric position. The Serizawa group has pioneered the use β-D-glucose derivatives with anomeric substituents as substrates for CDP (*Clostridium thermocellum*) (Figure 3a). Reactive substituents have been exploited to provide additional reactivity of the cellulose oligomers for post modification with other molecules, whereas non-reactive substituents have been tested as a means to control the self-assembly processes. Furthermore, in situ self-assembly using only plain substrates have been investigated under different reaction conditions (Figure 4C).

<u>Functionalised acceptors with reactive substituents</u>

In a proof-of-concept work, crystalline cellulose oligomers with surface-reactive azide groups, synthesised from α-D-Glc-1P and β-D-glucose 1-azide by CDP under aqueous and mild conditions, were postmodified with 1-ethynyl pyrene in DMF via click chemistry to produce a sheet-like 2D crystalline cellulose II (ca. 5 nm thickness) with broad fluorescence emission due to the pyrene units partially forming excimers on the nanomaterial surface. This result highlights the potential use of

CDP to generate well-defined cellulose-like materials bearing spectroscopic reporter groups incorporated in a site-specific manner [83]. In a similar approach involving surface-reactive substituents, highly crystalline 2D vinyl-cellulose nanosheets (ca. 6 nm thickness) were synthesised via enzymatic polymerisation of α-D-Glc-1P from the primer 2-(glucosyloxy)ethyl methacrylate and postmodified by covalent incorporation at different loadings into poly(ethylene glycol) (PEG) matrix through thiol-ene Michael addition to generate a series of nanocomposite hydrogels with enhanced mechanical strength [7]. Later, the Loos group extended the repertoire of vinyl-based oligocelluloses catalysed by CDP using five anomerically pure β-D-glucosyl primers elegantly synthesised from D-Glc/cellobiose and hydroxy-alkyl (meth)acrylates or (meth)acrylamides with commercial β-glucosidases. Although not focussing on post modification, the authors also suggested the potential use of these structures for further (co)polymerisation with different monomers in the context of property-tunable hydrogels [84]. In another study, Serizawa et al. investigated the protein adsorption properties of surface-aminated cellulose oligomers synthesised from α-D-Glc-1P and 2-aminoethyl β-D-glucoside. The resulting crystalline nanoribbons (ca. 5 nm thickness) were tested as solid support for specific proteins (BSA, cytochrome c, fibrinogen, IgG, lysozyme, transferrin, and trypsin) and revealed that only those negatively charged at pH 7.4 (fibrinogen, IgG and transferrin) were adsorbed onto the nanoribbons due to effective interaction with the primary amine on their surface. The aminated nanoribbons also showed no cytotoxicity against mammalian cells, thus demonstrating potential biomedical applications [85]. Most recently, functionalisedcellulose oligomers synthesised from α-D-Glc-1P and 2-azidoethyl β-D-glucoside by CDP were used to prepare functionalised cellulose paper. The paper was efficiently modified through heterogenous nucleation-based self-assembly (impregnation) of the oligomers bearing terminal azide dissolved inaqueous alkaline solution and subsequent neutralisation. Post modification of the fibre surface of cellulose paper was simply performed via click chemistry with alkynylated biotin for further high-sensitivity detection of IgG in a proof-of-concept diagnostic application. The successful noncovalent functionalisation and post modification of paper together with its long-term functional stability may boost the development of

sustainable devices [86]. Similarly, the Nidetzky group used reducing end thiol-modified nanocellulose, synthesised from α-D-Glc-1P and 1-thio-β-D-glucose via linear cascade with cellobiose phosphorylase (CBP, EC 2.4.1.20) and CDP, as a template for the directed assembly of silver nanoparticles in order to produce functional cellulose nanocomposite (Figure 3b). The resulting thiol-containing nanosheets of crystalline cellulose II (ca. 5 nm thickness) with high loading silver nanoparticles, selectively bound and well dispersed on the surface, were assessed against *Escherichia coli* and *Staphylococcus aureus* and showed excellent antibacterial activity. This study may expand the scope of the bottom-up approach towards the synthesis of functionalised nanocellulose with respect to metal nanoparticle composites [87].

Functionalised acceptors with non-reactive substituents

Since the poor colloidal stability of enzymatically synthesised nanocelluloses in aqueous solution restricts the formation of highly ordered structures, Serizawa et al. also studied the impact of β-D-glucosyl primers bearing hydrophilic [88] and hydrophobic [89] non-reactive substituents on self-assembly (Figure 3a). Oligo(ethylene glycol) (OEG) was initially selected due to its highly hydrated and dynamic nature in aqueous phases that could result in stability. Interestingly, the enzymatic reaction produced cellulose hydrogels consisting of crystalline nanoribbon networks with different structural features based on the OEG chain lengths. The cellulose DP shifted from 9 to 11 and the thickness of the nanoribbons increased from 5 to 6 nm with increase in the chain, which in turn decreased the width from several hundred to below 100 nm. These results suggested that the OEG chains modulated the self-assembling process of the cellulose oligomers possibly by colloidal stability of the nanoribbon precursors [88]. In a follow-up work, alkylated cellulose oligomers synthesised from alkyl β-D-glucoside primers varying in chain lengths (2, 4, 6, 8 carbons) self-assembled into different nanostructures such as nanoribbons, helical nanorods and distorted nanosheets. Surprisingly, the alkyl chain length affected not only the morphology but also the crystallinity of the material, probably by effective modulation of the intermolecular interactions of cellulose oligomers thus resulting in active assembly controlling. Shorter alkyl

chains (2 and 4 carbons) yielded hydrogels composed of long nanoribbons corresponding to cellulose II allomorph, whereas longer chains (6 and 8) generated dispersions composed of helical nanorods or distorted nanosheets resembling cellulose I allomorph [89].

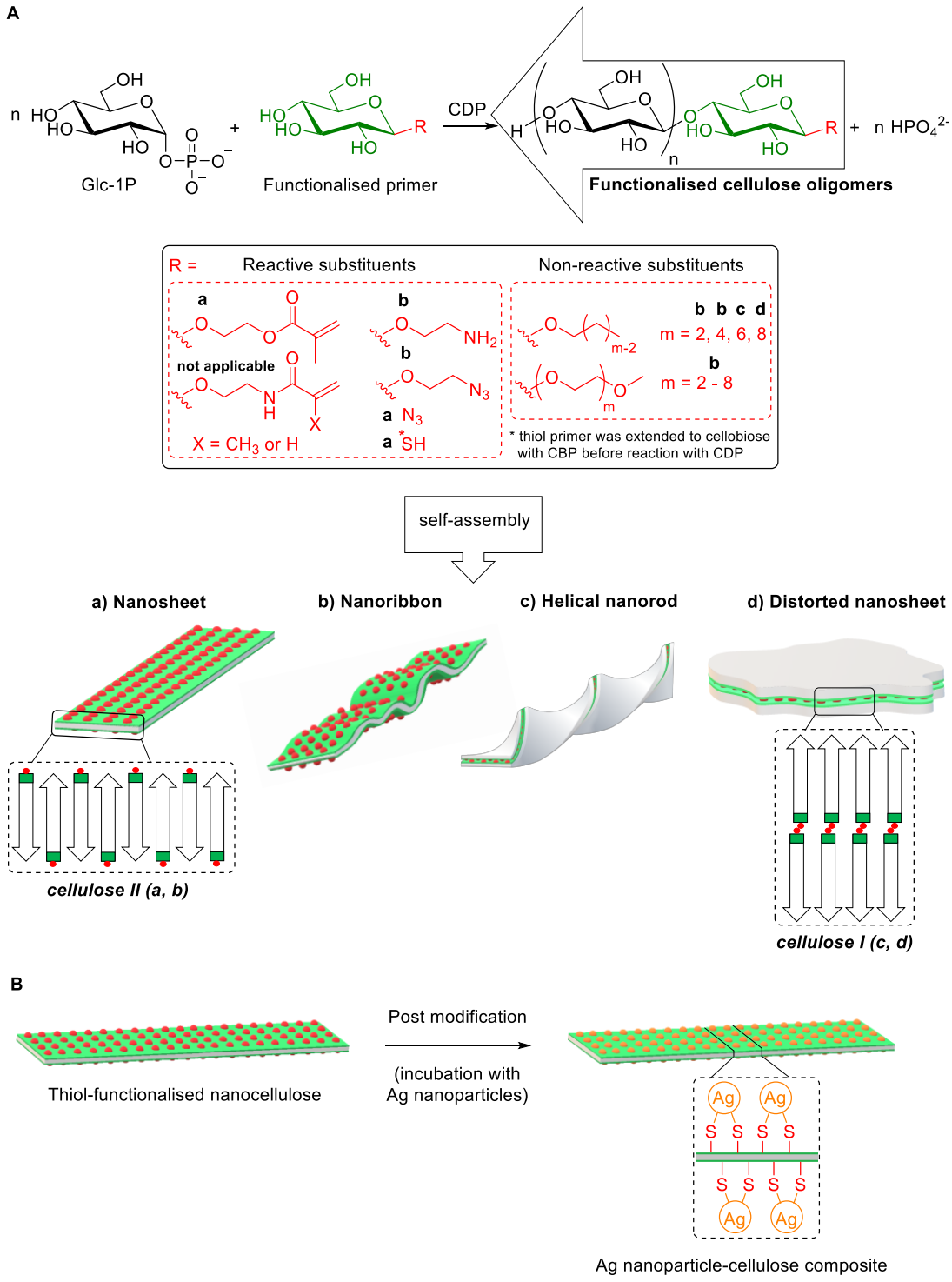


Figure 3. a Insoluble cellulose oligomers self-assembled into ordered nanostructures produced by cellodextrin phosphorylase (CDP)-catalysed synthesis from functionalised acceptors with reactive or non-reactivesubstituents at the

anomeric position. **b** Post modification example using thiol-functionalised cellulose oligomers in the presence of silver nanoparticles to produce nanocellulose composite (adapted from [87]).

Modified acceptors/donor with labelled atoms or probes

Cellulose oligomers containing selectively labelled atoms or site-specifically incorporated probes are important to report on local structure and environment, thus enabling their detailed characterisation and better understanding for broader application in advanced nanomaterials. Solid-state NMR analyses of cellodextrins corresponding to cellulose II nanosheets, enzymatically synthesised from ^{13}C-enriched D-Glc acceptors and α-D-Glc-1P (Figure 4a), revealed that the reducing-end units located on the surface tend to have β-anomeric configuration, which would be more sterically stable than alpha [90]. More recently, the Field group investigated the impact of single and multiple site-specific incorporation of fluorine into cellodextrin chains prepared from modified acceptors (2F-, 3F- and 6F-D-Glc) and α-D-Glc-1P with CBP as well as from modified donor (6F-α-D-Glc-1P) and D-cellobiose with CDP, respectively. As expected, the presence of a single fluorine at the reducing end of each chain had no substantial impact upon morphology and crystallinity as evidenced by the typical cellulose II precipitated nanosheets with average DP ca. 8. Conversely, the presence of multiple 6-deoxy-6-fluoro-D-glucose units yielded shorter nanosheets of unprecedented crystalline allomorph (Figure 4b) with higher average DP (ca. 10) and longer chains (up to DP 15). Furthermore, advanced solid-state NMR methods enabled deciphering the water-exposed and interior chemical environments for different carbon sites [91].

Non-functionalised substrates

The highly ordered hierarchical formation of nanostructures through in situ self-assembly of enzymatically synthesised insoluble cellulose oligomers is not only affected by chemical functionalisation, but also by different reaction conditions using natural substrates (Figure 4c). Inspired by the concept of molecular crowding, which occurs in intracellular environments with high concentration of different macromolecules and has considerable impact on the dynamics of biomolecular self-assembly [92], the Serizawa group synthesised cellulose oligomers from α-D-Glc-1P

and D-Glc in the presence of concentrated water-soluble polymers (dextran, poly(*N*-vinylpyrrolidone) - PVP, PEG and Ficoll) [93,94] to promote self-assembly of the cello-oligosaccharides. The reactions under in vitro macromolecular crowding resulted in stable hydrogels composed of well-grown crystalline nanoribbon networks of cellulose II, regardless of the polymer species, whereas the conventional rectangular sheet-like precipitates were formed in the absence of the polymers. It has been suggested that the aggregation of particles was possibly suppressed by increased solution viscosity and consequent decrease in diffusion rates of colloids, thus leading to dispersion stability for further growth into long nanoribbons networks to form a hydrogel. This concept was applied to produce a composite hydrogel via in situ self-assembly and cross linking of enzymatically synthesised cellulose oligomers in gelatin solution at 60 °C followed by cooling [93]. In a follow-up study, an equally robust double-network hydrogel was prepared by pH-triggered self-assembly of cellulose oligomers with gelatin. Insoluble cello-oligosaccharides (ca. DP 10) dissolved in alkaline solution (1 N NaOH) self-assembled upon pH decrease (7.4) in the presence of a warm acidic solution of gelatin to yield a hydrogel with improved stiffness, which was assumed to be based on entanglement between the networks of gelatin and the nanoribbon-shaped fibres of cellulose oligomers [95]. The pH neutralisation approach was also studied in the presence of a more complex mixture, such as serum-containing cell culture media, resulting in a physically cross-linked hydrogel of cellulose nanoribbon networks with anti-biofouling properties. The cells grew into spheroids (cell aggregates) and with no unfavourable sedimentation as the self-assembly progressed in a controlled manner for 3D cell culture, thus suggesting an attractive prospect towards biocompatible soft materials [96]. Considering that coexisting colloidal particles can also induce increase in viscosity and reduce stabilisation of colloidal dispersions, it has been hypothesised that colloidal particles could be used as crowding reagents towards nanoribbon network formation. In this sense, mechanically stable composite hydrogel with colloidal particles spatially immobilised (trapped) within cellulose nanoribbon network was produced by enzymatic synthesis and self-assembly of the oligomers in the presence of rod-like cellulose nanocrystals as model colloidal particles. The hybrid hydrogel proved to

be organic solvent-resistant and its stiffness was dependent on the concentration of the nanocrystals, suggesting physical cross-linking to the nanoribbons [97].

Besides the molecular crowding and colloidal particles approaches, the Serizawa group also investigated the influence of temperature [98] and organic solvents [99] upon in situ self-assembly of cello-oligosaccharides synthesised from α-D-Glc-1P and D-Glc (Figure 4c). Precipitated nanosheets were formed at higher temperatures (60 - 40 °C), whereas nanoribbon networks and dispersed nanosheets were produced by lowering the temperature to 30 °C and 20 °C, respectively. With this temperature shift, a decrease of DP from 10 to 8, and increase in crystallinity (from 52% to 66%) and similar dispersity index were observed. The temperature-directed assembly resulting into different cellulose II nanostructures was possibly driven by the synergy between reduced hydrophobic effect and concomitant induced self-crowding effect [98]. While rectangular precipitated nanosheets are traditionally obtained in aqueous buffer, reactions in mixture with organic solvents (DMSO, DMF, acetonitrile or ethanol) can yield hydrogels of well-grown nanoribbon networks. The solvent control over oligomerisation-induced self-assembly was rationalised through mechanistic studies based on the Kamlet-Taft parameters, which showed that the nanoribbon formation was only triggered by solvents with high β-values (hydrogen bond acceptor ability - DMSO and ethanol). This data suggests that the solvation resulting from hydrogen bonding between the organic solvent and the cello-oligosaccharides prevented aggregation and consequent precipitation towards nanosheet formation, thus allowing a higher-order structure to be formed. It was also noticed that cellulose II with similar average DP (8 - 10) was obtained in all reactions. However, the crystallinity was higher in the presence of organic solvent (ca. 64% vs 52%) due to higher uniformity of the chain lengths, evidenced by the lower dispersity index (ca. 1.2 vs 1.9) [99].

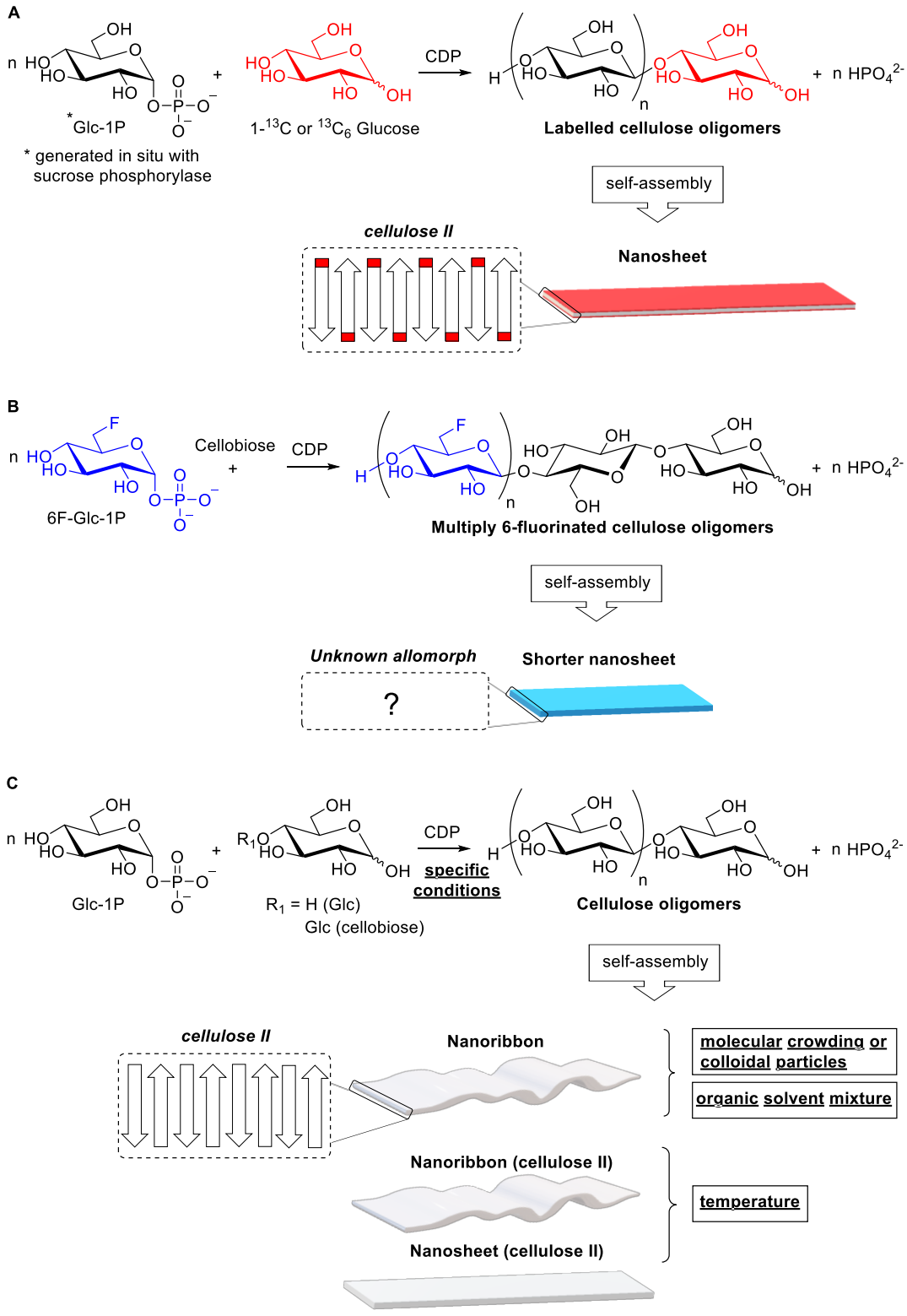


Figure 4. Insoluble cellulose oligomers self-assembled into ordered nanostructures produced by cellodextrin phosphorylase (CDP)-catalysed synthesis from selectively labelled acceptors (**A**), donor bearing site-specific probe (**B**), and non-functionalised substrates under specific reaction conditions (**C**).

A simple approach to produce hydrogel consisting of nanoribbon networks (cellulose II with average DP 7 - 8 and ca. 4.5 nm thickness) in aqueous buffer was performed using D-cellobiose [100,101] instead of conventional D-Glc as acceptor, which typically gives rise to precipitated nanosheets [102,103]. In this case, the distinct morphological outcome under the same reaction conditions was attributed to the effect of short-chain soluble oligosaccharides present in the supernatant of D-cellobiose reaction due to kinetic control. Soluble oligomers might accumulate once their rates of formation and self-assembly seemed to be similar. Conversely, no soluble oligomers were detected in the glucose reaction as the rate of cellobiose formation (rate-limiting step) was probably slower than the self-assembly rate [100,102]. In a follow-up study, higher concentration of D-Glc (200 and 500 mM) also resulted in nanoribbon networks hydrogel since the soluble oligomers possibly worked as dispersion stabilisers of the precursor particles, thus suppressing precipitation/aggregation and favouring growth[104].It is worth mentioning that previous use of D-cellobiose [105] resulted only in precipitated nanosheets because the reaction conditions were significantly different (lower concentration of reactants and enzyme). This emphasises the impact of setting suitable conditions in the process of generating tailor-made nanocellulosic materials.

4.1.2 Ingredients for animal and human nutrition (short-chain soluble oligosaccharides)

The state-of-the-art application of CDP also involves the synthesis of short-chain soluble oligosaccharides (DP ≤ 6) as means to access structurally defined model substrates as biological probes, focussing on the potential prebiotic effect for farm animals [106] and dietary fibres for humans [107]. Initially, a typical mg scale synthesis of short oligosaccharides has been studied by Loos et al. using α-D-Glc-1P and D-cellobiose in a constant molar ratio (20:1). The results showed a tendency towards the production of cellodextrins with reduced chain length in the presence of higher concentration of the acceptor substrate [108]. In this regard, Nidetzky et al. proposed a controlled biocatalytic synthesis of soluble cellodextrins using CBP (*Cellulomonas uda*) and CDP (*Clostridium cellulosi*) in a linear cascade reaction from α-D-Glc-1P and D-Glc (Figure 5a), which is economically more attractive and 7-fold

more soluble than D-cellobiose [109]. The optimised conversion of these substrates into soluble products with DP ranging from 3 to 6 (ca. 96 wt%) was mainly controlled by an appropriate molar ratio of α-D-Glc-1P/D-Glc (4:1) and in situ precipitation of released inorganic phosphate with Mg^{2+} ($MgCl_2$) to shift the equilibrium towards product formation. The activity ratio between both enzymes also proved to be an important parameter, showing best results with excess of CBP (3 U/mL) over CDP (2 U/mL) producing a final product concentration of 36 g/L within 2 hours with a conversion yield of 98% based on glucose. Although the disaccharide is a better substrate than glucose (ca. 50-fold), the two-enzyme reaction outcome is very similar in terms of product distribution and concentration when compared to the single step CDP conversion of D-cellobiose under the same conditions.

Recently, Nidetzky et al. also studied the integrated production of soluble oligomers via a three-enzyme phosphorylase cascade [110]. This biocatalytic process was developed to convert expedient substrates, such as sucrose and glucose, into short-chain cellodextrins using sucrose phosphorylase - SCP (*Bifidobacterium adolescentis*), CBP and CDP (Figure 5b). The iterative β-1,4-glycosylation of D-Glc from α-D-Glc-1P, generated in situ from sucrose and inorganic phosphate, was successfully achieved after optimisation of key parameters. The balance of the three phosphorylase activities (10:3:2 U/mL) along with the ratio sucrose/glucose, inorganic phosphate concentration and reaction time were crucial to DP control and increased concentration of soluble cellodextrins with DP 3 - 6 (40 g/L). Additionally, a convenient and efficient two-step procedure for product isolation was accomplished with high purity (≥ 95%) and yield (ca. 92%) via selective microbial conversion of the undesirable sugars (sucrose, fructose and glucose) using *S. cerevisiae*, followed by organic solvent precipitation. In a follow-up work, they exploited the cascade synthesis of soluble cello-oligosaccharides upon co-immobilisation of the three phosphorylases on solid support [111] (Figure 5c). The

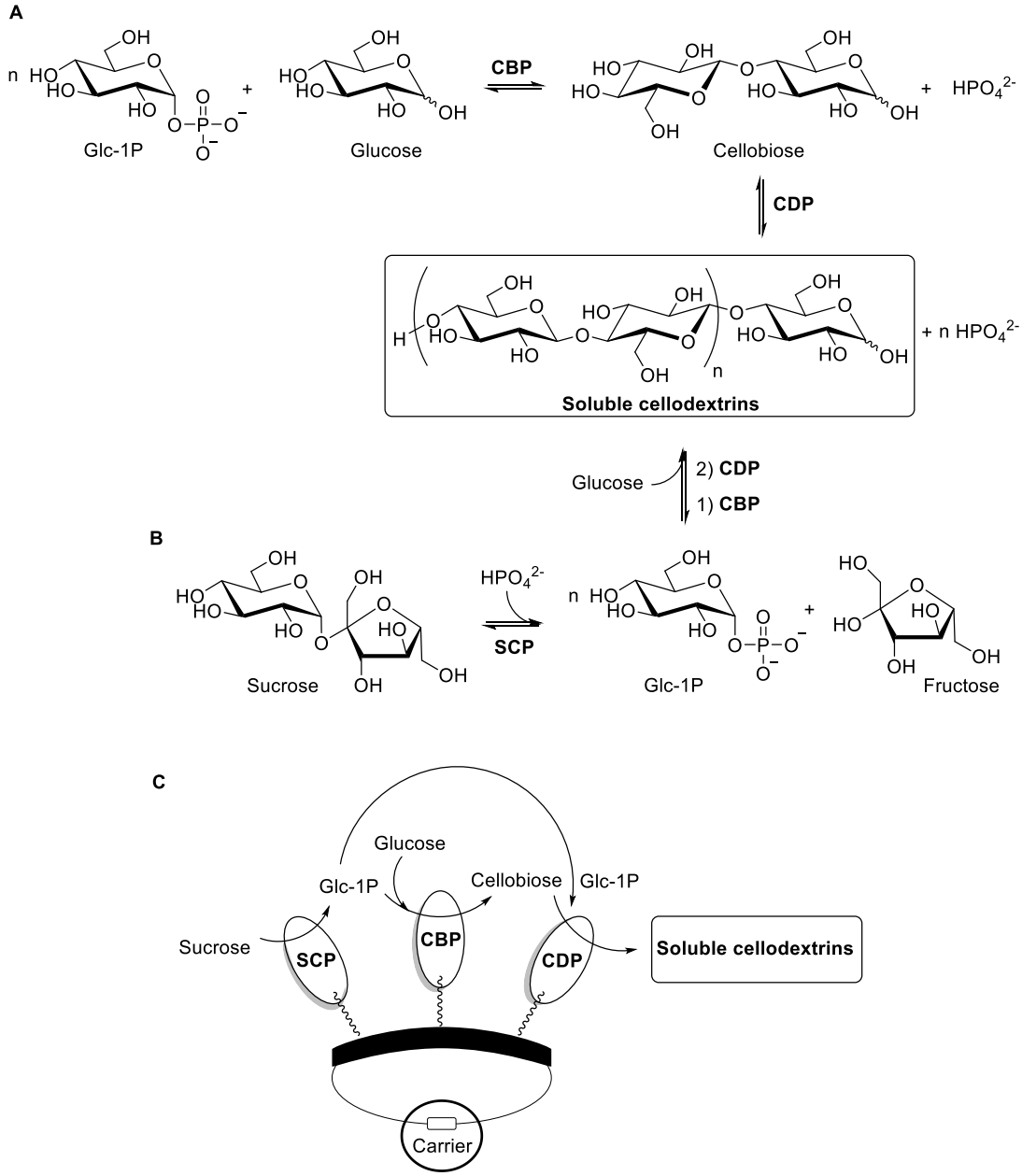


Figure 5. Controlled biocatalytic synthesis of soluble cellodextrins (DP ≤ 6, n= 1-4) using cellobiose phosphorylase (CBP) and cellodextrin phosphorylase (CDP) in a linear cascade reaction from α-D-Glc-1P and D-Glc (**A**), and via a three-enzyme phosphorylase cascade in solution (**B**) or on solid support (**C**) from sucrose and D-Glc including sucrose phosphorylase (SCP). (Part **C** was adapted from [111]).

chimeric enzymes harbouring the binding module were immobilised via ionic interaction on the same carrier to harness the effect of spatial proximity.As observed for the cascade reaction in solution, the phosphorylases on solid support also required a defined ratio of activities for an efficient DP-controlled synthesis.

Despite finding the optimal activity ratio and a good loading, the catalytic effectiveness was just above 50% and the molar yield concentration of soluble products was low (56%, 25 g/L) compared to the high sucrose conversion (90%). This was ascribed to the insoluble product formation (DP ≥ 8), which seems to occur at a higher rate when compared to the phosphorylases in solution. In this case, the reaction time was shortened to produce only the desired products but it also reduced their molar yield concentration to ca. 30% (12 g/L). The recyclability assessment revealed that these enzymes retained 85% of the overall initial activity after five cycles, though substantial release of CBP from the carrier was observed. Although promising, these results clearly show the challenges and limitations posed by a complex biocatalytic cascade performed on solid support.

Envisioning the scale up synthesis of such interesting soluble cello-oligosaccharides with potential industrial application, the Nidetzky group optimised their enzymatic production and assessed their growth stimulation among probiotic bacteria [112]. The one-pot biotransformation from sucrose and glucose using the three phosphorylases in solution was carried out in gram scale and optimised for a maximum soluble oligomers concentration of 93 g/L, which represents a 2.4-fold improvement compared to previous results [110]. The insoluble product formation was also reduced below 10 mol% after adjusting the enzyme activity ratio from 10:3:2 to 20:6:2 U/mL. The soluble products (DP 3 - 6) showed substantial growth stimulation (up to 4.1-fold with respect to the maximal cell density) for some probiotic strains when compared to known oligosaccharide prebiotics (inulin and trans-galacto-oligosaccharides) and D-cellobiose. Thus, these results support the importance of soluble cellulose oligomers as selective functional carbohydrates with considerable prebiotic potential.

Amongst the soluble cellodextrins, cellotriose is thought to be the most potent prebiotic. For instance, it is the preferred substrate for *Bifidobacterium breve*, a major probiotic bacterium in the human intestine [113]. Since the efficient synthesis of defined short-chain cello-oligosaccharides by wild type phosphorylases remains challenging, a recent study evaluated engineered CDP (*Clostridium*

cellulosi) and CBP (*Cellulomonas uda*) as possible biocatalysts for the enriched synthesis of cellotriose from α-D-Glc-1P and D-cellobiose [114]. Despite several site directed-mutations to disrupt a specific region of CDP active site to hampering further chain elongation, all mutants were still capable of elongating the chains towards higher DPs. A more promising result was achieved upon a directed mutation to optimise a well-known mutant (CBP OCP2) [115], which has five mutations located in the catalytic cleft that confer some degree of activity and affinity for D-cellobiose. The improved mutant (OCP2_M52R) showed 4-fold higher activity and affinity for D-cellobiose as well as reduction in further elongation. Consequently, cellotriose was synthesised with the highest purity (82%) and yield (73%) to date, and only trace amounts of cellotetraose and cellopentaose were observed. Interestingly, glucose can be also used as acceptor for cellotriose synthesis despite the new mutant has also exhibited higher affinity for the disaccharide. Therefore, this study highlights a positive perspective in engineering CBP for the defined production of cello-oligosaccharides longer than D-cellobiose.

D-cellobiose has been also a valuable and commercially attractive sugar with respect to zero-calorie sweeteners and as potential food additive. Although its synthesis from sucrose by SCP and CBP has been under implementation for industrial production [116], current research focussing on atom economy has been exploited via multi-enzyme one-pot biotransformation approaches with mathematical model assistance to predict optimal reaction conditions and enzyme loading ratio [117,118]. In this sense, D-cellobiose was synthesised from sucrose by three thermophilic enzymes with 60% yield within 10 h. The process involved phosphorolysis of sucrose by SCP (*Thermoanaerobacterium thermosaccharolyticum*) to form α-D-Glc-1P and fructose, which was isomerised by glucose isomerase (*Streptomyces murinus*) into D-Glc for further disaccharide synthesis by CBP (*Clostridium thermocellum*) (Figure 6a) [117]. More recently, a cost-effective platform for the in vitro synthesis of disaccharides from starch was successfully developed employing four enzymes. Starch was debranched with isoamylase to produce amylose for subsequent parallel conversion into D-Glc and α-D-Glc-1P with α-glucosidase and α-glucan phosphorylase, respectively, followed

by final catalysis with CBP (*Clostridium thermocellum*) (Figure 6b). Amongst other valuable disaccharides synthesised in this study (laminaribiose, trehalose and sophorose), D-cellobiose was produced within 8 h with yield higher than 80% from low and high concentrations of starch. Since in vitro multi-enzymatic cascades have emerged as a promising biomanufacturing platform, this process proved to be robust and could represent an alternative approach to the current disaccharide production [118]. Although not covered in this review, manno-oligosaccharides are also of interest in the production of nutraceuticals and pharmaceuticals due to theirprebiotic effects and many other biological properties [119]. In this regard, it is important to mentionthe relevance of β-1,4-D-mannan phosphorylase (GH130) as a cost-effective biocatalyst towards the in vitro synthesis of pure linear short β-1,4-mannan chains (DP 16) [120], thus emphasising its potential for biotechnological applications.

4.2 β-1,2-D-Glucan phosphorylases

In the presence of inorganic phosphate >100 g scale β-1,2-glucan synthesis was achieved enzymatically using a two-enzyme system of sucrose phosphorylase from *Bifidobacterium longum* subsp. *longum* and 1,2-β-oligoglucan phosphorylase from *Listeria innocua* expressed in *E. coli* [121]. A similar scale was also achieved utilising sophorose as the acceptor [122]. Few tools currently exist to elucidate the nature of proteins that interact with CβGs and consequently this difficulty is reflected in the literature, however advances are being made in the form of acid depolymerised oligosaccharide microarrays and derived probes to follow such interactions [123].

4.3 β-1,3-D-Glucan phosphorylases (laminaridextrin and laminaribiose phosphorylases)

The families GH94 and GH112 have been previously reported to contain inverting GPs with regio- and stereospecificity towards the synthesis of β-1,3-linked disaccharide or oligosaccharides. The in vitro preparation of disaccharide has been catalysed by bacterial laminaribiose phosphorylase (LBP, EC 2.4.1.31) [124,125] whereas crystalline oligosaccharides (DP 30) have been produced by a partially purified β-1,3-D-glucan phosphorylase extract from the microalga *Euglena gracilis*

[126]. In order to understand the glycobiology of *E. gracilis* and circumvent the laborious preparation of extract as source of catalyst, the Field group identified the algal β-1,3-glucan phosphorylase and uncovered a new family of phosphorylases (GH149). The candidate phosphorylase sequence (EgP1) was selected through proteomic analysis from cellular protein lysate and expressed in *E. coli* for in vitro characterisation and confirmation of its catalytic activity. Moreover, EgP1 orthologous sequences led to a bacterial metagenomic sequence that was expressed in *E. coli* (Pro_7066) and assessed to validate its function as β-1,3-D-glucan phosphorylase. Despite comparable activity in the synthesis of short oligosaccharides (DP up to 12) from D-Glc and α-D-Glc-1P, kinetic data suggested EgP1 preference for glucose and cellobiose whereas Pro_7066 showed similar catalytic efficiency for all tested acceptors (DP1 to DP6) [127]. In a continuous effort to identify and characterise new β-1,3-D-glucan phosphorylases, a phosphorylase sequence from heterokont algae *Ochromonas* spp (OcP1) was identified along with bacterial orthologs, thus leading to the identification of a new GH family (GH161). The activity of GH161 family members was established from a bacterial *GH161* gene sequence (*PapP*, *Paenibacillus polymyxa*) cloned and expressed in *E. coli*. Recombinant PapP showed neither synthetic activity in the presence of D-Glc and α-D-Glc-1P, nor phosphorolysis activity towards Glc-Glc disaccharides with different linkages. Nonetheless, phosphorolysis assays with DP3 to DP6 revealed its preference for longer substrates. These results strongly suggest that PapP can only operate on β-1,3 linear oligosaccharide acceptors with DP ≥ 2, evidencing its different specificity from GH149 glucan phosphorylases [128]. These studies expand the repertoire of GPs acting on β-1,3-D-glucans and provide information on substrate preference/specificity.

In order to gain understanding at the molecular level, the X-ray crystallographic structures of Pro_7066 (GH149) in the absence of substrate and in complex with laminarihexaose were solved in order to elucidate its chain length specificity for oligosaccharides. Although the overall domain organisation is similar to GH94, Pro_7066 enzyme contains two additional distinct domains flanking its catalytic region and a surface oligosaccharide binding site where laminarihexaose was

accommodated, which is distant from the catalytic site and may be involved in the recognition of longer substrates [129]. The crystallographic structure of the GH94 laminaribiose phosphorylase from *Paenibacillus* sp YM-1 (PsLBP) demonstrated its specificity for disaccharide. Interestingly, PsLBP was not only limited to act on D-laminaribiose as previously described [124], but also catalysed the synthesis of β-D-mannopyranosyl-1,3-D-glucopyranose from D-Glc and α-D-Man-1P, albeit with reduced catalytic efficiency (ca. 150-fold). Structural data of PsLBP in complex with α-D-Glc-1P and α-D-Man-1P, together with saturation transfer difference (STD) NMR studies, revealed a similar binding mode for α-D-Man-1P due to its close overlapping with α-D-Glc-1P. However, it was observed the loss of an important hydrogen bond between the axial hydroxy group at C2 and a key residue in the active site, thus possibly contributing to the low reaction turnover [130]. This study pioneered the molecular detailed recognition of an unnatural donor substrate by GPs as a means to provide background knowledge to harness their prospective as biocatalyst for synthetic applications. More recently, wild-type Pro_7066 and CDP showed reasonable tolerance to a variety of sugar 1- phosphates in a multi-milligram-scale synthesis of fragments of human milk oligosaccharides (HMOs) [131], which are currently a hot topic in enzymatic syntheses [132]. Kinetic data of these enzymes indicated general low efficiency (< 1%) for the unnatural donors α-D-Gal-1P, α-D-GlcN-1P and α-D-Man-1P compared to α-D-Glc-1P in the presence of both D-laminaribiose and D-cellobiose, though the transference of glucose onto both acceptors was effective. By producing mixed new β-1,4- and β-1,3-linkages from different donor substrates, the idea of 'one enzyme to one linkage' can no longer be applied to these phosphorylases in spite of their stereo- and regiospecific. Despite the natural tendency to generate long oligosaccharides, single turnover also was observed in some reactions. Considering that the lack of suitable enzymes hampers the synthesis of novel oligosaccharides, these results are highly encouraging as they uncover new features of GPs and start unfolding access to unique structures.

Although the synthetic potential of β-1,3-D-glucan phosphorylases towards long oligosaccharides remains underexplored compared toβ-1,4-D-glucan

phosphorylases, the biocatalytic synthesis of D-laminaribiose has been investigated in combination with other enzymes aiming at the industrial scale production in view of its biological importance and commercial applications. Many studies have demonstrated the outstanding in vitro activity of SCP as part of multi-enzyme synthetic processes in which downstream reactions catalysed by other phosphorylases use α-D-Glc-1P generated in situ through the conversion of expedient feedstock, thus leading to highly valuable sugars as already exemplified herein for β-1,4-linkage and broadly covered in recent reviews [23,133,134]. In this aspect, D-laminaribiose was produced from sucrose and D-Glc in one-pot reaction containing individually immobilised SCP and *E. gracilis* extract with enriched LBP activity (Figure 6c). The immobilised LBP onto Sepabeads retained over 50% activity when tested seven times, but the yield was low (20%). It has been suggested that the presence of other phosphorylases in the extract could generate longer oligosaccharides, which was confirmed by detection of laminaritriose [135]. The immobilised enzymes were also encapsulated with chitosan to improve thermal and operational stability. Significant increase of half-life (10-fold) and no activity loss after 12 times reuse were observed, but no yield improvement [136]. Additionally, the reportedly absence of an accessible recombinant LBP led the group to improve the LBP production in *E. gracilis* by testing different cultivation methods [137], although a recombinant preparation of LBP had been previously reported [124]. More recently, the overall improved immobilised two-enzyme system was used in a packed bed reactor for the first continuous production of D-laminaribiose. After optimisation of different parameters, the system exhibited operational stability during the course of 10 days and yielded over 46 g/L disaccharide maintaining the half-life of both biocatalysts [138]. Since the high production in continuous operation resulted in lower titre and purity, the system was further optimised by combining the adsorbent BEA zeolite in a consecutive packed bed column as a purification step. This approach increased the laminaribiose purity in 200-fold keeping a similar titre and yield [139].

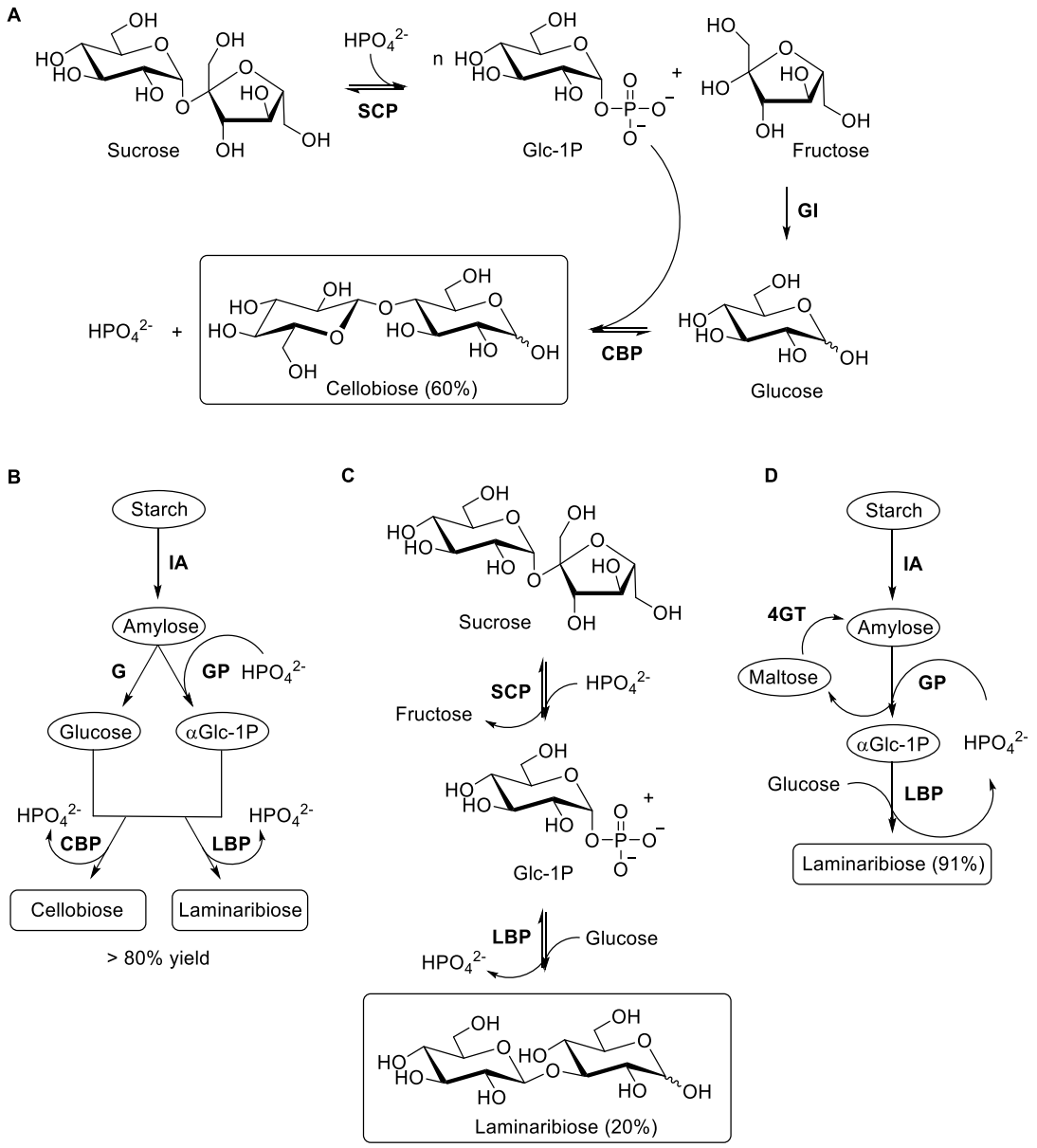


Figure 6. A) One-pot synthesis of D-cellobiose from sucrose catalysed by sucrose phosphorylase (SCP), glucose isomerase (GI) and cellobiose phosphorylase (CBP). **B**) Enzymatic platform for the synthesis of various disaccharides from starch employing isoamylase (IA), α-glucosidase (αG) and α-glucan phosphorylase (αGP) in parallel, and a disaccharide phosphorylase such as CBP or laminaribiose phosphorylase (LBP). **C**) D-laminaribiose synthesis from sucrose and D-Glc in one-pot reaction containing individually immobilised SCP and *E. gracilis* extract with enriched LBP activity. **D**) Cascade synthesis of D-laminaribiose from starch and glucose using IA, αGP, LBP and 4-α-glucanotransferase (4GT), which was used to recycle maltose into malto-oligosaccharides for continuous production of α-D-Glc-1P.

Despite all advantages of using sucrose as substrate and SCP as biocatalyst towards the synthesis of functional disaccharides, alternative multi-enzyme approaches have been also investigated using starch. Starch is cheaper than sucrose, the phosphorolysis produces more α-D-Glc-1P and the ratio α-D-Glc-1P/D-Glc can be tuned for optimisation[118]**.** In this sense, D-laminaribiose was successfully synthesised from starch and glucose by four thermophilic enzymes with 91% yield within 36 h. The process involved debranching of starch with isoamylase for conversion into α-D-Glc-1P by α-glucan phosphorylase and final catalysis by LBP (*Paenibacillus* sp YM-1) in the presence of external D-Glc. The maltose generated as by-product of α-glucan phosphorylase reaction was recycled into malto-oligosaccharides by 4-α-glucanotransferase for continuous production of α-D-Glc-1P (Figure 6d). High concentrations of substrates were also tested to demonstrate the industrial potential of this system, but it became a limitation to be addressed [140]. The follow-up study involving starch as the only substrate and a mathematical model to predict optimal conditions led to high production of D-laminaribiose (> 80%) from low and high concentrations of starch (Figure 6b). Some data indicated that D-Glc was slowly released from starch and remained at low concentration, thus minimising not only inhibition of the enzymes but also the effect of Maillard reaction during the process at high temperature [118].

5. Conclusions

The use of natural glycosyltransferases and unnatural glycoside hydrolase mutants has broadened the scope for β-glucan synthesis, which, in tandem with phosphorylases, offers great diversity in producing designer β-glucans. The reconstitution of glycosyltransferases in vitro has enabled the polymerisation of glucose to form a variety of cellulose morphologies, whilst the advancements seen with glycosynthases offer great potential for the design of β-glucans with tailored functionalities. The current applications of β-1,4- and β-1,3-D-glucan phosphorylases highlight their relevance along with other important carbohydrate-active enzymes. From solo catalysis to multi-enzymatic cascades and from long-chain insoluble to short-chain soluble oligosaccharides, these wild-type enzymes are capable of producing a broad range of well-defined natural and unnatural

structures in a controlled manner due to their tolerance to different reaction conditions and substrates. In particular, the oligomerisation-induced self-assembly of insoluble β-1,4 oligomers into highly ordered hierarchical nanostructures can be directed by simply tuning reaction conditions in the presence of natural substrates or using chemically functionalised substrates, thus representing a promising approach towards the bottom-up preparation of tailored cellulose-like materials. Considering what has been recently achieved, but also the prospect discovery of new phosphorylases [141], the great potential biocatalyst of such robust and versatile enzymes in academic research and industrial application becomes more evident.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the UK BBSRC via a personal fellowship (B/P011462/1), and the EPSRC via the GelEnz consortium (IUK 59000 442149) and DTP EP/N509565/1.

References

[1] M. Jayachandran, J. Chen, S.S.M. Chung, B. Xu, A critical review on the impacts of β-glucans on gut microbiota and human health, J. Nutr. Biochem. 61 (2018) 101–110.

[2] C.O. Tuck, E. Pérez, I.T. Horváth, R.A. Sheldon, M. Poliakoff, Valorization of biomass: Deriving more value from waste, Science. 337 (2012) 695–699.

[3] O. Rosales-Calderon, V. Arantes, A review on commercial-scale high-value products that can be produced alongside cellulosic ethanol, Biotechnol. Biofuels. 12 (2019) 1–58.

[4] S.J. Eichhorn, A. Dufresne, M. Aranguren, N.E. Marcovich, J.R. Capadona, S.J. Rowan, C. Weder, W. Thielemans, M. Roman, S. Renneckar, W. Gindl, S.

Veigel, J. Keckes, H. Yano, K. Abe, M. Nogi, A.N. Nakagaito, A. Mangalam, J. Simonsen, A.S. Benight, A. Bismarck, L.A. Berglund, T. Peijs, Review: Current international research into cellulose nanofibres and nanocomposites, J. Mater. Sci. 45 (2010) 1–33.

[5] C. Chen, Y. Kuang, S. Zhu, I. Burgert, T. Keplinger, A. Gong, T. Li, L. Berglund, S.J. Eichhorn, L. Hu, Structure–property–function relationships of natural and engineered wood, Nat. Rev. Mater. 5 (2020) 642–666.

[6] M.F. Moradali, B.H.A. Rehm, Bacterial biopolymers: from pathogenesis to advanced materials, Nat. Rev. Microbiol. 18 (2020) 195–210.

[7] J. Wang, J. Niu, T. Sawada, Z. Shao, T. Serizawa, A Bottom-Up Synthesis of Vinyl-Cellulose Nanosheets and Their Nanocomposite Hydrogels with Enhanced Strength, Biomacromolecules. 18 (2017) 4196–4205.

[8] A. Dufresne, Nanocellulose Processing Properties and Potential Applications, Curr. For. Reports. 5 (2019) 76–89.

[9] T. Owoyokun, C.M. Pérez Berumen, A.M. Luévanos, L. Cantú, A.C. Lara Ceniceros, Cellulose nanocrystals: Obtaining and sources of a promising bionanomaterial for advanced applications, Biointerface Res. Appl. Chem. 11 (2021) 11797–11816.

[10] E. Niinivaara, E.D. Cranston, Bottom-up assembly of nanocellulose structures, Carbohydr. Polym. 247 (2020) 116664.

[11] M.W. Breedveld, K.J. Miller, Cyclic β-glucans of members of the family Rhizobiaceae, Microbiol. Rev. 58 (1994) 145–161.

[12] A. Martirosyan, C. Pérez-Gutierrez, R. Banchereau, H. Dutartre, P. Lecine, M. Dullaers, M. Mello, S. Pinto Salcedo, A. Muller, L. Leserman, Y. Levy, G. Zurawski, S. Zurawski, E. Moreno, I. Moriyón, E. Klechevsky, J. Banchereau, S.K. Oh, J.P. Gorvel, Brucella β 1,2 Cyclic Glucan Is an Activator of Human and Mouse Dendritic Cells, PLoS Pathog. 8 (2012) e1002983.

[13] S. Javvadi, S.S. Pandey, A. Mishra, B.B. Pradhan, S. Chatterjee, Bacterial cyclic b-(1,2)-glucans sequester iron to protect against iron-induced toxicity, EMBO Rep. 19 (2018) 172–186.

[14] L.M. Papaspyridi, A. Zerva, E. Topakas, Biocatalytic synthesis of fungal β-glucans, Catalysts. 8 (2018) 274.

[15] T. Kimura, Natural products and biological activity of the pharmacologically active cauliflower mushroom Sparassis crispa, Biomed Res. Int. 2013 (2013)982317.

[16] A. Synytsya, M. Novák, Structural diversity of fungal glucans, Carbohydr. Polym. 92 (2013) 792–809.

[17] H. Xu, S. Zou, X. Xu, L. Zhang, Anti-Tumor effect of β-glucan from Lentinus edodes and the underlying mechanism, Sci. Rep. 6 (2016) 1–13.

[18] K. Lee, Y. Kwon, J. Hwang, Y. Choi, K. Kim, H.J. Koo, Y. Seo, H. Jeon, J. Choi, Synthesis and Functionalization of β-Glucan Particles for the Effective Delivery of Doxorubicin Molecules, ACS Omega. 4 (2019) 668–674.

[19] K.M.I. Bashir, J.S. Choi, Clinical and physiological perspectives of β-glucans: The past, present, and future, Int. J. Mol. Sci. 18 (2017).

[20] C.R. Santos, P.A.C.R. Costa, P.S. Vieira, S.E.T. Gonzalez, T.L.R. Correa, E.A. Lima, F. Mandelli, R.A.S. Pirolla, M.N. Domingues, L. Cabral, M.P. Martins, R.L. Cordeiro, A.T. Junior, B.P. Souza, É.T. Prates, F.C. Gozzo, G.F. Persinoti, M.S. Skaf, M.T. Murakami, Structural insights into β-1,3-glucan cleavage by a glycoside hydrolase family, Nat. Chem. Biol. 16 (2020) 920–929.

[21] B.A. Stone, Chemistry of β-glucans, Elsevier Inc., 2009.

[22] H. Zhang, F. Zhang, R. Yuan, Applications of natural polymer-based hydrogels in the food industry, Elsevier Inc., 2019.

[23] G. Pergolizzi, S. Kuhaudomlarp, E. Kalita, R.A. Field, Glycan Phosphorylases in Multi-Enzyme Synthetic Processes, Protein Pept. Lett. 24 (2017)696–709.

[24] C. Laroche, P. Michaud, New Developments and Prospective Applications for β (1,3) Glucans, Recent Pat. Biotechnol. 1 (2008) 59–73.

[25] L. Barsanti, V. Passarelli, V. Evangelista, A.M. Frassanito, P. Gualtieri, Chemistry, physico-chemistry and applications linked to biological activities of β-glucans, Nat. Prod. Rep. 28 (2011) 457–466.

[26] M. Del Cornò, S. Gessani, L. Conti, Shaping the innate immune response by dietary glucans: Any role in the control of cancer?, Cancers. 12 (2020) 155.

[27] A. Miyagawa, H. Yamamura, Synthesis of β-1,3-glucan mimics by β-1,3-glucan trisaccharyl monomer polymerization, Carbohydr. Polym. 227 (2020) 115105.

[28] B. Nidetzky, C. Zhong, Phosphorylase-catalyzed bottom-up synthesis of short-chain soluble cello-oligosaccharides and property-tunable cellulosic materials, Biotechnol. Adv. (2020) 107633.

[29] E.C. O'Neill, R.A. Field, Enzymatic synthesis using glycoside phosphorylases, Carbohydr. Res. 403 (2015) 23–37.

[30] R.A. Sheldon, D. Brady, M.L. Bode, The Hitchhiker's guide to biocatalysis: Recent advances in the use of enzymes in organic synthesis, Chem. Sci. 11 (2020) 2587–2605.

[31] R.M. Schmaltz, S.R. Hanson, C.H. Wong, Enzymes in the synthesis of glycoconjugates, Chem. Rev. 111 (2011) 4259–4307.

[32] J.B. McArthur, X. Chen, Glycosyltransferase engineering for carbohydrate synthesis, Biochem. Soc. Trans. 44 (2016) 129–142.

[33] B. Nidetzky, A. Gutmann, C. Zhong, Leloir Glycosyltransferases as Biocatalysts for Chemical Production, ACS Catal. 8 (2018) 6283–6300.

[34] L. Ban, N. Pettit, L. Li, A.D. Stuparu, L. Cai, W. Chen, W. Guan, W. Han, P.G. Wang, M. Mrksich, Discovery of glycosyltransferases using carbohydrate arrays and mass spectrometry, Nat. Chem. Biol. 8 (2012) 769–773.

[35] C.A.G.M. Weijers, M.C.R. Franssen, G.M. Visser, Glycosyltransferase-catalyzed synthesis of bioactive oligosaccharides, Biotechnol. Adv. 26 (2008) 436–456.

[36] J.M. Cote, E.A. Taylor, The glycosyltransferases of LPS core: A review of four heptosyltransferase enzymes in context, Int. J. Mol. Sci. 18 (2017) 2256.

[37] A. Derouaux, E. Sauvage, M. Terrak, Peptidoglycan glycosyltransferase substrate mimics as templates for the design of new antibacterial drugs, Front. Immunol. 4 (2013) 78.

[38] L. Mestrom, M. Przypis, D. Kowalczykiewicz, A. Pollender, A. Kumpf, S.R. Marsden, I. Bento, A.B. Jarzębski, K. Szymańska, A. Chruściel, D. Tischler, R. Schoevaart, U. Hanefeld, P.L. Hagedoorn, Leloir glycosyltransferases in applied biocatalysis: A multidisciplinary approach, Int. J. Mol. Sci. 20 (2019) 5263.

[39] K. Schmölzer, A. Gutmann, M. Diricks, T. Desmet, B. Nidetzky, Sucrose synthase: A unique glycosyltransferase for biocatalytic glycosylation process

development, Biotechnol. Adv. 34 (2016) 88–111.

[40] A. Gutmann, A. Lepak, M. Diricks, T. Desmet, B. Nidetzky, Glycosyltransferase cascades for natural product glycosylation: Use of plant instead of bacterial sucrose synthases improves the UDP-glucose recycling from sucrose and UDP, Biotechnol. J. 12 (2017) 1–10.

[41] M. Diricks, A. Gutmann, S. Debacker, G. Dewitte, B. Nidetzky, T. Desmet, Sequence determinants of nucleotide binding in Sucrose Synthase: Improving the affinity of a bacterial Sucrose Synthase for UDP by introducing plant residues, Protein Eng. Des. Sel. 30 (2017) 143–150.

[42] K. Schmölzer, M. Lemmerer, A. Gutmann, B. Nidetzky, Integrated process design for biocatalytic synthesis by a Leloir Glycosyltransferase: UDP-glucose production with sucrose synthase, Biotechnol. Bioeng. 114 (2017) 924–928.

[43] J.T. McNamara, J.L.W. Morgan, J. Zimmer, A molecular description of cellulose biosynthesis, Annu. Rev. Biochem. 84 (2015) 895–921.

[44] S. Turner, M. Kumar, Cellulose synthase complex organization and cellulose microfibril structure, Philos. Trans. R. Soc. A Math. Phys. Eng. Sci. 376 (2018) 9–16.

[45] D.P. Delmer, Cellulose biosynthesis: Exciting times for a difficult field of study, Annu. Rev. Plant Biol. 50 (1999) 245–276.

[46] M. Florea, H. Hagemann, G. Santosa, J. Abbott, C.N. Micklem, X. Spencer-Milnes, L. De Arroyo Garcia, D. Paschou, C. Lazenbatt, D. Kong, H. Chughtai, K. Jensen, P.S. Freemont, R. Kitney, B. Reeve, T. Ellis, Engineering control of bacterial cellulose production using a genetic toolkit and a new celluloseproducing strain, Proc. Natl. Acad. Sci. U. S. A. 113 (2016) E3431–E3440.

[47] J. Lai-Kee-Him, H. Chanzy, M. Müller, J.L. Putaux, T. Imai, V. Bulone, In vitro versus in vivo cellulose microfibrils from plant primary wall synthases: Structural differences, J. Biol. Chem. 277 (2002) 36931–36939.

[48] S.H. Cho, J. Du, I. Sines, V.G. Poosarla, V. Vepachedu, K. Kafle, Y.B. Park, S.H. Kim, M. Kumar, B.T. Nixon, In vitro synthesis of cellulose microfibrils by a membrane protein from protoplasts of the non-vascular plant Physcomitrella patens, Biochem. J. 470 (2015) 195–205.

[49] C. Cifuentes, V. Bulone, A.M.C. Emons, Biosynthesis of Callose and Cellulose by Detergent Extracts of Tobacco Cell Membranes and Quantification of the Polymers Synthesized in vitro, J. Integr. Plant Biol. 52 (2010) 221–233.

[50] P. Purushotham, S.H. Cho, S.M. Díaz-Moreno, M. Kumar, B.T. Nixon, V. Bulone, J. Zimmer, A single heterologously expressed plant cellulose synthase isoform is sufficient for cellulose microfibril formation in vitro, Proc. Natl. Acad. Sci. U. S. A. 113 (2016) 11360–11365.

[51] U. Römling, M.Y. Galperin, Bacterial cellulose biosynthesis: Diversity of operons, subunits, products, and functions, Trends Microbiol. 23 (2015) 545–557.

[52] O. Omadjela, A. Narahari, J. Strumillo, H. Mélida, O. Mazur, V. Bulone, J. Zimmer, BcsA and BcsB form the catalytically active core of bacterial cellulose synthase sufficient for in vitro cellulose synthesis, Proc. Natl. Acad. Sci. U. S. A. 110 (2013) 17856–17861.

[53] P.A. Penttilä, J. Sugiyama, T. Imai, Effects of reaction conditions on cellulose structures synthesized in vitro by bacterial cellulose synthases, Carbohydr. Polym. 136 (2016) 656–666.

[54] G.S. Bulmer, A.P. Mattey, F. Parmeggiani, R. Williams, H. Ledru, A. Marchesi, L.S. Seibt, P. Both, K. Huang, M.C. Galan, S.L. Flitsch, A.P. Green, J.M. van Munster, A promiscuous glycosyltransferase generates poly-β-1,4-glucan derivatives that facilitate mass spectrometry-based detection of cellulolytic enzymes, Org. Biomol. Chem. 19 (2021) 5529–5523.

[55] J.L.K. Him, L. Pelosi, H. Chanzy, J.L. Putaux, V. Bulone, Biosynthesis of (1→3)-β-D-glucan (callose) by detergent extracts of a microsomal fraction from Arabidopsis thaliana, Eur. J. Biochem. 268 (2001) 4628–4638.

[56] K. Okuda, L. Li, K. Kudlicka, S. Kuga, R.M. Brown, β-Glucan synthesis in the cotton fiber: I. Identification of β-1,4- and β-1,3-glucans synthesized in vitro, Plant Physiol. 101 (1993) 1131–1142.

[57] G. Billon-Grand, M.F. Marais, J.P. Joseleau, V. Girard, L. Gay, M. Fèvre, A novel 1.3-β-glucan synthase from the oomycete Saprolegnia monoica, Microbiology. 143 (1997) 3175–3183.

[58] S. Tagawa, T. Kondo, Secretion of a callose hollow fiber from herbaceous

plant protoplasts induced by inhibition of cell wall formation, J. Wood Sci. 64 (2018) 467–476.

[59] S. Matsuo, S. Tagawa, Y. Matsusaki, Y. Uchi, T. Kondo, Callose-synthesizing enzymes as membrane proteins of Betula protoplasts secrete bundles of β-1,3-glucan hollow fibrils under Ca2+-rich and acidic culture conditions, Holzforschung. 74 (2020) 725–732.

[60] L.F. Mackenzie, Q. Wang, R.A.J. Warren, S.G. Withers, Glycosynthases: Mutant glycosidases for oligosaccharide synthesis, J. Am. Chem. Soc. 120 (1998) 5583–5584.

[61] S.R. Marsden, L. Mestrom, D.G.G. McMillan, U. Hanefeld, Thermodynamically and Kinetically Controlled Reactions in Biocatalysis – from Concepts to Perspectives, ChemCatChem. 12 (2020) 426–437.

[62] P.M. Danby, S.G. Withers, Advances in Enzymatic Glycoside Synthesis, ACS Chem. Biol. 11 (2016) 1784–1794.

[63] G. Ramani, B. Meera, J. Rajendhran, P. Gunasekaran, Transglycosylating glycoside hydrolase family 1 β-glucosidase from Penicillium funiculosum NCL1: Heterologous expression in Escherichia coli and characterization, Biochem. Eng. J. 102 (2015) 6–13.

[64] M.I. Kusaykin, A.A. Belik, S.N. Kovalchuk, P.S. Dmitrenok, V.A. Rasskazov, V. V Isakov, T.N. Zvyagintseva, A new recombinant endo-1,3-β-d-glucanase from the marine bacterium Formosa algae KMM 3553: enzyme characteristics and transglycosylation products analysis, World J. Microbiol. Biotechnol. 33 (2017) 40.

[65] M.S. Mafa, H.W. Dirr, S. Malgas, R.W.M. Krause, K. Rashamuse, B.I. Pletschke, A Novel Dimeric Exoglucanase (GH5_38): Biochemical and Structural Characterisation towards its Application in Alkyl Cellobioside Synthesis, Molecules. 25 (2020) 746.

[66] S. Kobayashi, K. Kashiwa, T. Kawasaki, S.I. Shoda, Novel Method for Polysaccharide Synthesis Using an Enzyme: The First in Vitro Synthesis of Cellulose via a Nonbiosynthetic Path Utilizing Cellulase as Catalyst, J. Am. Chem. Soc. 113 (1991) 3079–3084.

[67] M.R. Hayes, J. Pietruszka, Synthesis of glycosides by glycosynthases,

Molecules. 22 (2017) 1434.

[68] M. Faijes, X. Pérez, O. Pérez, A. Planas, Glycosynthase Activity of Bacillus licheniformis 1,3-1,4-β-Glucanase Mutants: Specificity, Kinetics, and Mechanism, Biochemistry. 42 (2003) 13304–13318.

[69] M.R. Hayes, J. Pietruszka, Synthesis of glycosides by glycosynthases, Molecules. 22 (2017).

[70] J.A. Méndez-Líter, M. Nieto-Domínguez, B. Fernández De Toro, A. González Santana, A. Prieto, J.L. Asensio, F.J. Cañada, L.I. De Eugenio, M.J. Martínez, A glucotolerant β-glucosidase from the fungus Talaromyces amestolkiae and its conversion into a glycosynthase for glycosylation of phenolic compounds, Microb. Cell Fact. 19 (2020) 1–13.

[71] M. Hrmova, T. Imai, S.J. Rutten, J.K. Fairweather, L. Pelosi, V. Bulone, H. Driguez, G.B. Fincher, Mutated barley (1,3)-beta-D-glucan endohydrolases synthesize crystalline (1,3)-beta-D-glucans., J. Biol. Chem. 277 (2002) 30102–30111.

[72] X. Pérez, M. Faijes, A. Planas, Artificial mixed-linked β-glucans produced by glycosynthase-catalyzed polymerization: Tuning morphology and degree of polymerization, Biomacromolecules. 12 (2011) 494–501.

[73] T. Pozzo, M. Plaza, J. Romero-García, M. Faijes, E.N. Karlsson, A. Planas, Glycosynthases from Thermotoga neapolitana β-glucosidase 1A: A comparison of α-glucosyl fluoride and in situ-generated α-glycosyl formate donors, J. Mol. Catal. B Enzym. 107 (2014) 132–139.

[74] T. Pozzo, J. Romero-García, M. Faijes, A. Planas, E.N. Karlsson, Rational design of a thermostable glycoside hydrolase from family 3 introduces β-glycosynthase activity, Glycobiology. 27 (2017) 165–175.

[75] L. Ruzic, J.M. Bolivar, B. Nidetzky, Glycosynthase reaction meets the flow: Continuous synthesis of lacto-N-triose II by engineered β-hexosaminidase immobilized on solid support, Biotechnol. Bioeng. 117 (2020) 1597–1602.

[76] S. Fort, V. Boyer, L. Greffe, G.J. Davies, O. Moroz, L. Christiansen, M. Schülein, S. Cottaz, H. Driguez, Highly efficient synthesis of β(1 → 4)-oligo- and -polysaccharides using a mutant cellulase, J. Am. Chem. Soc. 122 (2000) 5429–5437.

[77] V. Codera, K.J. Edgar, M. Faijes, A. Planas, Functionalized Celluloses with Regular Substitution Pattern by Glycosynthase-Catalyzed Polymerization, Biomacromolecules. 17 (2016) 1272–1279.

[78] Z. Armstrong, F. Liu, H.M. Chen, S.J. Hallam, S.G. Withers, Systematic Screening of Synthetic Gene-Encoded Enzymes for Synthesis of Modified Glycosides, ACS Catal. 9 (2019) 3219–3227.

[79] H. Nakai, M. Kitaoka, B. Svensson, K. Ohtsubo, Recent development of phosphorylases possessing large potential for oligosaccharide synthesis, Curr. Opin. Chem. Biol. 17 (2013) 301–309.

[80] V. Puchart, Glycoside phosphorylases: Structure, catalytic properties and biotechnological potential, Biotechnol. Adv. 33 (2015) 261–276.

[81] F.N. Awad, Glycoside phosphorylases for carbohydrate synthesis: An insight into the diversity and potentiality, Biocatal. Agric. Biotechnol. 31 (2021) 101886.

[82] E.C. O'Neill, G. Pergolizzi, C.E.M. Stevenson, D.M. Lawson, S.A. Nepogodiev, R.A. Field, Cellodextrin phosphorylase from Ruminiclostridium thermocellum: X-ray crystal structure and substrate specificity analysis, Carbohydr. Res. 451 (2017) 118–132.

[83] Y. Yataka, T. Sawada, T. Serizawa, Enzymatic synthesis and post-functionalization of two-dimensional crystalline cellulose oligomers with surface-reactive groups, Chem. Commun. 51 (2015) 12525–12528.

[84] A. Adharis, D.M. Petrović, I. Özdamar, A.J.J. Woortman, K. Loos, Environmentally friendly pathways towards the synthesis of vinyl-based oligocelluloses, Carbohydr. Polym. 193 (2018) 196–204.

[85] T. Nohara, T. Sawada, H. Tanaka, T. Serizawa, Enzymatic synthesis and protein adsorption properties of crystalline nanoribbons composed of cellulose oligomer derivatives with primary amino groups, J. Biomater. Sci. Polym. Ed. 28 (2017) 925–938.

[86] M. Hanamura, T. Sawada, T. Serizawa, In-Paper Self-Assembly of Cellulose Oligomers for the Preparation of All-Cellulose Functional Paper, ACS Sustain. Chem. Eng. 9 (2021) 5684–5692.

[87] C. Zhong, K. Zajki-Zechmeister, B. Nidetzky, Reducing end thiol-modified

nanocellulose: Bottom-up enzymatic synthesis and use for templated assembly of silver nanoparticles into biocidal composite material, Carbohydr. Polym. 260 (2021) 117772.

[88] T. Nohara, T. Sawada, H. Tanaka, T. Serizawa, Enzymatic Synthesis of Oligo(ethylene glycol)-Bearing Cellulose Oligomers for in Situ Formation of Hydrogels with Crystalline Nanoribbon Network Structures, Langmuir. 32 (2016) 12520–12526.

[89] Y. Yataka, T. Sawada, T. Serizawa, Multidimensional Self-Assembled Structures of Alkylated Cellulose Oligomers Synthesized via in Vitro Enzymatic Reactions, Langmuir. 32 (2016) 10120–10125.

[90] Y. Kita, R. Kusumi, T. Kimura, M. Kitaoka, Y. Nishiyama, M. Wada, Surface structural analysis of selectively 13C-labeled cellulose II by solid-state NMR spectroscopy, Cellulose. 27 (2020) 1899–1907.

[91] P. de Andrade, J.C. Muñoz-García, G. Pergolizzi, V. Gabrielli, S.A. Nepogodiev, D. Iuga, L. Fábián, R. Nigmatullin, M.A. Johns, R. Harniman, S.J. Eichhorn, J. Angulo, Y.Z. Khimyak, R.A. Field, Chemoenzymatic Synthesis of Fluorinated Cellodextrins Identifies a New Allomorph for Cellulose-Like Materials, Chem. Eur. J. 27 (2021) 1374–1382.

[92] Y. Hata, T. Sawada, T. Serizawa, Macromolecular crowding for materials-directed controlled self-assembly, J. Mater. Chem. B. 6 (2018) 6344–6359.

[93] Y. Hata, T. Kojima, T. Koizumi, H. Okura, T. Sakai, T. Sawada, T. Serizawa, Enzymatic Synthesis of Cellulose Oligomer Hydrogels Composed of Crystalline Nanoribbon Networks under Macromolecular Crowding Conditions, ACS Macro Lett. 6 (2017) 165–170.

[94] Y. Hata, T. Sawada, T. Serizawa, Effect of solution viscosity on the production of nanoribbon network hydrogels composed of enzymatically synthesized cellulose oligomers under macromolecular crowding conditions, Polym. J. 49 (2017) 575–581.

[95] Y. Hata, T. Kojima, T. Maeda, T. Sawada, T. Serizawa, pH-Triggered Self-Assembly of Cellulose Oligomers with Gelatin into a Double-Network Hydrogel, Macromol. Biosci. 20 (2020) 1–8.

[96] T. Serizawa, T. Maeda, T. Sawada, Neutralization-Induced Self-Assembly of

Cellulose Oligomers into Antibiofouling Crystalline Nanoribbon Networks in Complex Mixtures, ACS Macro Lett. 9 (2020) 301–305.

[97] Y. Hata, T. Sawada, T. Sakai, T. Serizawa, Enzyme-Catalyzed Bottom-Up Synthesis of Mechanically and Physicochemically Stable Cellulose Hydrogels for Spatial Immobilization of Functional Colloidal Particles, Biomacromolecules. 19 (2018) 1269–1275.

[98] Y. Hata, T. Sawada, H. Marubayashi, S. Nojima, T. Serizawa, Temperature-Directed Assembly of Crystalline Cellulose Oligomers into Kinetically Trapped Structures during Biocatalytic Synthesis, Langmuir. 35 (2019) 7026–7034.

[99] Y. Hata, Y. Fukaya, T. Sawada, M. Nishiura, T. Serizawa, Biocatalytic oligomerization-induced self-assembly of crystalline cellulose oligomers into nanoribbon networks assisted by organic solvents, Beilstein J. Nanotechnol. 10 (2019) 1778–1788.

[100] T. Serizawa, Y. Fukaya, T. Sawada, Self-Assembly of Cellulose Oligomers into Nanoribbon Network Structures Based on Kinetic Control of Enzymatic Oligomerization, Langmuir. 33 (2017) 13415–13422.

[101] R. Pylkkänen, P. Mohammadi, S. Arola, J.C. De Ruijter, N. Sunagawa, K. Igarashi, M. Penttilä, In Vitro Synthesis and Self-Assembly of Cellulose II Nanofibrils Catalyzed by the Reverse Reaction of Clostridium thermocellum Cellodextrin Phosphorylase, Biomacromolecules. 21 (2020) 4355–4364.

[102] M. Hiraishi, K. Igarashi, S. Kimura, M. Wada, M. Kitaoka, M. Samejima, Synthesis of highly ordered cellulose II in vitro using cellodextrin phosphorylase, Carbohydr. Res. 344 (2009) 2468–2473.

[103] T. Serizawa, M. Kato, H. Okura, T. Sawada, M. Wada, Hydrolytic activities of artificial nanocellulose synthesized via phosphorylase-catalyzed enzymatic reactions, Polym. J. 48 (2016) 539–544.

[104] T. Serizawa, Y. Fukaya, T. Sawada, Nanoribbon network formation of enzymatically synthesized cellulose oligomers through dispersion stabilization of precursor particles, Polym. J. 50 (2018) 799–804.

[105] E. Samain, C. Lancelon-Pin, F. Férigo, V. Moreau, H. Chanzy, A. Heyraud, H. Driguez, Phosphorolytic synthesis of cellodextrins, Carbohydr. Res. 271 (1995) 217–226.

[106] Y. Uyeno, S. Shigemori, T. Shimosato, Effect of probiotics/prebiotics on cattle health and productivity, Microbes Environ. 30 (2015) 126–132.

[107] D. Mudgil, S. Barak, Composition, properties and health benefits of indigestible carbohydrate polymers as dietary fiber: A review, Int. J. Biol. Macromol. 61 (2013) 1–6.

[108] D.M. Petrović, I. Kok, A.J.J. Woortman, J. Ćirić, K. Loos, Characterization of Oligocellulose Synthesized by Reverse Phosphorolysis Using Different Cellodextrin Phosphorylases, Anal. Chem. 87 (2015) 9639–9646.

[109] C. Zhong, C. Luley-Goedl, B. Nidetzky, Product solubility control in cellooligosaccharide production by coupled cellobiose and cellodextrin phosphorylase, Biotechnol. Bioeng. 116 (2019) 2146–2155.

[110] C. Zhong, B. Nidetzky, Three-Enzyme Phosphorylase Cascade for Integrated Production of Short-Chain Cellodextrins, Biotechnol. J. 15 (2020) 1–9.

[111] C. Zhong, B. Duić, J.M. Bolivar, B. Nidetzky, Three-Enzyme Phosphorylase Cascade Immobilized on Solid Support for Biocatalytic Synthesis of Cello−oligosaccharides, ChemCatChem. 12 (2020) 1350–1358.

[112] C. Zhong, C. Ukowitz, K.J. Domig, B. Nidetzky, Short-Chain Cello-oligosaccharides: Intensification and Scale-up of Their Enzymatic Production and Selective Growth Promotion among Probiotic Bacteria, J. Agric. Food Chem. 68 (2020) 8557–8567.

[113] K. Pokusaeva, M. O'Connell-Motherway, A. Zomer, J. MacSharry, G.F. Fitzgerald, D. van Sinderen, Cellodextrin utilization by Bifidobacterium breve UCC2003, Appl. Environ. Microbiol. 77 (2011) 1681–1690.

[114] Z. Ubiparip, D.S. Moreno, K. Beerens, T. Desmet, Engineering of cellobiose phosphorylase for the defined synthesis of cellotriose, Appl. Microbiol. Biotechnol. 104 (2020) 8327–8337.

[115] M.R.M. De Groeve, L. Remmery, A. Van Hoorebeke, J. Stout, T. Desmet, S.N. Savvides, W. Soetaert, Construction of cellobiose phosphorylase variants with broadened acceptor specificity towards anomerically substituted glucosides, Biotechnol. Bioeng. 107 (2010) 413–420.

[116] F. Koch, T.J. Hässler, T., Kipping, Process for the enzymatic preparation of a product glucoside and of a co-product from an educt glucoside.

EP20150760467., 2016.

[117] C. Zhong, P. Wei, Y.H.P. Zhang, A kinetic model of one-pot rapid biotransformation of cellobiose from sucrose catalyzed by three thermophilic enzymes, Chem. Eng. Sci. 161 (2017) 159–166.

[118] S. Sun, X. Wei, X. Zhou, C. You, Construction of an Artificial in Vitro Synthetic Enzymatic Platform for Upgrading Low-Cost Starch to Value-Added Disaccharides, J. Agric. Food Chem. 69 (2021) 302–314.

[119] U.K. Jana, R.K. Suryawanshi, B.P. Prajapati, N. Kango, Prebiotic mannooligosaccharides: Synthesis, characterization and bioactive properties, Food Chem. 342 (2021) 128328.

[120] F. Grimaud, S. Pizzut-Serin, L. Tarquis, S. Ladevèze, S. Morel, J.L. Putaux, G. Potocki-Veronese, In Vitro Synthesis and Crystallization of β-1,4-Mannan, Biomacromolecules. 20 (2019) 846–853.

[121] K. Abe, M. Nakajima, M. Kitaoka, H. Toyoizumi, Y. Takahashi, N. Sugimoto, H. Nakai, H. Taguchi, Large-scale Preparation of 1,2-β-Glucan Using 1,2-β-Oligoglucan Phosphorylase, J. Appl. Glycosci. 62 (2015) 47–52.

[122] K. Kobayashi, M. Nakajima, H. Aramasa, S. Kimura, T. Iwata, H. Nakai, H. Taguchi, Large-scale preparation of β-1,2-glucan using quite a small amount of sophorose, Biosci. Biotechnol. Biochem. 83 (2019) 1867–1874.

[123] H. Zhang, A.S. Palma, Y. Zhang, R.A. Childs, Y. Liu, D.A. Mitchell, L.S. Guidolin, W. Weigel, B. Mulloy, A.E. Ciocchini, T. Feizi, W. Chai, Generation and characterization of β1,2-gluco-oligosaccharide probes from Brucella abortus cyclic β-glucan and their recognition by C-type lectins of the immune system, Glycobiology. 26 (2016) 1086–1096.

[124] M. Kitaoka, Y. Matsuoka, K. Mori, M. Nishimoto, K. Hayashi, Characterization of a bacterial laminaribiose phosphorylase, Biosci. Biotechnol. Biochem. 76 (2012) 343–348.

[125] T. Nihira, Y. Saito, M. Kitaoka, M. Nishimoto, K. Otsubo, H. Nakai, Characterization of a laminaribiose phosphorylase from Acholeplasma laidlawii PG-8A and production of 1,3-β-d-glucosyl disaccharides, Carbohydr. Res. 361 (2012) 49–54.

[126] Y. Ogawa, K. Noda, S. Kimura, M. Kitaoka, M. Wada, Facile preparation of

highly crystalline lamellae of (1→3)-β-d-glucan using an extract of euglena gracilis, Int. J. Biol. Macromol. 64 (2014) 415–419.

[127] S. Kuhaudomlarp, N.J. Patron, B. Henrissat, M. Rejzek, G. Saalbach, R.A. Field, Identification of Euglena gracilis β-1,3-glucan phosphorylase and establishment of a new glycoside hydrolase (GH) family GH149, J. Biol. Chem. 293 (2018) 2865–2876.

[128] S. Kuhaudomlarp, G. Pergolizzi, N.J. Patron, B. Henrissat, R.A. Field, Unraveling the subtleties of β-(1→3)-glucan phosphorylase specificity in the GH94, GH149, and GH161 glycoside hydrolase families, J. Biol. Chem. 294 (2019) 6483–6493.

[129] S. Kuhaudomlarp, C.E.M. Stevenson, D.M. Lawson, R.A. Field, The structure of a GH149 β-(1 → 3) glucan phosphorylase reveals a new surface oligosaccharide binding site and additional domains that are absent in the disaccharide-specific GH94 glucose-β-(1 → 3)-glucose (laminaribiose) phosphorylase, Proteins Struct. Funct. Bioinforma. 87 (2019) 885–892.

[130] S. Kuhaudomlarp, S. Walpole, C.E.M. Stevenson, S.A. Nepogodiev, D.M. Lawson, J. Angulo, R.A. Field, Unravelling the Specificity of Laminaribiose Phosphorylase from Paenibacillus sp. YM-1 towards Donor Substrates Glucose/Mannose 1-Phosphate by Using X-ray Crystallography and Saturation Transfer Difference NMR Spectroscopy, ChemBioChem. 20 (2019) 181–192.

[131] R.P. Singh, G. Pergolizzi, S.A. Nepogodiev, P. de Andrade, S. Kuhaudomlarp, R.A. Field, Preparative and Kinetic Analysis of β-1,4- and β-1,3-Glucan Phosphorylases Informs Access to Human Milk Oligosaccharide Fragments and Analogues Thereof, ChemBioChem. 21 (2020) 1043–1049.

[132] M. Faijes, M. Castejón-Vilatersana, C. Val-Cid, A. Planas, Enzymatic and cell factory approaches to the production of human milk oligosaccharides, Biotechnol. Adv. 37 (2019) 667–697.

[133] J. Franceus, T. Desmet, Sucrose phosphorylase and related enzymes in glycoside hydrolase family 13: Discovery, application and engineering, Int. J. Mol. Sci. 21 (2020) 2526.

[134] S. Sun, C. You, Disaccharide phosphorylases: Structure, catalytic mechanisms

and directed evolution, Synth. Syst. Biotechnol. 6 (2021) 23–31.

[135] C. Müller, T. Ortmann, A. Abi, D. Hartig, S. Scholl, H.J. Jördening, Immobilization and Characterization of E. gracilis Extract with Enriched Laminaribiose Phosphorylase Activity for Bienzymatic Production of Laminaribiose, Appl. Biochem. Biotechnol. 182 (2017) 197–215.

[136] C. Müller, D. Hartig, K. Vorländer, A.C. Sass, S. Scholl, H.J. Jördening, Chitosan-based hybrid immobilization in bienzymatic reactions and its application to the production of laminaribiose, Bioprocess Biosyst. Eng. 40 (2017) 1399–1410.

[137] A. Abi, C. Müller, H.J. Jördening, Improved laminaribiose phosphorylase production by Euglena gracilis in a bioreactor: A comparative study of different cultivation methods, Biotechnol. Bioprocess Eng. 22 (2017) 272–280.

[138] A. Abi, A. Wang, H.J. Jördening, Continuous Laminaribiose Production Using an Immobilized Bienzymatic System in a Packed Bed Reactor, Appl. Biochem. Biotechnol. 186 (2018) 861–876.

[139] A. Abi, D. Hartig, K. Vorländer, A. Wang, S. Scholl, H.J. Jördening, Continuous enzymatic production and adsorption of laminaribiose in packed bed reactors, Eng. Life Sci. 19 (2019) 4–12.

[140] S. Sun, X. Wei, C. You, The Construction of an In Vitro Synthetic Enzymatic Biosystem that Facilitates Laminaribiose Biosynthesis from Maltodextrin and Glucose, Biotechnol. J. 14 (2019) 1–8.

[141] S.S. Macdonald, A. Patel, V.L.C. Larmour, C. Morgan-Lang, S.J. Hallam, B.L. Mark, S.G. Withers, Structural and mechanistic analysis of a β-glycoside phosphorylase identified by screening a metagenomic library, J. Biol. Chem. 293 (2018) 3451–3467.

Chapter 4: A promiscuous glycosyltransferase generates poly-β-1,4-glucan derivatives that facilitate mass spectrometry-based detection of cellulolytic enzymes

This chapter is composed of one published article and supporting information.

G.S. Bulmer, A.P. Mattey, F. Parmeggiani, R. Williams, H. Ledru, A. Marchesi, L.S. Seibt, P. Both, K. Huang, M.C. Galan, S.L. Flitsch, A.P. Green, J.M. van Munster, A promiscuous glycosyltransferase generates poly-β-1,4-glucan derivatives that facilitate mass spectrometry-based detection of cellulolytic enzymes, Org. Biomol. Chem. 19 (2021) 5529–5533.

Foreword

This chapter presents the utility of the glycosyltransferase LgtB in detecting the activity of cellulolytic enzymes. LgtB demonstrated the ability to polymerise glucose, and was employed to produce imidazolium-based, tagged cello-oligosaccharides that could be used as probes for the eluciation of glycoside hydrolases and LPMO specificities. Derivitisation enabled the production of capped oligosaccharides capable of distinguishing enzymes with exo- or endo- activity.

Contribution

G.S. Bulmer designed and performed the majority of experiements. A.P. Mattey expressed and purified GOase and aided hydrazide derivatisations. F. Parmeggiani screened enzymes and investigated lgtB reaction conditions. R. Williams and H. Ledru synthesised the ITagged glycosides. A. Marchesi optimised SuSy cascade and screened enzymes. L.S. Seibt expressed and purified the LPMO panel. P. Both and K. Huang screened enzymes. M.C. Galan, S.L. Flitsch, A.P. Green and J. M. van Munster provided conceptualisation and supervision. G. S. Bulmer and J.M. van Munster wrote and edited the manuscript.

A promiscuous glycosyltransferase generates poly-β-1,4-glucan derivatives that facilitate mass spectrometry-based detection of cellulolytic enzymes

Gregory S. Bulmer,[a] Ashley P. Mattey,[a] Fabio Parmeggiani,[a,c] Ryan Williams,[b] Helene Ledru,[b] Andrea Marchesi,[a] Lisa S. Seibt,[a] Peter Both,[a] Kun Huang,[a] M. Carmen Galan,[b] Sabine L. Flitsch,[a] Anthony P. Green[a] and Jolanda M. van Munster*[a,d]

Abstract

Promiscuous activity of a glycosyltransferase was exploited to polymerise glucose from UDP-glucose via the generation of β-1,4-glycosidic linkages. The biocatalyst was incorporated into biocatalytic cascades and chemo-enzymatic strategies to synthesise cello-oligosaccharides with tailored functionalities on a scale suitable for employment in mass spectrometry-based assays. The resulting glycan structures enabled reporting of the activity and selectivity of cellulolytic enzymes**.**

Introduction

Cellulose, a linear polysaccharide consisting of β-1,4-linked glucose, is of critical importance in biotechnology, nutrition and microbial pathogenicity. As the major structural component of plant cell walls, it is exploited as a renewable resource in the bio-economy, enabling sustainable production of fuels and chemicals [1]. Cellulose and its oligosaccharides play key roles in health and disease, for example as dietary fiber [2] and have a broad range of applications as biosurfactants, nanomaterials and biogels [3]. The production of cellulose-based structures has therefore attracted much attention.

Chemical synthesis of cellulose derivatives is complex due to the required stereo- and regioselectivity [4]. Enzymatic synthesis of cellulose *in vitro* is challenging because the natural biosynthetic machinery consists of membrane-embedded, multi component systems [5]. Native and derivatised cellulose (oligosaccharides) have been generated from glucose-1-phosphate and cellobiose via the reversible reaction mechanism of cellodextrin phosphorylases, whereby reaction conditions can direct the degree of polymerization (DP) [6–8]. Furthermore, exploitation of enzymes that utiliseactivated glycosyl donors (e.g. glycosyl fluorides) such as glycosynthases, have opened up a raft of new options for synthesizing glycosides

[9–11]. However, such systems require unnatural donors or are biased towards production of very long oligomers. Despite these advances, it remains challenging to generate soluble cello-oligosaccharides from natural donors due to the limited availability of suitable biocatalysts.

Glycosyltransferases (GT) synthesise highly regio- and stereospecific glycosidic bonds between glycan acceptors and activated sugar donors [12]. Enzymatic synthesis of Glc-β-1,4-Glc linkages from nucleotide sugars has been restricted to plant and bacterial cellulose synthases. However, GTs can be promiscuous in acceptor and donor substrates, thereby providing biocatalysts with potentially exploitable side reactions with rates that in many cases were demonstrated to be sufficient for exploitation in glycoside synthesis [13–15]. Therefore, an GT able to accommodate both UDP-Glc as donor and a Glc-terminated structure as acceptor would in effect function as glucose polymerase (Fig. 1a), whereby the DP may be tuned via reaction optimisation or enzyme engineering.

Here we describe how the rational exploration of galactosyltransferase substrate promiscuity resulted in the identification of a broad-specificity biocatalyst that functions as a glucose polymerase *in vitro*. We demonstrate the synthesis of cello-oligosaccharides and their derivatives to a scale suitable for mass spectrometry-based detection and exemplify how these compounds facilitate the profiling of hydrolytic and oxidative biomass-degrading enzyme activities.

With the aim to identify a GT capable of generating β-1,4-linked glucose (Glc) oligosaccharides, we assembled a panel of five recombinantly expressed galactosyltransferases and screened it for promiscuous acceptance of Glc-based donor and acceptor substrates. Transfer of UDP-Glc to an acceptor with a terminal Glc-R motif would result in reaction products that can be re-used as an acceptor, thus enabling the desired Glc polymerisation. As the acceptor structure can potentially affect the generated glycosidic linkage type[16], enzymes with a variety of reported specificities were included.

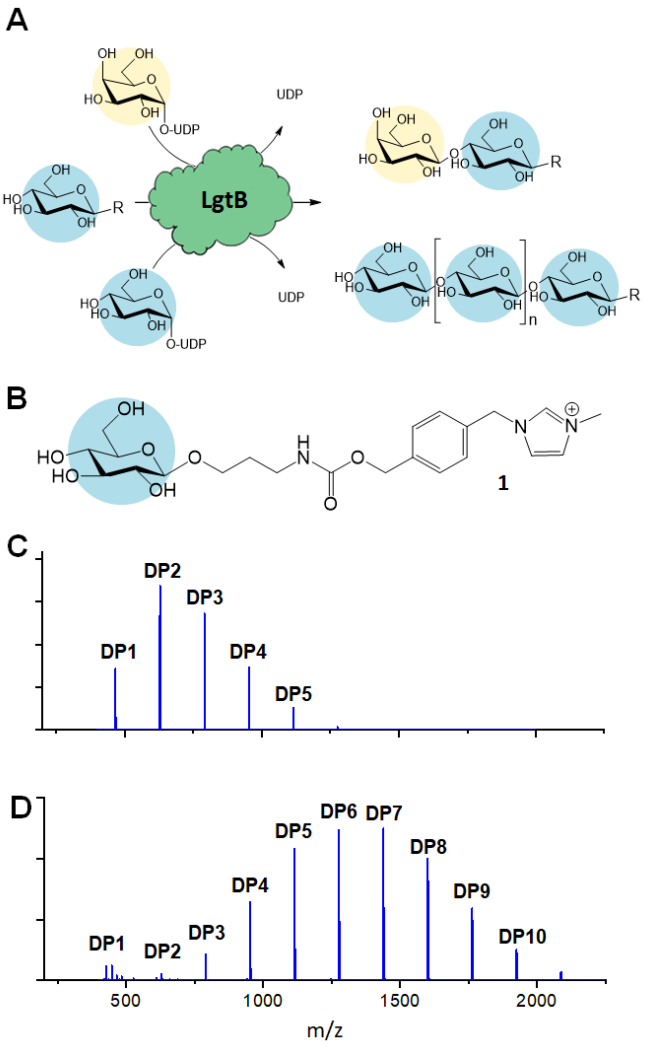


Figure 1. Promiscuous galactosyltransferase activity results in β-1,4-linked glucose polymerisation. **A**, Reactions catalysed by biocatalyst LgtB, blue and yellow circles represent Glc and Gal respectively, following symbol nomenclature for glycans [17,18]. **B**, LgtB acceptor substrate for LgtB, glucoside derivatized with ITag-1 (1). Incubation length and UDP-Glc concentration alters oligosaccharide length and ratio of oligosaccharides produced, **C**, MALDI-TOF spectraafter 2 d with Glc-ITag-1 (1), 1.5 mM UDP-Glc (low concentration) and **D**, after 7 d with 15 mM UDP-Glc (high concentration).

To monitor enzyme activity we employed a sensitive and fast assay based on glycosylation of sugar acceptors labelled with imidazolium-based probes (ITags), such as 4-(1-methyl-3-methyleneimidazolium)benzyl carbamate β-glucoside [15,19]Glc-ITag-1 (**1**) (Fig. 1b). Such cationic ITags generate strong signals in mass-spectrometry that dominate the analyte ionisation and can be used as soluble

handles that support immobilisation and purification of glycosides during chemical derivatisation [20,21].To aid subsequent assessment of the type of glycosidic linkage produced during polymerisation, we chemically synthesised the less complex, novel 4-(1-methyl-3-methyleneimidazolium) benzyl β-glucoside Glc-ITag-2 (**2**) (Fig. S1a) in 4 steps (Fig. S5) with a yield of 49%. The transfer of galactose (Gal) from UDP-Gal to ITagged acceptors was detected as an activity displayed by all tested panel members (Fig. S1, S2). In contrast, activity with UDP-Glc was detected only for *Homo sapiens* B4GALT4 and *Neisseria meningitidis* LgtB, with multiple reaction products due to polymerisation observed in the latter case. LgtB was therefore identified as the biocatalyst most suitable for our target activity, i.e. glucose polymerisation.

After incubation of LgtB with **1** or **2** and an excess of UDP-Glcat 37 °C, a range of ITagged glucose oligosaccharides with DPs of up to 7 and 9, respectively (Fig. S9) were observed. Adjustment of reaction conditions such as incubation time and donor concentration allowed direction of glucose polymerisation towards a desired product range (Fig. 1c, 1d, S6). In reactions optimised towards high DP products, a ≥ 99 % conversion of the initial acceptor (based on Maldi-TOF MS spectral intensity) into oligosaccharides of up to DP12 was detected. Chromatographic separation of the generated oligosaccharides proved challenging in our hands due to the chemical and structural similarities between the products and the limited amount of material available, and therefore the isolation and quantification of individual oligosaccharides and determination of their isolated yields could not be realised within the project constrains. 2D HSQC NMR analysis of purified ITagged oligosaccharides generated from (2) confirmed that the Glc residues were connected by β-1,4-glycosidic linkages (Fig. S8), demonstrating the strict selectivity in anomeric configuration and position of the glycoside linkage that is formed by LgtB.

LgtB was also able to polymerise Glc onto a broad range of other acceptor substrates (Table S1) including native cello-oligosaccharides and derivatives with reducing end conjugates e.g. Glc(n)-pNP (Fig. S4). This agrees with the reported broad acceptor substrate scope of this enzyme, which has been exploited for

chemo-enzymatic synthesis of β-1,4-linked galactosides incorporating e.g GlcNAc(-pNP), Man-pNP, Glc(-pNP) and various C2-derivatives [13–15,22]. No activity was found using Gal, xylose (Xyl), arabinose (Ara), lactose (Gal-β-1,4-Glc), trehalose (Glc-α-1,1-Glc) or UDP-Glc as acceptors. Acceptor substrates thus require an equatorial configuration of the C4 -OH group and the presence of a C6 -OH group while both the C1 and C2 substitutions are highly flexible. Limited transfer of Xyl and GalNAc but not GlcNAc from their respective UDP-conjugates to acceptors was also detected (Fig. S3). Taken together, LgtB was identified as a promiscuous biocatalyst with glucose polymerase activity that can be exploited for polymerization of glucose onto a broad range of acceptors on a scale suitable for detection by mass spectrometry.

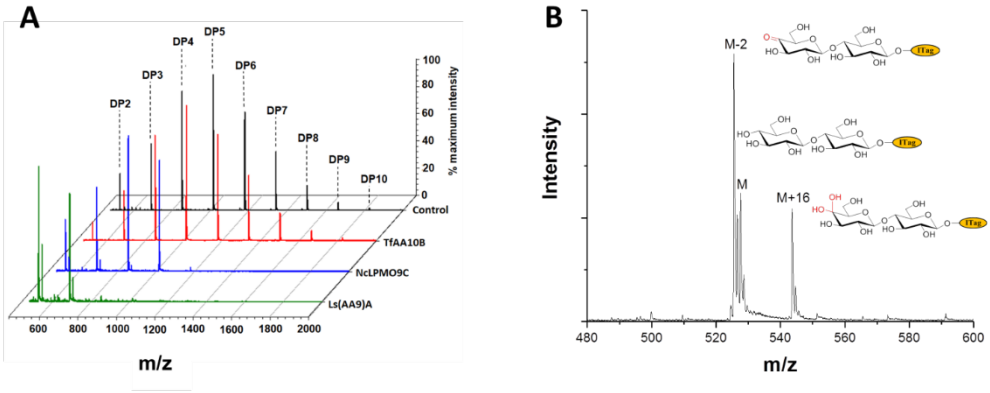


Figure 2. ITagged cello-oligosaccharides as probes for LPMO activity. **A**, selective activity on ITagged glucose oligosaccharide substrates by a panel of LPMOs. **B**, the products of *Nc*LPMO9C activity on ITagged cello-oligosaccharides include those with a mass corresponding to C4-ketone (M-2) and gemdiol (M+16) disaccharides, confirming C4-oxidative activity.

Polymerisation allows cello-oligosaccharide biosynthesis in biocatalytic cascades, such as those enabling regeneration of UDP-glucose. We combined LgtB in a one pot, two enzyme cascade (Fig. S7) with Solanum lycopersicum sucrose synthase (SLSUS6) [23], hereafter SuSy. In the presence of UDP, SuSy converts relatively inexpensive sucrose into fructose and UDP-Glc, this biocatalyst is therefore widely employed for the synthesis of nucleotide sugars and glucosides[24]. Incubation of SuSy with LgtB, sucrose and 1 or 2 resulted in formation of ITagged cello-

oligosaccharides (Fig. S9) with DP ranges and ratios similar to those observed using UDP-Glc directly. Thus, UDP-Glc generated via SuSy activity can be used as a donor substrate by LgtB, in a one-pot biocatalytic cascade that polymerises glucose from inexpensive sucrose.

To exemplify LgtB utility in mass-spectrometry based assays, we exploited ITagged glucose oligosaccharides generated via LgtB activity to rapidly profile activity of lytic polysaccharide monooxygenases (LPMOs), copper dependent enzymes that oxidatively degrade oligosaccharides and polysaccharides. LPMOs insert a single oxygen atom into the C1-H and/or C4-H bond of saccharide substrates, ultimately producing aldonic acids or 4-gemdiol-aldoses, respectively (Fig. S10) [25]. This activity enables more cost-effective production of biofuels from lignocellulose [26] and holds promise in carbohydrate functionalisation. Sensitive assays to profile libraries of LPMO variants with regard to activity on soluble oligosaccharides and selectivity for C1- vs C4- oxidation, are therefore of interest. A small panel of LPMOs was produced in *E. coli* as the expression host; *Lentinus similisLs*(AA9)A [25], *Neurospora crassaNc*LPMO9C [27] and *Thermobifida fuscaTf(*AA10)B [28] (see SI for detailed protocols). Biotransformations were performed with purified enzymes using an envelope of ITagged glucose oligosaccharides (DP 1-10) in the presence of H_2O_2 and ascorbate as a reducing agent. The distribution of labelled oligosaccharides was unchanged following incubation with *Tf(*AA10)B (Fig. 2a), confirming lack of activity towards the oxidation of soluble cello-oligosaccharides with DP < 10. Conversely, reactions with *Nc*(AA9)C and *Ls*(AA9)A led to complete consumption of oligosaccharides with DP ≥ 6 and 4, respectively (Fig. 2a), with concomitant formation of labelled C4-oxidised products (DP 2-4 and DP 2-3, respectively) with molecular weights of [M-2] (C4-ketone) and [M+16] (C4-gemdiol) (with M as the molecular weight of corresponding non-oxidised labelled carbohydrate) (Fig. 2b). These data are consistent with reported substrateprofiles. Notably, the ITag cationic nature avoids formation of sodium or potassium adducts [15] that complicate interpretation [29] of MS-based detection of LPMO products. Selective detection of endo-cellulase activity in glycoside hydrolase (GH)

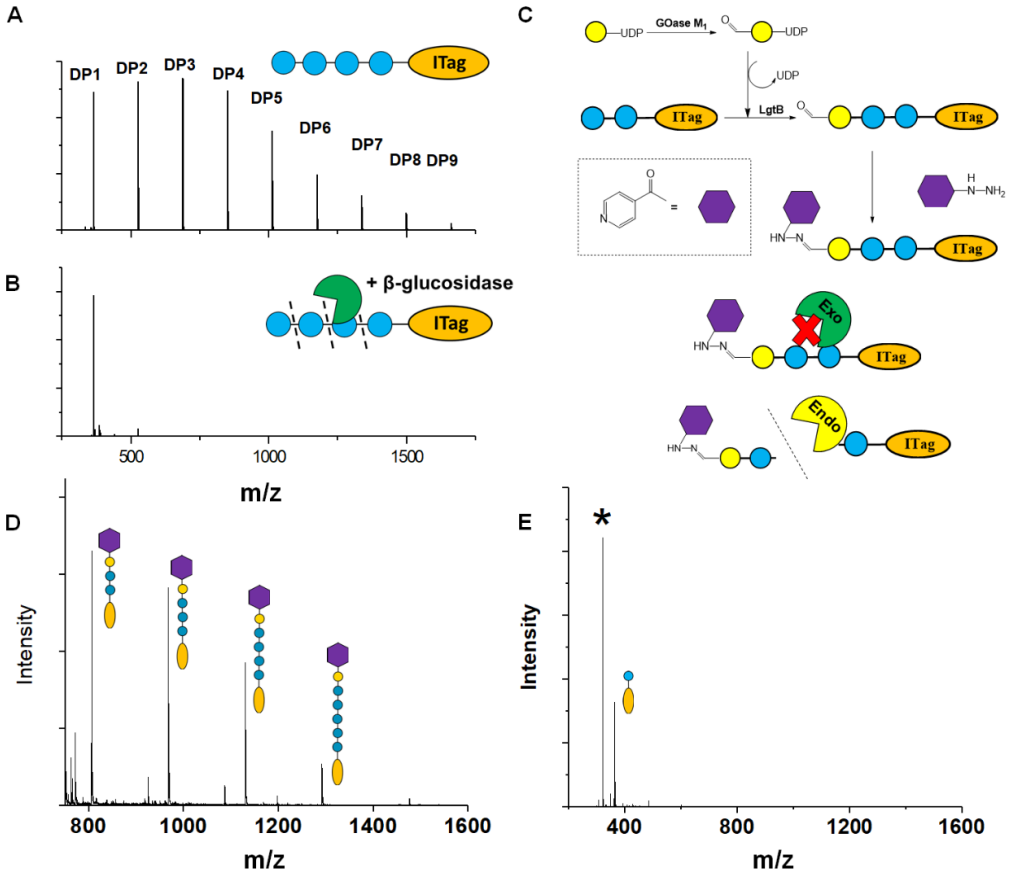


Figure 3. Chemo-enzymatic derivatisation of cello-oligosaccharides to selective endo-cellulase substrates. **A**, Representative MALDI-TOF MS spectra demonstrating ITagged cello-oligosaccharides after incubation without enzymes and **B**, with β-glucosidase from almond. **C**, Derivatisation strategy employed to protect LgtB product termini against hydrolytic activity; **D**, probes after oxidised galactose and nicotinic hydrazide ligation after incubation with β-glucosidase and β-galactosidase; **E**, same probes after incubation with *A. niger* cellulases. Asterix indicates a glycosylated hydrazide product.

preparations is in principle enabled via an enzyme substrate that cannot be hydrolysed from its termini. Exploiting LgtB, we generated such a 'blocked' substrate probe compatible with fast and sensitive detection of enzyme activity via MS. Initially we confirmed commercially available cellulolytic GHs degrade unmodified ITagged cello-oligosaccharides (Fig. 3a,b). 'Blocked' ITagged oligosaccharides were then generated via a chemo-enzymatic approach (See SI for detailed methods) based on the creation of a bio-orthogonal aldehydeon the non-reducing, Gal-capped termini of oligosaccharides. The aldehydes were subsequently derivatised with a nicotinic hydrazide group (Fig. S13b) via previously described

methodologies[30]. Derivatisation of the reducing end, here accomplished via ITag, was required to ensure generation of a homogenous end product (Fig. S12).

With the blocked substrates in hand, we confirmed their protection against exo-acting enzyme activity. Terminal galactose addition was sufficient to protect cello-oligosaccharides against β-glucosidase activity, while hydrazide derivatisation was required and sufficient to protect structures against a set of β-galactosidases (Fig. 3d, Fig. S14). Finally, we employed the blocked substrates to monitor endo-acting cellulases activity, and validated the presence of this activity by MALDI-TOF MS (Fig. 3e). The observed degradation profile was similar to that generated by the same enzyme incubated with unblocked oligosaccharides. We conclude these probes enable selective detection of endo-cellulase activity via MS.

Conclusions

In conclusion, LgtB has a broad substrate scope that includes polymerase activity with UDP-Glc as donor and Glc-terminated glycosides as acceptors. We successfully exploited this promiscuous activity to polymerise glucose via β-1,4-glycosidic linkages onto a variety of acceptors containing chemical handles, producing functionalised, tailored cello-oligosaccharide derivatives as detectible by mass spectrometry. We demonstrated the incorporation of LgtB-mediated glucose polymerisation into chemo-enzymatic derivatisation strategies, via production of probes to selectively detect endo-acting cellulolytic enzyme activities in a sensitive , MS-based assay. Further work regarding scalability of the LgtB reaction including determination of isolated yields should confirm if LgtB would be suitable as catalyst for glucose polymerization in reactions at preparative scale and upwards. As LgtB can be obtained via facile heterologous expression in E. coli, this biocatalyst for the synthesis of β-1,4-glycosidic linkages is very well suited to be engineered to accept desired non-natural substrates, which may open up novel routes to enzymatic synthesis of cellulose derivatives.

References

[1] C.O. Tuck, E. Pérez, I.T. Horváth, R.A. Sheldon, M. Poliakoff, Valorization of biomass: Deriving more value from waste, Science. 337 (2012) 695–699.

[2] Y. Uyeno, S. Shigemori, T. Shimosato, Effect of Probiotics/Prebiotics on Cattle Health and Productivity, Microbes Environ. 30 (2015) 126–132.

[3] D. Klemm, B. Heublein, H.-P. Fink, A. Bohn, Cellulose: Fascinating Biopolymer and Sustainable Raw Material, Angew. Chemie Int. Ed. 44 (2005) 3358–3393.

[4] A. Pardo-Vargas, M. Delbianco, P.H. Seeberger, Automated glycan assembly as an enabling technology, Curr. Opin. Chem. Biol. 46 (2018) 48–55.

[5] O. Omadjela, A. Narahari, J. Strumillo, H. Melida, O. Mazur, V. Bulone, J. Zimmer, BcsA and BcsB form the catalytically active core of bacterial cellulose synthase sufficient for in vitro cellulose synthesis, Proc. Natl. Acad. Sci. 110 (2013) 17856–17861.

[6] D.M. Petrović, I. Kok, A.J.J. Woortman, J. Ćirić, K. Loos, Characterization of Oligocellulose Synthesized by Reverse Phosphorolysis Using Different Cellodextrin Phosphorylases, Anal. Chem. 87 (2015) 9639–9646.

[7] C. Zhong, C. Luley-Goedl, B. Nidetzky, Product solubility control in cellooligosaccharide production by coupled cellobiose and cellodextrin phosphorylase, Biotechnol. Bioeng. 116 (2019) 2146–2155.

[8] E.C. O'Neill, G. Pergolizzi, C.E.M. Stevenson, D.M. Lawson, S.A. Nepogodiev, R.A. Field, Cellodextrin phosphorylase from Ruminiclostridium thermocellum: X-ray crystal structure and substrate specificity analysis, Carbohydr. Res. 451 (2017) 118–132.

[9] S. Fort, V. Boyer, L. Greffe, G.J. Davies, O. Moroz, L. Christiansen, M. Schülein, S. Cottaz, H. Driguez, Highly efficient synthesis of β(1 → 4)-oligo- and -polysaccharides using a mutant cellulase, J. Am. Chem. Soc. 122 (2000) 5429–5437.

[10] E. Brun, H. Brumer, L.F. MacKenzie, S.G. Withers, L.P. McIntosh, Assignment of selectively 13C-labeled cellopentaose synthesized using an engineered glycosynthase [2], J. Biomol. NMR. 21 (2001) 67–68.

[11] S. Kobayashi, K. Kashiwa, T. Kawasaki, S.I. Shoda, Novel Method for Polysaccharide Synthesis Using an Enzyme: The First in Vitro Synthesis of Cellulose via a Nonbiosynthetic Path Utilizing Cellulase as Catalyst, J. Am. Chem. Soc. 113 (1991) 3079–3084.

[12] D.-M. Liang, J.-H. Liu, H. Wu, B.-B. Wang, H.-J. Zhu, J.-J. Qiao, Glycosyltransferases: mechanisms and applications in natural product development, Chem. Soc. Rev. 44 (2015) 8350–8374.

[13] D. Namdjou, H. Chen, E. Vinogradov, D. Brochu, S.G. Withers, W.W. Wakarchuk, A β-1,4-Galactosyltransferase from Helicobacter pylori is an Efficient and Versatile Biocatalyst Displaying a Novel Activity for Thioglycoside Synthesis, ChemBioChem. 9 (2008) 1632–1640.

[14] K. Lau, V. Thon, H. Yu, L. Ding, Y. Chen, M.M. Muthana, D. Wong, R. Huang, X. Chen, Highly efficient chemoenzymatic synthesis of β1-4-linked galactosides with promiscuous bacterial β1-4-galactosyltransferases, Chem. Commun. 46 (2010) 6066–6068.

[15] K. Huang, F. Parmeggiani, H. Ledru, K. Hollingsworth, J. Mas Pons, A. Marchesi, P. Both, A.P. Mattey, E. Pallister, G.S. Bulmer, J.M. van Munster, W.B. Turnbull, M.C. Galan, S.L. Flitsch, Enzymatic synthesis of N-acetyllactosamine from lactose enabled by recombinant β1,4-

galactosyltransferases., Org. Biomol. Chem. 17 (2019) 5920–5924.

[16] L.L. Lairson, A.G. Watts, W.W. Wakarchuk, S.G. Withers, Using substrate engineering to harness enzymatic promiscuity and expand biological catalysis, Nat. Chem. Biol. 2 (2006) 724–728.

[17] A. Varki, R.D. Cummings, M. Aebi, N.H. Packer, P.H. Seeberger, J.D. Esko, P. Stanley, G. Hart, A. Darvill, T. Kinoshita, J.J. Prestegard, R.L. Schnaar, H.H. Freeze, J.D. Marth, C.R. Bertozzi, M.E. Etzler, M. Frank, J.F.G. Vliegenthart, T. Lütteke, S. Perez, E. Bolton, P. Rudd, J. Paulson, M. Kanehisa, P. Toukach, K.F. Aoki-Kinoshita, A. Dell, H. Narimatsu, W. York, N. Taniguchi, S. Kornfeld, Symbol nomenclature for graphical representations of glycans, Glycobiology. 25 (2015) 1323–1324.

[18] S. Neelamegham, K. Aoki-Kinoshita, E. Bolton, M. Frank, F. Lisacek, T. Lütteke, N. O'Boyle, N.H. Packer, P. Stanley, P. Toukach, A. Varki, R.J. Woods, Updates to the Symbol Nomenclature for Glycans guidelines, Glycobiology. 29 (2019) 620–624.

[19] I. Sittel, M.C. Galan, Imidazolium-labeled glycosides as probes to harness glycosyltransferase activity in human breast milk, Org. Biomol. Chem. 15 (2017) 3575–3579.

[20] M.C. Galan, A.T. Tran, C. Bernard, Ionic-liquid-based catch and release mass spectroscopy tags for enzyme monitoring, Chem. Commun. 46 (2010) 8968–8970.

[21] A.T. Tran, R. Burden, D.T. Racys, M. Carmen Galan, Ionic catch and release oligosaccharide synthesis (ICROS), Chem. Commun. 47 (2011) 4526–4528.

[22] O. Blixt, J. Brown, M.J. Schur, W. Wakarchuk, J.C. Paulson, Efficient preparation of natural and synthetic galactosides with a recombinant beta-

1,4-galactosyltransferase-/UDP-4'-gal epimerase fusion protein., J. Org. Chem. 66 (2001) 2442–8.

[23] X.C. Duan, H. Chen, F.F. Liu, L. Conway, S. Wei, Z.P. Cai, L. Liu, J. Voglmeir, One Assay for All: Exploring Small Molecule Phosphorylation Using Amylose-Polyiodide Complexes, Anal. Chem. 87 (2015) 9546–9550.

[24] X. Chen, J. Fang, J. Zhang, Z. Liu, J. Shao, P. Kowal, P. Andreana, P.G. Wang, Sugar nucleotide regeneration beads (superbeads): A versatile tool for the practical synthesis of oligosaccharides [12], J. Am. Chem. Soc. 123 (2001) 2081–2082.

[25] K.E.H. Frandsen, T.J. Simmons, P. Dupree, J.C.N. Poulsen, G.R. Hemsworth, L. Ciano, E.M. Johnston, M. Tovborg, K.S. Johansen, P. Von Freiesleben, L. Marmuse, S. Fort, S. Cottaz, H. Driguez, B. Henrissat, N. Lenfant, F. Tuna, A. Baldansuren, G.J. Davies, L. Lo Leggio, P.H. Walton, The molecular basis of polysaccharide cleavage by lytic polysaccharide monooxygenases, Nat. Chem. Biol. 12 (2016) 298–303.

[26] G. Vaaje-Kolstad, B. Westereng, S.J. Horn, Z. Liu, H. Zhai, M. Sørlie, V.G.H. Eijsink, An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides, Science. 330 (2010) 219–222.

[27] J.W. Agger, T. Isaksen, A. Várnai, S. Vidal-Melgosa, W.G.T. Willats, R. Ludwig, S.J. Horn, V.G.H. Eijsink, B. Westereng, Discovery of LPMO activity on hemicelluloses shows the importance of oxidative processes in plant cell wall degradation, Proc. Natl. Acad. Sci. U. S. A. 111 (2014) 6287–6292.

[28] Z. Forsberg, Å.K. Røhr, S. Mekasha, K.K. Andersson, V.G.H. Eijsink, G. Vaaje-Kolstad, M. Sørlie, Comparative study of two chitin-active and two cellulose-active AA10-type lytic polysaccharide monooxygenases, Biochemistry. 53 (2014) 1647–1656.

[29] V.G.H. Eijsink, D. Petrovic, Z. Forsberg, S. Mekasha, Å.K. Røhr, A. Várnai, B. Bissaro, G. Vaaje-Kolstad, On the functional characterization of lytic polysaccharide monooxygenases (LPMOs), Biotechnol. Biofuels. 12 (2019) 58.

[30] A.P. Mattey, W.R. Birmingham, P. Both, N. Kress, K. Huang, J.M. Van Munster, G.S. Bulmer, F. Parmeggiani, J. Voglmeir, J.E.R. Martinez, N.J. Turner, S.L. Flitsch, Selective Oxidation of N-Glycolylneuraminic Acid Using an Engineered Galactose Oxidase Variant, ACS Catal. 9 (2019) 8208–8212.

Supporting information

A promiscuous glycosyltransferase generates poly-β-1,4-glucan derivatives that facilitate mass spectrometry-based detection of cellulolytic enzymes

Gregory S. Bulmer,[a] Ashley P. Mattey,[a] Fabio Parmeggiani,[a,c] Ryan Williams,[b] Helene Ledru,[b] Andrea Marchesi,[a] Lisa S. Seibt,[a] Peter Both,[a] Kun Huang,[a] M. Carmen Galan,[b] Sabine L. Flitsch,[a] Anthony P. Green[a] and Jolanda M. van Munster*[a,d]

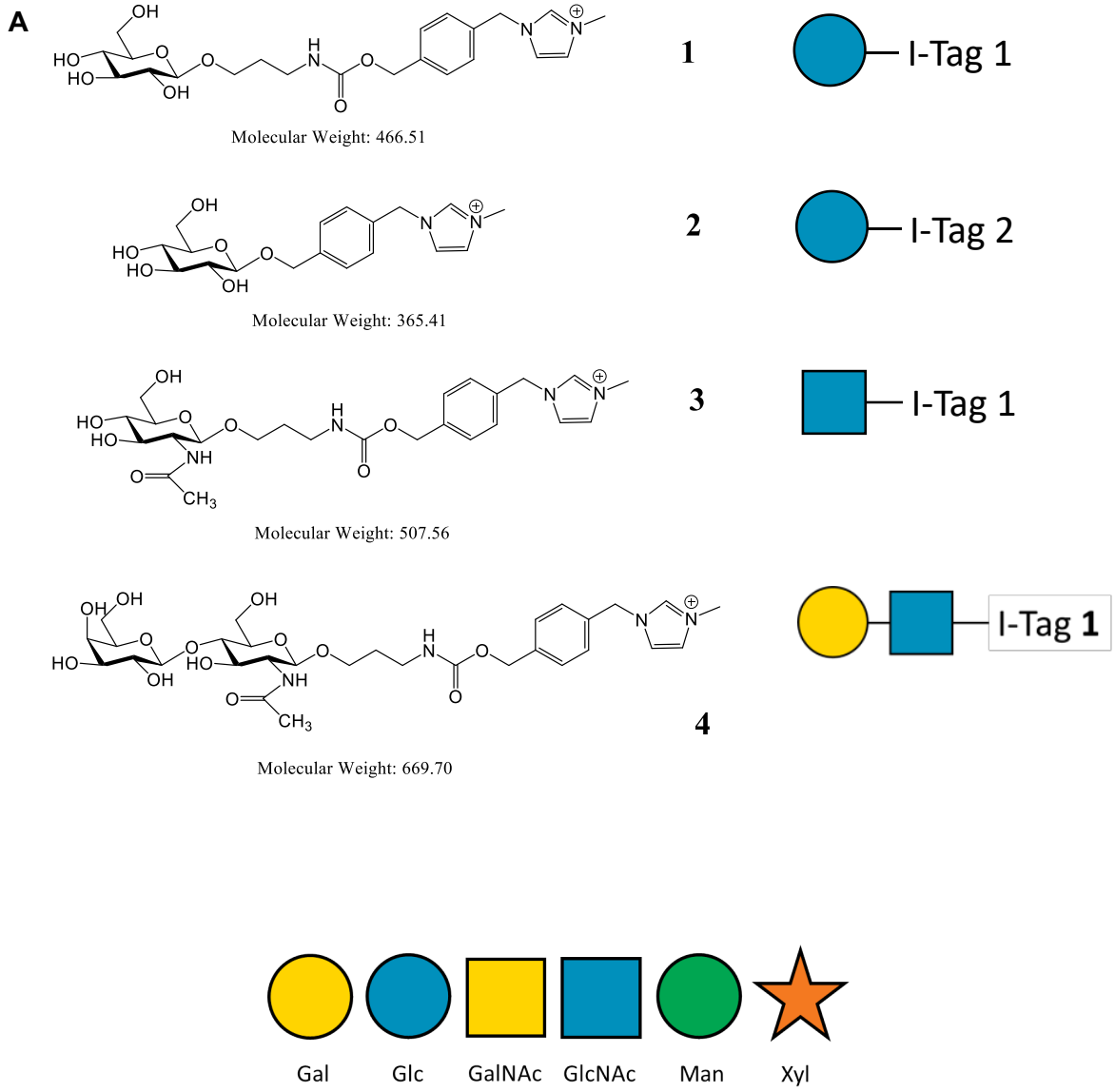
[a] Manchester Institute of Biotechnology (MIB) & School of Natural Sciences, The University of Manchester, 131 Princess Street, Manchester, M1 7DN, United Kingdom

[b] School of Chemistry, University of Bristol, Cantock's Close, Bristol, BS8 ITS, United Kingdom

[c] Department of Chemistry, Materials and Chemical Engineering "G. Natta", Politecnico di Milano, Via Mancinelli 7, 20131, Milano, Italy.

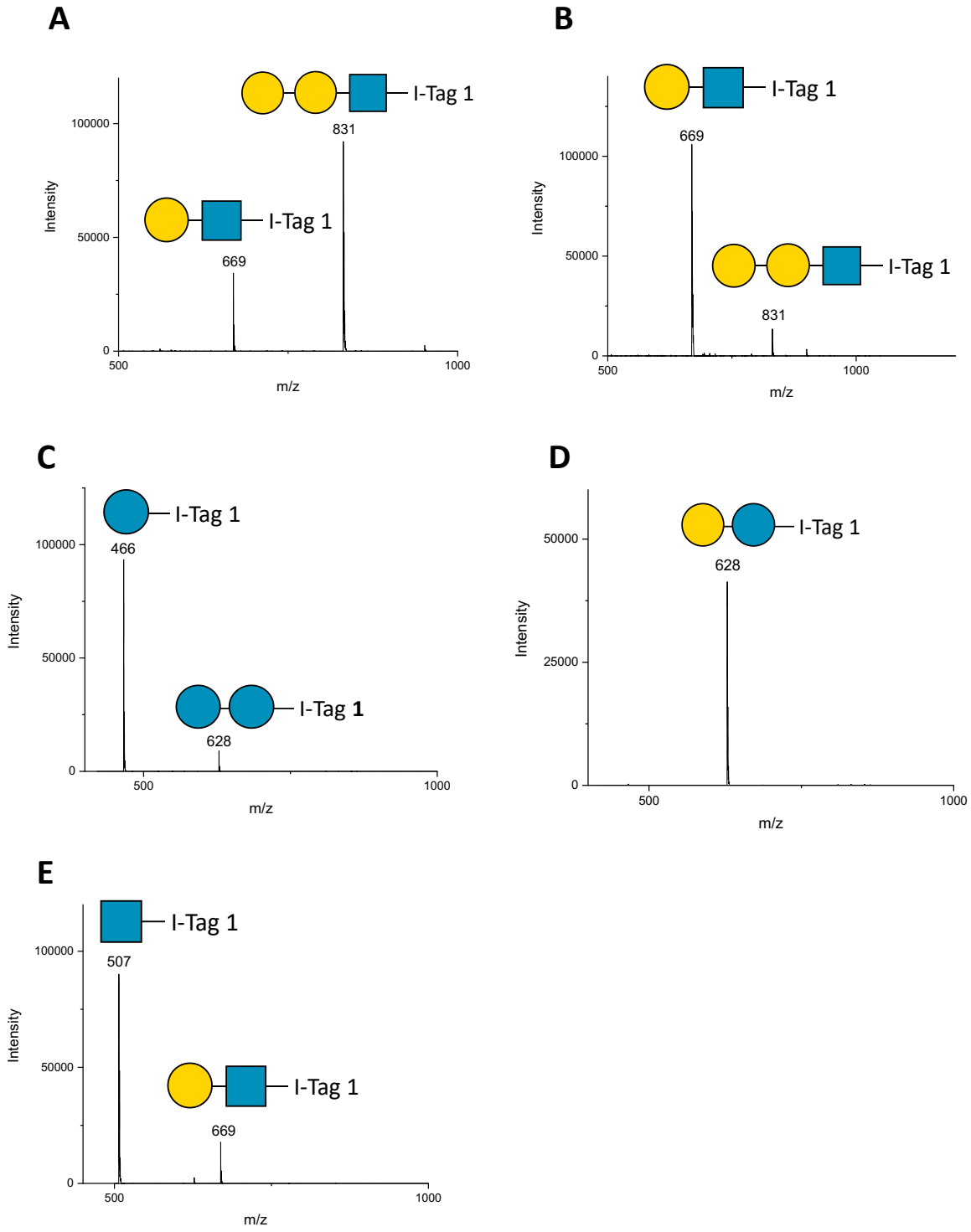
[d] Scotland's Rural College, West Mains Road, Kings Buildings, Edinburgh, EH9 3JG, United Kingdom

Contents

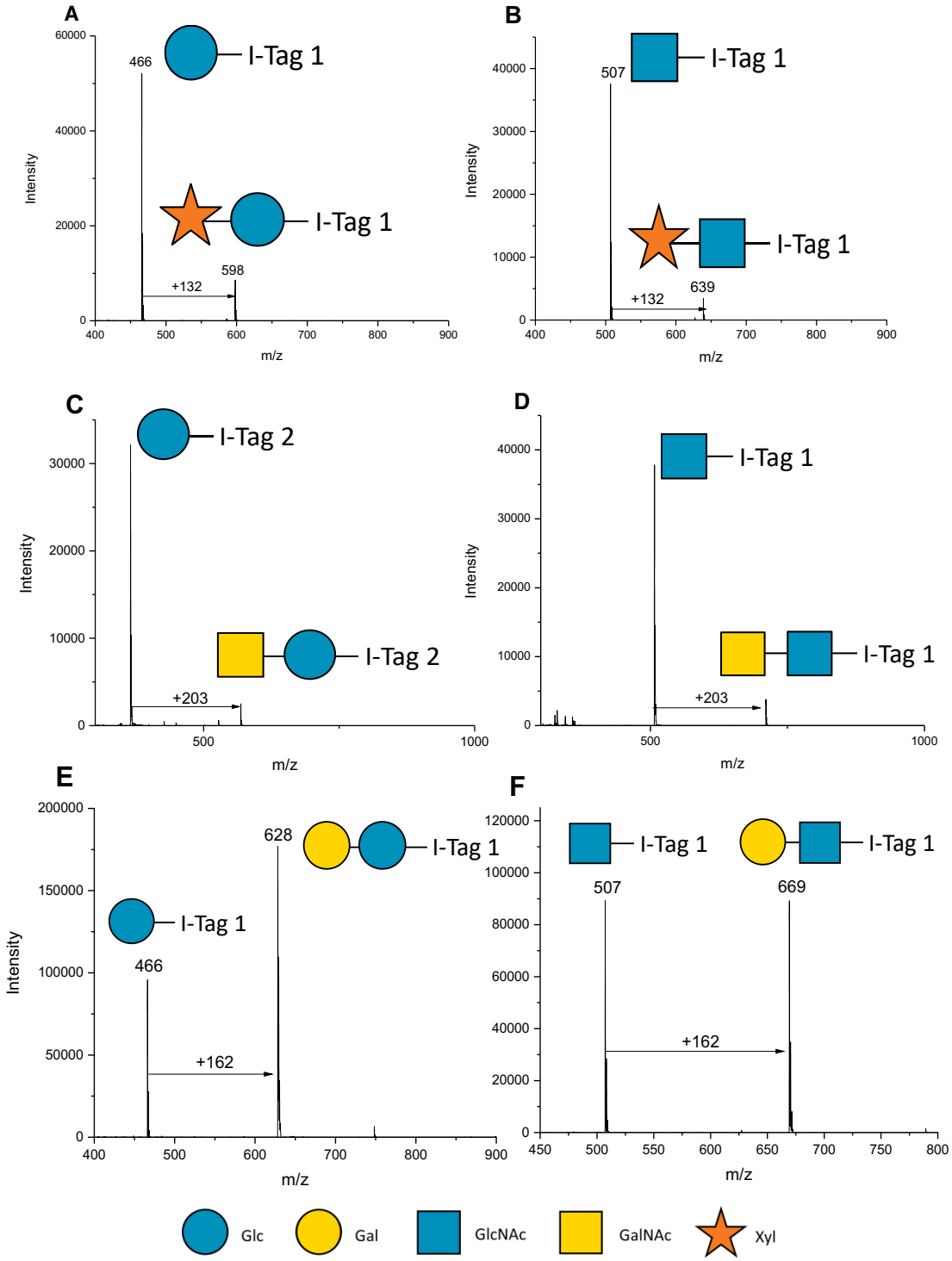


Supplementary figure 1 | ITagged oligosaccharides and SNFG nomenclature used in this study.

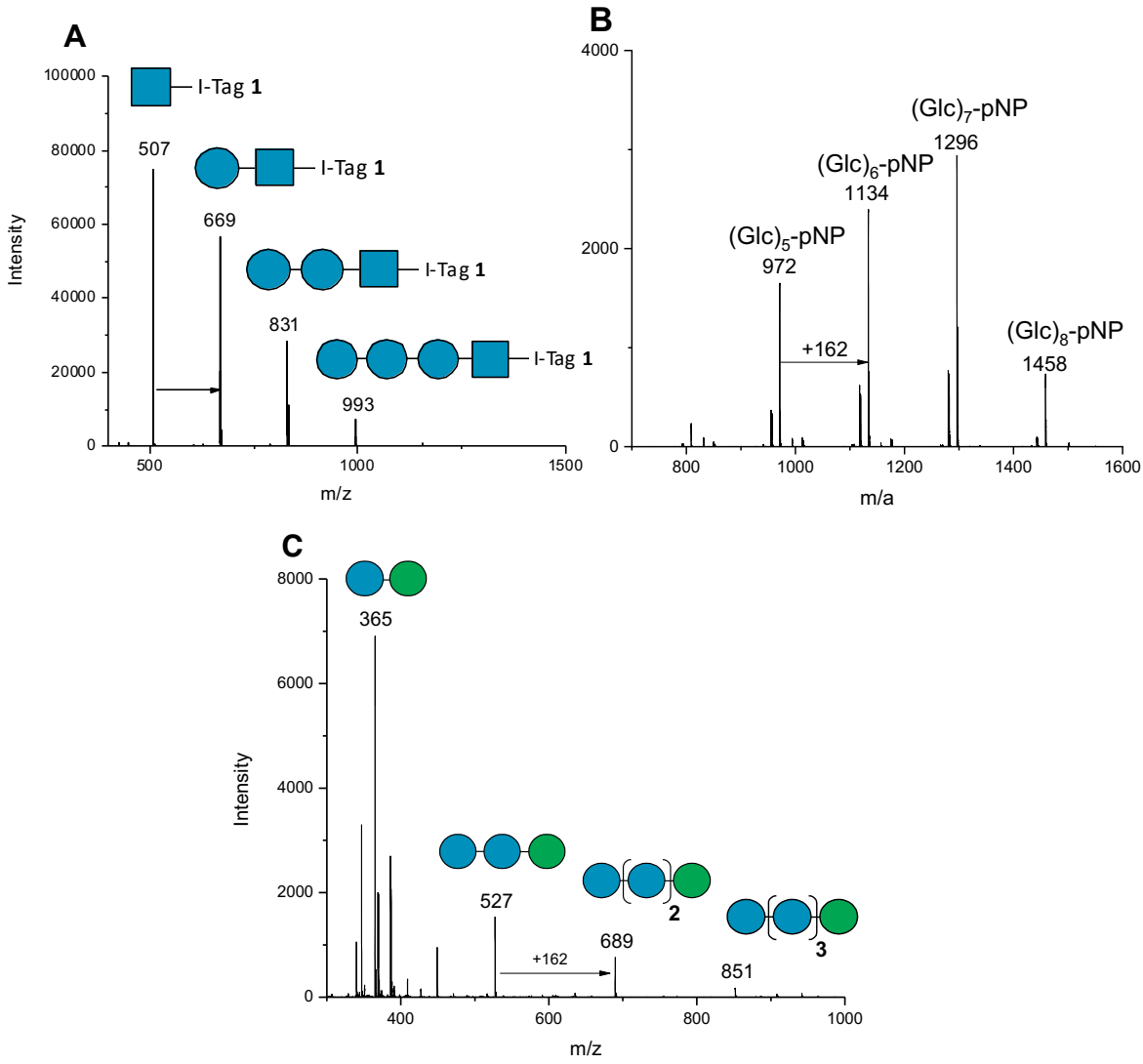
A, Above compounds are referred to as follows: Glc-ITag-1 (**1**), Glc-ITag-2 (**2**), GlcNAc-ITag-1 (**3**) and LacNAc-ITag-1 (**4**). Compounds **1**, **3** and **4** were synthesised as previously described[1]**B**, Monosaccharides as defined by the SNFG with accepted abbreviation.



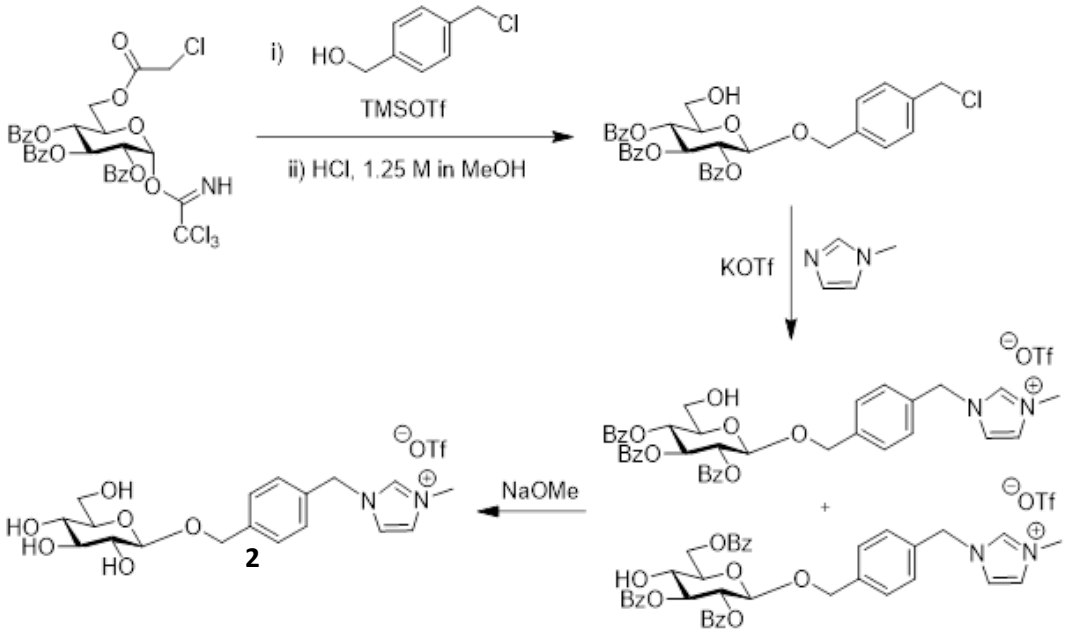
Supplementary figure 2| Galactosylation of ITagged glycosides during activity screening. Galactosylation activity of LgtC (**A**) and G0PH97 (**B**) against LacNAc-ITag-1 (**4**). A peak of *m/z* 831 was observed for both, corresponding to galactosylated LacNAc-ITag-1. B4GALT4 demonstrated both UDP-Glc (**C**) and UDP-Gal (**D**) transfer onto Glc-ITag-1 (**1**). LgtH (**E**) demonstrated transgalactosylation activity against GlcNAc-ITag-1 (**3**). A peak of *m/z* 669 was observed, corresponding to galactosylated GlcNAc-ITag-1 (LacNAc-ITag-1 **4**).



Supplementary figure 3 | Donor scope assessment of LgtB with ITagged acceptors.LgtB was able to transfer sugar from their respective UDP-conjugate when incubated with (**A** and **B**) UDP-Xyl; (**C** and **D**) UDP-GalNAc and (**E** and **F**) UDP-Gal onto ITag-Glc-1 (**1**) and ITag-GlcNAc-1 (**3**), respectively

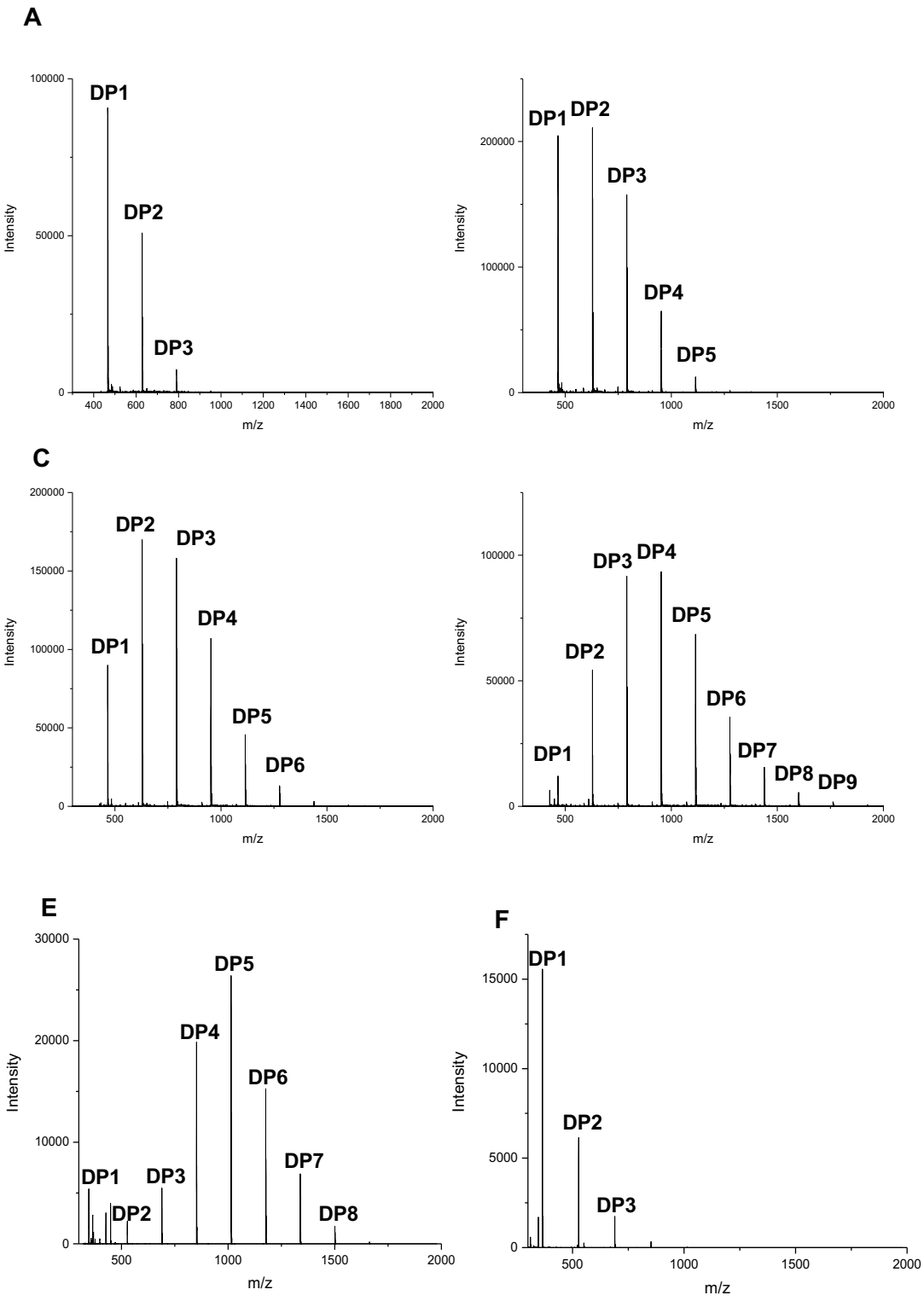


Supplementary figure 4 | Glucose polymerisation occurs onto a range of acceptors
Glc polymerisation by LgtB was observed via MALDI-TOF MS when using as acceptor substrates **A,** ITag-GlcNAc **3**, **B**, 4-Nitrophenyl β-D-glucose (Glc-pNP) and **C**, Mannose. Products are detected as [M] (for **A**) or [M+Na$^+$] (for **B**, **C**)

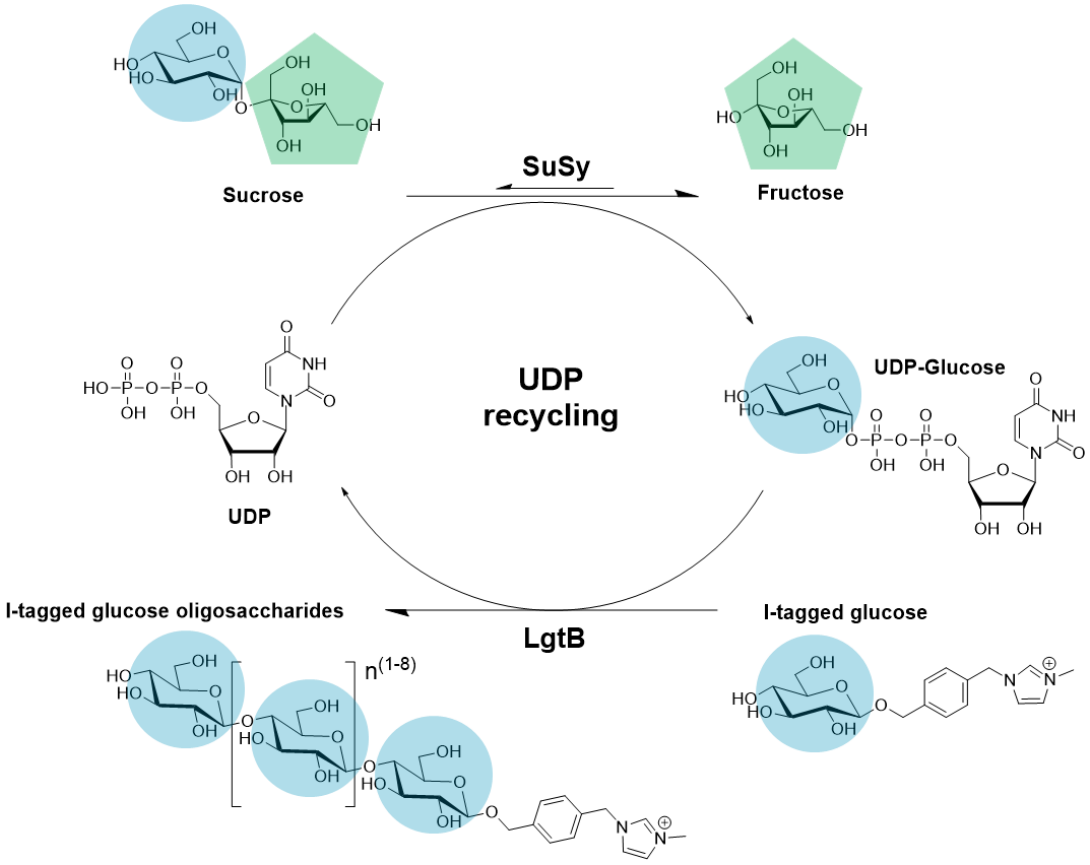


Supplementary figure 5 | Synthesis of 4-(1-Methyl-3-methyleneimidazolium)benzyl β-D-glucopyranoside trifluoromethanesulfonate (2)

Compound 2 was synthesised from 4-(chloromethyl)benzyl alcohol and glycosyl donor 2,3,4-tri-O-benzoyl-6-O-chloroacetyl-α-D-glucopyranosyl trichloroacetimidate in 4 steps with a yield of 49%, as described in detail in the methods section. Products were identified as follows: **IR** ν_{max}/cm^{-1} 3419br (OH), 3156w, 3113w, 2968w, 2929w, 1577, 1452, 1416, 1253s, 1225, 1160, 1076, 1028s, 758, 638, 574, 517; - 24˚ [*c* 1.08, MeOH].**^{1}H NMR** δ_H (500 MHz, Methanol-d_4) 8.95 (1 H, s, NC*H*N), 7.58 (1 H, d, *J* 2.0, NC*H*CHN), 7.56 (1 H, d, *J* 2.0, NCHC*H*N), 7.53 – 7.48 (2 H, m, H_{arom}), 7.43 – 7.38 (2 H, m, H_{arom}), 5.39 (2 H, s, NCH_2), 4.94 (1 H, d, *J* 12.3, (C-1)OC*H*H), 4.70 (1 H, d, *J* 12.3, (C-1)OCH*H*), 4.35 (1 H, dd, *J* 7.7, 0.9, H-1), 3.92 (3 H, s, NCH_3), 3.89 (1 H, dd, *J* 12.0, 2.1, H-6a), 3.68 (1 H, dd, *J* 11.8, 5.5, H-6b), 3.37 – 3.23 (4 H, m, H-2, H-3, H-4, H-5); **^{13}C NMR** δ_C (126 MHz, Methanol-d_4) 140.64 (4˚ C_{arom} (CH_2O(C-1))) , 134.44 (4˚ C_{arom} (CH_2N)), 129.91, 129.65 (C_{arom}), 125.21 (NCH*C*HN), 123.60 (N*C*HCHN), 103.41 (C-1), 78.10, 78.04, 75.11, 71.66 (C-2, C-3, C-4, C-5), 71.07 ((C-1)O*C*H_2), 62.77 (C-6), 53.83 (N*C*H_2), 36.52 (N*C*H_3); **m/z** (ESI-HRMS) $C_{18}H_{25}N_2O_6^+$ ([M – OTf]$^+$) calculated: 365.1707; found 365.1712; (TLC-MS- (ESI)) $CF_3O_3S^-$ ([OTf]$^-$) calculated 149.0; found 148.8.

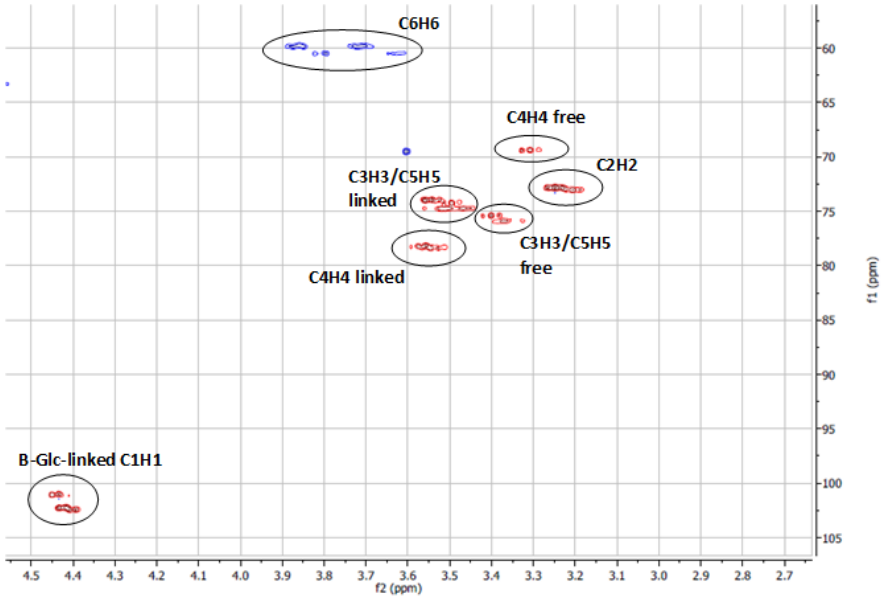


Supplementary figure 6 | Modification of reaction conditions drives oligosaccharide profile distribution. Product distribution broadens and degree of polymerisation increases during polymerisation of Glc onto Glc-ITag-1(**1**) after 96 h in response to LgtB concentration with **A** 0.17 mg ml^{-1} LgtB **B** 0.35 mg ml^{-1} LgtB, **C** 0.85 mg ml^{-1} LgtB, **D** 1.7 mg ml^{-1} LgtB. Starting concentration of acceptor affects oligosaccharide distribution, reactions analysed after 24 h with 1.7 mg ml^{-1} LgtB and **E** 0.2 mM Glc-ITag-2 (**2**)and **F** 0.5 mM **2**.



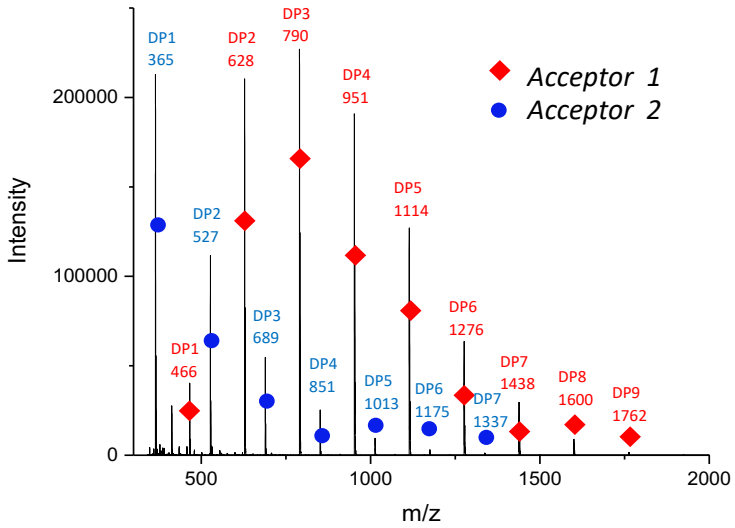
Supplementary figure 7 | Biocatalytic cascade enabling effective coupling of UDP-glucose regeneration with glucose polymerisation.

LgtB and sucrose synthase SuSy were utilised in a one-pot reaction for the production of ITagged cello-oligosaccharides.



Supplementary figure 8 | HSQC confirmed that ITagged oligosaccharides generated by LgtB consist of β1,4-linked glucose

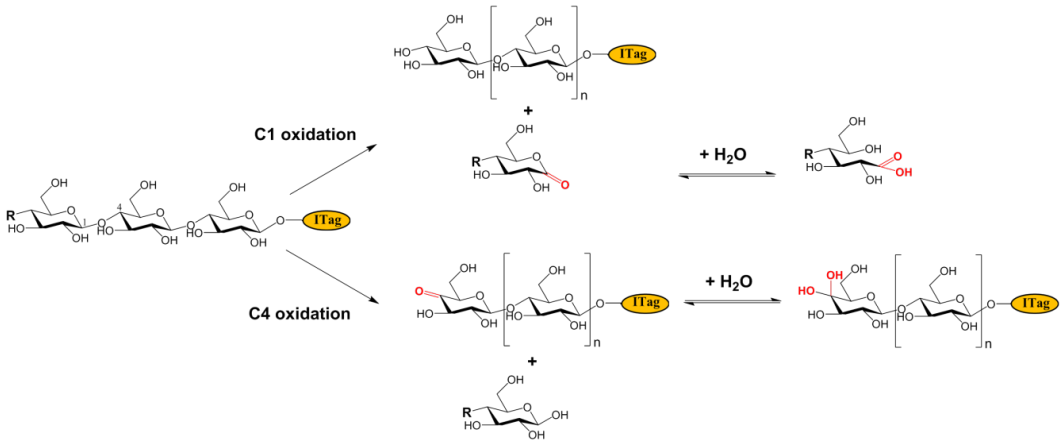
^{1}H-^{13}C gradient-selected sensitivity-enhanced multiplicity-edited HSQC of purified ITagged oligosaccharides generated by LgtB demonstrated the presence of β1,4-linked glucose.



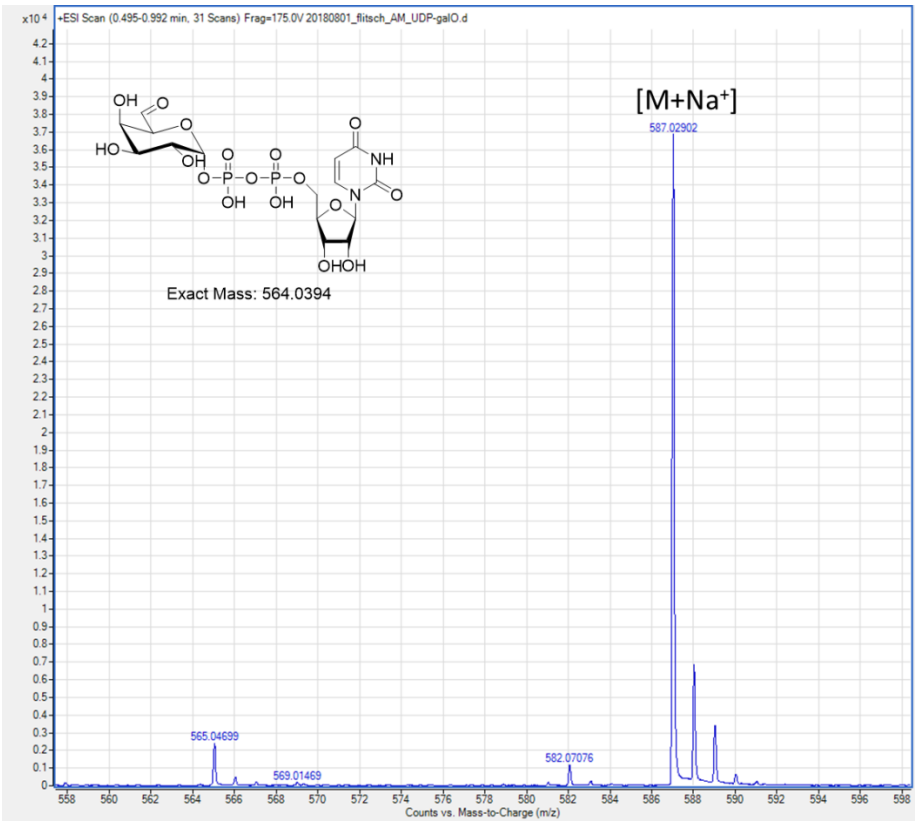
Supplementary figure 9 | SuSy cascade enables recycling of UDP for donor production

Superimposed MALDI-TOF spectra of oligosaccharide formation using the SuSy UDP recycling system, with Glc-ITag-1(**1**)and Glc-ITag-2 (**2**) as acceptors. Degree of polymerisation masses for **1** (red) as follows DP 1 (466), DP2 (628), DP3 (790), DP4 (952), DP5 (1114), DP6 (1276), DP7 (1438), DP8 (1600), DP9 (1762). Degree of

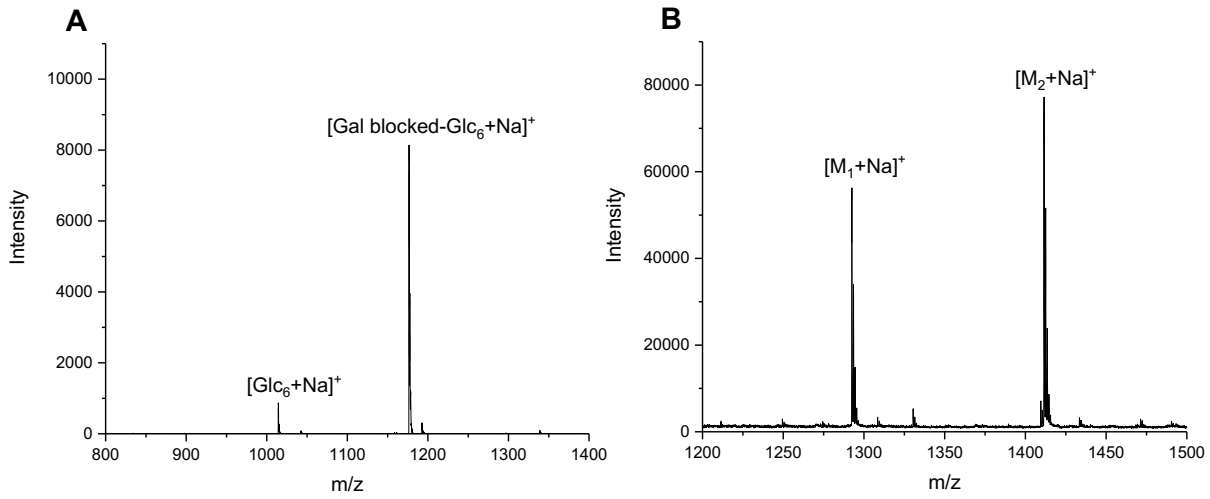
polymerisation masses for **2** (blue) as follows DP 1 (365), DP2 (527), DP3 (689), DP4 (851), DP5 (1013), DP6 (1175), DP7 (1337)



Supplementary figure 10 | LPMO reaction products
Scheme detailing the reaction products arising from C1- and C4 oxidation of ITagged cello-oligosaccharides

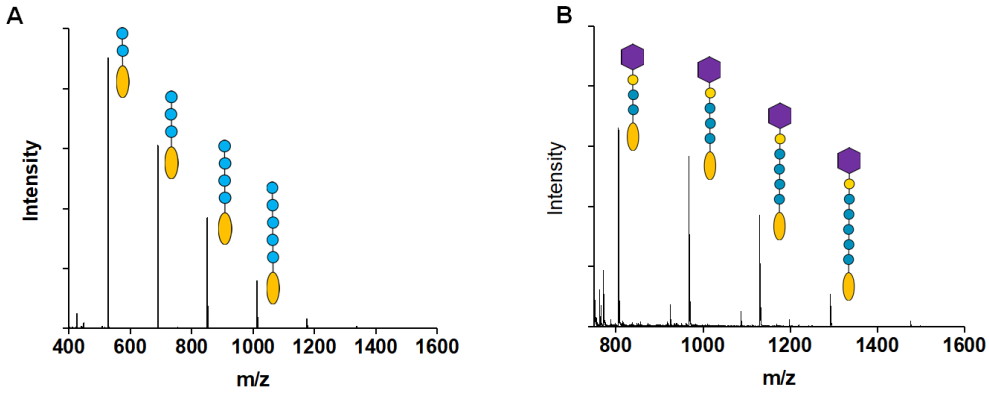


Supplementary figure 11 | Production of oxidised UDP-Gal
High-resolution mass spectrometry (HRMS) data confirming oxidised UDP-Gal has been produced via M_1 GOase.



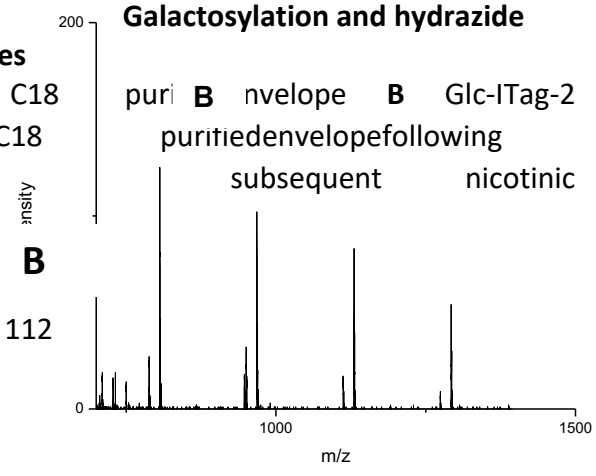
Supplementary figure 12| Galactosylation and hydrazide ligation of native cellooligosaccharide cellohexaose

A MALDI-TOF MS showing galactosylation of cellohexaose by LgtB (M = 1,175) with remaining cellohexaose substrate (M = 1,013). **B** MALDI-TOF MS showing sodiated adducts of single (M_1 = 1,270) and double (M_2 = 1,389) ligation of nicotinic hydrazide to the oxidised, galactose blocked cellohexaose. Double ligated species originate from nicotinic hydrazide ligating at the aldehyde of the Gal_{ox} as well as the aldehyde at the free reducing end. This procedure thus generated non-homogenous products with which it would prove difficult to assess the effectiveness of hydrolases



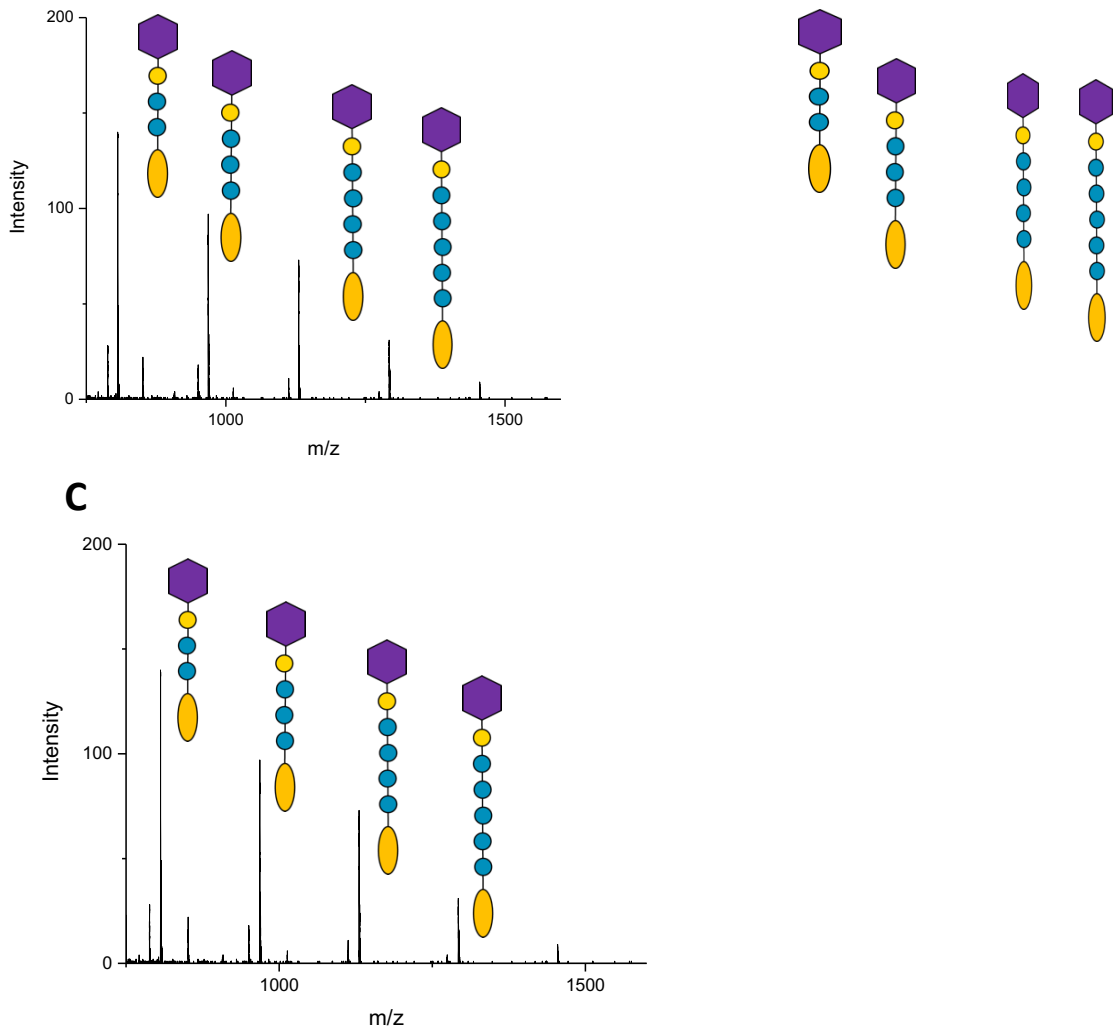
Supplementary figure 13| Galactosylation and hydrazide ligation of ITagged cello-oligosaccharides

A Glc-ITag-2 polymerised C18 puri **B** nvelope **B** Glc-ITag-2 polymerised C18 puritiedenvelopefollowing addition of oxidised Gal and subsequent nicotinic hydrazide ligation.



A

B



Supplementary figure 14| Confirmation of β-galactosidase activity blocking
MALDI-TOF MS showing nicotinic hydrazide ligated ITag-glucose oligosaccharides after incubation with β-galactosidases from glycoside hydrolase families **A** GH1 **B** GH42 and **C** GH50. Nicotinic hydrazide group represented by purple hexagon.

Supplementary Table 1 | LgtB activity on donor and acceptor substrates as detected by MALDI-TOF MS. (+) observed reaction products, (-) no observed reaction products.

Acceptors	**Donors**				
	UDP-Glc	UDP-Gal	UDP-Xyl	UDP-GalNAc	UDP-GlcNAc
Glc-ITag (**1**, **2**)	+	+	+	+	-
GlcNAc-ITag-1 (**3**)	+	+	+	+	-
Cellobiose (Glc-β1,4-Glc)	+	+		+	-
Glc	+	+			
GlcNAc	+	+			
GlcN	+	+			
β-GlcNAc-N3	+	+			
β-Glc-N3	+	+			
Glc-pNP	+	+			
Trehalose (Glc-α1,1-Glc)	-	-			
Man	+	+			
Sucrose (Glc-α1,2β-Fru)	+	+			
Cellotetraose-pNP	+	+			
Lactose (Gal-β1,4-Glc)	-	-			
Gal	-	-			
Xyl	-	-			
Ara	-	-			
UDP-Glc	-				

Materials and methods

Galactosyltransferase production

Galactosyltransferases LgtB (Uniprot: Q51116), LgtC (Uniprot: A0A3S5C3F9), LgtH (Uniprot: Q2TIJ3), B4GALT4 (Uniprot: B2RAZ5) and an uncharacterised homolog (Uniprot: G0PH97) were provided by Prozomix Lld, Haltwhistle, UK as purified protein suspended in $(NH_4)_2SO_4$. Samples were centrifuged (20000 g, 5 min, 4 °C), the supernatant was discarded and an equal volume of Tris buffer (20 mM pH 8.0) was added to resuspend the protein pellet. The protein sample was centrifuged again to remove precipitated protein and supernatant retained for analysis.

NMR of carbohydrates

NMR characterisation of ITagged glucose oligosaccharide linkages were performed using a ^{1}H-^{13}C gradient-selected sensitivity-enhanced multiplicity-edited HSQC on a Bruker AVIII 500 MHz spectrometer equipped with a QCI-F cryoprobe. Assignments were made based on the carbohydrate structure database (http://csdb.glycoscience.ru/database/).

Mass spectrometry of carbohydrates

Samples were prepared for analysis by MALDI-TOF via crystallisation with super DHB or THAP matrices. Super DHB (a 9:1 (w/w) mixture of 2,5-Dihydroxybenzoic acid and 2-hydroxy-5-methoxybenzoic acid) was prepared at 15 mg ml^{-1} in a mixture of 50 % (v/v) acetonitrile and 50 % (v/v) water containing 0.1 % trifluoroacetic acid (TFA). 2',4',6'-Trihydroxyacetophenone monohydrate (THAP) was prepared at 10 mg ml^{-1} in acetone. MALDI-TOF mass spectrometry was performed using the Bruker Ultraflex 3 in positive mode. MALDI-TOF was calibrated using peptide Calibration Standard II (Bruker) containing a range of peptides 757-3149 Da in size.

SuSy system for UDP recycling

Solanum lycopersicum Sucrose synthase (SLSUS6) was expressed in *E. coli* BL21(DE3) from a pET30 based vector containing an N-terminal Histag fusion as previously described[2]. Briefly, cells were cultured at 37 ºC, 250 rpm until reaching

an OD600 of 0.6, after which cultures where moved to 18 ºC and protein expression was induced via addition of 1 mM IPTG for 20 h. Cells were harvested and lysed, after which cleared cell lysates were loaded onto an immobilised metal affinity chromatography (IMAC, GE Healthcare) Ni^{2+}-charged column. After washing 30 column volumes with 50 mM Tris, 50 mM NaCl, pH 8.0, SLSUS6 was eluted via addition of 500 mM imidazole. SuSy UDP-sugar regeneration reactions were carried out in a total volume of 50 µl containing 50 mM MES buffer (pH 6.0),0.5 mM ITag substrate, 0.5 mM UDP, 35 mM sucrose, 5 mM $MgCl_2$, 2.8 µg SuSy and 20 µg LgtB. Reactions were incubated at 37 ^{o}C for 5 days and analysed by MALDI-TOF.

GalT screening panel

The panel of galactosyltransferases was screened in reactions as follows: 0.5 mM ITag acceptor (**1**-**4**), 1.5 mM UDP sugar donor, 10 mM $MgCl_2$, 10 mM $MnCl_2$, 0.1 mg ml^{-1} BSA, 50 mM Tris pH 8.0, and enzyme concentrations, one of the following: LgtH (0.3 mg ml^{-1}), G0PH97 (0.4 mg ml^{-1}), LgtC (0.1 mg ml^{-1}) and LgtB (1.7 mg ml^{-1}). Reactions were incubated at 37 ºC for 7 days to allow any potential polymerisation to occur. ITag acceptors Glc-ITag (**1**,**2**), GlcNAc-ITag-1 (**3**) and ITag-LacNAc-1 (**4**) were utilised in the screen along with UDP donors UDP-Glc and UDP-Gal.

Donor scope assessment and ITagged oligosaccharide production using LgtB

Final concentrations in ITag-glucose oligosaccharides experiments: 15 mM UDP-Glc, 1 mM ITag-glucose (**1**,**2**), 50 mM Tris HCl (pH 8.0), 10 mM $MgCl_2$, 10 mM $MnCl_2$, BSA 100 µg ml^{-1}, ~1.7 mg ml^{-1}LgtB (resuspended in dH_2O). Experiments assessing the activity of UDP-Xyl transfer onto ITag-glucose were as previous but with UDP-xylose 1mM final concentration, replacing UDP-glucose as the donor. ITag-glucose oligosaccharides were purified on a 5 ml C18 column using a gradient of methanol in water with concentrations ranging from 0 to 100% (v/v). All fractions were collected and analysed via MALDI-TOF MS, and those containing purified oligosaccharides were pooled.

Enzymatic digestions of ITagged-oligosaccharides

β-glucosidase from almond (49290, Sigma-Aldrich) was dissolved in water at 1 mg ml^{-1}. Cellulase from *Aspergillus niger* (C1184**,** Sigma-Aldrich) was dissolved in pH 5.0, sodium acetate buffer at 10mg ml^{-1}. Reactions were performed at pH 5.0, 37 °C for 2 h at concentrations of 0.5 mg ml^{-1} and 5 mg ml^{-1} β-glucosidase and cellulase respectively with 10 % (v/v) C18 column purified ITagged glucose oligosaccharides. LPMO reactions were set up as follows: 100 µM H_2O_2, 200 µM ascorbate and 5 µM LPMO in 25 mM Tris-HCl buffer pH 7.5 and performed at room temperature for 14 h.

LPMO production

The synthesised genes coding for *Neurospora crassa* NcLPMO9C (Uniprot: Q7SHI8) and *Thermobifida fuscaTf(*AA10)B (Uniprot: Q47PB9) were cloned into pET22b vector (encoding a pelB sequence for periplasmic secretion) and expressed using *E. coli* C43(DE3), via periplasmic secretion to obtain an N-terminal His. After growth at 37°C until OD600 of 0.6, protein expression was induced with 0.1mM IPTG at 25°C for 16 h, 200 rpm. *Lentinus similis* LsLPMO9A (Uniprot: A0A0S2GKZ1) was expressed as an E8K-vector construct in *E. coli DH5a* at 25 °C for 24 h, 200 rpm. Protein production was induced using 10 mM arabinose.

The cells were harvested by centrifugation (4000 rpm, 20 min, 4°C) and lysed in equilibration buffer (50 mM sodium phosphate (NaPi) buffer pH 8.0, containing 300 mM NaCl, 1 mg ml^{-1}lysozyme and 10 µg ml^{-1}DNase,) via sonication (10 min, 1s on 1s off, 40 % sonotrode power). Cell debris was removed using centrifugation at 20000 rpm 30 min at 4°C and cleared supernatant was applied to a pre-equilibrated 5 ml Strep-tactin superflow cartridge (Qiagen). The column was washed with equilibration buffer and the protein eluted using 5mM desthiobiotin. Fractions containing pure target protein, identified via SDS-PAGE, were pooled, concentrated and re-buffered into 25 mM Tris-HCl pH 7.5 using PD10 desalting columns. The protein was loaded with 5x molar excess $CuCl_2$ for 1 h at 4 °C and the excess Copper was then removed using PD10 desalting columns. The protein was stored at -80 °C, 25 mM Tris-HCl buffer pH 7.5. Activity of purified enzymes was confirmed via

detection of the LPMO side reaction that reduces O_2 to H_2O_2 in the absence of substrate, via a peroxidase-coupled assay as described previously[3].

Endo-substrate development

M_1 Galactose oxidase (GOase) was expressed and purified as previously described[4]. Oxidised UDP-Gal (UDP-Gal_{Ox}) was produced as follows: 0.1 mg ml^{-1} horseradish peroxidase (HRP), 1 mg ml^{-1} GOase M_1, 0.1 mg ml^{-1}catalase, 10 mM UDP-galactose, NaPi pH 7.4 reaction was incubated until completion at 25 ^oC, 250 rpm for 4 h filtered on a viva spin column to remove protein, desalted on a Sephadex G-10 column then freeze dried. Final concentrations in oxidised UDP-Gal blocking of cellohexaose were as follows: 10 mM UDP-Gal_{Ox}, 5 mM cellohexaose, 10 mM $MnCl_2$, 50 mM NaPi (pH 8), LgtB (1.7 mg ml^{-1}). Final concentrations in Gal_{Ox} blocking of ITagged glucose oligosaccharides were as follows: 10 mM UDP-Gal_{Ox}, 20 % (v/v) C18 column purified ITagged glucose oligosaccharides, 10 mM $MnCl_2$, 50 mM sodium phosphate (pH 8), LgtB (1.7 mg ml^{-1}). After incubation at 250 rpm, 37 ^oC for 16 h, reaction products were purified on a viva spin column (10000 MWCO). Subsequent hydrazide ligation required 20 equivalents of nicotinic hydrazide addition to the previous reaction mixes, which were then incubated at 30 ^oC, 250 rpm for 2 h.

Enzymatic galactosylation of native cello-oligosaccharides (Fig. S11a) followed by oxidation of the terminal Gal then nicotinic hydrazide ligation (Fig. S11b) was also demonstrated. Final concentrations in galactose capping of cellohexaose were as follows: 10 mM UDP-Gal, 5 mM cellohexaose, 10 mM $MnCl_2$, 50 mM NaPi (pH8), LgtB (1.7 mg ml^{-1}). Reaction was incubated at 250 rpm, 37 ^oC for 16 h. Reaction was then purified on a viva spin column. Product was then oxidised, final concentration during the oxidation of the galactose capped cellohexaose were as follows: 0.1 mg ml^{-1} HRP, 1 mg ml^{-1} GOase M_1, 0.1 mg ml^{-1} catalase, 0.5 mM galactosylated cellohexaose, NaPi pH 7.4. Reaction was incubated at 25 ^oC, 250 rpm for 4 h then treated with nicotinic hydrazide as in previous method.

Nicotinic hydrazide ligated ITagged cello-oligosaccharides were incubated with β-galactosidases to ensure inaccessibility of exo-active cellulolytic enzymes. β-Galactosidase 1A from Sulfolobus solfataricus (CZ04141), β-Galactosidase 42A from

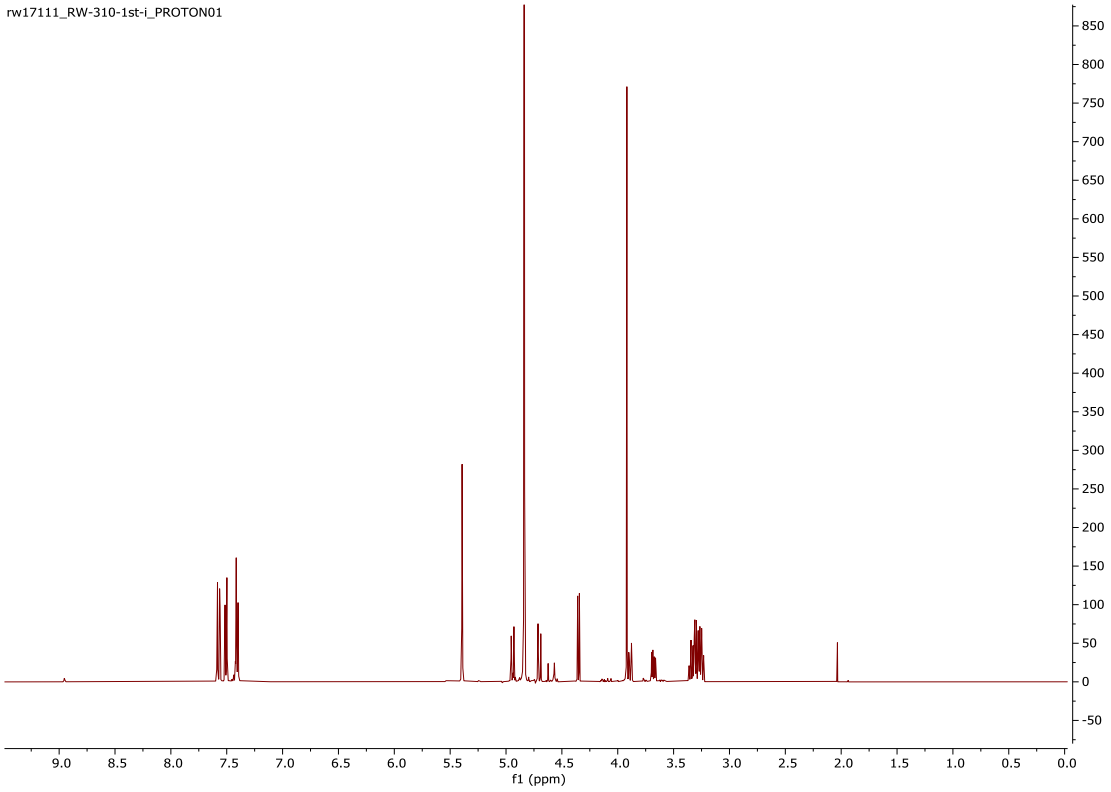
Caldicellulosiruptor saccharolyticus (CZ09741) and β-Galactosidase 50A from *Victivallis vadensis* (CZ05161) were purchased from NZYtech. Reactions with GH1 were incubated at 90 °C for 2 h in reactions containing: nicotinic hydrazide ligated ITagged cello-oligosaccharides (25 % v/v), 50 mM pH 5 sodium acetate buffer, 0.5 mg ml^{-1} GH1. Reactions with GH42 were incubated at 65 °C for 2 h, in reactions containing: nicotinic hydrazide ligated ITagged cello-oligosaccharides (25 % v/v), 50 mM pH 4 sodium acetate buffer, 0.5 mg ml^{-1} GH42. Reactions with GH50 were incubated at 40 °C for 2 h, in reactions containing nicotinic hydrazide ligated ITagged cello-oligosaccharides (25 % v/v), 50 mM pH 7 Tris buffer, 0.25 mg ml^{-1} GH50.

Synthesis of 4-(1-Methyl-3-methyleneimidazolium)benzyl β-D-glucopyranoside trifluoromethanesulfonate (2) (Glc-ITag-2)

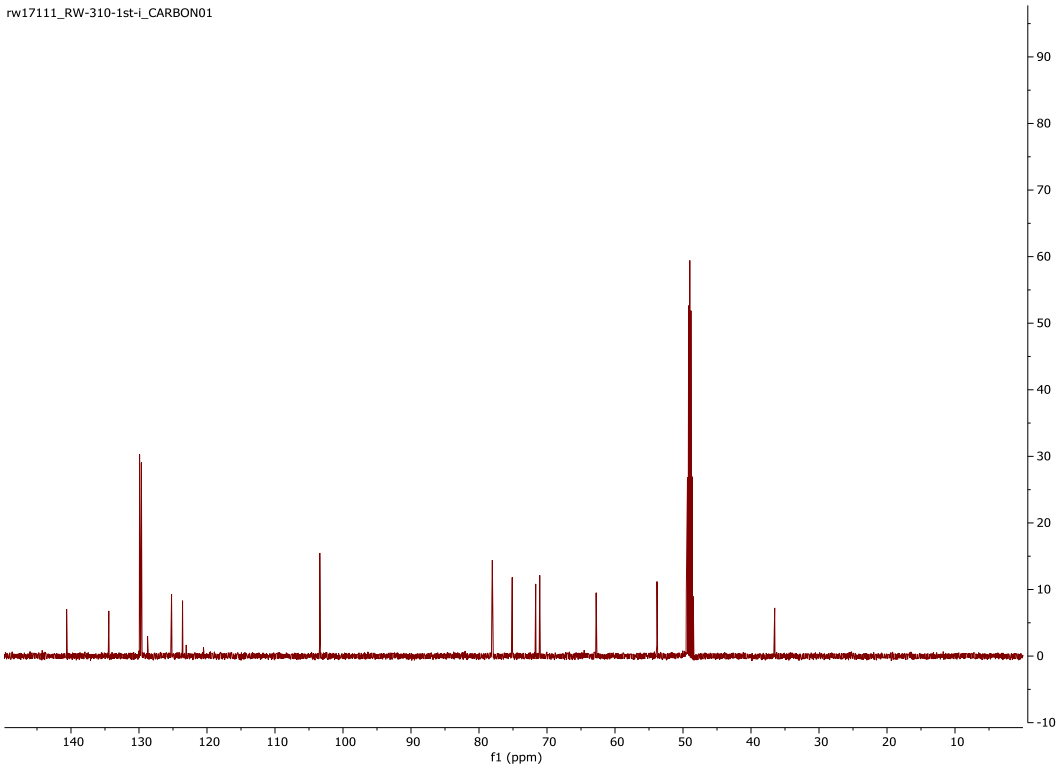
Glycosyl acceptor 4-(chloromethyl)benzyl alcohol (0.0689 g, 0.440 mmol, 1 eq) and glycosyl donor 2,3,4-tri-O-benzoyl-6-O-chloroacetyl-α-D-glucopyranosyl trichloroacetimidate (0.6277 g, 0.880 mmol, 2.0 eq) were placed in a dry vial and dried under vacuum for 30 min. 2.20 ml of anhydrous DCM was added to the donor/acceptor vial under nitrogen, resulting in a solution of volume 2.75 ml and therefore approximately 0.160 M in acceptor and 0.320 M in donor. A stock solution of TMSOTf (0.06 M in DCM) was made by dissolving TMSOTf (0.217 ml, 1.200 mmol) in 20 ml anhydrous DCM. 3.0 ml of this stock solution was used for this reaction. The flow microreactor and attached tubing was flushed with nitrogen. The donor/acceptor solution and TMSOTf solution were each taken up in a syringe and installed onto a syringe pump. The solutions were then injected into the microreactor (total internal volume of reactor chip and outlet tubing = 32.8 µL) at the desired flow rate corresponding to the residence time (15 seconds, 65.60 µl min^{-1} in each syringe for a combined flow rate of 131.20 µl min^{-1} in reactor zone) via the inlet tubing. The flow reaction was performed at RT. The mixture that flowed from the microreactor was dropped in a flask containing reagent grade DCM in air to quench the reaction. Reaction solution was collected for 41 min 30 sec, after which time the reaction mixture solvent was removed under reduced pressure. The crude product was dissolved in DCM (20 ml) and washed with water (8 ml), then the

water was extracted with a further portion of DCM (20 ml). The DCM fractions were collected, dried with magnesium sulfate, filtered and the solvent was removed under reduced pressure. The dried residue was dissolved in a minimal volume of DCM, then HCl, 1.25 M in MeOH (7.04 ml, 8.800 mmol) was added. The resulting solution was stirred for 16 h at RT in air, then diluted with DCM (15 ml) and water (10 ml) and product was extracted into the DCM phase. The aqueous phase was washed with a further DCM portion (15 ml) then DCM washings were combined, dried using magnesium sulfate, filtered and the solvent was removed under reduced pressure. The dried residue was washed with hexane (3 x 5 ml), then dried under reduced pressure for 1 h, before being dissolved in anhydrous MeCN (5 ml) under a nitrogen atmosphere. 1-Methyl imidazole (0.14 ml, 1.76 mmol) and potassium trifluoromethanesulfonate (0.3312 g, 1.76 mmol) were added and the resulting mixture was heated under reflux at 90 ˚C and stirred for 18 h, after which time TLC (DCM:MeOH 94:6) showed the reaction to be complete by MS. Solvent was removed under reduced pressure, then DCM (5 ml) and 1 M $HCl_{aq.}$ (5 ml) were added to the residue and product was extracted into the DCM phase. The aqueous phase was washed with DCM (2 x 5 ml), then the DCM portions were combined, dried using magnesium sulfate, filtered and the solvent was removed under reduced pressure. The crude residue was washed with neat Et_2O (5 ml) and DCM:Et_2O 5:95 (2 x 5 ml). The ITagged products were dissolved in methanol (2.2 ml) and sodium methoxide (50.0 µl, 0.22 mmol, 25 % wt in MeOH) was added. The solution was stirred at RT for 3 h after which time TLC-MS showed the reaction to be complete. The solution was then brought to pH 7 using 1 M $HCl_{aq.}$. Solvent was removed, then the residue was diluted with DCM and water and product was extracted into the aqueous phase. Water was removed under reduced pressure and the dried mixture was purified by reversed-phase HPLC (Water:MeCN) to yield the **ITag-glucose 2 (2)** (0.1110 g, 49 % over 4 steps) as a solid.

^{1}H NMR δ_H (500 MHz, Methanol-d_4) for ITag-glucose **2**



^{13}C NMR δ_C (126 MHz, Methanol-d_4) for ITag-glucose **2**



References

(1) Huang, K.; Parmeggiani, F.; Ledru, H.; Hollingsworth, K.; Mas Pons, J.; Marchesi, A.; Both, P.; Mattey, A. P.; Pallister, E.; Bulmer, G. S.; et al. Enzymatic Synthesis of N-Acetyllactosamine from Lactose Enabled by Recombinant B1,4-Galactosyltransferases. *Org. Biomol. Chem.***2019**, *17* (24), 5920–5924.

(2) Duan, X. C.; Chen, H.; Liu, F. F.; Conway, L.; Wei, S.; Cai, Z. P.; Liu, L.; Voglmeir, J. One Assay for All: Exploring Small Molecule Phosphorylation Using Amylose-Polyiodide Complexes. *Anal. Chem.***2015**, *87* (19), 9546–9550.

(3) Kittl, R.; Kracher, D.; Burgstaller, D.; Haltrich, D.; Ludwig, R. Production of Four Neurospora Crassa Lytic Polysaccharide Monooxygenases in Pichia Pastoris Monitored by a Fluorimetric Assay. *Biotechnol. Biofuels***2012**, *5* (1), 79.

(4) Herter, S.; McKenna, S. M.; Frazer, A. R.; Leimkühler, S.; Carnell, A. J.; Turner, N. J. Galactose Oxidase Variants for the Oxidation of Amino Alcohols in Enzyme Cascade Synthesis. *ChemCatChem***2015**, *7* (15), 2313–2317.

Chapter 5: Biochemical characterisation of a glycoside hydrolase family 43 β-D-galactofuranosidase from the fungus *Aspergillus niger*

This chapter is composed of one article in preparation and supporting information.

Gregory S. Bulmer, Fang Wei Yuen, Naimah Begum, Bethan S. Jones, Sabine L. Flitsch and Jolanda M. van Munster, Biochemical characterisation of a glycoside hydrolase family 43 β-D-galactofuranosidase from the fungus *Aspergillus niger*, in preparation

Foreword

This chapter presents the biochemical characterisation of XynD, a β-galactofuranosidase. Phylogenetic analysis of the GH43 subfamily 34 revealed clustering of known activities, which in tandem with molecular docking studies enabled the prediction of substrate specificity. XynD was assessed against a variety of Gal*f*-containing oligo- and polysaccharides with varying linkage types, activity was observed against pNP-Gal*f*.

Contribution

G.S. Bulmer designed and performed the majority of experiements. F. Yuen and N. Begum assisted with cloning, expression and characterisation of XynD. B. S. Jones assisted with molecular docking studies. S. L. Flitsch and J. M. van Munster provided conceptualisation and supervision. G. S. Bulmer and J. M. van Munster wrote and edited the manuscript.

Biochemical characterisation of a glycoside hydrolase family 43 β-D-galactofuranosidase from the fungus *Aspergillus niger*

Gregory S. Bulmer[1], Fang Wei Yuen[1], Naimah Begum[1], Bethan S. Jones[1], Sabine L. Flitsch[1] and Jolanda M. van Munster*[, 1,2]

[1] Manchester Institute of Biotechnology (MIB) & School of Natural Sciences, The University of Manchester, 131 Princess Street, Manchester, M1 7DN, United Kingdom

[2] Scotland's Rural College, West Mains Road, King's Buildings, Edinburgh, EH9 3JG, United Kingdom

*Corresponding author: Jolanda M. van Munster, Email: jolanda.van-munster@sruc.ac.uk

Abstract

β-D-Galactofuranose (Gal*f*) and its polysaccharides are found in bacteria, fungi and protozoa but do not occur in mammalian tissues, and thus represent a specific target for anti-pathogenic drugs. Understanding the enzymatic degradation of these polysaccharides is therefore of great interest, but the identity of fungal enzymes with exclusively galactofuranosidase activity has so far remained elusive. Here we describe the identification and characterisation of a galactofuranosidase from the industrially important fungus *Aspergillus niger*. Phylogenetic analysis of glycoside hydrolase family 43 subfamily 34 (GH43_34) members revealed the occurrence of three distinct clusters and, by comparison with specificities of characterised bacterial members, suggested a basis for prediction of enzyme specificity. Using this rationale, in tandem with molecular docking, we identified a putative β-D-galactofuranosidase from *A. niger* which was recombinantly expressed in *Escherichia coli*. The Gal*f*-specific hydrolase, encoded by *xynD* demonstrates maximum activity at pH 5, 25 °C towards 4-Nitrophenyl-β-galactofuranoside (*p*NP-β-Gal*f*), with a K_m of 17.9 ± 1.9 mM and V_{max} of 70.6 ± 5.3 µmol min^{-1}. The characterization of this first fungal GH43 galactofuranosidase offers further molecular insight into the degradation of Gal*f*-containing structures and may inform clinical treatments against fungal pathogens.

Introduction

The 5-membered ring form of galactose, β-D-galactofuranose (Gal*f*) is a key structural component of many pathogens, in which glycans containing the sugar can be highly immunogeneic [1]. Despite its presence spanning fungi, bacteria, protozoa, sponges and green algae [2] the monosaccharide is not present in mammalian tissue, offering a clear target for therapeutics [1–3]. The occurrence of Gal*f* differs greatly from its pyranoside ring form, which is found widely across mammalian biology [4]. Despite the importance of motifs containing this monosaccharide, knowledge is limited regarding enzymes that are active on glycans containing Gal*f*.

In many fungi Gal*f* is a key structural component of the cell wall and is present in secreted molecules, it has been shown to drive immunogenic responses in mammals [5–8]. Gal*f* containing polysaccharides are found in pathogens such as *Mycobacterium tuberculosis*, *Cryptococcus neoformans* and *Aspergillus fumigatus* which combined are responsible for over 2 million deaths worldwide per annum [5,9,10]. Gal*f* can constitute either the core or the branched sections of polysaccharides. In *M. tuberculosis* arabinogalactan (AG), a core of ~35 Gal*f* moieties with alternating β-1,5 and β-1,6 linkages is decorated with a variety of arabinose branches (Fig. 1a) [11–13]. *Mycobacterium* knockouts of the gene encoding UDP-galactopyranose mutase are unable to produce Gal*f* and fail to proliferate *in vitro*, showing that galactan synthesis is essential for replication of *Mycobacterium* [14]. In Aspergilli the predominant Gal*f*containing structure is galactomannan, a polysaccharide formed by a α-1,2/α-1,6-linked mannose backbone with Gal*f* side chains (Fig. 1b) [15–18]. The detection of this galactomannan is utilised to test for *Aspergillus* in a clinical setting [19]. Galactomannan plays a key role in cell integrity whereby *A. fumigatus* mutants lacking the ability to insert galactomannan into their cell walls suffer severe growth impairment [20]. Additionally Gal*f* is also found in a variety of glycoconjugates: such as O-antigens of *Escherichia coli* lipopolysaccharide [21], *Klebsiella pneumonia* galactan-I repeating unit [22], lipophosphoglycans and glycoinositolphospholipids of leishmaniosis-causing protozoa *Leishmania major* [23] and as part of

glycoinositolphospholipids and N- and O-glycans on proteins secreted by *A. niger* [24,25](Fig. 1c).

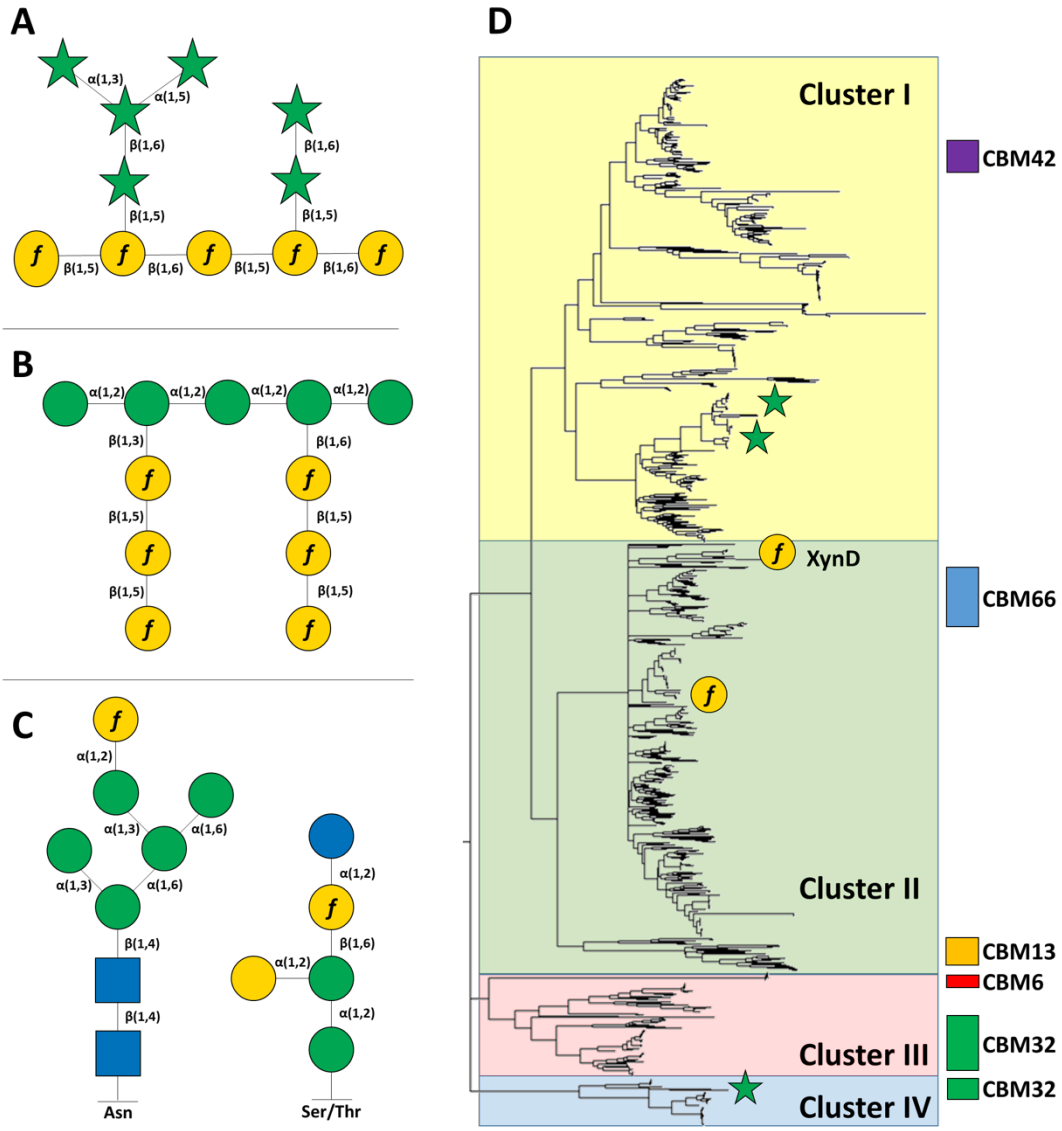


Figure 1: The presence of Galf throughout the fungal and bacterial kingdoms is not uniform and contains a variety of structures of both Galf backbones and Galf side chains, for example. **a** Structure of M. tuberculosis arabinogalactan fragment, **b** A. fumigatus galactomannan fragment, **c** N- and O-linked glycans. **d** Phylogenetic tree of GH43_34 family proteins with distinct clades visualised. Known activities are demonstrated as follows: β-galactofuranosidase (yellow circle), α-arabinofuranosidase (green star).

In a clinical setting, understanding the turnover of immunogenic sugars is of great importance to treating patients with invasive symptoms from various fungal infections. Additionally, comprehension of the metabolism of Gal*f* and polysaccharides containing it offer opportunities for industrial application. Fungi such as *A. niger* are often utilised as expression hosts for complex pharmaceuticals with human or other mammalian recipients [26,27]. Therefore, considering the metabolism of such immunogenic sugars as Gal*f* is important when designing such processes to avoid potential immune cross-reactivity.

Recently characterized bacterial β-galactofuranosidases (Gal*f*ases) begun to enter the literature, including Gal*f*ases specific for Gal*f* as well as bifunctional enzymes with activity on Gal*f* and Ara*f* [28–30]. However, research into fungal Gal*f*ases is considerably less advanced. A Gal*f*ase from *A. niger* has previously been employed as a tool for glycoconjugate analysis where proof of concept studies demonstrated the enzyme was able to remove Gal*f* from glycoproteins with *O*-linked (glucoamylase GAM-1) and *N*-linked (α-galactosidase A) Gal*f* containing glycans [31,32]. In addition the supernatant of *A. niger* cultures can hydrolyse the biologically relevant fungal galactomannan [33]. However, at a gene level the source of this activity lacks complete characterisation. *A. niger* α-L-arabinofuranosidases AbfA and AbfB from GH families 51 and 54 have a dual activity and can hydrolyse both pNp-α-Ara*f* and pNp-β-Gal*f* [33] but are not active on fungal galactomannan, indicating other β-galactofuranosidases remain uncharacterised. While undertaking the work described here, discovery of GH2 family β-galactofuranosidases GfgA and GfgB in *A. nidulans* has been reported, thus identifying the first specific fungal β-galactofuranosidases [34]. The biological function of Gal*f*ase enzymes *in situ* remain unclear, however, it has been suggested that the breakdown of galactomannan and other Gal*f* epitopes may be utilised as a carbon source, via degradation by Gal*f*ase, during carbon limitation to produce a source of galactose [32].

One reason for limited understanding of Gal*f*ases stems from the under examined nature of many large GH families. For example, members of GH43-34, comprised to our knowledge only 6 biochemically characterised bacterial proteins at the time of

writing, which all showed either β-D-galactofuranosidase or α-L-arabinofuranosidase activity [35–37]. As bi-functionality of Ara*f*ases and Gal*f*ases seems prevalent [28,33,34] and Ara*f* and Gal*f* are structurally similar we hypothesized that fungal Gal*f*ases were potentially harboured in GH families containing both Ara*f*ases and Gal*f*ases.

Therefore, in this work, we describe how we identified an *A. niger* GH43-34 enzyme, XynD, as a potential Gal*f*ase candidate via a combination of phylogenetic analysis, structural modelling and substrate docking. XynD is known to be expressed during *A. niger* interaction with plant biomass, however the biochemical function of XynD had yet to be elucidated. We express and characterise the enzyme, XynD, and demonstrate that the enzyme is specific for Gal*f*. This Gal*f*ase enzyme represents the first identified and characterized fungal GH43 family galactofuranosidase.

Results

Phylogenetic analysis of XynD

We aimed to elucidate the role and biochemical activity of the GH43-34 enzyme from *A. niger*. Members of this subfamily exhibit different biochemical specificities, as either galactofuranosidases, arabinofuranosidases or arabinanases. Therefore, we investigated whether biochemical activity could be inferred from phylogenetic analysis and comparison of characterised enzymes with known activities [35,38,39] in this subfamily. Catalytic domains of all fungal and bacterial entries from GH43-34 on the CAZy database [40] were selected for phylogenetic analysis. Construction of a phylogenetic tree via Maximum likelihood analysis revealed four distinct clusters (cluster I-IV) within the family (Fig. 1d). The distribution of sequences over the clusters does not follow the evolutionary relationships between species, with a mixture of fungal and bacterial sequences found in cluster I and II, while cluster III and IV comprised only bacterial members. For those organisms containing multiple GH43-34 members, different enzymes can be located in different clusters. Cluster I contains two arabinofuranosidases (ALJ58905.1 and AAO78780.1) whilst cluster II contains one biochemically characterised enzyme, a galactofuranosidase (ALJ48250.1). Based on this separation, we hypothesize that *A. niger* XynD

(An11g03120, GenBank: CAK40644.1), which is also found in cluster II, may have galactofuranosidase activity. Cluster IV contained one arabinofuranosidase (AAO78767.1) whilst cluster III contained no characterised members. Additionally through utilising the CAZy database we identified CBM (Carbohydrate Binding Module) protein domains that co-occur with GH43-34 members, namely CBM6, CBM13, CBM32, CBM42 and CBM66 (Fig. 1). See Fig. S1 for bootstrap values and fungal member locations.

Modelling and substrate docking

To investigate the potential activity of *A. niger* XynD in more detail we created a model of the protein's structure based on the crystal structure of GH43-34 arabinofuranosidase from *Bacteroides thetaiotaomicron* BT_3675 (PDB 3QZ4) [41]. Of the crystallised GH43 subfamily 34 members, 3QZ4 gave the greatest percentage coverage for modelling (89%), with 37% sequence identity. GH43 family enzymes display a five-bladed β-propeller fold [42], the presence of this fold was also predicted in the XynD model. The XynD model composes of a monomer 36.4 kDa in size and aligns with one of the monomers of the 3QZ4 dimer macromolecule. No kinetic or substrate binding data is available for 3QZ4. Therefore, using this model, we applied molecular docking simulations to gain insights into the functional binding of potential substrates in the active site. Docking studies showed energetically favourable binding of galactofuranose to the catalytic domain of XynD. Substrate docking suggested five residues involved in catalysis that were energetically favourable: Asp-28, Asp-152, Glu-199, Trp-219 and Arg-286. The first three of these (Fig. 2a) correspond with the catalytic Brønsted base, pKa modulator and catalytic acid that have been identified in GH43 family members [42]. Molecular docking of arabinose in the active site demonstrated that interaction between this sugar and residues Asp-152 and Glu-199 was lacking and thus suggested a specificity for galactofuranose as opposed to arabinofuranosidase or galactofuranosidase/ arabinofuranosidase dual activity. Analysis of the substrate binding site suggests a restricted binding pocket for galactofuranose (Fig. 2b).

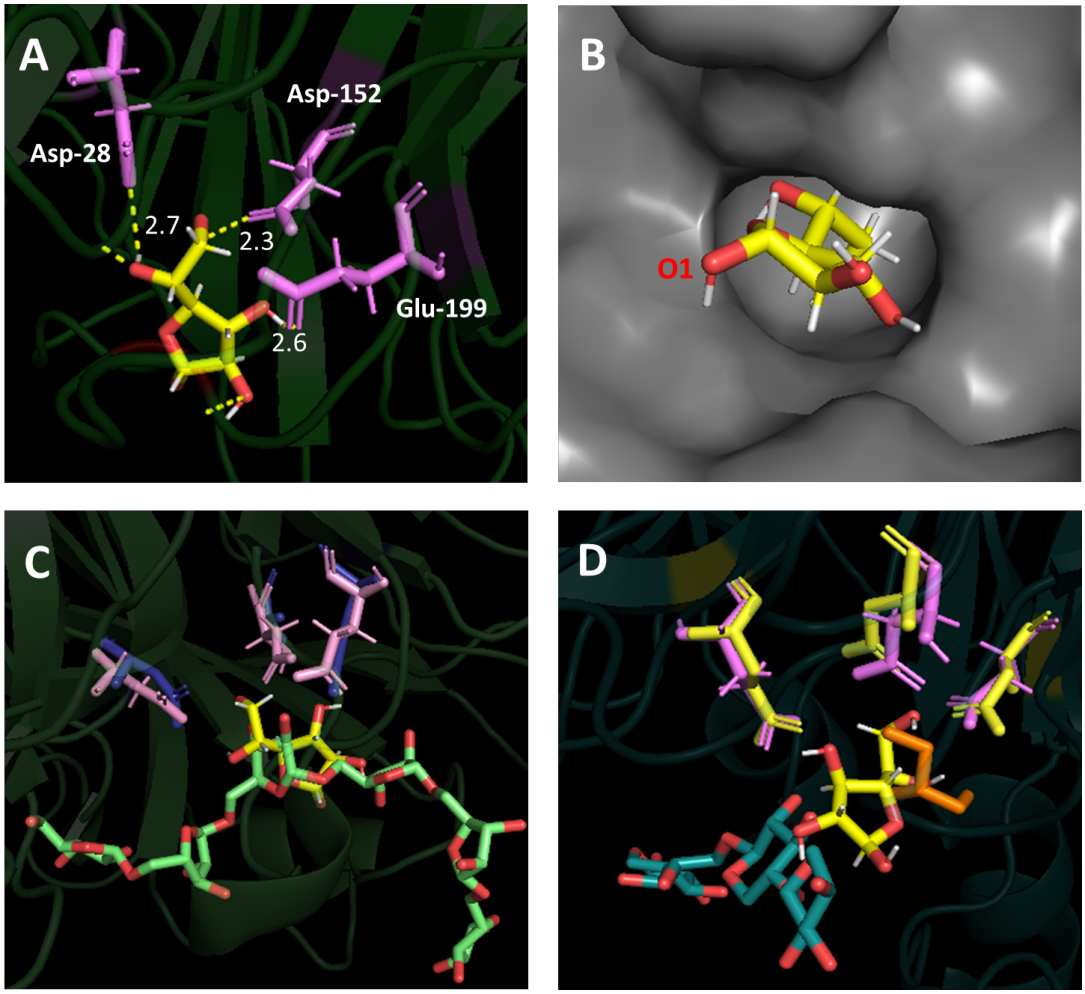


Figure 2: **a** Molecular docking (AutoDock VINA as implemented in YASARA) of galactofuranose (yellow) in a model of *A. niger* XynD generated using PDB 3QZ4 as a template. Catalytic residues Asp-28, Asp-152, Glu-199 are displayed in lilac with bonding lengths indicated in Å **b** XynD surface highlighting restricted binding pocket **c** GH43 *C. japonicus* alpha-L-arabinase Abf43A (PDB 1GYE) depicted with catalytic triad (blue) and arabinohexaose ligand (green) overlaid with catalytic triad of XynD (liliac) and Galf ligand (yellow) **d** *B. subtilis* arabinoxylan arabinofuranohydrolase BsAXH-m2,3 (PDB 3C7G) depicted with catalytic triad (yellow), glycerol (orange) and xylotetraose ligand (blue) overlaid with catalytic triad of XynD (liliac) and Galf ligand (yellow).

In order to understand the roles of residues in XynD the enzyme was compared with the crystal structures of related enzymes. Comparison of the XynD model with the *Cellvibrio japonicas* alpha-L-arabinanase Arb43A complexed with arabinohexaose (1GYE) (Fig. 2c) and *Bacillus subtilis* arabinoxylan arabinofuranohydrolase BsAXH-m2,3 in complex with xylotetraose (3C7G) (Fig. 2d),

suggested the catalytic triad was conserved. Differences were observed between these proteins in residues involved in hydrophobic stacking. BsAXH-m2,3 contains several residues involved in hydrophobic stacking: Phe-244 at the II subsite, Trp160 at the III (-1) subsite and Trp-101 at the IV (+1) glycerol containing subsite. XynD contains an equivalent residue for Phe-244 (XynD Trp-219) and Trp-101 (XynD Trp-86), but not for Trp-160. Alignment also identified Arg-286 as an equivalent residue to the BsAXH-m2,3 Arg-321 involved in +1 subsite stability. 1GYE shares the three conserved catalytic residues with XynD but the Phe-114 of 1GYE involved in the high binding affinity of the enzyme to arabinan has no corresponding residue in the XynD model. These results offer insight into the highly conserved nature of the catalytic residues, whilst the variation of other residues involved in binding is rather variable and may potentially play a role in enzyme specificity. Trp-219 is found exclusively in Gal*f*ase clusters (Fig. S2), whilst the majority of those in the Ara*f*ase clusters contain a Thr-219. This suggests a key role played by the residue in determining substrate specificity.

Analysis of the protein surface, pore size and accessibility to previously known subsites further confirms a suspected galactofuranosidase specificity of XynD. When superimposed against GH43 arabinoxylan arabinofuranohydrolase from BsAXH-m2,3 with bound xylotetraose (3C7G) (Fig. 3a) a loop on XynD appears to obstruct the I and II sites resulting in a far smaller and shallower binding cleft than BsAXH-m2,3 and suggests XynD is not able to bind and process larger arabinoxylan or derived oligosaccharides (Fig. 3b). Further comparison with Arb43A (1GYE) (Fig. 3c) again shows a far less open binding pocket for XynD when superimposed (Fig. 3d).

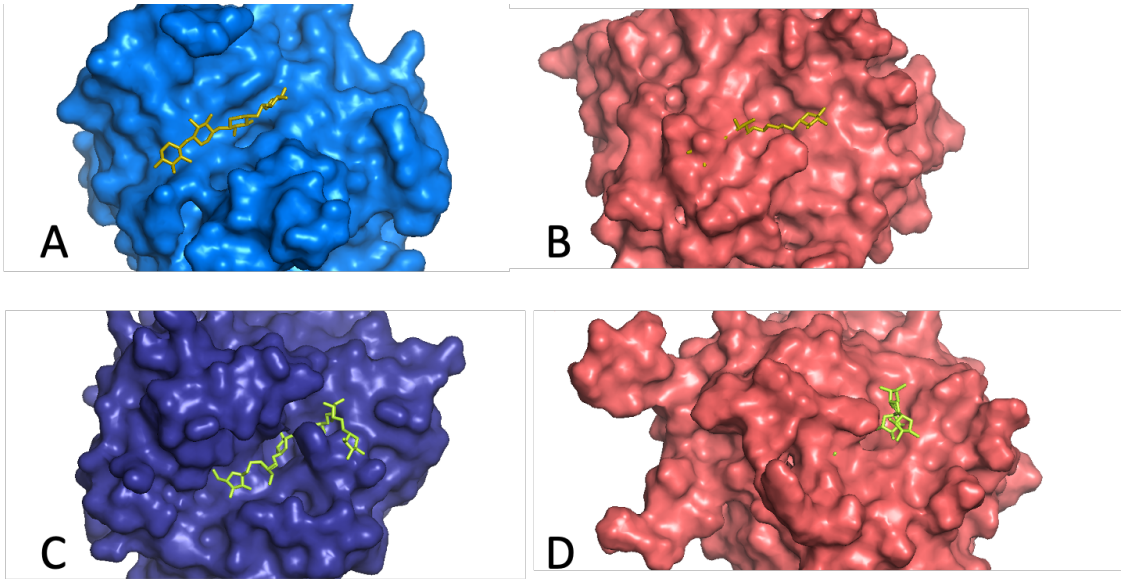


Figure 3: a Xylotetraose (X4) binding to *B. subtilis* arabinoxylan arabinofuranohydrolase BsAXH-m2,3 (PDB 3C7G) **b** Superimposition of XynD over 3C7G with X4 ligand **c** Arabinohexaose (A6) binding to *C. japonicus* alpha-L-arabinase Abf43A (PDB 1GYE) **d** Superimposition of XynD over 1GYE with A6 ligand.

Cloning, expression and purification of XynD

To confirm the activity of XynD experimentally, we heterologous expressed the protein. The predicted signal peptide (residues 1-24) is consistent with transcriptomics data [43] suggesting that XynD is a secreted protein and that XynD contains four putative N-glycosylation sites. We therefore attempted expression in *Pichia pastoris*, based on a codon optimised cDNA sequence and the pPICZα expression system, but initial trails were unsuccessful. Expression was achieved in *Escherichia coli* Rosetta (DE3) as C-terminally His-tagged XynD without its signal peptide from a pET21a plasmid. The protein was purified via Ni-NTA resin affinity chromatography followed by anion exchange chromatography with a final yield of 7.7 mg l^{-1} culture. Analysis of the purified protein on SDS-PAGE (Fig. S3) indicates that the recombinant enzyme possesses an apparent mass of approximately 35 kDa, which is consistent with theoretical molecular weight of 36.3 kDa.

Substrate specificity

The activity of XynD was assessed against a panel of ten 4-nitrophenyl-conjugated saccharides. After incubation of 0.5 mg XynD with 1 mM substrate at 25 °C for 2 h, activity was observed only against pNP-β-Gal*f*, while no hydrolysis was detected of pNP-β-Gal*p*, pNP-β-(1,5)-Galf_3, pNP-α-Ara*f*, pNP-β-Glc, pNP-α-Glc, pNP-β-Xyl, pNP-β-Xyl_2, pNP-β-GlcNAc, pNP-βlactose. This analysis confirms that XynD is a galactofuranosidase.

Monosaccharide analysis of XynD incubation with galactomannan derived from *A. niger* strain N402, via HPAEC-PAD, detected no release of Gal*f* and thus suggests XynD is not directly involved in degradation of intact galactomannan (Fig. S4). This was confirmed via analysis of the presence of the Gal*f* epitope in galactomannan via the Platelia Aspergillus galactomannan ELISA assay, the abundance of Gal*f* was identical before and after treatment of galactomannan with XynD.

To further test the substrate specificity of XynD, A variety of potential substrates were tested. No XynD activity was detected against the β-(1,5)-linked oligosaccharides 4MU-Galf_3, 4MU-Galf_4, 4MU-Galf_5, 4MU-Galf_6, a gift of the Oka group [44]. Incubation of XynD with Octyl β-D-Galactofuranosyl-(1,6)-β-D-galactofuranosyl-(1,5)-β-D-galactofuranoside and Octyl β-D-Galactofuranosyl-(1,5)-β-D-galactofuranosyl-(1,6)-β-D-galactofuranoside, a gift from the Lowary group [3] revealed no observable activity (by MALDI-ToF MS and TLC analysis) against these trisaccharides (Fig. S5). Together these results indicate that XynD is not active on β-1,5- or β-1,6-linked Gal*f* residues.

N-linked glycans on *A. niger* proteins have been described to contain terminal α-(1,2)-linked Gal*f* moities [45] as well as β-linked Gal*f* *residues*. To assess activity of XynD on such oligosaccharides, N-linked glycans (Fig. S6) were isolated from Transglucosidase L "Amano" (Amano) as described previously [45] and purified, and incubated with XynD. Analysis by MALDI-TOF MS (Fig. S7) revealed no observable XynD activity against the glycans. However, as the glycans consisted of a range of structures with mass differences corresponding to one pyranose residue,

assessement of activity was not straight forward. The glycan pool was therefore co-digested with jack bean α-mannosidase which cleaves terminal α-(1,2), α-(1,3) and α-(1,6)-mannoses, thus simplifying the glycan structures. A main product structure with a mass of 1095 was detected, corresponding with the mass expected for the sodiated retainment of Gal*f*-α-(1,2)-Man-α-(1,3)-Man-α-(1,6)-Man-β-(1,4)-GlcNAc-β-(1,4)-GlcNAc, thus confirming that glycan structures terminated with Gal*f* (or potentially other non-mannose residues) were obtained. However, no reaction products were observed after incubation of this substrate with XynD, confirming XynD is not active on these Gal*f* N-glycans.

We conclude that the natural substrate of XynD may either be a disaccharide or consist of a glycan with a different linkage type from β-(1,5), β-(1,6) or α-(1,2), for example such as β-(1,2)-linked Gal*f* as found in glycoinositolphospholipids.

Biochemical characterisation

Using Gal*f*-pNP as a model substrate, we assessed the optimum reaction conditions and kinetic parameters for XynD. The enzyme exhibited its highest activity at pH 4-5, as assessed after incubation of purified XynD at 25 °C for 1 h in a range of citrate-phosphate buffers at pH 3-9 (Fig. 4a). At pH >5 the enzyme showed substantial reduction in activity. This is similar to a previously described *A. niger* β-D-galactofuranosidase that demonstrated optimal activity at pH 3-4 [32]. XynD temperature optimum was 25 °C (Fig. 4b) with substantial decrease in activity observed after incubation at temperatures greater than 30 °C. The temperature stability of XynD decreased when incubated above 30 °C for 1 h (Fig. S8).

The specific activity of XynD against *p*NP-β-Gal*f* at 25 °C, pH 5 was 0.23 μmol min^{-1} mg^{-1}. The activity appears to follow Michaelis-Menten kinetics, displaying a K_m of 17.9 ± 1.9 mM, V_{max} of 70.6 ± 5.3 μM min^{-1} and K_{cat} of 0.14 s^{-1} (Fig. S9).

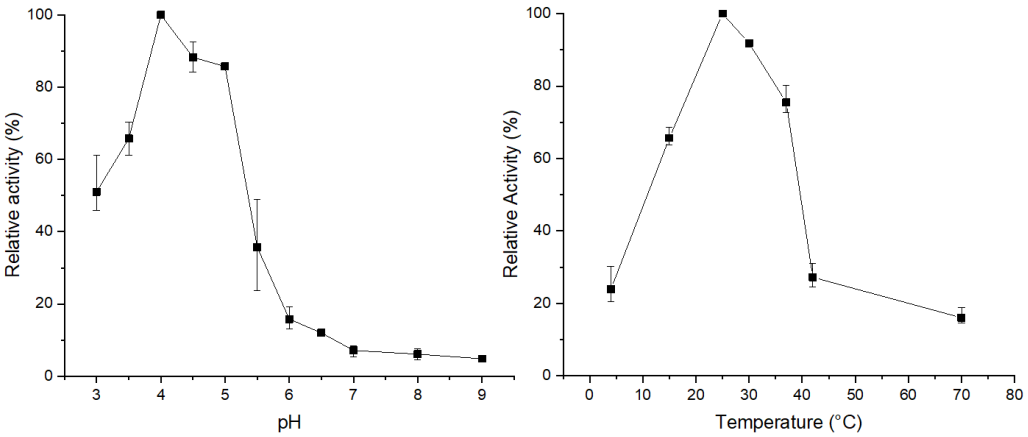


Figure 4: Optimisation of reaction conditions for pNP-Galf hydrolysis by XynD **a** Relative activity at pH 3-9 **b** Relative activity at temperatures 4-70 °C.

Discussion

Gal*f* containing glycans such as galactomannan or O-glycans are important as drug targets and immunogenic motifs in the production of pharmaceuticals due to the absence of the Gal*f* monosaccharide in mammalian tissues [23]. There has therefore been a longstanding interest in elucidation of the enzymatic mechanism contributing to the turnover of such glycans[1].

Galactomannan degrading activity of an unidentified enzyme produced by *A. niger* has been previously reported [32]. Subsequent investigation to identify this source of activity has revealed a variety of Gal*f* active enzymes within the *Aspergillus* genus. *A. niger* AbfA and AbfB both have dual activity against pNP-Gal*f* and pNP-Ara*f* yet do not degrade galactomannan [33]. More recently, *A. nidulans* exo-acting Gal*f*ases GfgA and GfgB have been identified, and have been shown to hydrolyze both β-(1,5) and β-(1,6) linkages, whilst *A. fumigatus* Afu2g14520 demonstrates dual activity against pNP-Gal*f* and pNP-Ara*f,* with higher activity against the latter [34]. Here we demonstrated that XynD encodes an galactofuranosidase active against pNP-Gal*f.* The enzyme is not responsible for galactomannan degradation and lacks activity against β-(1,5), β-(1,6)- and α-(1,2) linked Gal*f* substrates. Instead its activity suggests glycoinositolphospholipidsmay be the natural substrate.

XynD has previously been annotated as a putative xylosidase in the literature and our identification of its galactofuranosidase activity highlights how detailed

phylogenetic analysis combined with biochemical characterisation is critical to understand the roles of glycoside hydrolases. Our phylogenetic analysis of the subfamily, followed by mapping of biochemically characterised members, demonstrated how GH43-34 is grouped into multiple clusters containing characterised members with arabinofuranosidase activity. Assignment of Cluster I members as arabinofuranosidases was further supported by association of a subset of members from this cluster with CBM42, which are commonly associated with GH54 arabinofuranosidases and bind arabinofuranose present in arabinoxylan [46]. Only cluster II contained bacterial members with galactofuranosidase activity and our structural modelling, substrate docking and biochemical characterisation confirmed our hypothesis that as fungal member of this cluster, XynD, had galactofuranose activity.

Previously studies of GH family 43 reveal only the catalytic residues (Asp, Glu, and its pK_a-modulating Asp) are well conserved [42], and our molecular docking studies confirm the binding pose of galactofuranose in XynD is stabilized to these from local Asp, Glu, Trp and Arg residues. In XynD, many potential subsites are completely occluded by surrounding loops between separate beta-strands and thus explains the lack of catalysis demonstrated against large oligo- and polysaccharides. The variety of specificities observed in the GH43 family stem from the variable orientation of which the substrate can sit within the deep pocket of the active site, thus resulting in a diverse and complex family [47,48].

Our modelling of the XynD structure combined with sequence alignments highlighted how an amino acid in the active site of the GH43-34, is consistently different between the arabinofuranosidases and galactofuranosidases and could affect the substrate specificity in the GH43-34 subfamily. This W219, whose equivalent is a T in arabinofuranosidases, is positioned to affect sugar binding in the subsite and we hypothesised that mutation of this site may switch enzyme activity from a galactofuranosidase to and arabinofuranosidase. Attempts to a form a W219T mutant of XynD were unsuccessfull due to a lack of protein production from a mutant plasmid.

XynD was the first fungal member of GH family 43-34 to be recombinantly expressed and biochemically characterised. We showed XynD demonstrates hydrolytic activity against pNP-Gal*f.* Interestingly, despite the substantial structural similarities between Gal*f* andAra*f,* XynD does not show activity against pNP-Ara*f* and is thus thereby differs from other characterised *A. niger* galactofuranosidases that exhibit dual activity [33]. An unidentified *A. niger* enzyme with Gal*f*ase activity has been studied [32]. The enzyme is substantially larger than XynD at ~80 kDa (after N-glycan removal by EndoH treatment), although this difference could be accounted for by the presence of O-linked glycans. The unidentifed Gal*f*ase exhibits a lower a K_m (4 mM) in comparison to that of XynD (17.9 mM), and therefore likely is a different enzyme from XynD.

XynD displays an optimum pH between pH 4 and pH 5, and concurs with previously described recombinant galactofuranosidases with varying optimum pHs of 3-7 [28,30,34,49,50]. Optimal temperature of characterised galactofuranosidases vary, for example, the Galf-ase derived from Streptomyces sp. JHA26 lost its activity when incubated at a temperature over 45 °C whilst Streptomyces sp. JHA19 galactofuranosidases had an optimal temperature at 60 °C [50,51]. Since XynD was expressed in *E. coli* no post translational modifications, such as glycosylation, will have occurred. Therefore, lack of post translational modifications should be considered when evaluating parameters such as activity, pH optimum and temperature optimum.

To gain more insight in the biological role of XynD, we assessed the conditions under which *xynD* expression and XynD secretion has been reported. Gene co-expressing networks, based on 155 transcriptomics experiments [52] and available via FungiDB, highlighted expression of *xynD* is positively correlated with 110 genes, and gene ontology (GO) enrichment analysis using REViGO [53] (Fig. S10) revealed that these are enriched for transmembrane transporter activity (GO:0022857), oxidoreductase activity (GO:0016491), hydrolase activity, acting on glycosyl bonds (GO:0016798) and hydrolase activity with activity on carbon-nitrogen (but not peptide) bonds (GO:0016810). This positive correlation suggest *xynD* is expressed during carbohydrate degradation. In *A. niger*, *xynD* expression has previously been

reported to be induced in response to plant-derived polysaccharides and lignocellulose, despite lack of putative binding sites in its promoter for the master regulator XlnR. Low levels of expression *xynD* are observed on a variety of easily metabolised carbon sources (glucose, sucrose, fructose and sugar beet pulp) and expression is suspected to be subject to carbon catabolite repression, as supported by 11 binding sites for carbon catabolite repressor CreA in the promotor region of *xynD* [54] and other genes undergoing such repression [55]. Whilst *xynD* is not expressed during growth on maltose in the fungal wild-type [55], mutants with an inactivated *amyR* (encoding a regulator of starch degradation induced by maltose[56]) display upregulation of *xynD* transcription [55]. Similarly, *xynD* expression on arabinose is low in the wild-type, but an *araR* disruption mutant showed upregulation of *xynD* transcription [55].In the *amyR* [57]and *araR* [58]mutants the expression of sugar transporters and metabolic enzymes that enable growth on maltose and arabinose, respectively, will be negatively affected and thus the observed expression patternis consistent with a response to carbon limitation and concurs with the expression response to starvation in other *A. niger* galactofuranosidases [32].

In conclusion, this study revealed further insight into the degradation of galactofuranose containing biomolecules with the identification of a GH43-34 *A. niger* galactofuranosidase. Understanding this vast but largely undescribed subfamily could play a key role in developing knowledge of glycosylation in a variety of fungal products, such as pharmaceuticals.

Experimental procedures

Phylogenetic analysis of XynD

GH 43-34 domains were identified using the Pfam data base [59] then subsequently aligned and trimmed to just the catalytic domain using BioEdit [60]. Evolutionary analyses were conducted in MEGA X[61]. The evolutionary history was inferred by using the Maximum Likelihood method and JTT matrix-based model [62]. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbour-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the JTT

model, and then selecting the topology with superior log likelihood value. The tree was drawn to scale, with branch lengths measured in the number of substitutions per site. This analysis involved 963 amino acid sequences. There were a total of 494 positions in the final dataset.

Computational docking studies

Enzyme modelling and molecular docking simulation was performed using the YASARA software from YASARA Biosciences GmbH. As there is no PDB structure of XynD a model was produced using *Bacteroides thetaiotaomicron* endo-1,4-beta-xylanase D (PDB 3QZ4) as a template and was subsequently used to generate the receptor for simulations. In all cases, the raw crystal structures were first prepared for simulation using the YASARA „Clean“ script, which detects and amends crystallographic artefacts and assigns protonation states. Before docking, the energy minimization script of the YASARA software package was applied to the protein in order to obtain the most likely starting structure. For energy minimization, a simulation cell was defined around the whole protein and the cell was filled with water molecules. For docking, the water was removed afterwards with a variety of arabinose and galactofuranose based docking ligands. Ligands were minimised in vacuo using the energy minimization script of the YASARA software package. The docking simulation itself was performed using the dock_run.mcr macro using VINA as the docking method and the AMBER03 force field. Appropriate simulation cells were defined for the respective docking simulations. The created receptor-ligand complex structures were further processed using the PyMOL software from Schrodinger LLC involving the identification of polar contacts between the ligand and the receptor, as well as determination of bond distances.

Construction of expression vectors

The gene *xynD* (An11g03120, GenBank: CAK40644.1), with an 86 bp intron removed, was synthesised by GeneArt gene synthesis (Thermo Fisher scientific) with codon optimisation for expression in *Pichia pastoris*. Construct included 5’ and 3’ complementary overhangs for ligation into pPICZαA using NEBuilder HiFi DNA Assembly (New England Biolabs) to form pPICZαA_GH43_34. The coding sequence

of *xynD* gene was amplified from pPICZαA_GH43_34 using primers 5' CGCCATATGCACCCACAACAGAACTTG-3' (forward) and 5'-CGCCTCGAGAGACAAAGTTCTACCCTCGACAC-3' (reverse). The primers contained restrictions sites *Nde*I and *Xho*I (underlined) for forward and reverse primers, respectively. The PCR conditions were as follows: initial denaturation 95 °C (5 min), [95 °C (30 s) denaturation, 58 °C (30 s) annealing and 72 °C (30 s) extension] for 20 cycles, final extension 5 min. PCR amplifications were conducted using Phusion DNA polymerase. The resulting products were run on 1 % agarose gel then extracted using QIAquick Gel Extraction Kit (Qiagen). Restriction enzymes and Phusion DNA polymerase were purchased from New England BioLabs Inc and used according to the manufacturer's instructions. The amplified fragment was cloned into linearized (NdeI and XhoI) pET21a to form pET21a_43_34. Correct insertion of the gene in the expression vector was confirmed by sequencing. The amino acid sequence of the recombinant XynD is as follows: MHPQQNLLATTTSNTKAGNPVFPGWYADPEARLFNAQYWIYPTYSADYSEQTFFDAFSSPDL LTWTKHPTILNITNIPWSTNRAAWAPSVGRKLRSSANAEEEYDYFMYFSVGDGTGIGVAKSTT GKPEGPYEDVLGEPLVNGTVYGAEAIDAQIFQDDDGRNWLYFGGWSHAVVVELGEDMISLK GDYLEITPEGYVEGPWMLKRNGIYYYMFSVGGWGDNSYGVSYVTADSPTGPFSSTPKKILQG NDAVGTSTGHNSVFTPDGQDYYIVYHRRYVNDTARDHRVTCIDRMYFNEAGEILPVNITLEGV EGRTLS(LEHHHHHH-)

Protein expression and purification

E. coli strain Rosetta™(DE3) carrying pET21a_43_34 was inoculated into LB media supplemented with 100 ug ml^{-1} ampicillin and grown at 37 ^{o}C, 250 rpm until OD_{600} ~0.25. Cells were then induced with 1mM IPTG final concentration and grown at 22 ^{o}C, 250 rpm for 16 h. Cells were harvested by centrifugation (4000 rpm, 15 min, 4 °C), suspended in His-buffer (Tris 50 mM, NaCl 250 mM, imidazole 20 mM, pH 8) and lysed by sonication (20s on, 20s off, 5 min). Cell debris was removed by centrifugation (4000 rpm, 15 min, 4 °C). Protein was purified by metal affinity chromatography on a Ni^{2+}-NTA column (Qiagen). Cell free extract was loaded onto column followed by two wash steps with wash buffer (Tris 50 mM, NaCl 250 mM, imidazole 20 mM, pH 8) followed by application of elution buffer (Tris 50

mM, NaCl 250 mM, imidazole 200 mM, pH 8), all fractions were retained for SDS-PAGE analysis. The his-tagged purified protein was then further purified using anion exchange chromatography employing a HITrapQ HP column 1 mL, whereby the protein was eluted with a gradient of 20 mM Tris-HCl buffer, 0.5M NaCl, pH 7.2. Eluted proteins were analysed by 4—15% sodium dodecyl sulfate polyacrylamide (SDS-PAGE) gel electrophoresis (Mini-PROTEAN® TGX™ Precast Gels by Bio-rad) using broad range molecular weight markers (PageRuler™ Plus Prestained Protein Ladder) as standard, and the protein bands were visualised by staining with Coomassie Brilliant Blue G250. The protein concentration was determined by Bradford method using bovine serum albumin (BSA) as a standard based on the protocol provided by Sigma-Aldrich (96 Well Plate Assay Protocol). All experiments presented in this manuscript were performed with XynD purified by anion exchange chromatography unless otherwise stated.

Mass spectrometry of carbohydrates

Samples were prepared for analysis by MALDI-TOF via crystallisation with super DHB or THAP matrices. Super DHB (a 9:1 (w/w) mixture of 2,5-Dihydroxybenzoic acid and 2-hydroxy-5-methoxybenzoic acid) was prepared at 15 mg ml^{-1} in a mixture of 50 % (v/v) acetonitrile and 50 % (v/v) water containing 0.1 % trifluoroacetic acid (TFA). 2',4',6'-Trihydroxyacetophenone monohydrate (THAP) was prepared at 10 mg ml^{-1} in acetone. MALDI-TOF mass spectrometry was performed using the Bruker Ultraflex 3 in positive mode. MALDI-TOF was calibrated using peptide Calibration Standard II (Bruker) containing a range of peptides 757-3149 Da in size.

Substrate specificity

XynD (0.5 mg ml^{-1}) was incubated overnight at 25 °C in 50 mM citric acid-phosphate pH 5 with pNP-labelled substrates at 1 mM. All experiments were performed in triplicate. Reactions were halted via a 1:1 volume addition of Na_2CO_3 1M with pNP release measured at OD_{405}. The following substrates were used: pNP-β-D-galactofuranoside (pNP-β-D-Gal*f*; Carbosynth), pNP-α-L-arabinofuranoside (pNP-α-L-Ara*f*; Megazyme), pNP-β-galactopyranoside (pNP-β-Gal*p*; Sigma-Aldrich), pNP-β-glucopyranoside (pNP-β-Glc*p*; Sigma-Aldrich), pNP-α-glucopyranoside (pNP-α-Glc*p*; Sigma-Aldrich), pNP-N-acetyl-β-D-glucosaminide

(pNP-β-D-GlcNAc; Sigma-Aldrich), pNP-lactose (pNP-Lac; Sigma-Aldrich), pNP-xylopyranoside (pNP-Xyl*p*; Sigma-Aldrich), pNP-xylobiose (pNP-β-Xyl_2, Carbosynth), pNP-β-(1,5)-Gal*f*-β-(1,5)-Gal*f*-β-(1,5)-Gal*f*(pNP-β-(1,5)-Galf_3); gift of Dr. T. Oka.

XynD (0.5 mg ml^{-1}) was incubated overnight at 25 °C in 50 mM citric acid-phosphate pH 5 with the following 4MU-labelled substrates at 1 mM: 4MU-Galf_3, 4MU-Galf_4, 4MU-Galf_5, 4MU-Galf_6, (gifts of Dr. T. Oka, Sojo University, Japan). Reactions were halted by the addition of stop solution (1M sodium hydroxide, 1M glycine, pH 10) with released 4-MU detected after excitation at 365 nm and emission at 440 nm.

XynD (0.5 mg ml^{-1}) was incubated overnight with 1 mM final concentration Octyl β-D-Galactofuranosyl-(1,6)-β-D-galactofuranosyl-(1,5)-β-D-galactofuranoside and Octyl β-D-Galactofuranosyl-(1,5)-β-D-galactofuranosyl-(1,6)-β-D-galactofuranoside (gifts of Dr. T. Lowry, The Institute of Biological Chemistry, Academia Sinica, China) in 50 mM citric acid-phosphate buffer. Reaction were then monitored by MALDI-TOF MS.

The α-(1,2)-Gal*f*-containing N-linked glycans were released from 1 mg Transglucosidase L "Amano" (Amano) by using PNGase F (New England Biolabs) under denaturing conditions. Reactions were then filtered via Vivaspin 500 (Sigma-Aldrich) to obtain the glycan products. Samples were then purified using a 1 ml Supelclean™ ENVI-Carb™ SPE Tube. The column was prepared with 1 ml of each solution in a sequential order: 75% acetonitrile + 0.1% trifluoroacetic, 1 M NaOH, water, 30% acetic acid, water, 75% acetonitrile + 0.1% trifluoroacetic. The sample was then applied and washed with 3 column volumes of water followed by subsequent elution with 3 mL 25% acetonitrile + 0.1 % trifluoroacetic acid. The N-linked glycans were then dried and resuspended in water. Glycans were incubated with final concentation 1 mg ml^{-1} XynD in 50 mM citric acid-phosphate buffer, pH 5 overnight, then analysed by MALDI-TOF MS. Additionally glycans were co-incubated with 1 mg ml^{-1} XynD and 1 mg ml^{-1} α-Mannosidase from *Canavalia ensiformis* (Sigma-Aldrich).

Activity of XynD on *A. niger* galactomannan, a kind gift from Dr A. Ram (The University of Leiden, the Netherlands), was assessed via overnight incubations containing final concentration 0.5 mg ml^{-1} XynD and 1 mg ml^{-1}*A. niger* N402 galactomannan in 50 mM citric acid-phosphate buffer at 25 °C. Reactions were then filtered via Vivaspin 500 (Sigma-Aldrich) to remove protein from the reactions. Galactomannan, hydrolysed via incubation in trifluoroacetic acid at 100 °C for 5 h acted as a positive control for detection of constituant monosaccharides XynD activity was assessed via detection of released galactose by HPAEC-PAD as described in detail below, and via Platelia™ Aspergillus Ag (Bio-Rad), as detailed by the manufacturer.

HPAEC-PAD analysis

High performance anion exchange chromatography (HPAEC) was performed to analyze monosaccharide release by XynD, based on a published method[63]. A Thermo Scientific Dionex ICS-6000 HPAEC system with pulse amperometric detection (PAD) controlled by Chromeleon software version 7.1 was used. A CarboPac PA1 guard (2 mm x 50 mm, Thermo Fisher Scientific) and a CarboPac PA1 analytical column (250 mm × 2 mm, Thermo Fisher Scientific) for monosaccharide analysis were employed. Monosaccharides were eluted isocratically using H_2O and 1M NaOH in a ratio of 80:20 for 30 min with a flow rate of 0.25 ml min^{-1}.

Identification of pH and temperature optimum and stability

To identify the optimal pH for activity, 1 mM pNP-β-Galfwas incubated with XynD (0.5 mg ml-1) for 1 h at 25 °C in 50 mM citric acid-phosphate and Tris-HCl buffers at a range of pH 4-9. Reactions were halted via a 1:1 volume addition of 1M NaCO3, after which pNP release as measured at OD_{405}. To identify the optimal reaction temperature, reactions were incubated in 50 mM pH 5 citric acid-phosphate buffer at a range of temperatures from 4-90 °C. All experiments performed in triplicate and enzyme activity was linear over time under the conditions described. pH stability was determined via pre-incubation of the enzyme (0.3 mg ml^{-1}) for 1 h at pH 3-9 before assaying residual activity on pNP-Galf under standard condition reactions at pH 5.

Co-expression analysis of XynD

A. niger microarray data from FungiDB was used to elucidate co-expression networks. Spearman scores of 0.5 of greater are considered significantly co-expressed. 1241 / 110 genes showed a positive correlation with the *xynD* gene (Spearman coefficient ≥ 0.5 / ≥ 0.75) and 1476 / 70 genes a negative correlation (Spearman coefficient ≥−0.5 / ≥−0.75).

Acknowledgements

We thank Dr A. Ram of The University of Leiden, the Netherlands, for providing us with *A. niger* galactomannan, Dr T. Lowry of The Institute of Biological Chemistry, Academia Sinica, China for providing us with the (1,5),(1,6) and (1,6),(1,5) galactofuranoside trisaccharides and Professor T. Oka of Sojo University, Japan, for the pNP and 4-MU based galactofuranosides. We are grateful to Dr W. Finnigan and R. Sung of The University of Manchester, UK for their assistance with bioinformatics and HPAEC-PAD analysis, respectively.

Funding and additional information

This study was funded by the BBSRC via BB/P011462/1 (JvM), the EPSRC via DTP EP/N509565/1 supporting GB, an undergraduate bursary from the British Mycological Society supporting FWY and a Learning Through Research Internship from the University of Manchester supporting NB. BSJ and SLF acknowledge funding by the European Research Council under grant agreement no. 788231-ProgrES-ERC-2017-ADG.

Conflicts of interest

The authors declare that they have no conflicts of interest.

References

[1] C. Marino, A. Rinflerch, R.M. De Lederkremer, Galactofuranose antigens, a

target for diagnosis of fungal infections in humans, Futur. Sci. OA. 3 (2017) FSO119.

[2] P. Peltier, R. Euzen, R. Daniellou, C. Nugier-Chauvin, V. Ferrières, Recent knowledge and innovations related to hexofuranosides: structure, synthesis and applications, Carbohydr. Res. 343 (2008) 1897–1923.

[3] G.C. Completo, T.L. Lowary, Synthesis of galactofuranose-containing acceptor substrates for mycobacterial galactofuranosyltransferases, J. Org. Chem. 73 (2008) 4513–4525.

[4] M.R. Richards, T.L. Lowary, Chemistry and biology of galactofuranose-containing polysaccharides, ChemBioChem. 10 (2009) 1920–1938.

[5] C. Heiss, M.L. Skowyra, H. Liu, J.S. Klutts, Z. Wang, M. Williams, D. Srikanta, S.M. Beverley, P. Azadi, T.L. Doering, Unusual galactofuranose modification of a capsule polysaccharide in the pathogenic yeast Cryptococcus neoformans, J. Biol. Chem. 288 (2013) 10994–11003.

[6] J.P. Latge, Galactofuranose containing molecules in Aspergillus fumigatus, Med. Mycol. 47 (2009) 104–109.

[7] B. Tefsen, A.F.J. Ram, I. Van Die, F.H. Routier, Galactofuranose in eukaryotes: Aspects of biosynthesis and functional impact, Glycobiology. 22 (2012) 456–469.

[8] W. Morelle, M. Bernard, J.P. Debeaupuis, M. Buitrago, M. Tabouret, J.P. Latgé, Galactomannoproteins of Aspergillus fumigatus, Eukaryot. Cell. 4 (2005) 1308–1316.

[9] P. Glaziou, K. Floyd, M.C. Raviglione, Global Epidemiology of Tuberculosis, Semin. Respir. Crit. Care Med. 39 (2018) 271–285.

[10] F. Gsaller, P. Hortschansky, T. Furukawa, P.D. Carr, B. Rash, J. Capilla, C. Müller, F. Bracher, P. Bowyer, H. Haas, A.A. Brakhage, M.J. Bromley, Sterol Biosynthesis and Azole Tolerance Is Governed by the Opposing Actions of SrbA and the CCAAT Binding Complex, PLoS Pathog. 12 (2016)e1005775.

[11] X. Xue, R.B. Zheng, A. Koizumi, L. Han, J.S. Klassen, T.L. Lowary, Synthetic polyprenol-pyrophosphate linked oligosaccharides are efficient substrates for mycobacterial galactan biosynthetic enzymes, Org. Biomol. Chem. 16 (2018) 1939–1957.

[12] S.K. Angala, J.M. Belardinelli, E. Huc-Claustre, W.H. Wheat, M. Jackson, The cell envelope glycoconjugates of Mycobacterium tuberculosis, Crit. Rev. Biochem. Mol. Biol. 49 (2014) 361–399.

[13] M. Jankute, S. Grover, H.L. Birch, G.S. Besra, Genetics of Mycobacterial Arabinogalactan and Lipoarabinomannan Assembly, Microbiol. Spectr. 2 (2014).

[14] F. Pan, M. Jackson, Y. Ma, M. McNeil, Cell wall core galactofuran synthesis is essential for growth of mycobacteria, J. Bacteriol. 183 (2001) 3991–3998.

[15] J. Engel, P.S. Schmalhorst, F.H. Routier, Biosynthesis of the fungal cell wall polysaccharide galactomannan requires intraluminal GDP-mannose, J. Biol. Chem. 287 (2012) 44418–44424.

[16] A.L. Matveev, V.B. Krylov, L.A. Emelyanova, A.S. Solovev, Y.A. Khlusevich, I.K. Baykov, T. Fontaine, J.P. Latgé, N. V. Tikunova, N.E. Nifantiev, Novel mouse monoclonal antibodies specifically recognize Aspergillus fumigatus galactomannan, PLoS One. 13 (2018) e0193938 .

[17] T. Fontaine, C. Simenel, G. Dubreucq, O. Adam, M. Delepierre, J. Lemoine, C.E. Vorgias, M. Diaquin, J.P. Latgé, Molecular organization of the alkali-

insoluble fraction of Aspergillus fumigatus cell wall, J. Biol. Chem. 275 (2000) 27594–27607.

[18] J.P. Latge, H. Kobayashi, J.P. Debeaupuis, M. Diaquin, J. Sarfati, J.M. Wieruszeski, E. Parra, J.P. Bouchara, B. Fournet, Chemical and immunological characterization of the extracellular galactomannan of Aspergillus fumigatus, Infect. Immun. 62 (1994) 5424–5433.

[19] J. D'Haese, K. Theunissen, E. Vermeulen, H. Schoemans, G. De Vlieger, L. Lammertijn, P. Meersseman, W. Meersseman, K. Lagrou, J. Maertens, Detection of galactomannan in bronchoalveolar lavage fluid samples of patients at risk for invasive pulmonary aspergillosis: Analytical and clinical validity, J. Clin. Microbiol. 50 (2012) 1258–1263.

[20] L. Muszkieta, T. Fontaine, R. Beau, I. Mouyna, M.S. Vogt, J. Trow, B.P. Cormack, L.-O. Essen, G. Jouvion, J.-P. Latgé, The Glycosylphosphatidylinositol-Anchored DFG Family Is Essential for the Insertion of Galactomannan into the β-(1,3)-Glucan–Chitin Core of the Cell Wall of Aspergillus fumigatus , MSphere. 4 (2019) e00397-19.

[21] M. Linnerborg, A. Weintraub, G. Widmalm, Structural studies of the O-antigen polysaccharide from the enteroinvasive Escherichia coli O164 cross-reacting with Shigella dysenteriae type 3, Eur. J. Biochem. 266 (1999) 460–466.

[22] B.R. Clarke, D. Bronner, W.J. Keenleyside, W.B. Severn, J.C. Richards, C. Whitfield, Role of Rfe and RfbF in the initiation of biosynthesis of D-galactan I, the lipopolysaccharide O antigen from Klebsiella pneumoniae serotype O1, J. Bacteriol. 177 (1995) 5411–5418.

[23] M. Seničar, P. Lafite, S. V. Eliseeva, S. Petoud, L. Landemarre, R. Daniellou, Galactofuranose-related enzymes: Challenges and hopes, Int. J. Mol. Sci. 21

(2020) 3465.

[24] L. Wang, U.K. Aryal, Z. Dai, A.C. Mason, M.E. Monroe, Z.X. Tian, J.Y. Zhou, D. Su, K.K. Weitz, T. Liu, D.G. Camp, R.D. Smith, S.E. Baker, W.J. Qian, Mapping N-linked glycosylation sites in the secretome and whole cells of aspergillus niger using hydrazide chemistry and mass spectrometry, J. Proteome Res. 11 (2012) 143–156.

[25] M. Goto, Protein O-glycosylation in fungi: Diverse structures and multiple functions, Biosci. Biotechnol. Biochem. 71 (2007) 1415–1427.

[26] S. Boecker, S. Grätz, D. Kerwat, L. Adam, D. Schirmer, L. Richter, T. Schütze, D. Petras, R.D. Süssmuth, V. Meyer, Aspergillus niger is a superior expression host for the production of bioactive fungal cyclodepsipeptides, Fungal Biol. Biotechnol. 5 (2018) 4.

[27] T.C. Cairns, C. Nai, V. Meyer, How a fungus shapes biotechnology: 100 years of Aspergillus niger research, Fungal Biol. Biotechnol. 5 (2018) 13.

[28] E. Matsunaga, Y. Higuchi, K. Mori, N. Yairo, T. Oka, S. Shinozuka, K. Tashiro, M. Izumi, S. Kuhara, K. Takegawa, Identification and characterization of a novel galactofuranose-specific β-D-galactofuranosidase from Streptomyces species, PLoS One. 10 (2015) e0137230.

[29] E. Matsunaga, Y. Higuchi, K. Mori, N. Yairo, S. Toyota, T. Oka, K. Tashiro, K. Takegawa, Characterization of a PA14 domain-containing galactofuranose-specific β-D-galactofuranosidase from Streptomyces sp., Biosci. Biotechnol. Biochem. 81 (2017) 1314–1319.

[30] L. Shen, A. Viljoen, S. Villaume, M. Joe, I. Halloum, L. Chêne, A. Méry, E. Fabre, K. Takegawa, T.L. Lowary, S.P. Vincent, L. Kremer, Y. Guérardel, C. Mariller, The endogenous galactofuranosidase GlfH1 hydrolyzes

mycobacterial arabinogalactan, J. Biol. Chem. 295 (2020) 5110–5123.

[31] G.L.F. Wallis, R.J. Swift, F.W. Hemming, A.P.J. Trinci, J.F. Peberdy, Glucoamylase overexpression and secretion in Aspergillus niger: Analysis of glycosylation, Biochim. Biophys. Acta - Gen. Subj. 1472 (1999) 576–586.

[32] G.L.F. Wallis, F.W. Hemming, J.F. Peberdy, An extracellular β-galactofuranosidase from Aspergillus niger and its use as a tool for glycoconjugate analysis, Biochim. Biophys. Acta - Gen. Subj. 1525 (2001) 19–28.

[33] B. Tefsen, E.L. Lagendijk, J. Park, M. Akeroyd, D. Schachtschabel, R. Winkler, I. Van Die, A.F.J. Ram, Fungalα-arabinofuranosidases of glycosyl hydrolase families 51 and 54 show a dual arabinofuranosyl- and galactofuranosyl-hydrolyzing activity, Biol. Chem. 393 (2012) 767–775.

[34] E. Matsunaga, Y. Tanaka, S. Toyota, H. Yamada, T. Oka, Y. Higuchi, K. Takegawa, Identification and characterization of β-D-galactofuranosidases from Aspergillus nidulans and Aspergillus fumigatus, J. Biosci. Bioeng. 131 (2021) 1–7.

[35] M. Wu, N.P. McNulty, D.A. Rodionov, M.S. Khoroshkin, N.W. Griffin, J. Cheng, P. Latreille, R.A. Kerstetter, N. Terrapon, B. Henrissat, A.L. Osterman, J.I. Gordon, Genetic determinants of in vivo fitness and diet responsiveness in multiple human gut Bacteroides, Science.350 (2015) aac5952.

[36] M.A. Mahowald, F.E. Rey, H. Seedorf, P.J. Turnbaugh, R.S. Fulton, A. Wollam, N. Shah, C. Wang, V. Magrini, R.K. Wilson, B.L. Cantarel, P.M. Coutinho, B. Henrissat, L.W. Crock, A. Russell, N.C. Verberkmoes, R.L. Hettich, J.I. Gordon, Characterizing a model human gut microbiota composed of members of its two dominant bacterial phyla, Proc. Natl. Acad. Sci. U. S. A. 106 (2009) 5859–5864.

[37] W. Helbert, L. Poulet, S. Drouillard, S. Mathieu, M. Loiodice, M. Couturier, V. Lombard, N. Terrapon, J. Turchetto, R. Vincentelli, B. Henrissat, Discovery of novel carbohydrate-active enzymes through the rational exploration of the protein sequences space, Proc. Natl. Acad. Sci. U. S. A. 116 (2019) 6063–6068.

[38] A. Cartmell, J. Muñoz-Muñoz, J.A. Briggs, D.A. Ndeh, E.C. Lowe, A. Baslé, N. Terrapon, K. Stott, T. Heunis, J. Gray, L. Yu, P. Dupree, P.Z. Fernandes, S. Shah, S.J. Williams, A. Labourel, M. Trost, B. Henrissat, H.J. Gilbert, A surface endogalactanase in Bacteroides thetaiotaomicron confers keystone status for arabinogalactan degradation, Nat. Microbiol. 3 (2018) 1314–1326.

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

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

[41] J.C. for S.G. (JCSG), PDB ID: 3QZ4 Crystal structure of an Endo-1,4-beta-xylanase D (BT_3675) from Bacteroides thetaiotaomicron VPI-5482 at 1.74 A resolution, (2011).

[42] D. Nurizzo, J.P. Turkenburg, S.J. Charnock, S.M. Roberts, E.J. Dodson, V.A. McKie, E.J. Taylor, H.J. Gilbert, G.J. Davies, Cellvibrio japonicus α-l-arabinanase 43a has a novel five-blade β-propeller fold, Nat. Struct. Biol. 9 (2002) 665–668.

[43] G.P. Borin, C.C. Sanchez, E.S. de Santana, G.K. Zanini, R.A.C. dos Santos, A. de Oliveira Pontes, A.T. de Souza, R.M.M.T.S. Dal'Mas, D.M. Riaño-Pachón, G.H. Goldman, J.V. de C. Oliveira, Comparative transcriptome analysis reveals different strategies for degradation of steam-exploded sugarcane bagasse by Aspergillus niger and Trichoderma reesei, BMC Genomics. 18 (2017) 501.

[44] Y. Chihara, Y. Tanaka, M. Izumi, D. Hagiwara, A. Watanabe, K. Takegawa, K. Kamei, N. Shibata, K. Ohta, T. Oka, Biosynthesis of β-(1,5)-Galactofuranosyl Chains of Fungal- Type and, MSphere. 5 (2020) 1–19.

[45] T. Takayanagi, A. Kimura, S. Chiba, K. Ajisaka, Novel structures of N-linked high-mannose type oligosaccharides containing alpha-D-galactofuranosyl linkages in Aspergillus niger alpha-D-glucosidase, Carbohydr. Res. 256 (1994) 149–158.

[46] A. Miyanaga, T. Koseki, Y. Miwa, Y. Mese, S. Nakamura, A. Kuno, J. Hirabayashi, H. Matsuzawa, T. Wakagi, H. Shoun, S. Fushinobu, The family 42 carbohydrate-binding module of family 54 α-L- arabinofuranosidase specifically binds the arabinofuranose side chain of hemicellulose, Biochem. J. 399 (2006) 503–511.

[47] C. Brüx, A. Ben-David, D. Shallom-Shezifi, M. Leon, K. Niefind, G. Shoham, Y. Shoham, D. Schomburg, The Structure of an Inverting GH43 β-Xylosidase from Geobacillus stearothermophilus with its Substrate Reveals the Role of the Three Catalytic Residues, J. Mol. Biol. 359 (2006) 97–109.

[48] E. Vandermarliere, T.M. Bourgois, M.D. Winn, S. Van Campenhout, G. Volckaert, J.A. Delcour, S. V. Strelkov, A. Rabijns, C.M. Courtin, Structural analysis of a glycoside hydrolase family 43 arabinoxylan arabinofuranohydrolase in complex with xylotetraose reveals a different binding mechanism compared with other members of the same family, Biochem. J. 418 (2009) 39–47.

[49] E. Matsunaga, Y. Higuchi, K. Mori, N. Yairo, S. Toyota, T. Oka, K. Tashiro, K. Takegawa, Characterization of a PA14 domain-containing galactofuranose-specific β- <scp>d</scp> -galactofuranosidase from *Streptomyces* sp., Biosci. Biotechnol. Biochem. 81 (2017) 1314–1319.

[50] M. Seničar, L. Legentil, V. Ferrières, S. V. Eliseeva, S. Petoud, K. Takegawa, P. Lafite, R. Daniellou, Galactofuranosidase from JHA 19 Streptomyces sp.: subcloning and biochemical characterization, Carbohydr. Res. 480 (2019) 35–41.

[51] E. Matsunaga, Y. Higuchi, K. Mori, N. Yairo, S. Toyota, T. Oka, K. Tashiro, K. Takegawa, Characterization of a PA14 domain-containing galactofuranose-specific β-D-galactofuranosidase from Streptomyces sp., Biosci. Biotechnol. Biochem. 81 (2017) 1314–1319.

[52] P. Schäpe, M.J. Kwon, B. Baumann, B. Gutschmann, S. Jung, S. Lenz, B. Nitsche, N. Paege, T. Schütze, T.C. Cairns, V. Meyer, Updating genome annotation for the microbial cell factory Aspergillus Niger using gene co-expression networks, Nucleic Acids Res. 47 (2019) 559–569.

[53] F. Supek, M. Bošnjak, N. Škunca, T. Šmuc, Revigo summarizes and visualizes long lists of gene ontology terms, PLoS One. 6 (2011) e21800.

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

[55] B.S. Gruben, M.R. Mäkelä, J.E. Kowalczyk, M. Zhou, I. Benoit-Gelber, R.P. De

Vries, Expression-based clustering of CAZyme-encoding genes of Aspergillus niger, BMC Genomics. 18 (2017) 900.

[56] P.A. Vankuyk, J.A.E. Benen, H.A.B. Wösten, J. Visser, R.P. De Vries, A broader role for AmyR in Aspergillus niger: Regulation of the utilisation of d-glucose or d-galactose containing oligo- and polysaccharides, Appl. Microbiol. Biotechnol. 93 (2012) 285–293.

[57] H. Zhang, S. Wang, X. xiang Zhang, W. Ji, F. Song, Y. Zhao, J. Li, The amyR-deletion strain of Aspergillus niger CICC2462 is a suitable host strain to express secreted protein with a low background, Microb. Cell Fact. 15 (2016) 68.

[58] J.E. Kowalczyk, R.J.M. Lubbers, M. Peng, E. Battaglia, J. Visser, R.P. De Vries, Combinatorial control of gene expression in Aspergillus Niger grown on sugar beet pectin, Sci. Rep. 7 (2017) 12356.

[59] S. El-Gebali, J. Mistry, A. Bateman, S.R. Eddy, A. Luciani, S.C. Potter, M. Qureshi, L.J. Richardson, G.A. Salazar, A. Smart, E.L.L. Sonnhammer, L. Hirsh, L. Paladin, D. Piovesan, S.C.E. Tosatto, R.D. Finn, The Pfam protein families database in 2019, Nucleic Acids Res. 47 (2019) D427–D432.

[60] T.A. Hall, BIOEDIT: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/ NT., Nucleic Acids Symp. Ser. 41 (1999) 95–98.

[61] S. Kumar, G. Stecher, M. Li, C. Knyaz, K. Tamura, MEGA X: Molecular evolutionary genetics analysis across computing platforms, Mol. Biol. Evol. 35 (2018) 1547–1549.

[62] D.T. Jones, W.R. Taylor, J.M. Thornton, The rapid generation of mutation data matrices from protein sequences, Bioinformatics. 8 (1992) 275–282.

[63] J. Park, B. Tefsen, M. Arentshorst, E. Lagendijk, C.A. van den Hondel, I. van Die, A.F. Ram, Identification of the UDP-glucose-4-epimerase required for galactofuranose biosynthesis and galactose metabolism in A. niger, Fungal Biol. Biotechnol. 1 (2014) 6.

Supplementary information

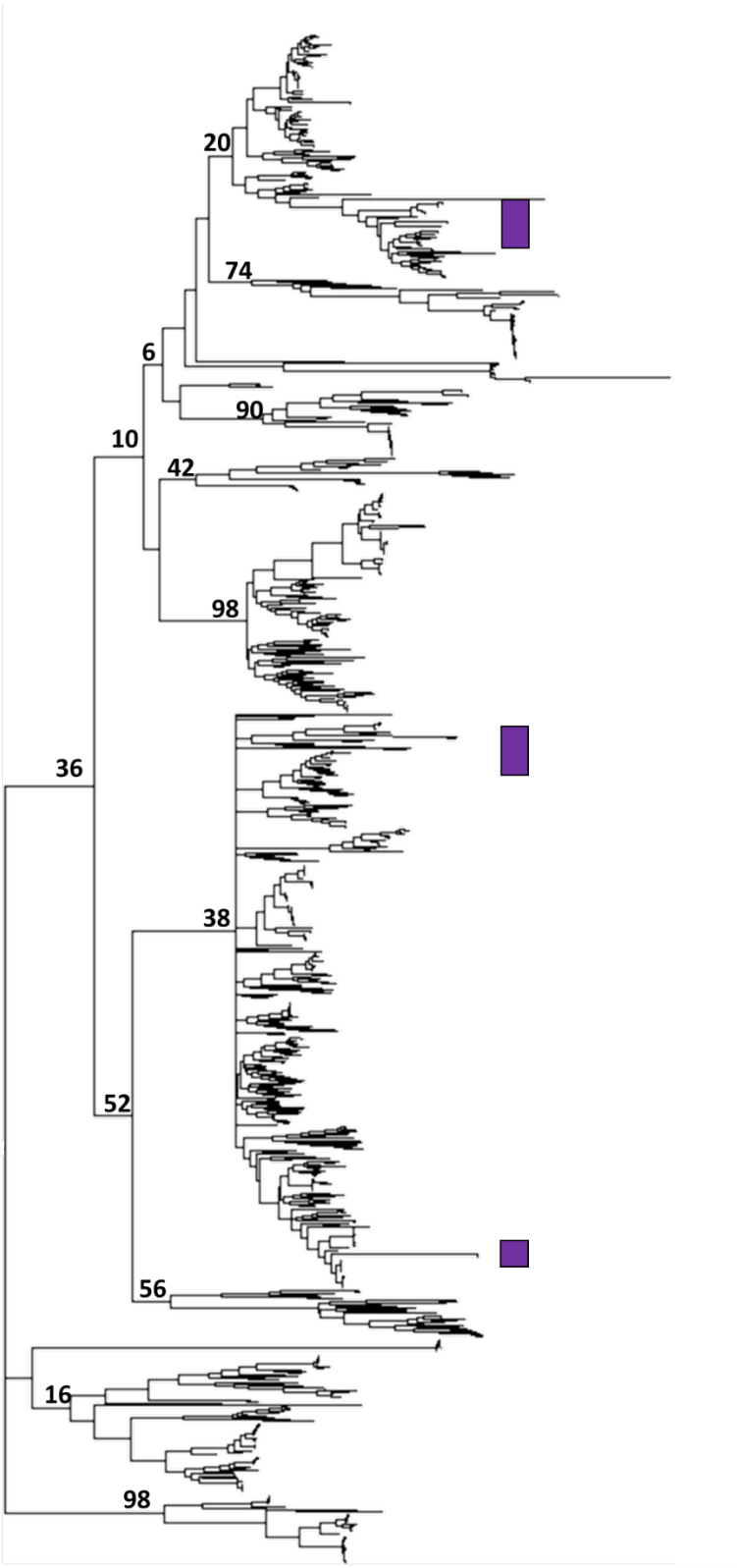
Biochemical characterisation of a glycoside hydrolase family 43 β-D-galactofuranosidase from the fungus *Aspergillus niger*

Gregory S. Bulmer[1], Fang Wei Yuen[1], Naimah Begum[1], Bethan S. Jones[1], Sabine L. Flitsch[1] and Jolanda M. van Munster*[, 1,2]

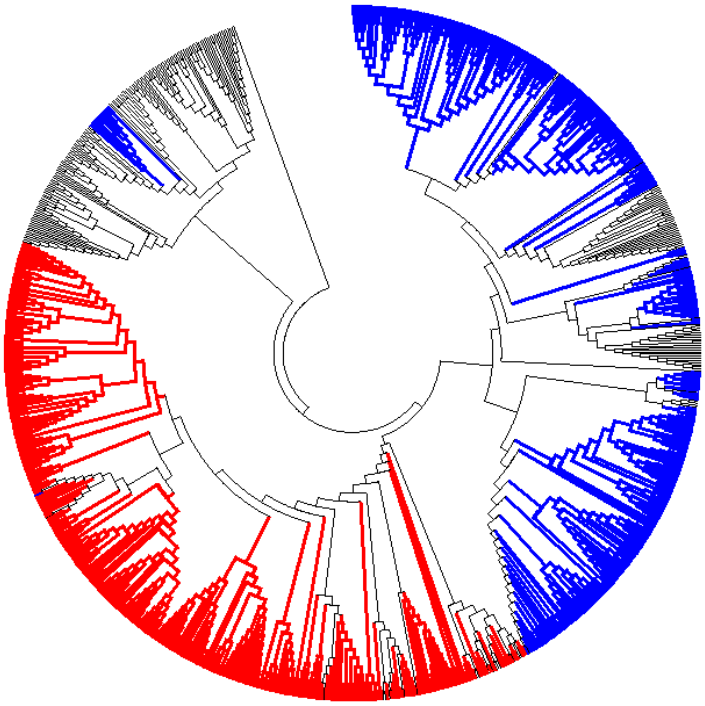
[1] Manchester Institute of Biotechnology (MIB) & School of Natural Sciences, The University of Manchester, 131 Princess Street, Manchester, M1 7DN, United Kingdom

[2] Scotland's Rural College, West Mains Road, King's Buildings, Edinburgh, EH9 3JG, United Kingdom

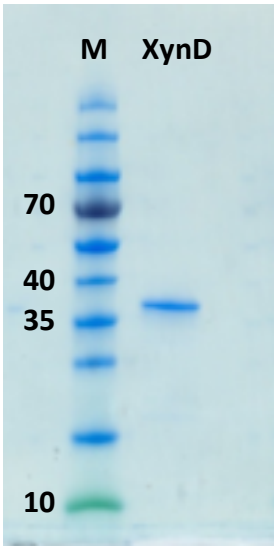
Contents



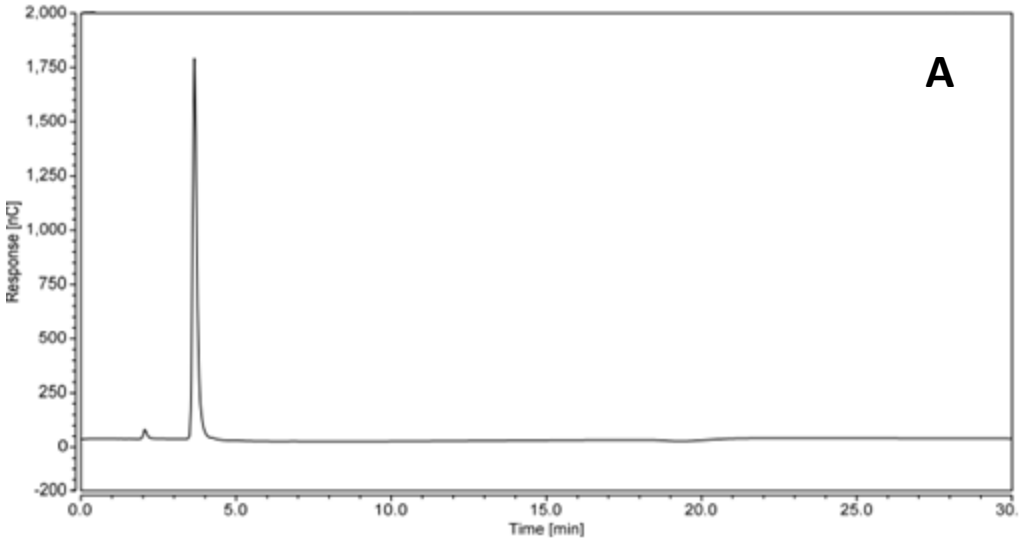
Supplementary figure 1: Phylogenetic tree of GH43 subfamily 34, annotated with boostrap values and locations of fungal sequences highlighted (purple).

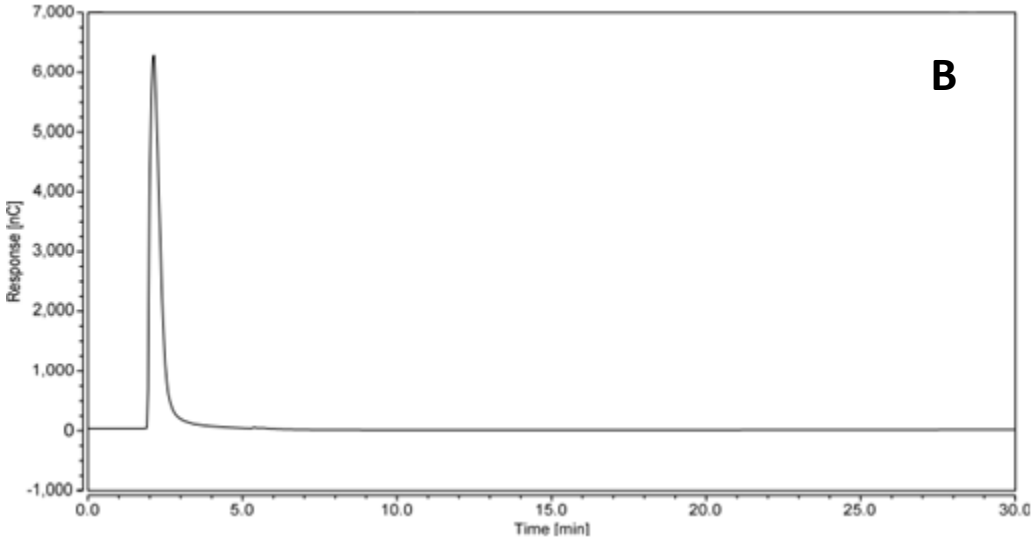


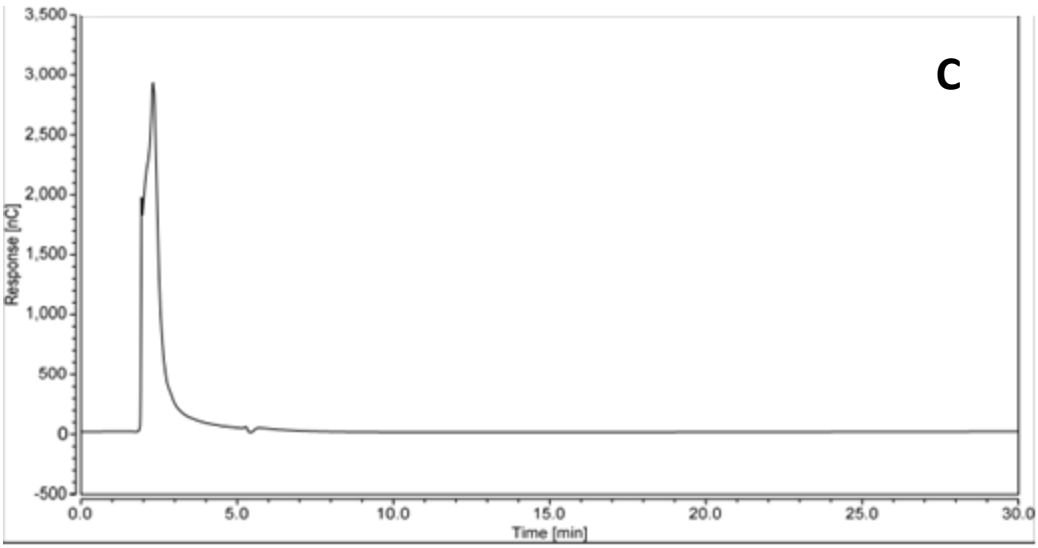
Supplementary figure 2: Phylogenetic tree of GH43 subfamily 34 with amino acid identity at position 219 highlighted: W (red), T (blue) or other (grey)

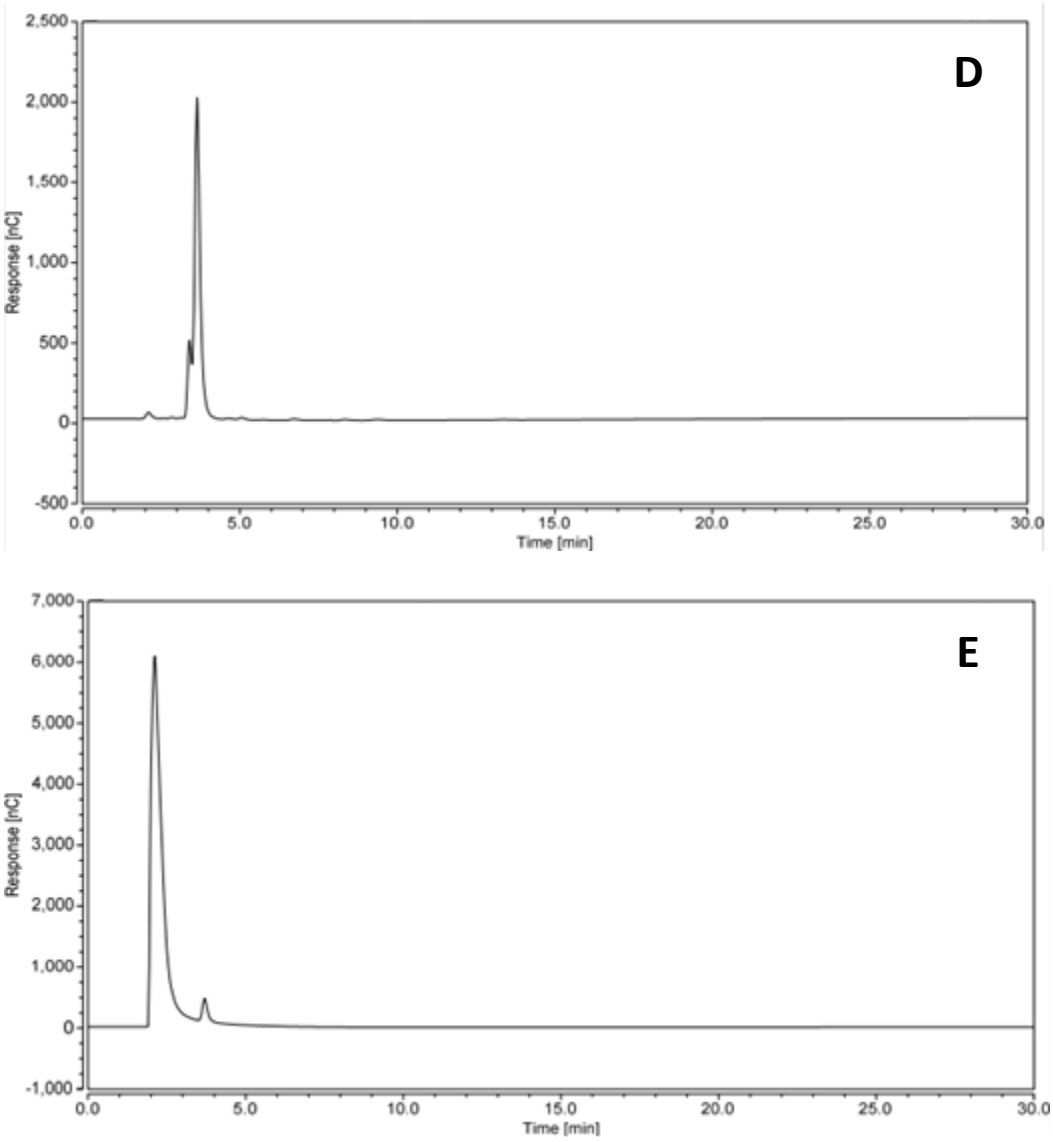


Supplementary figure 3: SDS-PAGE analysis of heterologously expressed XynDafter His-tag purification. M: Molecular weight Marker.

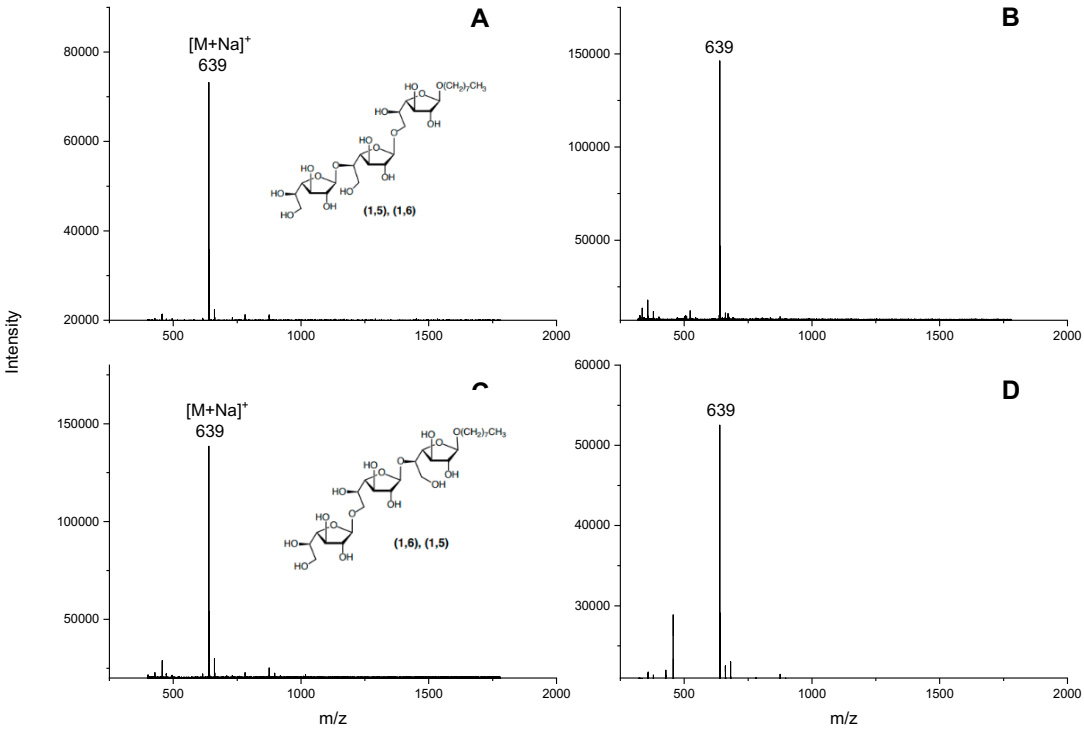

A
Response [nC]
Time [min]


B
Response [nC]
Time [min]

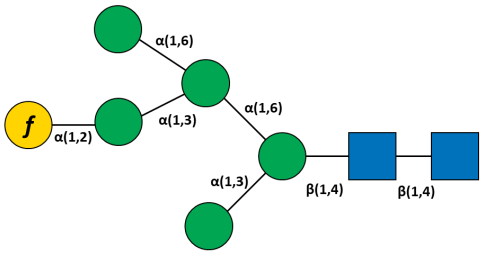

C
Response [nC]
Time [min]



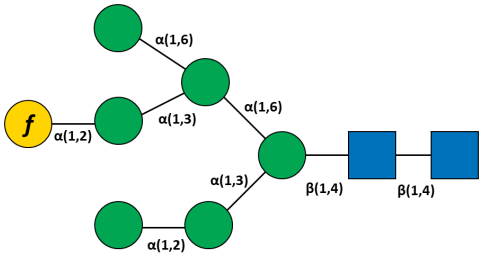
Supplementary figure 4: High-Performance Anion Exchange Chromatography (HPAEC) analysis of XynD activity on A. niger N402 galactomannan. **A** 1 mM Galactose **B** *A. niger* N402 galactomannan after incubation with XynD **C***A. niger* N402 galactomannan incubation negative control (incubation with protein fraction as obtained after purification of an *E. coli* culture expressing an empty pET 21a vector under conditions identical to those used to obtain XynD) **D** Acid-hydrolysed *A. niger* N402 galactomannan **E** *A. niger* N402 galactomannan empty vector incubation spiked with galactose



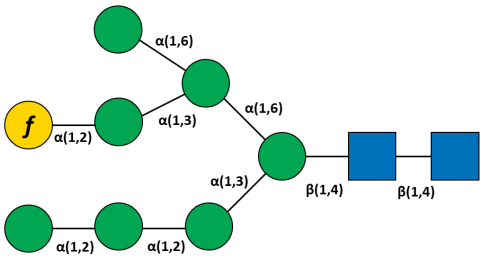
Supplementary figure 5: Assessment of XynD activity on β-1,5/β-1,6 Galf trisaccharides via MALDI-TOF MS. **A** Octyl β-D-Galactofuranosyl-(1,5)-β-D-galactofuranosyl-(1,6)-β-D-galactofuranoside (M = 616) **B** Octyl β-D-Galactofuranosyl-(1,5)-β-D-galactofuranosyl-(1,6)-β-D-galactofuranoside after incubation with XynD **C** Octyl β-D-Galactofuranosyl-(1,6)-β-D-galactofuranosyl-(1,5)-β-D-galactofuranoside (M = 616) **D** Octyl β-D-Galactofuranosyl-(1,6)-β-D-galactofuranosyl-(1,5)-β-D-galactofuranoside after incubation with XynD. Expected mass of 477 if Gal*f* residue removal (not observed).



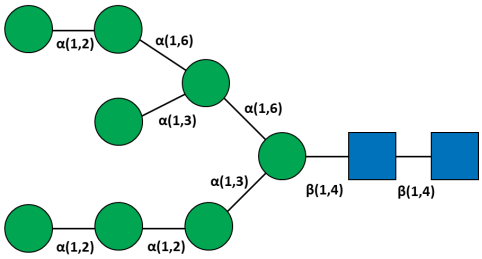
M = 1396
$[M+Na]^+$ = 1419



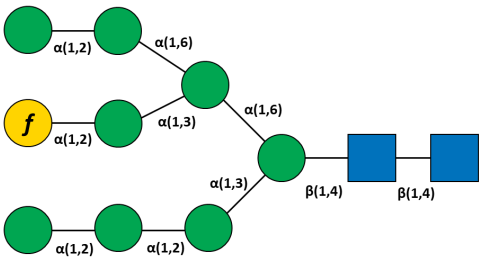
M = 1558
$[M+Na]^+$ = 1581



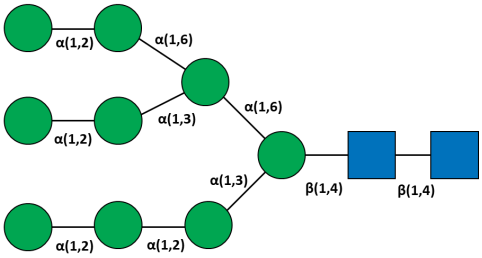
M = 1720
$[M+Na]^+$ = 1743



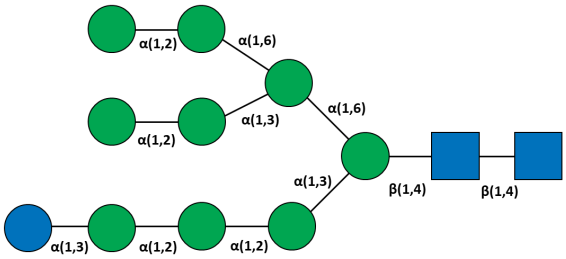
M = 1720
$[M+Na]^+$ = 1743



M = 1882
$[M+Na]^+$ = 1905

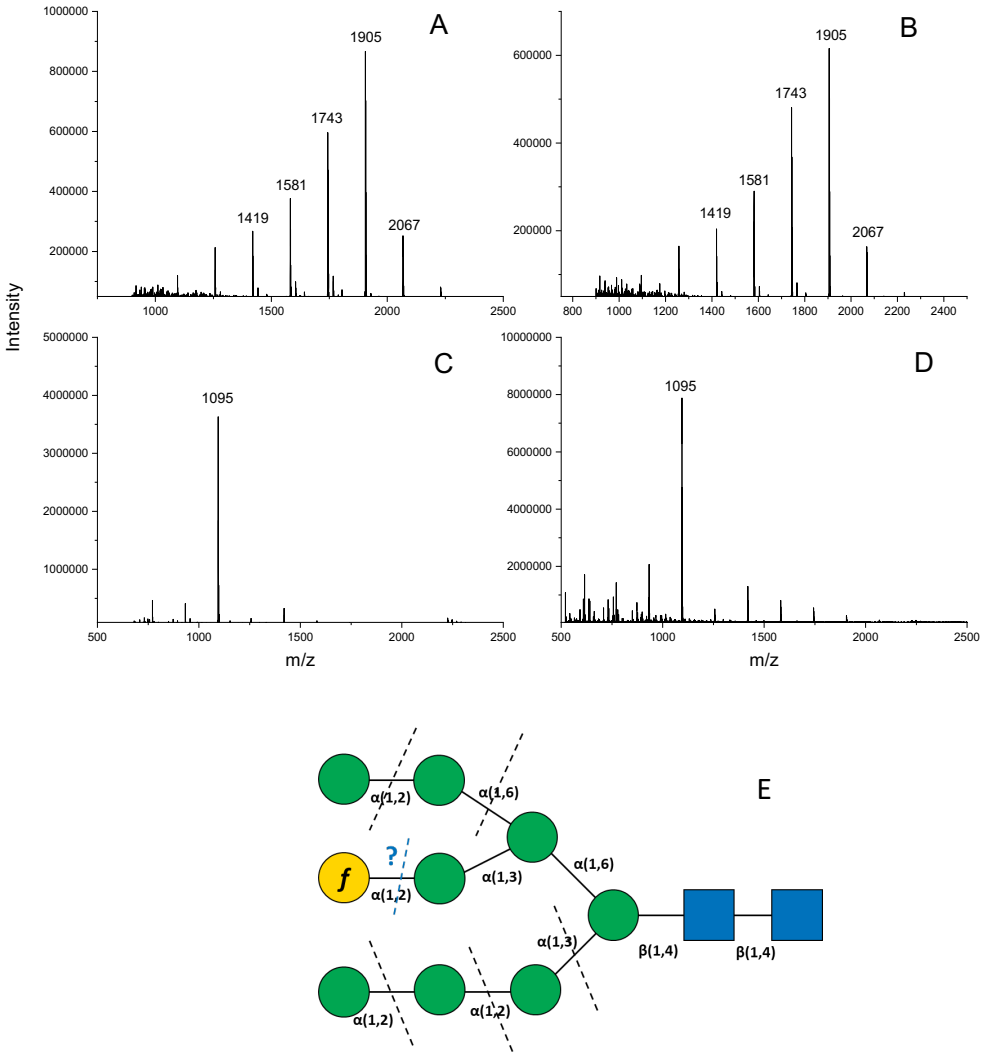


M = 1882
$[M+Na]^+$ = 1905

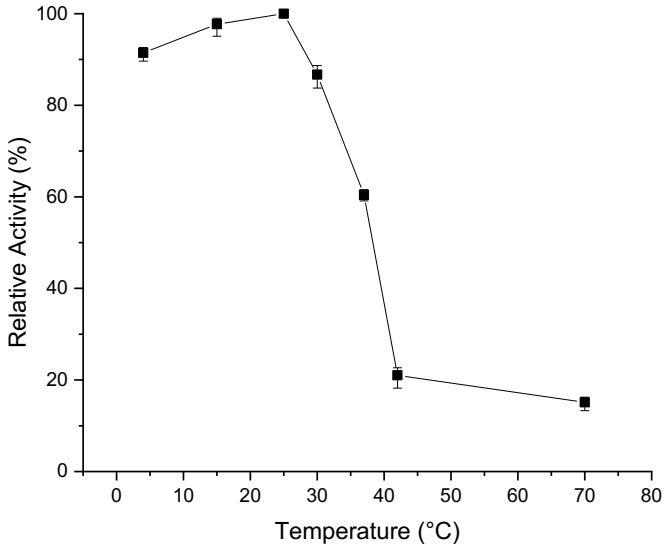


M = 2044
$[M+Na]^+$ = 2067

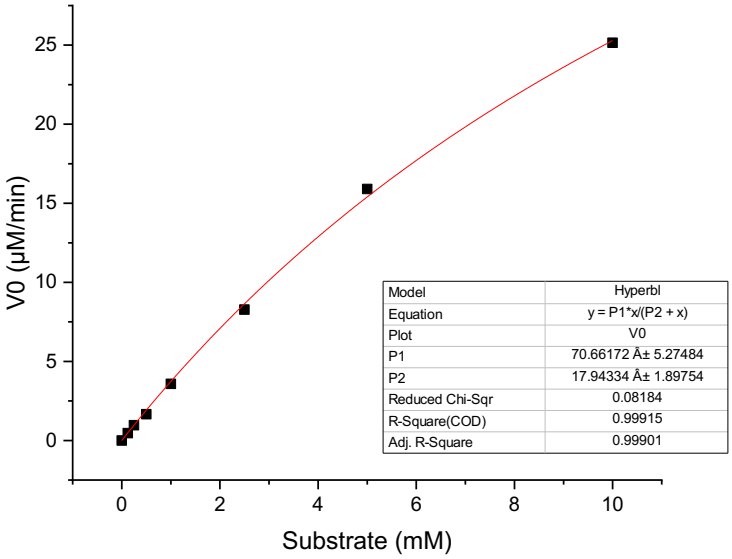
Supplementary figure 6: N-linked glycans produced by incubation of Transglucosidase L "Amano" (Amano) with PNGase F, as described in [1].



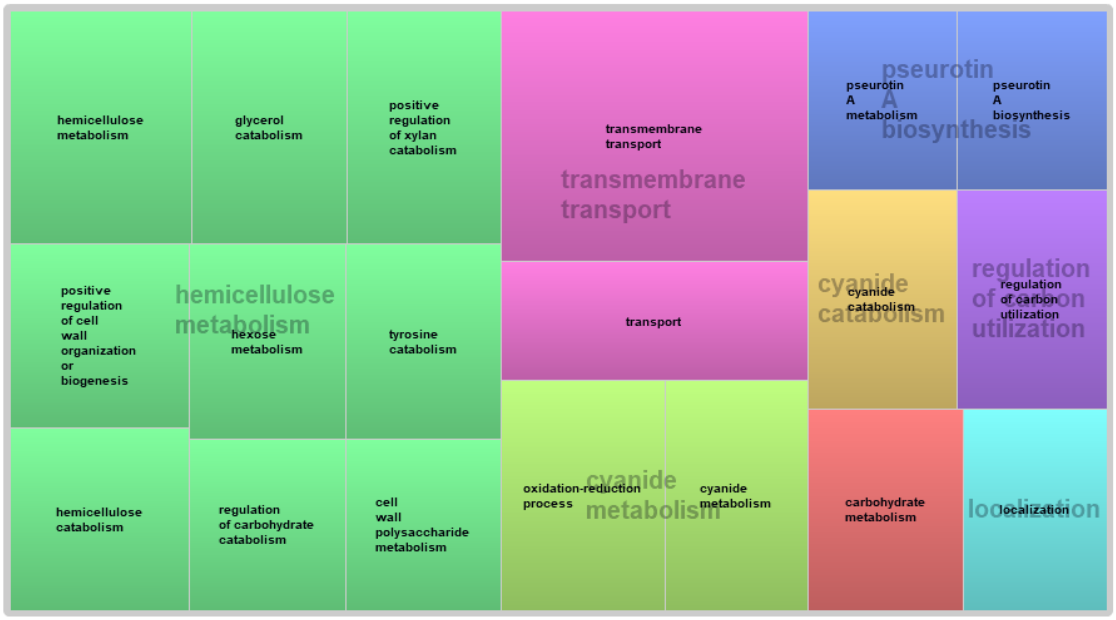
Supplementary figure 7: MALDI-TOF MS of α-1,2 containing-N-linked glycans incubated with XynD. **A** Purified Amano Transglucosidase N-linked glycans **B** Purified Amano Transglucosidase N-linked glycans XynD incubation. No shift of -162, consistent with terminal Gal*f* removal was observed. **C** Purified Amano Transglucosidase N-linked glycansco-incubated withmannosidase and XynD **D** Purified Amano Transglucosidase N-linked glycansincubated with mannosidase**. E** Diagram demonstrating enzymatic activity against N-linked glycan, black dashed lines representing exo-mannosidase activity, blue dashed line and question mark representing potential Gal*f*ase activity. Lack of Gal*f*ase activity blocks mannosidase activity against Gal*f*-terminating chain, leaving the hexasaccharide detected in panels C and D. Conversely, Gal*f*ase activity would enable degradation of the glycan to the trisaccharide Man-GlcNAc-GlcNAc core ($[M+Na]^+$= 609).



Supplementary figure 8: Temperature stability of XynD. After 1 h incubations at indicated temperatures, residual activity of XynD was measured under standard reaction conditions (pH 5, 25 °C).



Supplementary figure 9: Michaelis-Menten Kinetics of XynD using pNP-Gal*f* as the substrate.



Supplementary figure 10: Revigo TreeMap of GO terms (biological processes ontology) positively correlating with *xynD* expression

References

(1) Takayanagi, T.; Kimura, A.; Chiba, S.; Ajisaka, K. Novel Structures of N-Linked High-Mannose Type Oligosaccharides Containing Alpha-D-Galactofuranosyl Linkages in Aspergillus Niger Alpha-D-Glucosidase. *Carbohydr. Res.* 1994, *256* (1), 149–158.

Chapter 6: Utilisation of a thermotolerant feruoyl esterase from *Thermobacillus xylanilyticus* that processes pre-biotic ferulated xylooligosaccharides

This chapter is composed of one article in preparation and supporting information.

C. Garbelotti, G. S. Bulmer, R. J. Ward and J. M. van Munster, Utilisation of a thermotolerant feruoyl esterase from *Thermobacillus xylanilyticus* that processes pre-biotic ferulated xylooligosaccharides, manuscript in preparation

Foreword

This chapter describes the biochemical characterisation of the feruoyl esterase Tx-Est1 and its application in the processing of pre-biotic oligosaccharides. Following optimisation for pH and temperature, activity against ferulated xylooligosaccharides was demonstrated. The ability of Tx-Est1 to release ferulic acid from a variety of polymeric natural substrates was then investigated.

Contribution

C. Garbelotti cloned, expressed and biochemically characterised (kinetic studies, CD, DSC) Tx-Est1. G.S. Bulmer performed activity assays on oligosaccharides/ polymeric natural substrates and analysed reaction products via MALDI-TOF MS and HPLC. R. J. Ward provided conceptualisation and supervision. J.M. van Munster produced wheat bran oligosaccharides, conducted NMR analysis and provided conceptualisation and supervision. C. Garbelotti, G. S. Bulmer, R. J. Ward and J. M. van Munster wrote and edited the manuscript.

Utilisation of a thermotolerant feruoyl esterase from *Thermobacillus xylanilyticus* that processes pre-biotic ferulated xylooligosaccharides

Authors

Carolina Garbelotti[1], Gregory S. Bulmer[2], Richard J. Ward[1*] and Jolanda M. van Munster[2,3*]

Affiliations

[1] Departamento de Química, Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, SP CEP 14040-901, Brazil

[2] Manchester Institute of Biotechnology (MIB) & School of Natural Sciences, The University of Manchester, 131 Princess Street, Manchester, M1 7DN, United Kingdom

[3] Scotland's Rural College, West Mains Road, Kings Buildings, Edinburgh, EH9 3JG, United Kingdom

* Corresponding authors; rjward@fmrp.usp.br ; jolanda.van-munster@sruc.ac.uk

Abstract

Ferulic acid has anti-oxidant and antimicrobial properties that are of interest to the food industry and can be released from natural plant fibres using feruloyl esterases. Here we report the biochemical characterization of the feruloyl esterase from *Thermobacillus xylanilyticus* (Tx-Est1). The specific activity of purified recombinant Tx-Est1 with ethyl ferulate was 29.2 ± 2.9 U mg^{-1}, with a K_m 0.09 ± 0.02 mM and a catalytic efficiency (K_{cat}/K_m) 393.7 ± 9.8 s^{-1} mM^{-1}. The catalytic temperature and pH optima were 60 °C and 7.5, which correlated with a melting temperature of 63 and 60 °C measured by CD and DSC, respectively, and Tx-Est1 retains 70% activity after 25 h at 40 °C. MALDI-TOF MS revealed that Tx-Est1 released all ferulic acid decorations from xylooligosaccharides from wheat bran arabinoxylan with DP4 - DP13, and from DP6-8 containing two ferulic acid groups. Furthermore, HPLC demonstrated that ferulic acid release from destarched wheat bran by Tx-EST1 was strongly potentiated by co-incubation with the GH11 xylanase from *T. lanuginosus*. These catalytic and functional properties suggest that Tx-Est1 can be employed for the production of a high-value compound from agricultural waste streams or as

part of the saccharification process of complex, heterogeneous plant polysaccharides.

Introduction

The enzymatic treatment of food for modulation of taste, texture and handing properties during processing is a well-established and safe technology that improves food quality for both human and animal consumption[1]. The exceptional regio- and stereospecificity of enzyme catalysis offers the possibility of precise modification of food components on a large scale, and the food industry makes widespread use of glycosyl hydrolases, lipases, proteases and oxidoreductases in a diverse variety of applications[2]. Microbial enzymes are particularly useful to the food industry due to the ease and scalability of production using renewable resources, low environmental impact and their reduced cost[3].

The hydrolysis of long acyl chains in oils and fats by lipases has been widely exploited to alter the properties of food emulsions[4] and to enhance taste in dairy products[5], however the use of other esterases to specifically process short acyl and other ester-linked food constituents has received less attention. Feruloyl esterase, or ferulic acid esterase (FAE, EC 3.1.1.73), hydrolyses the ester linkage between the ferulic acid (FA, or 4-hydroxy-3-methoxycinnamic acid) and the carbohydrate backbone in branched arabinoxylans and pectins, polysaccharides that are typically found in the cell walls and seed husks of many commelinid plants that are routinely consumed as dietary fibre. Ferulic acid is considered to be a bioactive compound, and due to its phenolic nature has been identified as a biological antioxidant that may offer protection against neurodegenerative diseases, diabetes, thrombosis and cancer[6]. Synergism in intestinal microflora involving FAE producing bacterial strains improve the release of FA from dietary fibre and formation of prebiotic dietary polyphenols [1], and FAE activity has been shown to influence the probiotic effect of FA release by microorganisms [2].

Ferulic acid and many of its derivatives also have antimicrobial properties, which has led to efforts to develop their use as natural food preservatives[7] and as a natural cross-linking agent in edible synthetic biofilms for food packaging, thereby

providing and antimicrobial and antioxidant coating and reducing the use of non-biodegradable plastic films. The crosslinking of FA and derivatives to form adducts such as 5-5'-, 8-5'-, 8-8'-, and 8-O-4'-diferuloyl linkages [3] between arabinoxylan or pectin polysaccharides that limit the enzymatic digestion of the lignocellulose [4]. Treatment of lignocellulose with FAE improves enzymatic saccharification for bioenergy use [5], and enzymatic removal of FA can increase silage digestibility in ruminant feedstocks whilst simultaneously decreasing FA levels to non-toxic levels for rumen microorganisms [6]. In addition, vanillin and related compounds such as vanillic acid and vanillyl alcohol are high-value natural product derivatives of FA which find applications as flavouring agents in foods and beverages [7].

Due to the increasing interest in FA and its uses, FAEs have attracted attention as valuable biotechnological tools, and the recent revisions have expanded the FAE enzyme classification system to include 13 subfamilies taking into consideration a wealth of sequence and functional data [8]. Although fungi have been the main focus for biotechnological applications of FAEs [9], prokaryotes are a diverse and potentially rich source for novel enzymes with catalytic and thermostable properties that are compatible with applications in the food industry. The majority of bacterial FAEs studied to date are derived from *Bacillus spp*. and *Lactobacillus spp*. [10] and present low amino acid similarity with the fungal enzymes. A thermostable FAE, denominated as Tx-Est1, has previously been described from the Gram-positive bacterium *Thermobacillus xylanilyticus* [11]. The Tx-Est1 demonstrates a high optimum catalytic temperature and catalytic profile that may suitable for biotechnological applications and uses methyl ferulate or methyl sinapinate as preferred synthetic substrates[11] and is therefore compatible with Class A FAEs. Here we aimed to assess the potential of this enzyme for application in the processing of complex food-derived carbohydrates with prebiotic properties such as cinnamoyl-polysaccharides and oligosaccharides. We describe a more complete biochemical characterization of the enzyme accompanied by an evaluation of its FAE activity against a synthetic substrate and a range of oligosaccharides derived from wheat bran arabinoxylan.

Materials and methods

Cloning expression and purification of feruloyl esterase Tx-Est1

The coding region from gene *tx-est1* from *Thermobacillus xylanilyticus (*Genbank GU592999.1) was synthesized and cloned into expression vector pET28a using restriction enzymes NdeI and HindIII. The vector was transformed to chemically competent *Escherichia coli* BL21 (DE3) Star and expressed via induction with 0. 5 mM of Isopropyl β-D-1-thiogalactopyranoside (IPTG). Enzyme purification was performed by affinity chromatography with nickel resin (HisLink, Promega), whereby bound protein was washed with purification buffer (TRIS 50 mM, NaCl 100 mM, pH 8.0) supplemented with 5, 10 and 40 mM imidazole, and finally eluted with purification buffer supplemented with 250mM imidazole. Purified enzyme was dialyzed in deionized water to remove imidazol before use.

Enzyme activity assays on model substrate

Activity of Tx-Est1 on ethyl ferulate (Sigma-Aldrich) was assessed in an assay based on spectrophotometric detection of conversion of the ester to ferulic acid [12], incubating 10 µL of 0.05 mg ml^{-1} enzyme at 60 °C with 100 mM ethyl ferulate in McIlvine (citrate-phosphate 0.2 M) pH 7.4 buffer. Kinetics data was calculated using non-linear fitting for Michaelis-Menten (Origin 9).To determine the optimal pH and temperature, the reaction was performed with ethyl ferulate in McIlvine buffer at pH values ranging from 4.0 to 8.0, and temperatures ranging from 40 to 70 °C, respectively. Thermostability data was produced by incubating the enzyme at different temperatures (35 to 60 °C) and measuring its activity against ethyl ferulate in intervals (0 to 1800 minutes or until enzyme had less than 20% of residual activity). Enzyme activity was followed at 350 and 286 nm after 2 min.

Circular dichroism spectroscopy

CD thermal denaturation was performed in Jasco J-815 CD Spectrometer (JASCO Corporation, Easton, MD, USA) equipped with a Peltier temperature control and using a 1 mm path length quartz cell. Spectra were recorded in 4 scans from 260 to 190 nm, with a scanning speed of 100 nm min-1, spectral bandwidth of 2 nm and temperature from 20 to 75 °C in 5 °C intervals, with a 1°C per minute rate. Enzyme

was in a 0.14 mg ml^{-1} solution in water and 0.157 mg ml^{-1} solution in McIlvine (citrate-phosphate 0.2 M) pH 7.4 buffer. Protein unfolding or melting temperatures (T_m) were obtained by fitting the ellipticity values at 222 nm for each temperature to the Boltzmann equation.

Differential scanning calorimetry

Differential scanning calorimetry (DSC) experiments were performed on a Nano-DSC II – Calorimetry Sciences Corporation, CSC (Lindon, Utah, USA). Enzyme was at 1.05 mg ml^{-1} in water, McIlvine (citrate-phosphate 0.2 M) pH 7.4 buffer or in the presence of 20 mM ethyl ferulate in buffer, references were the same solutions without protein. Heating scans were recorded from 10 to 100 °C with a 1.0 °C/min rate. Protein unfolding or melting temperatures (T_m) for each condition were obtained measuring the maximum of the transition peaks and corresponding enthalpies were calculated by integrating the area under the transition peaks.

Processing of sugar cane

Sugar cane (SP-3280) was processed as described previously [13]. Briefly, sugar cane culms were harvested and dried at 60 °C for 48 h and milled to 20 mesh. The powder was subjected to six extractions with 50 ml g^{-1} of biomass of ethanol 80 % (v/v) at 80 °C for 20 min and the alcohol insoluble residue (AIR) was washed with deionized water and dried at 60 °C for 24 h. AIR fraction was treated twice with 50 ml g^{-1} ammonium oxalate 0.5 M (pH 7.0) at 80 °C for 3 h each with stirring. The oxalate-extracted solid was recovered by filtration and extracted with 40 ml g^{-1} of 3 % (m/v) sodium chlorite in 0.3 % (v/v) acetic acid at room temperature for 1 h. The chlorite-extracted cell wall was then filtrated from the solution, thoroughly washed with deionized water, and dried at 60 °C for 24h.

Production of oligosaccharide enzyme substrates

Wheat bran (Holland & Barrett) was ground to a powder with pestle and mortar. The powder was destarched via incubation of 39 g bran in 500 ml 50 mM phosphate buffer pH 6.5 with 180 µl (3.9 mg) α-amylase from *Bacillus licheniformis* (Termamyl 120, with 1124 units/mg protein, Sigma Aldrich A3403)(10 µg amylase/g bran) for 30 min at 80 °C, after which no starch was detectable by incubation with lugol

iodine solution. The suspension was centrifuged for 10 min at 4000*g*, washed with an equal volume of water and dried ON at 50 °C, to yield 22.4 g of destarched material.

Ten grams of destarched bran was incubated with 0.5 g Driselase (Sigma Aldrich D9515) in 1 l water at 35 °C at 100 rpm. Similarly, 10 g was incubated with 2 g xylanase from *Thermomyces lanuginosus* (Sigma Aldrich X2753) in 1 l 20 mM phosphate buffer pH 7, 35 °C at 100 rpm. This xylanase is reported [14] to release a pentasaccharide end-product which has a α-1,3-linked arabinose decoration at the pre-terminal non-reducing end xylose (β-D-Xyl*p*-(1–4)-[α-L-Ara*f*-(1–3)]-β-D-Xyl*p*-(1–4)-β-D-Xyl*p*-(1–4)-β-D-Xyl*p*), consistent with reported GH11 ability to release oligosaccharides with decorations on the pre-terminal non-reducing-end xylose[15] but not, as with GH10 enzymes, on the terminal non-reducing xylose. After incubation for 24 h, incubations were heated to 100 °C to inactivate enzymes, and the supernatant was collected by centrifugation as above. Bran residue recovered from Driselase and Xylanase treatment was 6.07 and 6.76 g respectively, indicating solubilisation of 39 % and 32 % of start material.

Oligosaccharides with aromatic residues were enriched from the supernatant via capture on ~ 10 ml Amberlite XAD4 resin, washed with 50 ml water, followed by elution with 50 % ethanol and 90 % ethanol. Driselase-derived oligosaccharides were obtained as two pools of 233 mg and 166 mg, and xylanase-derived oligosaccharides as pool of 234 mg. Obtained oligosaccharide pools were further separated on a Bio-Gel P2, fine (Bio-Rad) column (1.5 x 75 cm) at room temperature, ran in water at 0.5 ml min^{-1}. Fractions of 2-5 ml volume were collected, in which oligosaccharide presence was verified by MALDI-TOF MS. Fractions from identical enzyme treatments and with similar mass-profiles were pooled, obtained fractions were freeze dried and analysed by MALDI-TOF MS and NMR.

NMR of oligosaccharides

Oligosaccharide fractions were dissolved in 650 µl deuterium oxide and analysed by NMR. A Bruker Advance 400 NMR spectrometer, operating at at 400.13 MHz, 298.1 K, was used to record ^{1}H and ^{13}C NMR spectra. 1H NMR spectra were recorded with 30-degree pulses, a data acquisition time of 4.09 s and a relaxation delay of 1 s, with spectral width of 20.03 ppm. Data were Fourier transformed, baseline corrected and integrated using Bruker TOPSPIN software. Spectra were approximately referenced using the deuterium lock, followed by subsequent accurate referencing in the ^{1}H-dimension by setting the reducing end α-xylose resonance to 5.184 ppm (generating chemical shifts equivalent to those referenced relative to internal acetone (2.225 ppm)) [16,17]. Data were Fourier transformed, baseline corrected and integrated using Bruker TOPSPIN software. 2D 1H-13C-HSQC (Heteronuclear Single Quantum Coherence Spectroscopy) NMR experiments were performed using the Bruker pulse program hsqcedetgpsisp2.3, which enables multiplicity edited HSQC using echo/antiecho detection and gradient pulses. Spectral assignment was based on matching 1D and 2D data to reports in literature for esterified and non-esterified arabinoxylan oligosaccharides [17–21] .

Enzyme activity assays on oligosaccharides and polymeric natural substrates

Oligosaccharides, at 0.3 mg ml^{-1}, were incubated with 0.1 mg ml^{-1} Tx-Est1in 50 mM citrate-phosphate buffer pH 6.5 at 60 $^{\circ}C$ for 30 min. Substrates and reaction products were analysed via MALDI-TOF Mass Spectrometry.

Polymeric substrates were incubated at 1% w/v with 0.1 mg ml^{-1} Tx-Est1, 50 mM citrate-phosphate buffer pH 6.5 at 60 $^{\circ}C$ overnight. Polymers were as follows: chlorite extracted sugar cane, ethanol extracted sugar cane, wheat bran, wheat straw, acid de-branched arabinoxylan (Wheat Flour; 26% arabinose, P-ADWAX22, Megazyme), Arabinoxylan (Wheat Flour; Medium Viscosity, Megazyme,) and Arabinoxylan (Wheat Flour; Insoluble, P-WAXYI, Megazyme). Reactions co-/incubated with *Thermomyces lanuginosus* endo-xylanase (X2753, Sigma Aldrich) contained a final concentration of 8 mg ml^{-1}. Reactions were then filtered and analysed by HPLC.

MALDI-TOF Mass Spectrometry

Reaction mixtures (0.5 µl) were spotted on target plates, then mixed with super-DHB (0.5 µl, 15 mg ml^{-1} in 50 % acetonitrile + 0.1 % TFA) and dried in ambient temperature. The product was analysed in positive mode on a Bruker Ultraflex or Autoflex IIIMALDI-TOF instrument using the following MS parameters: laser intensity 40 %, matrix suppression <300Da, reflector intensity 3.2x, detection range 400 to 1900 m/z.

HPLC of enzyme reaction products

Reverse Phase HPLC was carried out on an Agilent 1260 Infinity II Series system equipped with a G1379A degasser and G1312A binary pump, G1316A temperature-controlled compartment and a diode array detector. A Kinetex® 5µm, C18 100 Å, 50 x 2.1 mm column was used as a stationary phase. Mobile phase A: Water + 0.1 % trifluoroacetic acid. Mobile phase B: Acetonitrile + 0.1 % trifluoroacetic acid. Flow rate: 0.6 mL/min. Gradient: 0-5 min isocratic 20% B, 5-15 min linear gradient 20-80% B, 15-25 min isocratic 20% B.

Results and discussion

Optimal reaction conditions and kinetic characterisation of Tx-Est1

The Tx-Est1 from *Thermobacillus xylanilyticus* is a feruloyl esterase [11] that is amenable to facile heterologous expression. We produced Tx-Est1 as N-terminal 6xHis fusion protein after expressing its encoding gene via a pET28a-based construct in *E. coli*. After purification of Tx-Est1 to homogeneity via nickel affinity chromatography, the enzyme was obtained with a yield of 5 mg per litre culture. Activity of the Tx-Est1 enzyme towards the synthetic substrate ethyl ferulate exhibited specific activity of 29.2 ± 2.9 U mg^{-1}, a K_m 0.09 ± 0.02 mM (Fig.1a) and catalytic efficiency (K_{cat}/K_m) 393.7 ±9.8 s^{-1} mM^{-1}. The affinity of the enzyme for ethyl ferulate is similar to the affinity reported [11] for methyl ferulate (0.11 mM), methyl-sinapinate (0.10 mM) and higher than that reported for methyl-*p*-coumarate (1.28 mM). The enzyme had noticeably higher catalytic efficiency for ethyl ferulate, with catalytic efficiency increased 3, 2 and 43-fold, respectively, from

reported values. In combination with a previously reported capacity to release diferulate (mainly 5-5' diferulate and 8-O-4 diferulate) from wheat straw [11], the catalytic properties identified here place the Tx-Est1 in the Class A FAEs following the functional ABCD classification system [22], which is still the preferred nomenclature for bacterial FAEs since the bacterial enzymes are currently not included in the sequence based sub-family classification system.

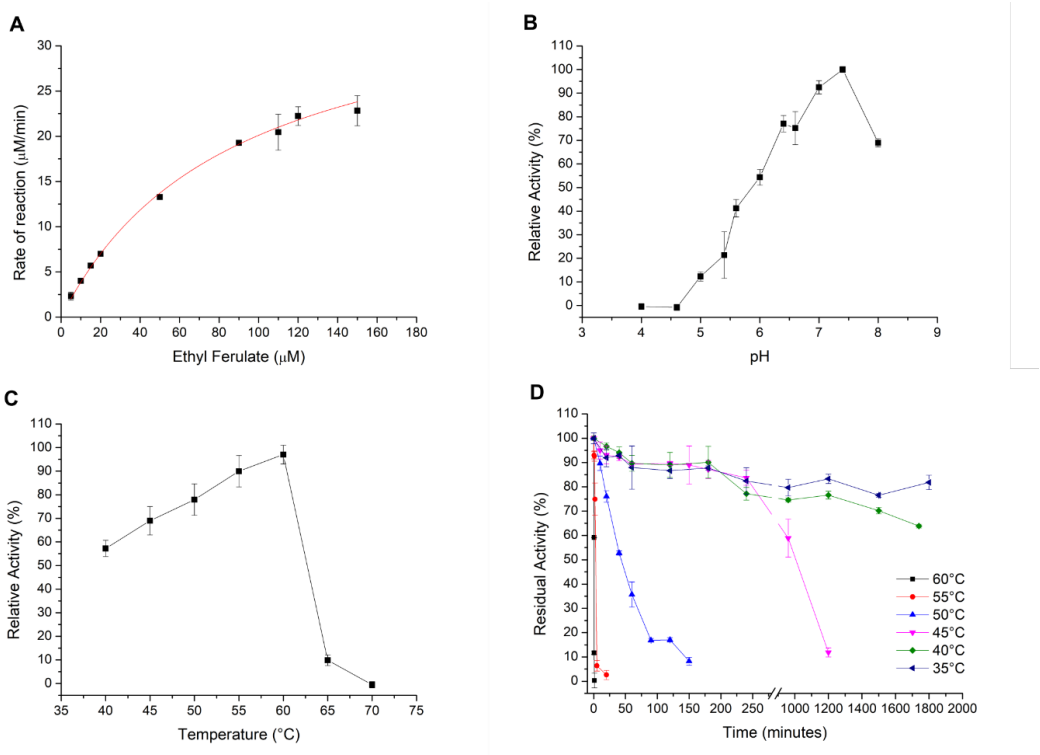


Figure 1. Biochemical characterisation of Tx-Est1 with model substrate ethyl ferulate. **a** Kinetics against ethyl ferulate, **b** Optimum pH, **c** Optimum temperature, **d** Thermal stability.

As Tx-Est1 had the highest catalytic efficiency for ethyl ferulate, the effect of reaction conditions on the activity of Tx-Est1 towards this substrate was assessed. The enzyme was found to have optimal activity at pH 7.4 (Fig. 1b) and 60 °C (Fig. 1c), while it maintained more than 50 % of activity at pH between 6.0 to 8.0 and 40 to 60 °C (Fig. 1d). The protein exhibited good stability in temperatures below 45 °C, keeping more than 70 % of activity after 25 hours and preserved more than 80 % of activity after a week at room temperature (from 20 – 30°C) or after 6 months at 4°C. The low stability at 60-50 °C but optimal temperature at 60 °C is explained by its denaturation at around 54 °C when in aqueous solution but increased structure

stability in the presence of the buffer and substrate. Addition of buffer and substrate changes the unfolding temperature to 60 and 63 °C, respectively, as shown in the CD spectroscopy denaturation and DSC experiments (Supplementary Figure S2 and Table S1).

Production and characterisation of feruloyl arabinoxylan oligosaccharides

To assess the substrate scope of Tx-Est1 its activity was tested towards more complex, and natural substrates. The enzyme was previously shown to be active against feruloylated arabino-xylotetraose [11], XFAXX, from wheat bran (xyl*p*-β1,4-[5-O-FE-Ara*f*-α1,3]-xyl*p*-β1,4-xyl*p*-β1,4-xyl*p*), whichhas a ferulated arabinose decorating the second xylose from the non-reducing end [20,23]. Here, we prepared a panel of cinnamoyl oligosaccharide substrates, via enzymatic digestion of destarched wheat bran, followed by chromatography-based fractionation in a procedure based on [18,19]. Results for 1D ^{1}H and 2D ^{1}H-^{13}C HSQC NMR experiments with the oligosaccharide panel (Table 1, Table S2) were indicative of oligosaccharides with xylopyranose backbones with, dependent on the fraction, single α-1,3-linked and/or double α-1,2/α-1,3-linked arabinofuranose decoration. For all fractions, presence of ferulic acid was observed, as indicated by diagnostic signals of C7/H7, C8/H8 and OMe groups (Table S2). Hereby the proportion of decorated arabinoses, calculated as signal intensity for Ara*f*-C1 associated ^{1}H signal intensity proportional to the FA-derived aromatic ^{1}H signal, varied between 8-104 % (Table 1). For the samples with an esterification ratio of >0.2, arabinose-esterification with ferulic acid was confirmed in the HSQC via diagnostic Ara*f*-C5/H5 signals. Fraction JvM006 and JvM014 represent the most defined substrates, which have 1 and 2 main esterified Ara*f* environments, respectively. NMR signals for JvM014 as expected after GH11 digestion [15] include those corresponding with [20] the presence of esterified oligosaccharide XFAXX.

Table 1. Obtained oligosaccharide fractions, a summary of their characteristics as determined by NMR and MALDI-TOF MS, and ability to function as substrates for Tx-Est1. DP degree of polymerisation; NA, not observed; FA, ferulic acid; Ara*f* arabinofuranosidase; Ara*f*-*s*3 and Ara*f*-*d*2,3 refer to single and double substitution of xylose with α-1,2 and/or α-1,3-linked Ara*f.*

Oligo-saccharide fraction	Identified decorations (NMR)			Esterification	(DP MALDI-TOF MS)			Tx-Est1 substrate
	Ara*f*-*s*3	Ara*f*-*d*2,3	Ara*f*-FA	Ara*f*-FA/Ara*f*	DP	DP-FA	DP-2FA	
JvM001	NA	yes	yes	0.1	5-7	7-10	NA	yes
JvM002	NA	yes	yes	0.1	6-9	9-12	NA	yes
JvM003	NA	NA	yes	0.2	NA	4-8	NA	yes
JvM005	NA	NA	yes	0.7	NA	4-6	NA	yes
JvM006	NA	NA	yes	1.0	NA	4-5	NA	yes
JvM010	yes	yes	yes	0.1	7-10	10-13	NA	yes
JvM012	yes	yes	yes	0.1	6	4-10	NA	yes
JvM013	yes	NA	yes	0.2	NA	7-9	NA	yes
JvM014	NA	NA	yes	0.7	NA	5-8	NA	yes
JvM017	NA	NA	yes	1.0	NA	4-5	6-8	yes

Broad substrate scope of Tx-Est1 on feruloyl oligosaccharides

MALDI-TOF analysis of the fractions revealed a variety of ferulic acid-decorated oligosaccharides ranging from DP 4 – DP 13 (Table 1, Fig. 2 and Fig. S3-9) which dominated the MS spectrum, in addition to non-decorated oligosaccharides in selected fractions. Additionally, fraction JvM017 (Table 1, Fig. S10) contained DP6, DP7 and DP8 with two ferulic acid groups. The observed *m/z* and Ara*f* decoration ratio suggest that these oligosaccharides contain two esterified Ara*f* each rather than di-ferulates, but it should be noted multiple resonances in the aromatic region of the NMR spectra have not been assigned. Incubation of the ferulated xylooligosaccharides with Tx-Est1 revealed activity against all esterified substrates, with signals corresponding to ferulic acid-esterified oligosaccharides no longer visible after 30 minutes incubation. Interestingly Tx-Est1 showed a wide activity

scope against all ferulated oligosaccharides including those with additional mono- and di-Ara*f* substitutions, thus indicating that this enzyme is able to accommodate a wide range of substituted soluble substrates without experiencing considerable steric hindrance.

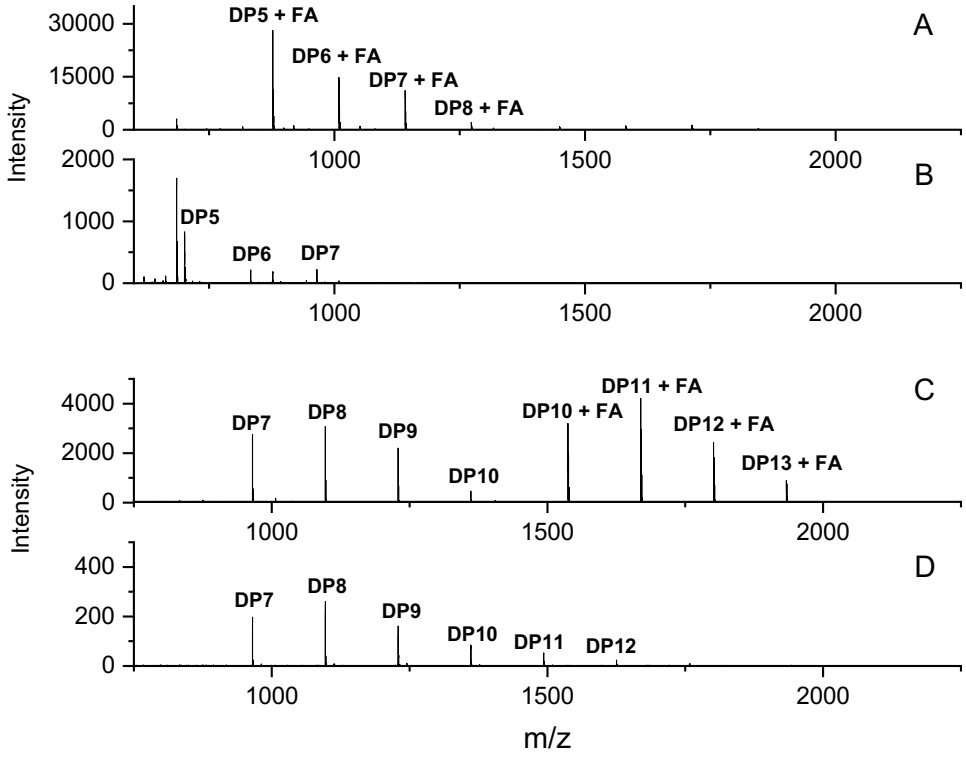


Figure 2. MALDI-TOF MS spectra of oligosaccharide fractions JvM014 (**a,b**) and JvM010 (**c,d**) before (**a**/**c**) and after incubation (**b**/**d**) with Tx-Est1. Reactions performed in 50 mM pH 6.5 citric acid-phosphate buffer at 60 °C for 30 min. Degree of polymerisation (DP) as indicated, FA, ferulic acid.

Activity of Tx-Est1 on complex plant polysaccharides

The activity of Tx-Est1 was assessed against a variety of complex plant polysaccharides and cell wall matrices to ascertain its utility in the processing of food ingredients and agricultural waste streams. Release of FA was detected after incubation of the esterase with acid-debranched wheat arabinoxylan and insoluble Wheat Arabinoxylan, whereas, unexpectantly, no activity was observed after incubation with medium viscosity wheat arabinoxylan. No activity was observed against wheat straw or chlorite treated sugarcane (Fig. S11). Ferulic acid release

was detected from incubation of wheat bran with Tx-Est1 (67 μg g^{-1} wheat bran, Fig. 3d), whilst co-incubation with a GH11 xylanase (Fig. 3e) revealed a substantial increase in ferulic acid release (1.0 mg g^{-1} wheat bran) thus suggesting synergism of Tx-Est1 with other degradative enzymes. Such ferulic acid release values (when co-incubated) are consistent with the order of magnitude seen with other feruoyl esterases [24]. Incubation of wheat bran with either xylanase alone, or with a combination of xylanase and heat-inactivated Tx-Est1, revealed no detectable ferulic acid. Addition of the digestive enzyme mixture Driselase also increased the level of ferulic acid released by Tx-Est1 (1.3 mg g^{-1} wheat bran), but in contrast to previous reports [25] the Driselase alone also displayed esterase activity (1.0 mg g^{-1} wheat bran).

This data indicates that the feruloyl esterase is capable of releasing ferulic acid directly from its location in the complex matrix of cell wall polymers of wheat bran, and potentially that of sugar cane. Furthermore, together with the activity against free oligosaccharides as demonstrated by the mass spectrometry data (Fig. 2), the feruloyl esterase activity is potentiated by xylanases that release a range of ferulated oligosaccharides, suggesting that activity of Tx-Est1 on the complex cell wall matrices is limited by solubility of or access to its substrate. This is further supported by the differences in the lignin content of the substrates tested here; as sugar cane (bagasse), wheat straw, and wheat bran have been reported to contain 25-32%, 21%, and 3% lignin, respectively [20,26], Tx-Est1 released the highest levels of ferulic acid from the substrate with the lowest lignin content.

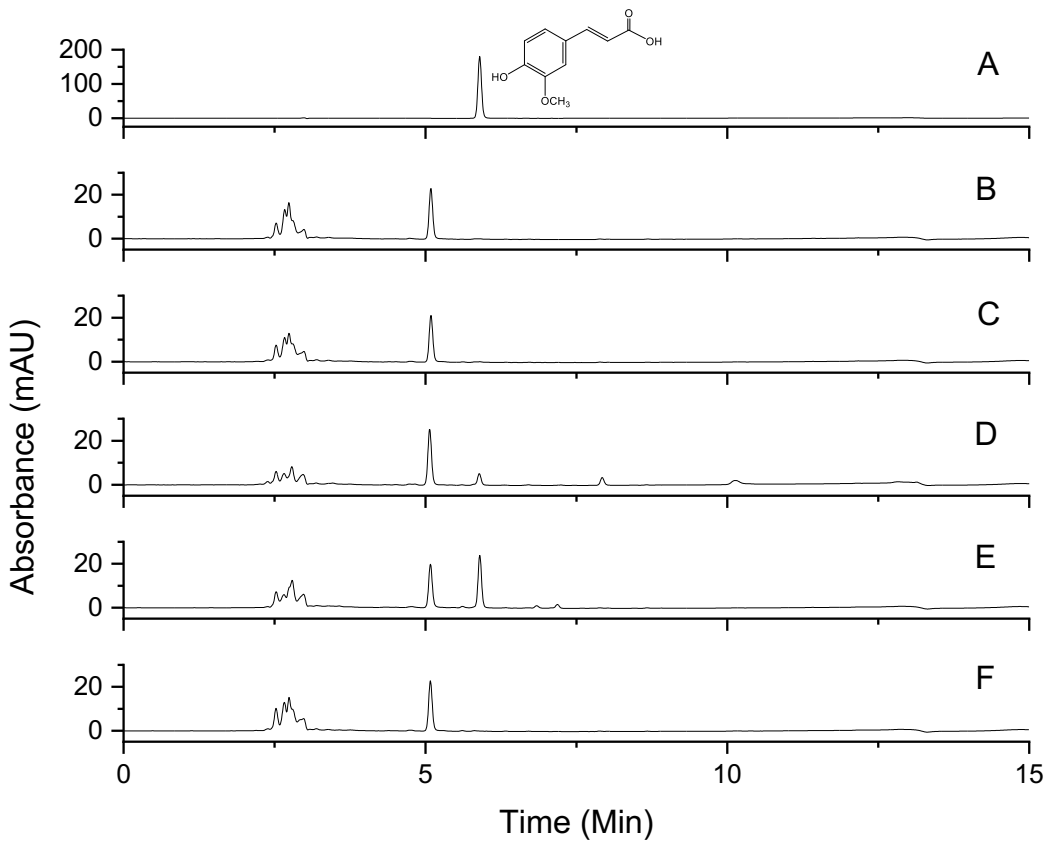


Figure 3. HPLC profiles of Tx-Est1 against wheat bran. **a** Ferulic Acid 0.5 mM, **b** Wheat Bran 1 % (w/v) negative control, **c** Wheat Bran 1 % (w/v) incubated with *T. lanuginosus* xylanase, **d** Wheat Bran 1 % (w/v) incubated with TxEst1,**e**Wheat Bran 1 % (w/v) co-incubated with TxEst1 and *T. lanuginosus* xylanase,**f**Wheat Bran 1 % (w/v) co-incubated with denatured TxEst1 and *T. lanuginosus* xylanase.

Conclusions

The bioactive nature of ferulic acid, coupled with its antioxidant and preservative capabilities, makes it an important ingredient throughout the food industry. Enzymatic, greener and more cost-effective routes for its production are therefore of great interest for industries utilising this compound in their products. Consequently, feruloyl esterases will play an important role in the development of such technologies. In this study, we expressed and biochemically characterised the feruloyl esterase Tx-Est1 and assessed its activity against a variety of different carbohydrates present in agricultural by-products. The enzyme was highly efficient at removing ferulic acid from a variety of length DP xylooligosaccharides, in addition

to showing activity against wheat bran. When utilised with a commercially available xylanase, the ferulic acid release showed a marked improvement in comparison to Tx-Est1 alone and highlights a potential synergism between this esterase and other proteins involved in the deconstruction of plant polysaccharides. Our data suggests Tx-Est1 can be employed for the production of a high-value compound from agricultural waste streams or as part of the saccharification process of complex, heterogeneous plant polysaccharides.

CRediT authorship contribution statement

Carolina Garbelotti: Investigation, Methodology, Visualization, Writing - original draft. Gregory Stuart Bulmer: Investigation, Methodology, Visualization, Writing - original draft. Richard John Ward: Conceptualization, Funding acquisition, Methodology, Supervision, Writing - original draft, review & editing. Jolanda Margaretha Van Munster: Conceptualization, Funding acquisition, Methodology, Supervision, Writing - original draft, review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This study was funded by the BBSRC via BB/P011462/1 (JvM), FAPESP-SPRINT research grant 2018/14030-2 (RJW, JvM), FAPESP doctorate fellowship 2018/25664-2 (CVG), the National Institute of Science and Technology of Bioetanol (INCT-Bioetanol) (FAPESP 2011/57908-6 and 2014/50884-5, CNPq 574002/2008-1 and 465319/2014-9) (Brazil),and the EPSRC via DTP EP/N509565/1.

References

[1] S. Raveendran, B. Parameswaran, S.B. Ummalyma, A. Abraham, A.K. Mathew, A. Madhavan, S. Rebello, A. Pandey, Applications of Microbial Enzymes in Food Industry, Food Technol. Biotechnol. 56 (2018) 16.

[2] M. Russo, A. Marquez, H. Herrera, C. Abeijon-Mukdsi, L. Saavedra, E. Hebert, P. Gauffin-Cano, R. Medina, Oral administration of Lactobacillus fermentum CRL1446 improves biomarkers of metabolic syndrome in mice fed a high-fat diet supplemented with wheat bran, Food Funct. 11 (2020) 3879–3894.

[3] S. Mathew, T.E. Abraham, Ferulic Acid: An Antioxidant Found Naturally in Plant Cell Walls and Feruloyl Esterases Involved in its Release and Their Applications, Crit. Rev. Biotechnol. 24 (2008) 59–83.

[4] John H. Grabber, and John Ralph, R.D. Hatfield, Ferulate Cross-Links Limit the Enzymatic Degradation of Synthetically Lignified Primary Walls of Maize, J. Agric. Food Chem. 46 (1998) 2609–2614.

[5] D.W.S. Wong, Feruloyl esterase, Appl. Biochem. Biotechnol. 2006 1332. 133 (2006) 87–112.

[6] D.E. Akin, W.S. Borneman, L.L. Rigsby, S.A. Martin, p-Coumaroyl and feruloyl arabinoxylans from plant cell walls as substrates for ruminal bacteria, Appl. Environ. Microbiol. 59 (1993) 644–647.

[7] T. Koseki, Y. Ito, S. Furuse, K. Ito, K. Iwano, Conversion of ferulic acid into 4-vinylguaiacol, vanillin and vanillic acid in model solutions of shochu, J. Ferment. Bioeng. 82 (1996) 46–50.

[8] A. Dilokpimol, M.R. Mäkelä, M.V. Aguilar-Pontes, I. Benoit-Gelber, K.S. Hildén, R.P. de Vries, Diversity of fungal feruloyl esterases: updated phylogenetic classification, properties, and industrial applications, Biotechnol. Biofuels. 9 (2016) 231.

[9] A. Dilokpimol, M.R. Mäkelä, S. Varriale, M. Zhou, G. Cerullo, L. Gidijala, H. Hinkka, J.L.A. Brás, P. Jütten, A. Piechot, R. Verhaert, K.S. Hildén, V. Faraco, R.P. de Vries, Fungal feruloyl esterases: Functional validation of genome mining based enzyme discovery including uncharacterized subfamilies, N. Biotechnol. 41 (2018) 9–14.

[10] J. Donaghy, P.F. Kelly, A.M. McKay, Detection of ferulic acid esterase production by Bacillus spp. and lactobacilli, Appl. Microbiol. Biotechnol. 50 (1998) 257–260.

[11] H. Rakotoarivonina, B. Hermant, B. Chabbert, J.P. Touzel, C. Remond, A thermostable feruloyl-esterase from the hemicellulolytic bacterium Thermobacillus xylanilyticus releases phenolic acids from non-pretreated plant cell walls, Appl. Microbiol. Biotechnol. 90 (2011) 541–552.

[12] M.-C. Ralet, C.B. Faulds, G. Williamson, J.-F. Thibault, Degradation of feruloylated oligosaccharides from sugar-beet pulp and wheat bran by ferulic acid esterases from Aspergihs niger t, Carbohydr. Res. 263 (1994) 257–269.

[13] A.P. de Souza, D.C.C. Leite, S. Pattathil, M.G. Hahn, M.S. Buckeridge, Composition and Structure of Sugarcane Cell Wall Polysaccharides: Implications for Second-Generation Bioethanol Production, BioEnergy Res. 2012 62. 6 (2012) 564–579.

[14] L. Lama, A. Tramice, I. Finore, G. Anzelmo, V. Calandrelli, E. Pagnotta, G. Tommonaro, A. Poli, P. Di Donato, B. Nicolaus, M. Fagnano, M. Mori, A. Impagliazzo, A. Trincone, Degradative actions of microbial xylanolytic activities on hemicelluloses from rhizome of Arundo donax, AMB Express. 4 (2014) 1–9.

[15] P. Biely, S. Singh, V. Puchart, Towards enzymatic breakdown of complex plant xylan structures: State of the art, Biotechnol. Adv. 34 (2016) 1260–1274.

[16] J.M. van Munster, B. Thomas, M. Riese, A.L. Davis, C.J. Gray, D.B. Archer, S.L. Flitsch, Application of carbohydrate arrays coupled with mass spectrometry to detect activity of plant-polysaccharide degradative enzymes from the fungus Aspergillus niger, Sci. Rep. 7 (2017) 43117.

[17] F.J.M. Kormelink, R.A. Hoffmann, H. Gruppen, A.G.J. Voragen, J.P. Kamerling, J.F.G. Vliegenthart, Characterisation by 1H NMR spectroscopy of oligosaccharides derived from alkali-extractable wheat-flour arabinoxylan by digestion with endo-(1 → 4)-β-d-xylanase III from Aspergillus awamori, Carbohydr. Res. 249 (1993) 369–382.

[18] R.R. Schendel, A. Becker, C.E. Tyl, M. Bunzel, Isolation and characterization of feruloylated arabinoxylan oligosaccharides from the perennial cereal grain

intermediate wheat grass (Thinopyrum intermedium), Carbohydr. Res. 407 (2015) 16–25.

[19] M. Bunzel, E. Allerdings, V. Sinwell, J. Ralph, H. Steinhart, Cell wall hydroxycinnamates in wild rice (Zizania aquatica L.) insoluble dietary fibre, Eur. Food Res. Technol. 2002 2146. 214 (2002) 482–488.

[20] C. Lequart, J.M. Nuzillard, B. Kurek, P. Debeire, Hydrolysis of wheat bran and straw by an endoxylanase: Production and structural characterization of cinnamoyl-oligosaccharides, Carbohydr. Res. 319 (1999) 102–111.

[21] R.A. Hoffmann, B. Leeflang, M. de Barse, J. Kamerling, J. Vliegenthart, Characterisation by 1H-n.m.r. spectroscopy of oligosaccharides, derived from arabinoxylans of white endosperm of wheat, that contain the elements -4)[alpha-L-Araf-(1-3)]-beta-D-Xylp-(1- or -4)[alpha- L-Araf-(1-2)][alpha-L-Araf-(1-3)]-beta-D-Xylp-(1-, Carbohydr. Res. 221 (1991) 63–81.

[22] V.F. Crepin, C.B. Faulds, I.F. Connerton, Functional classification of the microbial feruloyl esterases, Appl. Microbiol. Biotechnol. 2004 636. 63 (2003) 647–652.

[23] C. Rémond, I. Boukari, G. Chambat, M. O'Donohue, Action of a GH 51 α-l-arabinofuranosidase on wheat-derived arabinoxylans and arabino-xylooligosaccharides, Carbohydr. Polym. 72 (2008) 424–430.

[24] Z. Xu, J. Kong, S. Zhang, T. Wang, X. Liu, Comparison of Enzyme Secretion and Ferulic Acid Production by Escherichia coli Expressing Different Lactobacillus Feruloyl Esterases, Front. Microbiol. 11 (2020) 568716.

[25] P. Kroon, M. Garcia-Conesa, I. Fillingham, G. Hazlewood, G. Williamson, Release of ferulic acid dehydrodimers from plant cell walls by feruloyl esterases, J. Sci. Food Agric. 79 (1999) 428–434.

[26] J. Zeng, Z. Tong, L. Wang, J.Y. Zhu, L. Ingram, Isolation and structural characterization of sugarcane bagasse lignin after dilute phosphoric acid plus steam explosion pretreatment and its effect on cellulose hydrolysis, Bioresour. Technol. 154 (2014) 274–281.

Supplementary information

Utilisation of a thermotolerant feruoyl esterase from *Thermobacillus xylanilyticus* that processes pre-biotic ferulated xylooligosaccharides

Carolina Garbelotti[1], Gregory S. Bulmer[2], Richard J. Ward[1*] and Jolanda M. van Munster[2,3*]

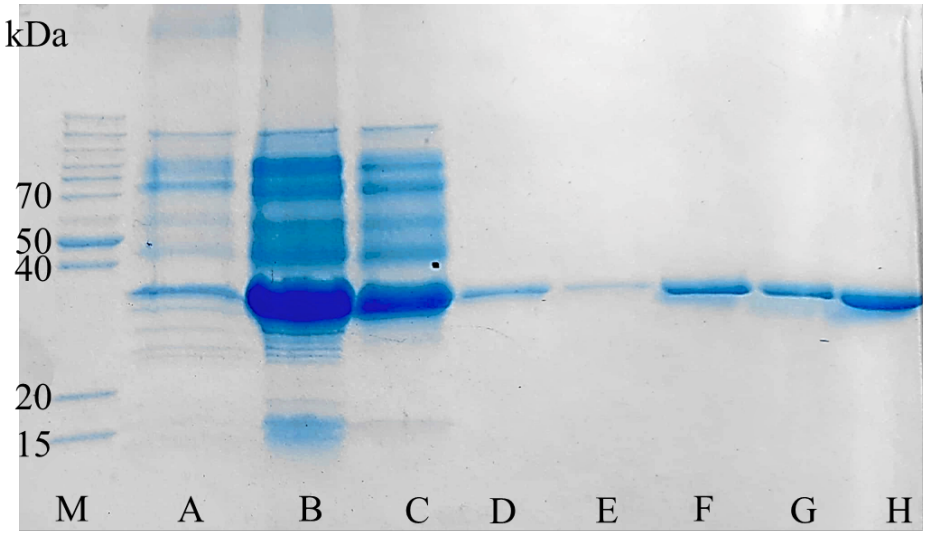
Affiliations

[1] Departamento de Química, Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, SP CEP 14040-901, Brazil

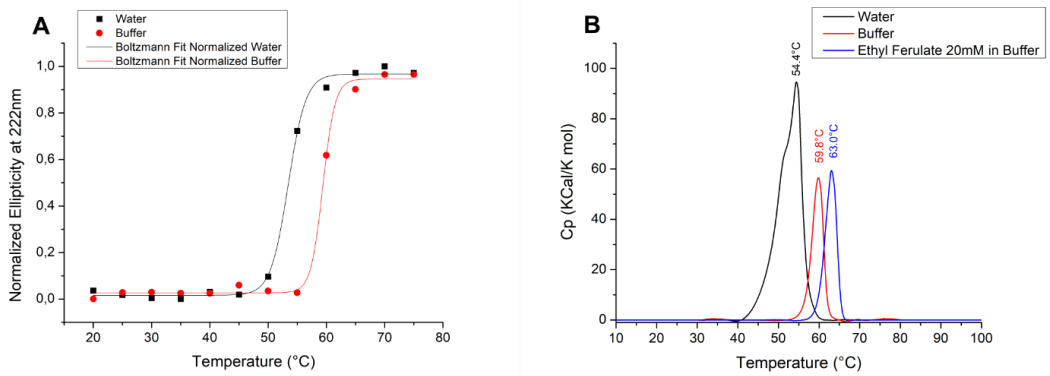
[2] Manchester Institute of Biotechnology (MIB) & School of Natural Sciences, The University of Manchester, 131 Princess Street, Manchester, M1 7DN, United Kingdom

[3] Scotland's Rural College, West Mains Road, Kings Buildings, Edinburgh, EH9 3JG, United Kingdom

Contents



Supplementary figure 1: SDS-PAGE Expression and Purification of Feruloyl Esterase Tx-Est1in *E. coli* BL21 STAR**. M** Molecular weight standard; **A** Expression T_0; **B** Expression 16h; **C** Flow Through purification; **D** Wash with 5 mM imidazole; **E** Wash with 10 mM imidazole; **F** Wash with 40 mM imidazole; **G** Eluted with 250 mM imidazole; **H** Pure protein after dialysis and buffer change from TRIS 50 mM, NaCl 100 mM pH 8.0 with 250 mM imidazole to deionized water.



Supplementary figure 2: **A** Boltzmann fit of normalized circular dichroism spectroscopy denaturation data of Tx-Est1 in water (black) and McIlvine (citrate-phosphate 0.2 M) pH 7.4 buffer (red). **B** Differential scanning calorimetry transition peaks for Tx-Est1 in water (black line), McIlvine (citrate-phosphate 0.2 M) pH 7.4 buffer (red line) and ethyl ferulate 20 mM in buffer (blue line). Annotations above peaks show melting temperatures obtained at maximum.

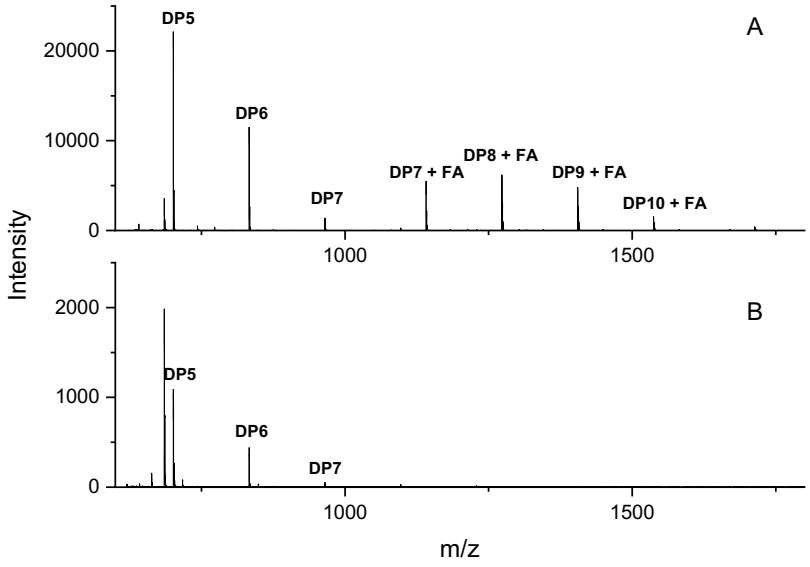
Supplementary Table 1: Thermodynamic parameters obtained from CD spectroscopy and DSC data analysis.

	CD spectroscopy data		*DSC data*		
	Enzyme concentration (mg/mL)	***Tm (°C)***	***Enzyme concentration (mg/mL)***	***Tm (°C)***	***ΔH (Kcal/mol)***
Tx-Est1 in water	0.140	53.45±0.21	1.05	54.4	636.12
Tx-Est1 in McIlvine pH 7.4 buffer	0.157	59.40±0.26	1.05	59.8	208.33
Tx-Est1 with ethyl ferulate 20mM in McIlvine pH 7.4 buffer	-	-	1.05	63.0	213.4

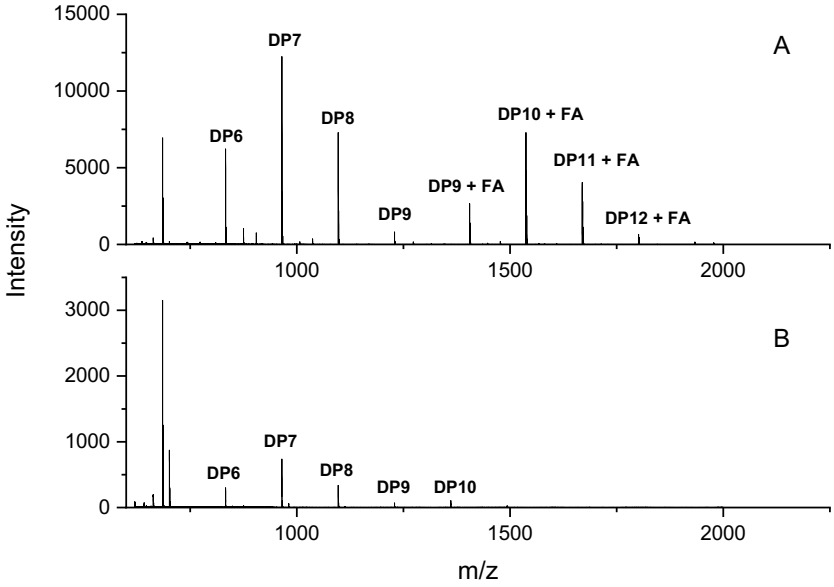
		JvM001		JvM002		JvM003		JvM005		JvM006	
		^{1}H	^{13}C	^{1}H	^{13}C	^{1}H	^{13}C	^{1}H	^{13}C	^{1}H	^{13}C
C1(-H)	α-Xyl*p*^{r}	5.184 (3.4)	92.0	5.182 (3.2)	91.9	5.184 (3.6)	92.0	5.184 (3.6)	92.0	5.184 (3.6)	92.0
	α-unidentified-*p*^{r}							5.059 (3.6)	91.8		
	β-Xyl*p*^{r}	4.585 (8.0)	96.5	4.584 (7.2)	96.4	5.585 (7.6)	96.5	4.584 (8.0)	96.5	4.575 (8.0)	96.5
								4.574 (8.0)			
	β-Xyl*p*^{N}-A*f*-*O*-2,3	4.639 (7.2)	99.9	4.638 (7.2)	99.9						
	β-Xyl*p*	4.596 (7.6)	100.0								
	β-Xyl*p*^{i}	4.50-4.42	101.6	4.43-4.51	101.6	4.51-4.43	101.8	4.419 (7.6)	101.6	4.501 (7.6)	101.5
	β-Xyl*p*^{i}							4.500 (7.2)	101.6	4.420 (8.0)	101.7
	α-Ara*f*-X^{N}-*O*-2d	5.226	108.6	5.225	108.6						
	α-Ara*f*-X^{N}-*O*-3d	5.275	108.1	5.273	108.0						
	α-Ara*f* (X^{N}-*O*-2d)			5.245	108.6						
	α-Ara*f*(X^{N}-*O*-3d)			5.294	108.0						
	α-Ara*f*	5.239	108.5			5.237	108.6				
	α-Ara*f*	5.246	108.5			5.244	108.6				
	α-Ara*f*(-FA)	5.367	108.1	5.420	107.3	5.405	107.9	5.410	107.9	5.270	108.2
	α-Ara*f*(-FA)	5.352	108.1			5.38-5.28	108.2-108.4	5.361	107.9	5.233	108.6
	α-Ara*f*(-FA)							5.342	107.9		
	α-Ara*f*(-FA)							5.30-5.26	108.2		
	α-Ara*f*(-FA)							5.24-5.21	108.6		
	α-Ara*f*(-FA)*							5.117	108.5		
	α-Ara*f*-FA							5.396	107.9	5.396	107.9
C2(-H)	FA^{t}	7.330	111.5	7.324	NA	7.275	111.4	7.271	111.4	7.282	111.4
		7.307	111.8	7.347				7.267		7.241	
C5(-H)	FA^{t}	6.93-6.97	115.9	6.979 (8.5)	NA	6.923 (7.6)	115.8	6.902 (8.0)	115.9	6.922 (8.4)	115.8
								6.885 (8.0)		6.890 (8.4)	
	α-Ara*f*-FA							4.34, 4.49	64.2	4.34, 4.49	64.2
	α-Ara*f*	3.72, 3.80	61.2	3.73, 3.81	61.1	3.72, 3.81	61.1				
C6(-H)	FA^{t}	7.230	123.8	7.261	NA	7.185	123.9	7.19	123.8	7.175	123.6
		7.255				7.205		7.17		7.196	
C7(-H)	FA^{t}	7.77 (16.0)	147.2	7.78 (16.4)	NA	7.69-7.76	146.9	7.70 (16.0)	146.9	7.70 (16.0)	146.8
								7.71 (16.0)			
C8(-H)	FA^{t}	6.465	113.8	6.484	NA	6.42-6.50	113.8	6.431 (16.0)	113.6	6.445 (16.0)	113.9
		6.505		6.520				6.445 (16.0)			
		6.541		6.556							
OCH3	FA^{t}	3.917	56.0	3.924	56.0	3.899	55.9	3.887	55.9	3.894	55.9
						3.980	56.0			3.860	

		JvM010		JvM012		JvM013		JvM014		JvM017	
		^{1}H	^{13}C	^{1}H	^{13}C	^{1}H	^{13}C	^{1}H	^{13}C	^{1}H	^{13}C
C1(-H)	α-Xyl*p*^{r}	5.184 (3.6)	92.0	5.184 (3.6)	91.9	5.184 (3.6)	91.9	5.184 (3.6)	92.0	5.184 (3.6)	91.9
	β-Xyl*p*^{r}	4.584 (8.0)	96.5	4.584 (8.0)	96.4	4.584 (8.0)	96.5	4.584 (8.0)	96.5	4.584 (8.0)	96.5
	β-Xyl*p*^{N}-A*f*-*O*-2,3	4.635 (7.2)	99.9	4.638 (7.2)	99.9						
	β-Xyl*p*^{i}	4.53-4.41	101.5	4.53-4.42	101.6	4.54-4.41	101.6	4.53-4.40	101.6	4.52-4.41	101.6
	β-Xyl*p*^{i}							4.35-4.28	101.7		
	α-Ara*f*-X^{N}-*O*-3s	5.397	107.6	5.397	107.5	5.397	107.6				
	α-Ara*f*-X^{N}-*O*-3s	5.391	107.6								
	α-Ara*f*-X^{N}-*O*-2d	5.223	108.7	5.225	108.6						
	α-Ara*f*-X^{N}-*O*-3d	5.273	108.1	5.273	108.0						
	α-Ara*f*(-FA)	5.365	107.9	5.360	107.9	5.361	107.7	5.389	107.8		
	α-Ara*f*(-FA)	5.29-5.33	108.1			5.316	108.5	5.29-5.36	108.0-108.2		
	α-Ara*f*(-FA)					5.291	108.5				
	α-Ara*f*-FA							5.406	107.8	5.396	107.9
C2(-H)	FA^{t}	7.32	111.5	7.33	111.6	7.34	111.4	7.222	111.3	7.293, 7.297	111.4
C5(-H)	FA^{t}	6.99-6.92	115.8	6.97-6.91	116.0	6.96-6.92	115.8	6.905 (8.0)	115.6	6.941 (8.0)	115.6
	α-Ara*f*-FA							4.34, 4.48	64.2	4.35, 4.54	64.2
	α-Ara*f*	3.73, 3.80	61.2	3.72, 3.78	61.2	3.75, 3.79	61.3	3.73, 3.79	61.3		
C6(-H)	FA^{t}	7.27-7.20	123.8	7.27-7.20	124.1	7.26-7.20	123.9	7.134 7.154	123.6	7.21-7.17	123.5
C7(-H)	FA^{t}	7.79-7.70	147.0	7.80-7.71	NR	7.80-7.72	147.1	7.65 (16.0)	146.7	7.74-7.69	146.6
C8(-H)	FA^{t}	6.54-6.43	113.8	6.54-6.45	113.4	6.53-6.45	NR	6.401 (15.6)	113.9	6.465 (16.0)	114.2
OCH3	FA^{t}	3.92	55.9	3.91	56.0	3.91 3.92	55.9	3.87 3.90	55.9 56.0	3.90 4.06	55.9 55.1

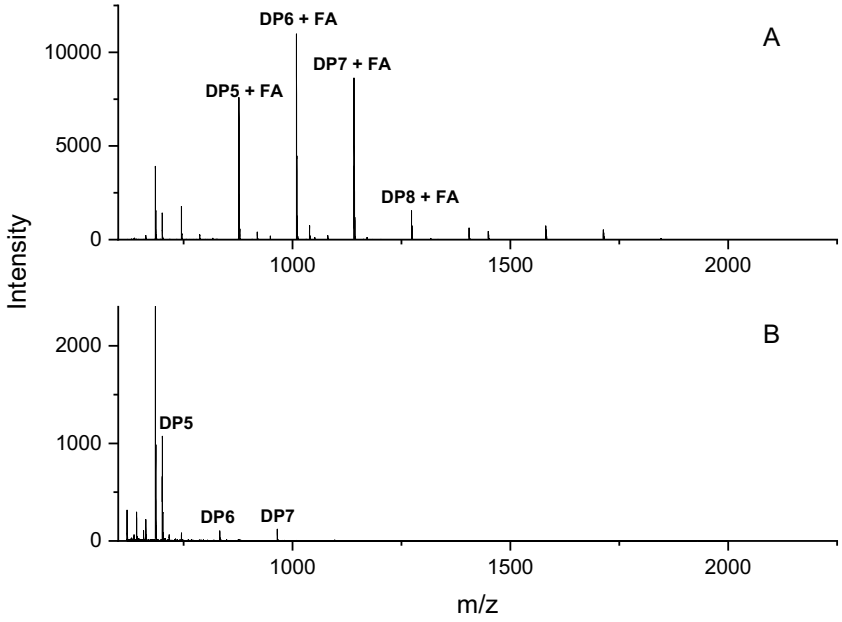
Supplementary Table 2. Assignment of chemical shifts of NMR resonances. Chemical shifts (δ) of ^{1}H and ^{13}C in p.p.m., coupling constants (*J*) in Hz. Data for ^{1}H and ^{13}C are derived from 1D ^{1}H NMR or 2D ^{1}H-^{13}C HSQC respectively, except α-Ara*f* C5-linked ^{1}H, which were derived from HSQC. Xyl*p*^{r} indicates the reducing end xylopyranose, Xyl*p*^{i} the internal unsubstituted and single substituted xylose backbone residues, β-Xyl*p*^{N}-A*f*-*O*-2,3 indicates Xyl*p* with double α1,2 and α1,3 substitution, whereby α-Ara*f*-X^{N}-*O*-3s and α-Ara*f*-X^{N}-*O*-2d/3d represent single (3s) and double (2 and 3d) arabinofuranose decoration on unidentified Xyl*p* via α-1,3 or α-1,2 and α-1,3-linkage respectively. FA^{t} indicates trans-Ferulic Acid, α-Ara*f*-FA and α-Ara*f*(-FA) indicate Ara*f* esterified with FA, and potentially esterified with FA, respectively. * tentative assignment, NA not identified and NR not resolvable. Fractions JvM005 and JvM006 have an additional unidentified aromatic resonances at ^{1}H 7.19 ppm, ^{13}C 130.8 ppm. JvM017 has the following unidentified aromatic resonances (^{1}H/^{13}C δ in ppm):7.19/119.5; 7.28/122.5; 7.30/124.9;7.34/138.9;d7.52 (8.0)/111.9; d7.72/118.5;8.15-8.23/



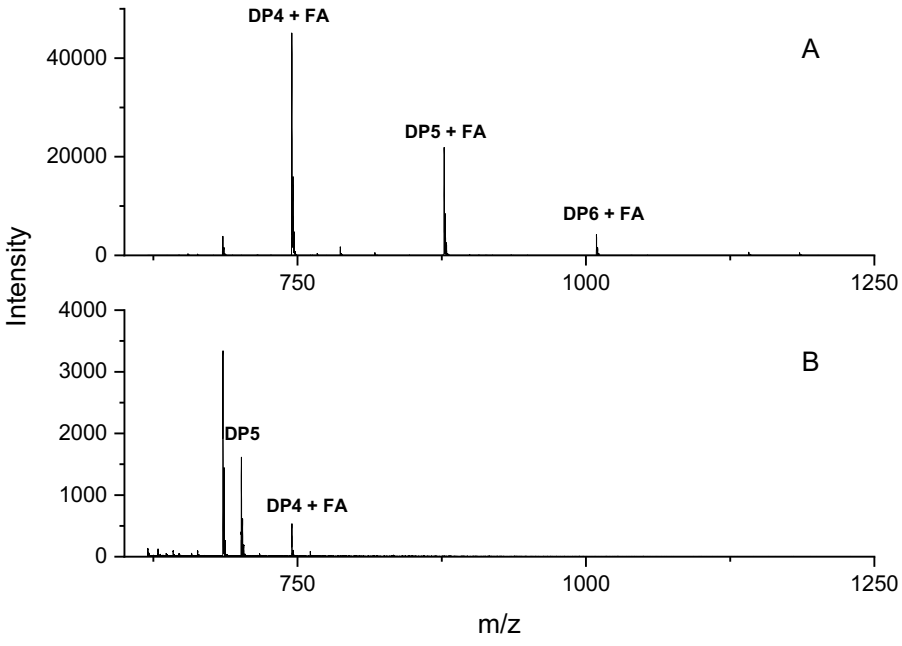
Supplementary figure 3: MALDI-TOF MS spectra of oligosaccharide fractions JvM011before (**A**) and after incubation (**B**) with Tx-Est1. Reactions performed in 50 mM pH 6.5 citric acid-phosphate buffer at 60 °C for 30 min. Degree of polymerisation (DP) as indicated, FA, ferulic acid.



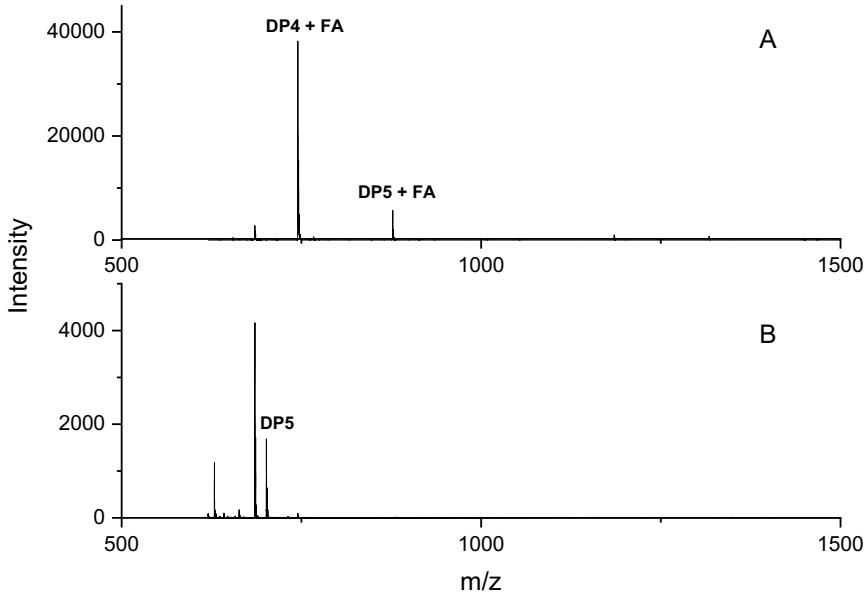
Supplementary figure 4: MALDI-TOF MS spectra of oligosaccharide fractions JvM002before (**A**) and after incubation (**B**) with Tx-Est1. Reactions performed in 50 mM pH 6.5 citric acid-phosphate buffer at 60 °C for 30 min.



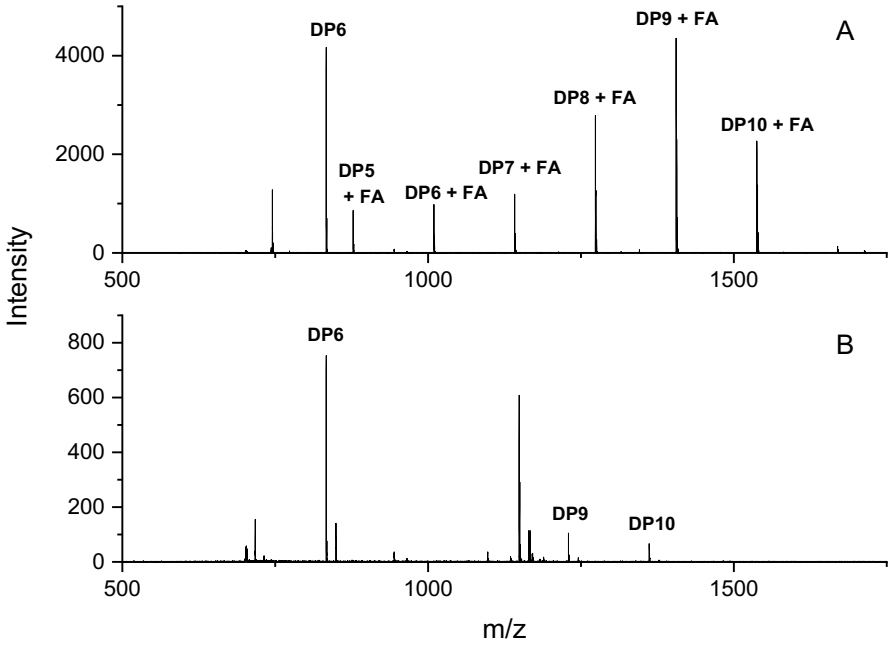
Supplementary figure 5: MALDI-TOF MS spectra of oligosaccharide fractions JvM003before (**A**) and after incubation (**B**) with Tx-Est1. Reactions performed in 50 mM pH 6.5 citric acid-phosphate buffer at 60 °C for 30 min.



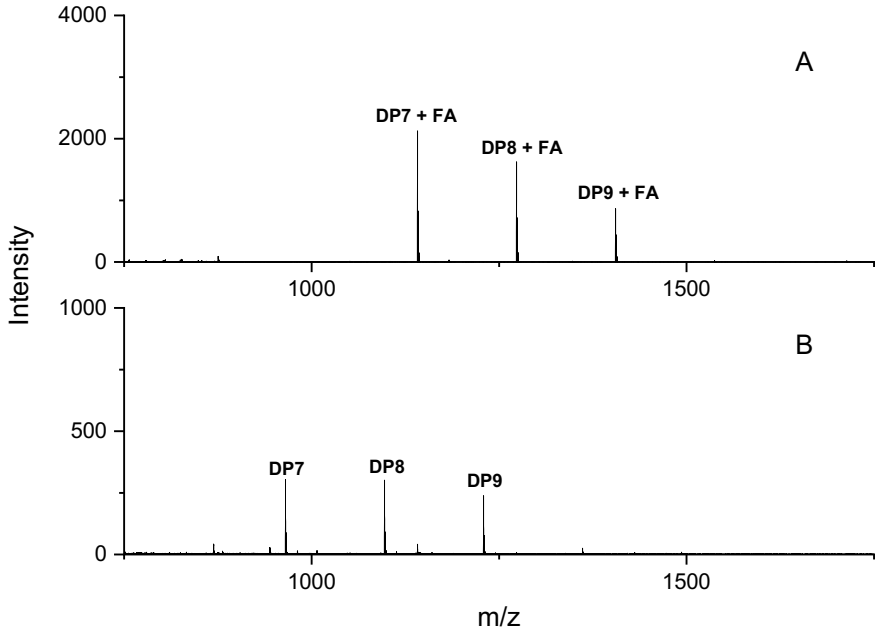
Supplementary figure 6: MALDI-TOF MS spectra of oligosaccharide fractions JvM005before (**A**) and after incubation (**B**) with Tx-Est1. Reactions performed in 50 mM pH 6.5 citric acid-phosphate buffer at 60 °C for 30 min.



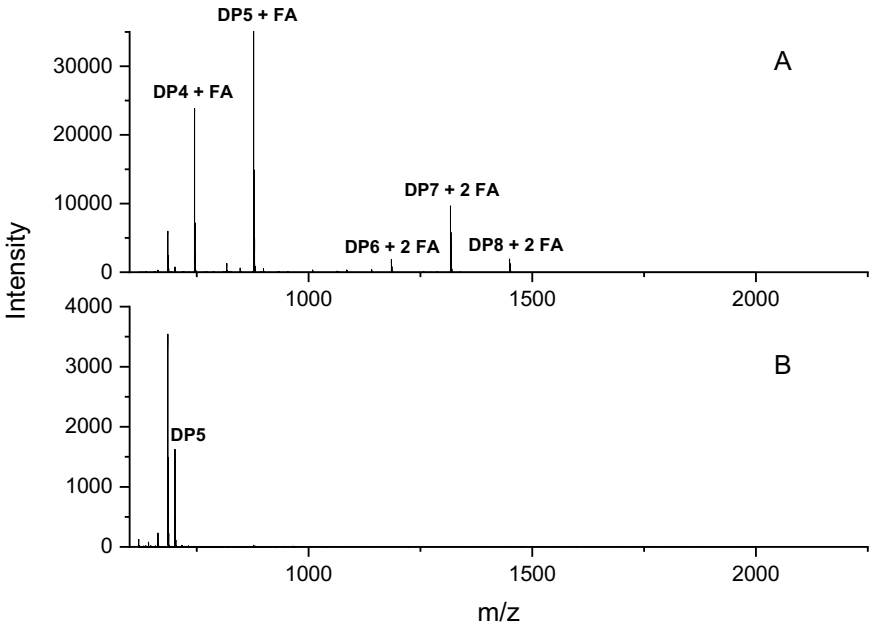
Supplementary figure 7: MALDI-TOF MS spectra of oligosaccharide fractions JvM006 before (**A**) and after incubation (**B**) with Tx-Est1. Reactions performed in 50 mM pH 6.5 citric acid-phosphate buffer at 60 °C for 30 min.



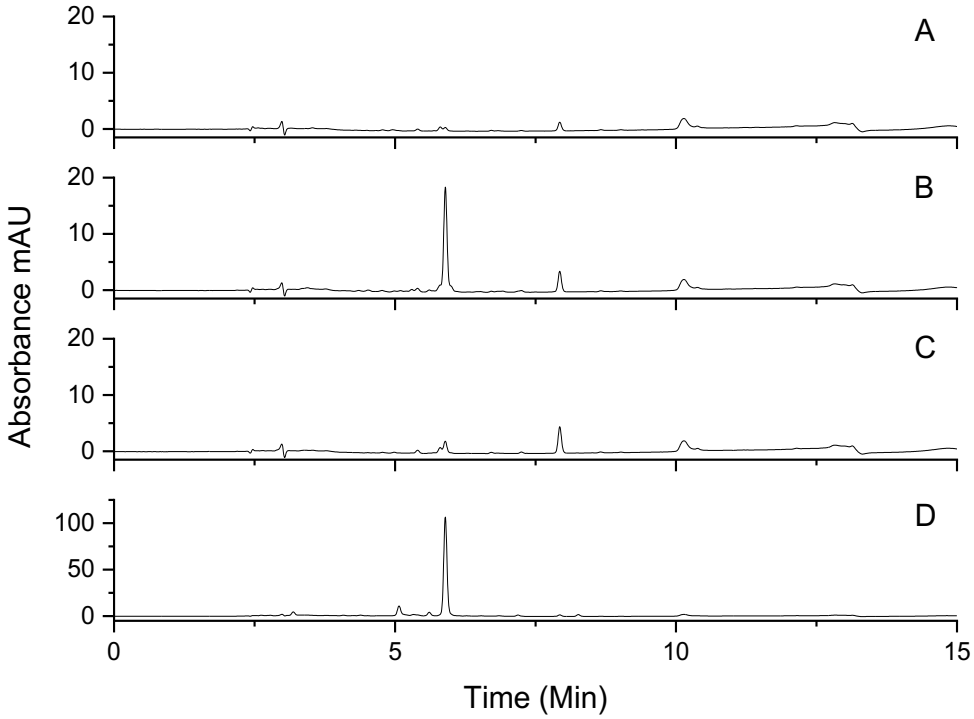
Supplementary figure 8: MALDI-TOF MS spectra of oligosaccharide fractions JvM012before (**A**) and after incubation (**B**) with Tx-Est1. Reactions performed in 50 mM pH 6.5 citric acid-phosphate buffer at 60 °C for 30 min.



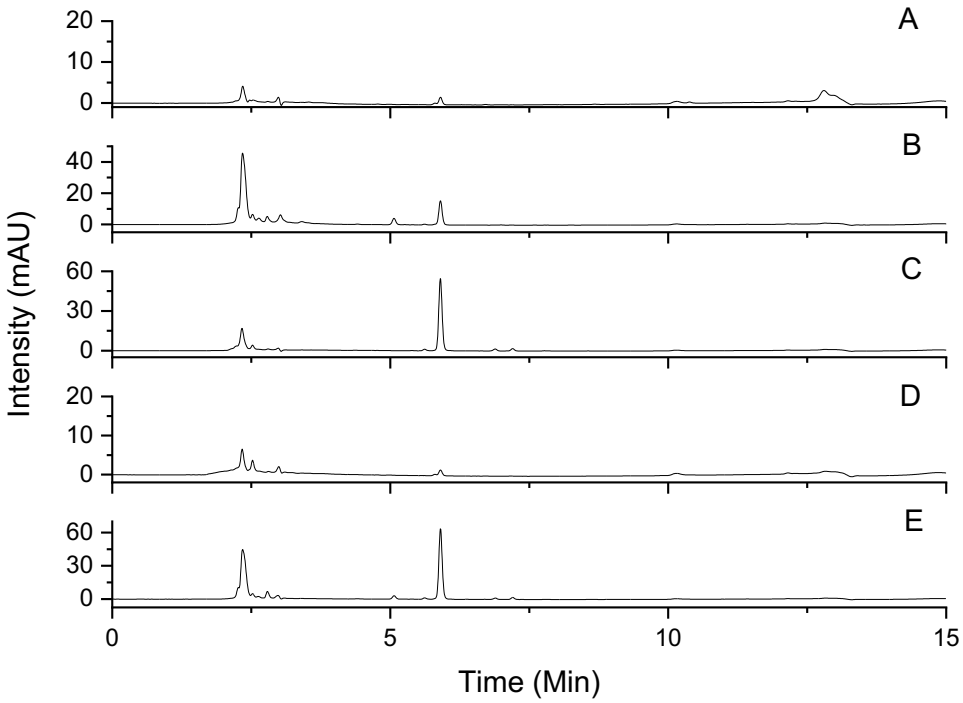
Supplementary figure 9: MALDI-TOF MS spectra of oligosaccharide fractions JvM013 before (**A**) and after incubation (**B**) with Tx-Est1. Reactions performed in 50 mM pH 6.5 citric acid-phosphate buffer at 60 °C for 30 min.



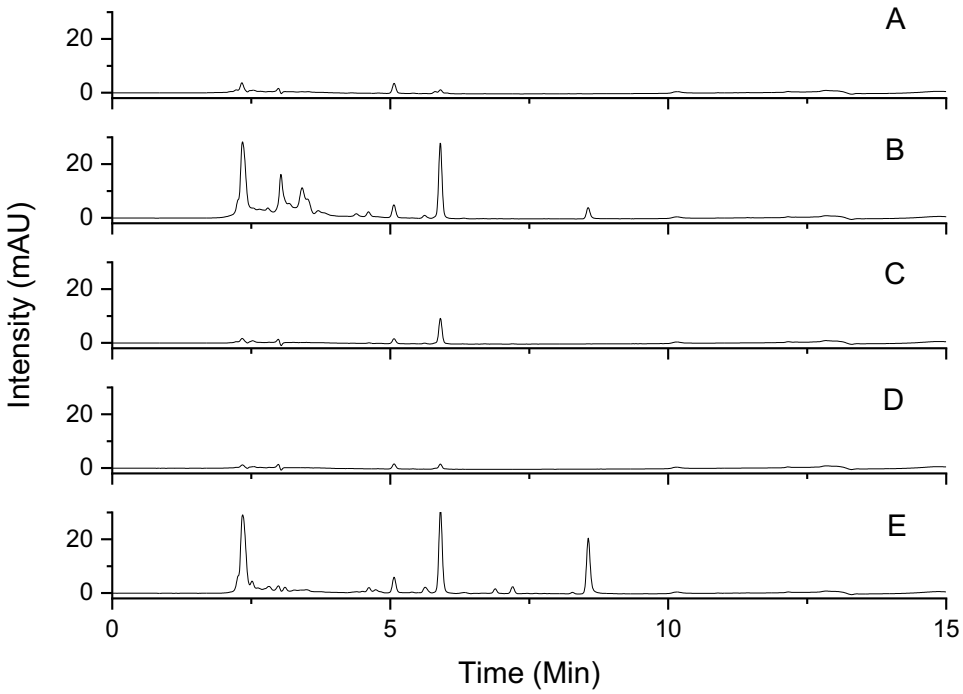
Supplementary figure 10: MALDI-TOF MS spectra of oligosaccharide fractions JvM017before (**A**) and after incubation (**B**) with Tx-Est1. Reactions performed in 50 mM pH 6.5 citric acid-phosphate buffer at 60 °C for 30 min.



Supplementary figure 11: HPLC of Tx-Est1 incubation with insoluble wheat arabinoxylan. **A** Insoluble wheat arabinoxylan 1 % (w/v) negative control; **B** Insoluble wheat arabinoxylan 1 % (w/v) incubated with TxEst1; **C** Insoluble wheat arabinoxylan 1 % (w/v) incubated with heat inactivated TxEst1; **D** Insoluble wheat arabinoxylan co-incubated with TxEst1 and *T. lanuginosus* xylanase.



Supplementary figure 12: HPLC of Tx-Est1 incubation with Wheat arabinoxylan(acid debranched). **A** Acid debranched wheat arabinoxylan 1 % (w/v) negative control; **B** Acid debranched wheat arabinoxylan 1 % (w/v) incubated with *T. lanuginosus* xylanase; **C** Acid debranched wheat arabinoxylan 1 % (w/v) incubated with TxEst1; **D** Acid debranched wheat arabinoxylan 1 % (w/v) incubated with heat inactivated TxEst1; **E** Acid debranched wheat arabinoxylan co-incubated with TxEst1 and *T. lanuginosus* xylanase.



Supplementary figure 13: HPLC of Tx-Est1 incubation with destarched wheat bran. **A** Destarched Wheat bran 1 % (w/v) negative control; **B** Destarched Wheat bran 1 % (w/v) incubated with *T. lanuginosus* xylanase; **C** Destarched Wheat bran 1 % (w/v) incubated with TxEst1; **D** Destarched Wheat bran 1 % (w/v) incubated with heat inactivated TxEst1; **E** Destarched Wheat bran co-incubated with TxEst1 and *T. lanuginosus* xylanase.

Chapter 7: Conclusions and outlook

The highly abundant nature of agricultural biomass offers a vast source of lignocellulosic material for the production of fuels, materials and other high-value compounds. However, the variation in agricultural waste streams, in tandem with the heterogeneous nature of the plant cell wall (which varies by plant species, tissue type and age) poses problems to its large-scale conversion into renewable products. Consequently, there has been a concerted effort to understand and exploit the ability of microorganisms for the enzymatic deconstruction of the plant cell wall. The characterisation of CAZymes are therefore essential in order to use them as biotechnological tools for the processing and lignocellulose and its components. The overall aim of this thesis was to investigate the application of such enzymes in biotechnology and expand upon the knowledge of CAZymes activities.

The utilisation of enzymes for biocatalytic applications is a growing part of a global enzyme industry that constituted a market share of $ 8.18 billion in 2015 [1]. The growing nature of biocatalysis for the production of defined products is exemplified in chapter 3. The synthesis of β-glucans and cellulose by enzymatic means demonstrates the numerous ways that disparate types of enzymes can produce functionally similar products. Although the natural abilities of GHs and GTs can be harnessed for β-glucan production, the production of glycosynthases through minimal active site mutation highlights the utility of enzyme engineering in adding completely new and scientifically important activities. Furthermore, the multiple application of these (non)functionalised products in fields such as green energy, food science and nanotechnology show how CAZymes will play important roles in a future bio-based economy.

In order to facilitate the rapid discovery and subsequent directed evolution required for making enzymes useful tools in industry, high-throughput methods for the quick assessment of initial enzyme activities are required. Glycosyltransferases can synthesise highly regio- and stereospecific glycosidic bonds between glycan acceptors and activated sugar donors, making them excellent candidates for

biocatalysis. In chapter 4, imidazolium-based tags (ITags) were employed to enable the screening of glycosyltransferases in order to identify a candidate for glucose polymerisation. The excellent ionisation of the ITag enabled the sensitive, rapid detection of reaction products between a variety of labelled glycosides and UDP donors. One GT, LgtB showed a broad donor scope with the ability to transfer UDP-Glc, UDP-Gal, UDP-GalNAc and UDP-Xyl. Additionally LgtB was able to transfer Glc and Gal to a variety of functionalised and non-functionalised acceptors, such as those containing a pNP group. However, of most interest in this study, was the glucose polymerisation activity of LgtB. This polymerisation activity within in vitro reactions has previously only been confined to complex, membrane-bound systems that have been extracted from the cell wall of plants or bacteria, as opposed to simple heterologous expression within *E. coli*. The β-1,4-linked, ITagged cello-oligosaccharides produced by LgtB can be used as substrates for cellulolytic enzymes in a sensitive MALDI-TOF based assay, reporting both the activity and selectivity of cellulolytic enzymes such as GHs and LPMOs. As no crystal structure exists for LgtB, or a closely related enzyme, no studies into molecular modelling experiments were undertaken. Going forward, crystallography of LgtB would therefore be of great interest to understand the catalysis underpinning both the substrate promiscuity and glucose polymerisation ability described herein. Such work would also enable mutagenesis work to identify key residues, and perhaps engineer the protein towards specific DPs or to increase yields beyond those of analytical scale magnitude.

The ability to perform molecular docking proved useful in the characterisation of XynD, described in chapter 5. XynD is upregulated during the degradation of lignocellulose by *A. niger* and was initially hypothesised as an arabinoxylan active enzyme. Phylogenetic analysis of the subfamily in which XynD resided showed clear clustering of known activities, and enabled the prediction of its β-galactofuranosidase ability— that was supported by modelling and molecular docking studies. Phylogenetic analysis also showed that an amino acid in the active site is consistently different between arabinofuranosidases (T) and galactofuranosidases (W) and suggests a key role in substrate specificity. To

examine this in future work, mutation of this one residue (W219 in XynD) to T could potentially alter the specificity from Gal*f*ase to Ara*f*ase and show how these enzymes have diverged within their evolutionary history. A wide variety of potential natural substrates were tested against XynD, including β-(1,5), β-(1,6) and α-(1,2)-linked Gal*f*-containing structures, with activity demonstrated against pNP-Gal*f*. This therefore suggests that the natural substrate of the enzyme is a disaccharide and/or a structure with a differing linkage type to those tested. The potential biological role of XynD may be as a secondary enzyme in the deconstruction of the cell wall during division, active on smaller glycans released during cell division. This could be investigated by incubating XynD against pre-digested (chemically or biologically) samples of the *A. niger* cell wallfrom different morphological stages, to gain insight in how this enzyme might act temporally.

Chapter 6 focussed on the assessment of Tx-Est1 and its ability to augment ferulic acid release from wheat bran. Improved biochemical assessment revealed temperature and pH optima in Tx-Est1 as well as circular dichroism data revealing a melting temperature of ~ 60 ˚C. Tx-Est1 was active against both single and double ferulated oligosaccharides, and released ferulic acid from plant biomass at levels similar to other bacterial feruloyl esterases. As such, Tx-Est1 represents a thermotolerant candidate for the production of dietary fibre, via its processing of ferulated xylooligosaccharides and could play a useful role in the food and drink industry. Our studies predominantly focussed on wheat bran derived oligosaccharides and other plant polymers, however to fully appreciate the role of this enzyme against wheat bran analysis of its interaction with intact plant biomass would be useful. Time-of-Flight Secondary Ion Mass Spectrometry (ToF-SIMS) is a surface-sensitive analytical method that would enable analysis of the mass changes on the surface of wheat bran associated with the interaction of Tx-Est1. Such a method would inform what chemical changes occur to the bran surface during ferulic acid release. Labelling of the protein followed by microscopy could also inform where on the wheat bran surface Tx-Est1 localises and suggest potential macromolecular structures to which the enzyme shows a preference.

The future of research regarding the structure, function and biocatalytic utilisation of CAZymes is one of great opportunity and major data acquisition. The 21st century has heralded a revolution in genetics and biotechnology and has removed many barriers to once laborious, expensive and time-intensive experiments. Next generation sequencing has enabled the rapid deconvolution of genes and genomes which, when coupled with now routine technologies such as CRISPR, has enabled genetic modification to reach an industrial scale. Whilst DNA synthesis and sub-cloning has opened up new possibilities in generating numerous protein mutants in a matter of days. Artificial intelligence programs, such as the AlphaFold [2], show great promise in predicting the structure of proteins to the same degree of accuracy as X-ray crystallography or cryo-EM and hints at the dominance of computational biology over traditional lab-based techniques. The aforementioned technologies will be fundamental in understanding CAZymes at both an individual and meta-analytical level, enabling quicker characterisation and modification of CAZymes and maximising their employment in the valorisation of plant lignocellulose or production of high-value chemicals.

References

[1] A. Pellis, S. Cantone, C. Ebert, L. Gardossi, Evolving biocatalysis to meet bioeconomy challenges and opportunities, N. Biotechnol. 40 (2018) 154–169.

[2] M.A. Pak, K.A. Markhieva, M.S. Novikova, D.S. Petrov, I.S. Vorobyev, E.S. Maksimova, F.A. Kondrashov, D.N. Ivankov, Using AlphaFold to predict the impact of single mutations on protein stability and function, BioRxiv. (2021) 2021.09.19.460937.

Appendix

Selective oxidation of N-glycolyl neuraminic acid using an engineered galactose oxidase variant

Ashley P. Mattey, William R. Birmingham, Peter Both, Nico Kress, Kun Huang, Jolanda M. van Munster, Gregory S. Bulmer, Fabio Parmeggiani, Josef Voglmeir, Juana E. R. Martinez, Nicholas J. Turner, Sabine L. Flitsch, **ACS Catal**. 2019, **9**, 8208–8212.

Abstract

N-Glycolylneuraminic acid (Neu5Gc) is a common cell surface ligand in animals which is not biosynthesized in humans, but can be acquired in human tissue from dietary sources such as red meat. It is important to understand the relevance of this potentially immunogenic glycan on human health, and selective detection methods are needed that can distinguish Neu5Gc from its biosynthetic precursor common in humans, i.e. N-acetylneuraminic acid (Neu5Ac). Here, we demonstrate that Neu5Gc can be selectively oxidised by an engineered variant of galactose oxidase without any reaction towards Neu5Ac. Oxidation of Neu5Gc itself allowed for the full spectroscopic characterization of the aldehyde product. In addition, we show that Neu5Gc is also oxidised when part of a typical animal oligosaccharide motif and when attached to a protein-linked N-glycan. Oxidation of Neu5Gc introduces bioorthogonal functionality that can be exclusively labelled. We demonstrate that in combination with sialidase mediated hydrolysis, this two-enzyme system can provide a useful tool for the selective detection of Neu5Gc in complex biological samples such as the biopharmaceutical alpha acid glycoprotein.

Enzymatic synthesis of N-acetyllactosamine from lactose enabled by recombinant β1,4-galactosyltransferases

Kun Huang, Fabio Parmeggiani, Helene Ledru, Kristian Hollingsworth, Jordi Mas Pons, Andrea Marchesi, Peter Both, Ashley P. Mattey, Edward Pallister, Gregory S. Bulmer, Jolanda M. van Munster, W. Bruce Turnbull, M. Carmen Galan and Sabine L. Flitsch **Org. Biomol. Chem.**, 2019,**17**, 5920-5924

Abstract

Utilising a fast and sensitive screening method based on imidazolium-tagged probes, we report unprecedented reversible activity of bacterial β1,4-galactosyltransferases to catalyse the transgalactosylation from lactose to N-acetylglucosamine to form Nacetyllactosamine in the presence of UDP. The process is demonstrated by the preparative scale synthesis of pNP-β-LacNAc from lactose using β1,4-galactosyltransferase NmLgtB-B as the only biocatalyst.