Analysis of fungal high-mannose structures using CAZymes

Kołaczkowski, Bartłomiej M.; Jørgensen, Christian I.; Spodsberg, Nikolaj; Stringer, Mary A.; Supekar, Nitin T.; Azadi, Parastoo; Westh, Peter; Krogh, Kristian B.R.M.; Jensen, Kenneth

Published in:
Glycobiology

DOI:
10.1093/glycob/cwab127

Publication date:
2022

Document Version
Peer reviewed version

Citation for published version (APA):
Kołaczkowski, B. M., Jørgensen, C. I., Spodsberg, N., Stringer, M. A., Supekar, N. T., Azadi, P., Westh, P., Krogh, K. B. R. M., & Jensen, K. (2022). Analysis of fungal high-mannose structures using CAZymes. *Glycobiology*, *32*(4), 304-313. https://doi.org/10.1093/glycob/cwab127

Analysis of fungal high mannose structures using CAZymes

Kołaczkowski, Bartłomiej M.; Jørgensen, Christian I.; Spodsberg, Nikolaj; Stringer, Mary A.; Supekar, Nitin T.; Azadi, Parastoo; Westh, Peter; Krogh, Kristian B. R. M.; Jensen, Kenneth

Published in:
Glycobiology

Link to article, DOI:
10.1093/glycob/cwab127

Publication date:
2022

Document Version
Peer reviewed version

Analysis of fungal high mannose structures using CAZymes

Bartłomiej M. Kołaczkowski[1], Christian I. Jørgensen[2], Nikolaj Spodsberg[2], Mary A. Stringer[2], Nitin T. Supekar[3], Parastoo Azadi[3], Peter Westh[4], Kristian B. R. M. Krogh[2], Kenneth Jensen[*,2]

[1] Roskilde University, Department of Science and Environment, Universitetsvej 1, Building 28, 4000 Roskilde, Denmark

[2]Novozymes A/S, Biologiens Vej 2, 2800 Kongens Lyngby, Denmark

[3]Complex Carbohydrate Research Center, University of Georgia, Athens, 30602 Georgia, United States

[4]Technical University of Denmark, Department of Biotechnology and Biomedicine, Building 224, 2800 Kongens Lyngby, Denmark

[*]To whom correspondence should be addressed: Tel: +45 30770529; e-mail: khjn@novozymes.com

Running head:

Analysis of fungal high mannose structures using CAZymes

Keywords: *Aspergillus oryzae* / cellobiohydrolase / fungal glycoproteins / glycoside hydrolase / *O*-glycosylation

Supplementary data included:

Figure S1 – S8 and Table SI-SII

Abstract

Glycoengineering ultimately allows control over glycosylation patterns to generate new glycoprotein variants with desired properties. A common challenge is glycan heterogeneity, which may affect protein function and limit the use of key techniques such as mass spectrometry. Moreover, heterologous protein expression can introduce non-native glycan chains which may not fulfil the requirement for therapeutic proteins. One strategy to address these challenges is partial trimming or complete removal of glycan chains, which can be obtained through selective application of exo-glycosidases. Here, we demonstrate an enzymatic *O*-deglycosylation toolbox of a GH92 α-1,2-mannosidase from *Neobacillus novalis*, a GH2 β-galactofuranosidase from *Amesia atrobrunnea* and the jack bean α-mannosidase. The extent of enzymatic *O*-deglycosylation was mapped against a full glycosyl linkage analysis of the *O*-glycosylated linker of cellobiohydrolase I from *Trichoderma reesei* (*Tr*Cel7A). Furthermore, the influence of deglycosylation on *Tr*Cel7A functionality was evaluated by kinetic characterization of native and *O*-deglycosylated forms of *Tr*Cel7A. This study expands structural knowledge on fungal *O*-glycosylation and presents a ready-to-use enzymatic approach for controlled *O*-glycan engineering in glycoproteins expressed in filamentous fungi.

Introduction

Filamentous fungi are known to be an attractive platform for heterogenous gene expression due to their ability to secrete large amounts of proteins in the growth media (Cherry and Fidantsef 2003). One common attribute of proteins produced by fungi is glycosylation. This post-translational modification occurs by attachment of a glycan chain either to the nitrogen atom of an asparagine residue in the consensus glycosylation motif Asn-X-Ser/Thr (*N*-glycosylation) or

to the oxygen atom of a serine or a threonine residue (*O*-glycosylation) (Reily et al. 2019). In humans and higher eukaryotes, *O*-linked glycosylation is typically of the mucin- or α-dystroglycan-type (Hopkins et al. 2015), whereas in filamentous fungi, the *O*-glycosylation core is composed of mannose chains, optionally capped with a terminal galactofuranose (t-Gal*f*) residues as demonstrated for some glycoproteins produced in *Aspergillus* spp (Goto 2007).

Fungal expression hosts have been widely applied in the production of protein-based therapeutics for use in humans due to the overall ease of scale-up, high titers and the ability to preserve glycosylation. However, non-native glycosylation of a heterologously produced protein can impact its half-life, cellular targeting, immunogenicity or therapeutic activity (Gerngross 2004; Gupta and Shukla 2018). Glycosylation also serves important roles in the function of industrial enzymes, including modification of activity (Adney et al. 2009; Rubio et al. 2019), proteolytic- and thermal stability (Amore et al. 2017) and substrate affinity (Kołaczkowski et al. 2020). In these cases, precise identification of the purified glycoprotein and the composition of its glycoforms are required. Mass spectrometry (MS) is well suited for this purpose, however, result interpretation may be challenging due to unresolved signal in mass spectra imposed by glycan heterogeneity. This heterogeneity is especially challenging for electrospray ionization-mass spectrometry (ESI-MS), therefore the removal of glycans is often a prerequisite to obtain a lower number of possible glycoforms and higher ionization efficiency (Balaguer and Neusüss 2006).

Multiple strategies for glycoengineering of heterologous proteins are available, such as site-directed mutagenesis of consensus glycosylation motifs, manipulation of genes encoding glycosylation pathways or changing the growth medium (Gupta and Shukla 2018). After glycoprotein production, the undesired glycans can be removed using endo- or exo-glycosidases, without disruption of the protein backbone. Enzymatic removal of *N*-linked glycans is well

established with enzymes like endoglycosidase H (endo H) or peptide-*N*-glycosidase F (PNGase F) that hydrolyze the glycosidic bond between two *N*-acetylglucosamine (GlcNAc) units or the innermost GlcNAc and the asparagine residue, respectively (Freeze and Kranz 2010). Modification of the fungal *O*-glycosylation pattern is more challenging as there is no universal endo-type enzyme. Rather, this must be approached with glycan 'trimming' by application of exo-glycosidases. The most commonly applied enzyme is a jack bean α-mannosidase (JBM) with broad specificity towards α-1,2/α-1,3/α-1,6 glycosidic bonds between mannose (Man) residues. This enzyme is commercially available as a crude cellular extract, which often exhibits low efficiency or unwanted proteolytic side-activity (Hopkins et al. 2015). Recently, α-mannosidase from *Bacteroides* spp. have been demonstrated to be active on α-linked mannosyl residues present in high-mannose *N*-glycans that are often structurally similar to *O*-glycans from filamentous fungi. (Cuskin et al. 2015; Briliūtė et al. 2019). To expand the variety of enzymes targeting glycosidic bonds in *O*-glycan chains, an array of exo-glycosidases was screened on a well-described model glycoprotein produced in the filamentous fungus *Aspergillus oryzae*. Specifically, we used the cellobiohydrolase from *Trichoderma reesei* (*Tr*Cel7A). *Tr*Cel7A harbors a peptide linker with a high number of serine and threonine residues for which the presence and structure of *O*-glycans have been characterized, (Harrison et al. 1998; Amore et al. 2017). We identified an enzymatic toolbox composed of three enzymes capable of reducing *O*-glycosylation in fungal expressed glycoproteins: a GH92 α-1,2-mannosidase from *Neobacillus novalis* (*Nn*GH92), a GH2 β-galactofuranosidase from *Amesia atrobrunnea (Aa*GH2*)* and a jack bean α-mannosidase.

Results

*Tr*Cel7A wild-type (WT) and a variant *Tr*Cel7A Δ*N*-glyc, with the three known *N*-glycosylation sites removed by the point mutations (N45Q, N270Q, N384Q), were successfully expressed in *Aspergillus oryzae,* and subsequently purified and characterized as described elsewhere (Kołaczkowski et al. 2020). To determine the extent and composition of *O*-glycans in the linker region of *Tr*Cel7A, *O*-glycomics analysis was performed. *O*-linked glycans were released from *Tr*Cel7A Δ*N*-glyc by reductive β-elimination. In this method, *O*-glycan release was catalyzed by addition of NaOH. In the resulting alkaline environment, the oligosaccharides are prone to degradation by "peeling" reactions targeting the reducing end of the released *O*-glycans. To prevent unwanted degradation, a reducing agent, $NaBH_4$, was employed to bind and stabilize the reducing end. (Shajahan, Heiss, et al. 2017; Shajahan, Supekar, et al. 2017; Shajahan et al. 2020). Among the released *O*-glycans, only hexose (Hex) residues were identified, and these were comprised of Hex_{1-14}. The dominant species comprised 95% of the total *O*-glycans which consisted of Hex_1-Hex_5 with Hex_2 as the most prevalent *O*-glycan chain (Figure S1). Monosaccharide composition analysis indicated the presence of mannose (Man) and galactose (Gal) in a ratio of 1.3:1. Based on glycosyl linkage analysis, the main *O*-glycans components were terminal mannopyranose (t-Man, 45%), 2-linked mannopyranose at the reducing terminus (red-2-Man, 31%), 4-linked galactopyranose (4-Gal*p*, 12%) and terminal galactofuranose (t-Gal*f*, 5%) residues (Figure S2, Table SI). The reduction (by $NaBH_4$) of the released *O*-glycans in the β-elimination reaction resulted in conversion of their proximal ("reducing end") residues to alditols. As a result, since alditols are not in ring form, the subsequent permethylation reaction (here using red-2-Man as an example) installed methyl groups on O-1 and O-5, as well as on O-3, O-4, and O-6, but not on O-2 of this 2-linked Man residue. This is in contrast to any residues in the pyranose ring form, which all end up with acetyl groups on O-1 and O-5, regardless of whether they come from the reducing end or not. The fact that the partially methylated alditol acetate

(PMAA) of red-2-Man contained 5 methyl groups and only 1 acetyl group caused it to elute much earlier than the other PMAAs in this study which had at least 2 acetyl groups. It also led to characteristic changes in the mass spectrum (Fig S2, Table SI). The structure of the individual PMAAs giving rise to the proposed linkage positions can be found in Table SI.

Enzymes with putative activity towards the identified monosaccharide residues and their glycosyl linkages (Table SI) were screened for their ability to trim *O*-glycans from the linker region of *Tr*Cel7A. These enzymes were selected based on their sequence homology to enzymes with activity towards α-linked mannopyranosyl and galactosyl residues already reported in literature (Goto 2007; Cuskin et al. 2015; Matsunaga et al. 2015). This led to the identification of a putative α-1,2-mannosidase (*Nn*GH92) and a putative β-galactofuranosidase (*Aa*GH2). Purity of the studied enzymes were evaluated by SDS-PAGE analysis (Figure S3).

According to the CAZy database classification (Lombard et al. 2014), the GH92 family is comprised of exo-acting α-mannosidases with specificity towards α-1,2-, α-1,3-, α-1,6-linked mannooligosaccharides. These linkages are found in α-mannan, a polysaccharide attached as a *N*-glycan to the yeast cell wall proteins (Orlean 2012). Multiple genes encoding GH92 α-mannosidases have been identified in the human gut bacterium *Bacteroides thetaiotaomicron* (Zhu et al. 2010; Cuskin et al. 2015), supporting its ability to grow on yeast *S. cerevisiae* α-mannan (Cuskin et al. 2015). In synergy with other α-mannan-degrading enzymes, this bacterium deconstructs yeast α-mannan to use it as a carbon source. Other bacterial GH92 α-mannosidases have been characterized which are involved in *N*-glycan processing (Robb et al. 2017; Li et al. 2020). *Neobacillus novalis* is a common soil bacterium having a gene encoding a GH92 α-1,2-mannosidase, suggesting that this bacterium might be also involved in processing of high-mannose type glycans. Characterization of *Nn*GH92 by high-performance anion-exchange

chromatography (HPAEC) (see details in the SI) confirmed its α-1,2-mannosidase activity, as well as some degree of α-1,3-mannosidase activity if incubation time was prolonged significantly (17 hours). No activity was seen towards α-1,6-mannobiose (Figure S4).

GH2 is a much more diverse family with multiple activities ascribed to it, including β-galactofuranosidases and β-galactopyranosidases. The only characterized β-D-galactofuranosidase is a bacterial GH2 from *Streptomyces* spp., although homologs from filamentous fungi have been retrieved by blast searches (Matsunaga et al. 2015). β-galactofuranosidase activity has been reported for the filamentous fungi *Aspergillus* spp. (Wallis et al. 2001) and *Penicillium fellutanum* (Miletti et al. 1999), although no protein or gene responsible for this activity has been identified . Gal*f* is a component of galactomannoproteins side chains (galactofurans), found in the filamentous fungal cell wall, and *N*- and *O*-glycans of extracellular enzymes (Komachi et al. 2013). The GH2 used in this study is from the filamentous fungus *Amesia atrobrunnea*, and although its natural function has not been elucidated, it is likely involved in reducing the Gal*f*-containing glycoconjugates in certain stress-induced or unfavorable cultivation conditions.

Inspired by the published characterization of the *Streptomyces* spp. β-D-galactofuranosidase (Matsunaga et al. 2015), activity of *Aa*GH2 towards 4-nitrophenyl β-D-galactofuranoside, 4-nitrophenyl β-D-galactopyranoside and 4-nitrophenyl α-L-arabinofuranoside was determined (Figure S5). No activity was seen towards β-D-galactopyranoside, suggesting that *Aa*GH2 is not a β-D-galactopyranosidase, an otherwise common hydrolase activity in the GH2 family. In contrast to the GH51 and GH54 α-arabinofuranosidases from *Aspergillus niger* (Tefsen et al. 2012), *Aa*GH2 activity was demonstrated only on β-D-galactofuranoside and *Aa*GH2 did not show dual specificity for the structural similar α-L-arabinofuranoside (Figure S5), indicating that

*Aa*GH2 may be Gal*f* specific, like its counterpart from *Streptomyces* spp. A previous study using site-directed mutagenesis identified several aspartic acid and glutamic acid residues in the catalytic domain of *Streptomyces* spp. GH2 as having a critical function for its β-D-galactofuranosidase activity (Matsunaga et al. 2015). These residues are conserved in *Aa*GH2, which further supports its annotation as a β-D-galactofuranosidase (Figure S6).

To demonstrate the ability of *Nn*GH92 and *Aa*GH2 to hydrolyze *O*-linked glycans in fungal glycoproteins, *Tr*Cel7A WT was selected as a model system. *N*-linked glycans were removed from *Tr*Cel7A by treatment with endo H (Roche), either as a separate reaction or in combination with *Nn*GH92 and *Aa*GH2. The reaction mixtures were incubated for 19 h at 37 °C and then analyzed by SDS-PAGE (Figure 1A) and LC-MS (Figure 1B-F).

Each enzymatic deglycosylation step with endo H, *Nn*GH92 and *Aa*GH2, or a combination of both enzymatic treatments resulted in the reduction of the glycan content in *Tr*Cel7A. This was observed as a decrease in apparent molecular weight (Coomassie blue staining), with lower band intensity (acid-Schiff staining) by SDS-PAGE analysis (Figure 1A), and by changes in the mass profiles of intact protein LC-MS (Figure 1B-E). In this analysis, the mass profile of TrCel7A corresponded to the mass of the protein backbone and total number of sugar residues attached to it. However, intact protein LC-MS does not provide any information about the length or complexity of the *O*-glycan chains attached to *Tr*Cel7A. In *Tr*Cel7A, the deconvolution of the mass spectrum was only possible after the removal of *N*-glycans (Figure 1B-C), as the heterogenous nature of *N*-glycosylation gives a large multitude of glycoforms. Endo H treated *Tr*Cel7A demonstrated a range of *O*-linked glycoforms with the highest intensity peak corresponding to the protein with the expected 3×*N*-acetylhexosamine (HexNAc) positioned at the *N*-glycosylation sites and 19×Hex (Figure 1C). Because endo H is active only on *N*-linked

glycans, the remaining 19×Hex most likely correspond to *O*-glycans (Freeze and Kranz 2010). Further refinement could be achieved by the additional combined activity of *Nn*GH92 and *Aa*GH2, which resulted in a mass spectrum with a major peak corresponding to a glycoform with 3×HexNAc and 9×Hex (Figure 1E). *Nn*GH92 and *Aa*GH2 also lowered the glycan heterogeneity of *Tr*Cel7A, leading to the presence of one dominant glycoform (Figure 1E). The enzymatic treatment with either *Nn*GH92 or *Aa*GH2 alone did not reach the same level of deglycosylation as when both enzymes acted simultaneously (Figure S7). Other conditions were tested and neither pH 7.0 nor higher enzyme dose improved the overall *O*-deglycosylation activity of *Nn*GH92 and *Aa*GH2 (Figure S8).

The remaining 9×Hex (Figure 1E) may be distributed as single Man residues directly attached to the 2×Ser and 7×Thr residues in the *Tr*Cel7A linker region by Man-α-O-Ser/Thr linkages (Harrison et al. 1998; Amore et al. 2017), making them inaccessible to the *Nn*GH92. In literature, the GH38 (either jack bean or lysosomal) α-mannosidases have been described as being able of hydrolyzing this linkage (Gomathinayagam and Hamilton 2014; Hopkins et al. 2015). To test this hypothesis, two commercially available JBMs were applied to the *N*-deglycosylated *Tr*Cel7A, either alone or in combination with *Nn*GH92 and *Aa*GH2. On its own, JBM GKX-5010 (Agilent) was able to further deglycosylate the target protein, as indicated by the appearance of a predominant peak corresponding to the *Tr*Cel7A backbone with 3×HexNAc residues and 2×Man residues (Figure 1F). Interestingly, the combined action of JBM GKX-5010 with *Nn*GH92 and *Aa*GH2 resulted in an even more uniform glycan population, harboring fewer glycoforms. This is clearly illustrated by a reduction in higher mass glycoforms following the predominant peak of *Tr*Cel7A with 3×HexNAc + 2×Hex (Figure 1G). The addition of *Nn*GH92 and *Aa*GH2 in combination with endo H and JBM GKX-5010 (Figure 1G) lowered the number of higher mass

glycoforms when compared to the treatment solely with endo H and JBM GKX-5010 (Figure 1F). Moreover, in Figure 1G, the intensity of peaks corresponding to the *Tr*Cel7A backbone with 3×Hex to 5×Hex were increased, most likely pointing at an accumulation of a subset of glycan structures, resistant to hydrolysis by the enzymatic mixture presented in this study.

The other JBM product (M6944, Sigma) showed inferior *O*-deglycosylation effectiveness compared to the JBM GKX-5010 (Figure S7). After prolonged contact with JBM, the occurrence of multiple bands in SDS-PAGE suggested truncation of *Tr*Cel7A, maybe reflecting the presence of proteolytic enzymes (Gomathinayagam and Hamilton 2014) acting on the flexible, deglycosylated linker.

After the enzymatic deglycosylation of *Tr*Cel7A, the hydrolysates were analyzed using PA1/Dionex. The products released from *Tr*Cel7A by *Nn*GH92 and *Aa*GH2 were mannose and galactose (Figure 2). The combined use of *Aa*GH2 and *Nn*GH92 provided a higher total release of mannose compared to the sole use of the *Nn*GH92 (Figure 2). This indicated that a subset of the *O*-glycan chains was shielded from the *Nn*GH92, most likely by a steric hindrance from the t-Gal*f*. This prevented *Nn*GH92 from acting on these *O*-glycan chains, even during prolonged incubation. The observed release of mannose by *Nn*GH92 alone points at a population of *O*-linked glycans chains with terminal Man residues.

The impact of reduced *O*-glycosylation on *Tr*Cel7A was investigated by steady-state kinetics using the model cellulosic substrate Avicel (Figure 3). This demonstrated that enzymatic deglycosylation by *Nn*GH92 and *Aa*GH2 did not change the kinetic profile of *Tr*Cel7A as compared to the native glycosylated *Tr*Cel7A. In particular, the changes in the Michaelis constant (K_M) were not significantly different and only a minor difference in turnover frequency (expressed as V_{max}/E_0) was observed for the different glycoforms.

Discussion

In this study, the target glycoprotein, *Tr*Cel7A, was enzymatically deglycosylated through the removal of *N*-linked glycans by endo H and *O*-linked glycans by *Nn*GH92, *Aa*GH2 and JBM GKX-5010. In addition, *O*-glycans were investigated by *O*-glycomics, including reductive β-elimination, methylation and GC-MS analysis. This lead to the identification of the component glycan residues and their linkage positions, and elucidation of glycosidic linkages in the *O*-glycan chains (Shajahan, Heiss, et al. 2017). Based on these results, the predominant *O*-glycan structures in *Tr*Cel7A are summarized in Figure 4. From the glycosyl linkage analysis (Shajahan, Heiss, et al. 2017) of reduced glycans, three types of residues are commonly identified: residues at the reducing terminus, internal residues and residues at the non-reducing terminus. Residues at the reducing terminus are directly linked to the protein backbone, while residues at the non-reducing terminus are positioned as the end-capping monosaccharides in the glycan chain. If a residue is not identified at a terminal position, then this residue is located in the internal part of the glycan chain between non-reducing and reducing ends. The combination of endo H, *Nn*GH92 and *Aa*GH2 resulted in a major homogenous population of *Tr*Cel7A with a glycoform comprised of 3×HexNAc residues and 9×Hex residues (Figure 1E). The α-1,2-mannosidase activity of *Nn*GH92 (Figure S4) suggests that Man residues in *O*-glycans are connected mainly by α-1,2-glycosidic bonds. This is well reflected in the glycosyl linkage analysis that indicates red-2-Man (directly attached to the protein backbone) is linked to t-Man through an α-1,2-glycosidic bond, in the predominant *O*-glycan chain with DP of 2 (structure A, Figure 4). Moreover, *Nn*GH92 can most likely target glycosidic bonds of the internal 2,3-linked Man (structure C, Figure 4) based on its partial activity on α-1,3-mannobiose during prolonged reactions (17 h, Figure S4).

The activity of *Aa*GH2 indicated the presence of t-Gal*f* residues, which was also confirmed by the glycosyl linkage analysis (Table SI). The t-Gal*f* residues are presumably linked by β-1,6-glycosidic bonds as shown for other *O*-linked glycans from glycoproteins expressed in *Aspergillus* spp. (Goto 2007). t-Gal*f* residues are unlikely to exist as DP2 glycans as it would require t-Gal*f* to be linked through an α/β-1,2-glycosidic bond to red-2-Man, a linkage not previously reported in *O*-glycosylated proteins from *Aspergillus* spp. (Goto 2007; Komachi et al. 2013). Therefore, the t-Gal*f* residues appear to be primarily present in longer chains (DP>2), linked to internal residues as presented in structures D and F in Figure 4. As none of the internal residues were 5-linked (Table SI), this indicates that t-Gal*f* were linked through β-1,6-glycosidic bond.

The presence of the *O*-glycan linkage 4-linked galactopyranosyl (4-Gal*p*, Table SI) was not found by enzymatic hydrolysis using well-characterized β-galactopyranosidases (Table SII). This strongly suggests that *O*-glycans containing 4-Gal*p* are extended or capped with a β-1,4- or α-1,4-linked Man (structure B, Figure 4), presenting a glycosidic bond which could not be hydrolyzed by the *Nn*GH92 or the JBM GKX-5010. The presence of a capping residue attached to 4-Gal*p* glycan would create a glycan conformation that is not likely to be accommodated in the active site of the β-galactopyranosidase and therefore no β-galactopyranosidase activity was observed. Despite the significant presence of 4-Gal*p* (12%), all the internal residues can be present only in longer glycan chains with (DP>2) (12% relative abundance, Figure S1). Therefore, 4-Gal*p* and the other internal residues may have a rather low impact on the overall *O*-glycan heterogeneity.

The enzymatic hydrolysis of *O*-glycans with *Nn*GH92 and *Aa*GH2 led to the presence of a *Tr*Cel7A linker with the remaining Man residues directly linked to the Ser and Thr residues,

which can be further removed by the action of JBM GKX-5010 (Figure 1G). This treatment also exposed low abundant *O*-glycan structures, like B, E and F (Figure 4), which were not hydrolyzed by the *Nn*GH92, *Aa*GH2 and JBM GKX-5010. The appearance of high intensity peaks in Figure 1G may represent an enrichment of these non-hydrolyzed *O*-glycan structures including DP3 chains containing 4-Gal*p* (structure B, Figure 4).

The enzymatic toolbox defined by this study provides the necessary deglycosylation activities to reduce *O*-glycan heterogeneity and examine different profiles of *O*-glycosylation in other glycoproteins expressed in filamentous fungal hosts such as *Aspergillus* spp. (Figure 5). Most studies on the relationship between activity and glycosylation rely on variants with point mutations that prevent the addition of selected glycans (Beckham et al. 2012; Kołaczkowski et al. 2020). A new toolbox of enzymes that trim *O*-linked glycans under non-denaturing conditions allows for the characterization of proteins in their native form rather than using an 'all or nothing approach' as previously discussed by Beckham *et al*. (Beckham et al. 2012). Individual application of either *Nn*GH92, *Aa*GH2 or JBM may allow better control over the type and length of *O*-glycan structures. Having a homogenous glycan population, as demonstrated in Figure 1E, would help to investigate the functional changes between a protein with one major glycoform and a protein with a complex mixture of glycoforms. Modification of the amount of mannose and galactose present in the *Tr*Cel7A linker could simplify the functional studies of linker glycans in *Tr*Cel7A and potentially verify the computational work on this topic (Payne et al. 2013; Amore et al. 2017).

MD simulations have shown that *O*-glycosylation of the linker region of *Tr*Cel7A may modulate linker binding affinity to cellulose (Amore et al. 2017). In another study (Badino et al. 2017), site-directed mutagenesis of a Thr rich region (TTTT to PPGP) of the linker region limited the

extent of *O*-glycosylation. This resulted in a higher turnover, but reduced substrate affinity as compared to *Tr*Cel7A with its native *O*-glycosylation pattern. It has been speculated that interaction between linker glycans and cellulose appear to depend on the location of *O*-glycosylation sites and that changes in location may change structure and dynamics of the linker (Badino et al. 2017). The functional impact of posttranslational *O*-deglycosylation of *Tr*Cel7A by *Nn*GH92 and *Aa*GH2, was investigated by steady-state kinetics using the model cellulosic substrate Avicel (Figure 3) (Kołaczkowski et al. 2020). The derived K_m and V_{max} of the native and deglycosylated *Tr*Cel7A (Figure 3) did not reveal any major differences at the tested conditions. This suggests that the reduced DP and high homogeneity of *O*-glycans in the *O*-deglycosylated *Tr*Cel7A (Figure 1E) do not influence kinetics. Since the majority of *O*-glycans are located in the linker region of *Tr*Cel7A (Harrison et al. 1998; Amore et al. 2017) and the activity of *Nn*GH92 and *Aa*GH2 reduce the extent of *O*-glycans to DP1, one could speculate that the presence of a single Man residue at each glycosylation site is sufficient to maintain the proposed linker substrate interactions as suggested by (Amore et al. 2017; Badino et al. 2017). This may indicate that the extent of *O*-glycosylation plays a lesser role in *Tr*Cel7A kinetics than the impact of site occupancy at every glycosylation site in the linker region.

Complete removal of *O*-glycosylation may not be beneficial as *O*-glycosylation has been shown to protect the protein structure against proteolytic attack. This has been demonstrated (Amore et al. 2017) by an attempt to express *Tr*Cel7A with a non-glycosylated linker (Thr/Ser clusters in linker mutated to Ala residues). The deglycosylated enzyme was found to lose the CBM module, indicating a proteolytic susceptibility of the linker during enzyme expression. It has been proposed that glycosylation provides a steric hindrance around the peptide backbone (Solá and Griebenow 2009). Therefore, it is likely that at least one monosaccharide residue attached to each

glycosylation site in the exposed peptide, like *Tr*Cel7A linker, can provide a sufficient level of protein stability. Thus, enzymatic *O*-deglycosylation with *Nn*GH92 and *Aa*GH2 can be advantageous, as it simplifies the overall glycosylation profile while retaining some degree of protein stability. Most commercial exo-glycosidases are suited for mammalian-type *O*-glycan modification and the selection of enzymes for glycoproteins produced in filamentous fungi remains limited. The enzymes described in this study provide a way to trim *O*-glycan structures attached to glycoproteins produced by filamentous fungi. This will be relevant in studies on the influence of glycosylation on biochemical properties. Moreover, a low glycan heterogeneity will enable an analysis of glycoproteins in top-down MS and improve deconvolution of mass spectra. In proteomic studies with bottom-up MS, protein coverage can also be enhanced by increasing the population of non-glycosylated peptides.

Materials and Methods

Cloning, expression, and purification of *Aa*GH2 and *Nn*GH92

The data for GH2 β-galactofuranosidase from *Amesia atrobrunnea* (*Aa*GH2) was deposited in the GenBank at NCBI under accession number MW316479. The *Aa*GH2 was expressed as an extracellular enzyme in *Aspergillus oryzae* in a similar setup as described elsewhere (Borch et al. 2014). The fermentation broth was sterile filtrated and then 1.8 M ammonium sulphate was added. After filtration on 0.22 µm PES filter (Nalge Nunc International, Nalgene labware cat#595-4520), the filtrate was loaded onto a Phenyl Sepharose™ 6 Fast Flow column (high sub) (GE Healthcare, Piscataway, NJ, USA) equilibrated in 25 mM HEPES pH 7.0 (buffer A) with 1.8 M ammonium sulphate. After loading, unbound material was washed with 5 column volume (CV) of equilibration buffer, and bound proteins were eluted in buffer A. The fractions were pooled and

applied to a Sephadex™ G-25 (medium) (GE Healthcare, Piscataway, NJ, USA) column equilibrated in buffer A. The fractions were applied to a SOURCE™ 15Q (GE Healthcare, Piscataway, NJ, USA) column equilibrated in buffer A, and bound proteins were eluted with a linear gradient from 0-1000 mM sodium chloride over 10 CV. Fractions were analyzed by SDS-PAGE, and fractions containing the enzyme were combined.

The data for GH92 α-1,2-mannosidase from *Neobacillus novalis* (*Nn*GH92) was deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under accession number LR963497.1. The *Nn*GH92 expressed as an extracellular enzyme in *Bacillus subtilis* in a similar setup as described elsewhere (Jensen et al. 2010) with the following modifications. The native signal peptide was replaced by the Alcalase signal peptide followed by a histidine tag (6xHis + PR) resulting in a N-terminal sequence of MKKPLGKIVASTALLISVAFSSSIASAHHHHHHPR. The fermentation broth was sterile filtrated and then 500 mM NaCl was added and adjusted to pH 7.5/NaOH. The sample was loaded onto a Ni-Sepharose™ 6 Fast Flow column (GE Healthcare, Piscataway, NJ, USA) equilibrated in 50 mM HEPES, pH 7.5 with 500 mM NaCl (buffer A). After loading, the column was washed with 10 CV of buffer A, and bound proteins were eluted with 500 mM imidazole in buffer A. The fractions containing the enzyme were pooled and applied to a Sephadex™ G-25 (medium) (GE Healthcare, Piscataway, NJ, USA) column equilibrated and eluted in 50 mM HEPES pH 7.5. Fractions were analyzed by SDS-PAGE, and fractions containing the enzyme were combined.

Cloning, expression and purification of *Tr*Cel7A WT and variant *Tr*Cel7A Δ*N*-glyc

Trichoderma reesei Cel7A (*Tr*Cel7A) enzymes were cloned and expressed in *Aspergillus oryzae* as described earlier (Borch et al. 2014). The changes in *N*-glycosylation were introduced by point mutations of asparagine to glutamine in the consensus *N*-glycosylation motifs. This resulted in a

variant *Tr*Cel7A Δ*N*-glyc with the following mutations (N45Q, N270Q, N384Q). Numbering corresponds to the protein sequence of P62694 (UniProt entry). Enzyme purification was performed according to the protocol described elsewhere (Sørensen et al. 2015).

Analysis of released *O*-glycans from *Tr*Cel7A Δ*N*-glyc

O-glycans were released by reductive β-elimination reaction. The solution containing the peptides and glycopeptides was made alkaline with 250 µL 50 mM NaOH, and 250 µL of a 76 mg/mL solution of sodium borohydride in 50 mM NaOH. The resulting mixture was incubated at 45 °C for 18 h. The samples were cooled, neutralized by 10 % acetic acid, passed through Dowex H^+ resin column, lyophilized and borates were removed under a stream of nitrogen. The released, reduced glycans were permethylated for structural characterization by mass spectrometry as reported previously (Shajahan, Heiss, et al. 2017; Shajahan, Supekar, et al. 2017; Shajahan et al. 2020). Briefly, the dried eluate was dissolved with dimethyl sulfoxide and methylated by using methyl iodide on DMSO/NaOH mixture. The reaction was quenched with water and the reaction mixture was extracted with methylene chloride and dried. The permethylated glycans were dissolved in methanol and analyzed using ESI-MS.

Glycosyl linkage analysis

For determination of sugar linkages, partially methylated alditol acetates were prepared from the released, reduced and permethylated *O*-glycans. Briefly, permethylated glycans were hydrolyzed with 2 M TFA at 121 ˚C for 2 h, followed by reduction with $NaBH_4$ in 50 mM NaOH (400 µL, 10 mg/mL) and acetylation with acetic anhydride/TFA (1:1, v/v) at 45 °C for 25 min. The partially methylated alditol acetates thus obtained were analyzed by GC-MS.

Enzyme assay conditions

To confirm the theoretical activity of *Nn*GH92, 66 nM of the enzyme was incubated with 1 mM mannobiose solubilized in the assay buffer (50 mM MES pH 6.0, 50 mM NaCl, 2 mM $CaCl_2$, 0.01 % Triton X-100) for 1 hour. Then the enzymes were inactivated with 0.15 M NaOH and the reaction hydrolysates were analyzed on a Dionex HPAEC (as described later).

*Aa*GH2 activity was tested on three different 4-nitrophenyl-linked substrates: α-L-arabinofuranoside β-D-galactopyranoside, β-D-galactofuranoside. Reaction mixture was composed of 20 µL of diluted enzymes and 120 µL 1 mM substrate dissolved in the assay buffer. The reaction mixture was incubated for 15 min at 37 °C and then quenched with 100 µL 1 M glycine pH 10 buffer. The absorption of the released 4-nitrophenol was measured at 405 nm using a plate reader (Spectra Max 3; Molecular Devices, Sunnyvale, CA, USA).

Standard deglycosylation mixture contained 300 µg of *Tr*Cel7A dissolved in 50 mM sodium acetate pH 5.0 buffer with addition of 200 mM NaCl and 2 mM $CaCl_2$ incubated with 7.5 µg of *Nn*GH92 and/or *Aa*GH2. The reaction mixture had a final volume of 250 µL. Commercially available JBMs (M6944, Sigma and GKX-5010, Agilent) were applied as recommended by the manufacturer.

Intact protein mass spectrometry (LC-MS)

Enzymatically deglycosylated enzymes were analyzed for their intact molecular weight using a MAXIS II electrospray mass spectrometer (Bruker Daltonik GmbH, Bremen, Germany). The samples were diluted to 0.1 mg/mL in 50 mM sodium acetate pH 5.0 buffer with addition of 200 mM NaCl and applied to an AdvanceBio Desalting-RP column (Agilent Technologies). Samples were eluted from the column with an acetonitrile linear gradient from 5 to 95% (v/v) and introduced to the electrospray source with a flow of 0.4 mL/min by an Ultimate 3000 LC system

(Thermo Fisher Scientific). Data analysis was performed with DataAnalysis version 4.3 (Bruker Daltonik GmbH, Bremen, Germany). The molecular weight of the samples was calculated by deconvolution of the raw data in the range 30 000 to 70 000 Da.

Dionex HPAEC

Products of mannobiose and *Tr*Cel7A *O*-linked glycans hydrolysis were analyzed on a Dionex ICS-3000 HPAEC-PAD (Thermo Fisher Scientific, Waltham, MA) and separated on a CarboPac PA-1 column. Samples were eluted using the following multistep gradient program at a flow rate 0.8 mL min^{-1}: 15 mM NaOH (0–4.5 min), 35 mM NaOH (4.5–7 min), 25 mM NaOAc + 75 mM NaOH (7–10 min), 175 mM NaOAc + 75 mM NaOH (10–20 min), 15 mM NaOH (20-30 min). As standards, glucose, mannose, galactose, were used. The sugars were purchased from Sigma. Data were analyzed using the instrument software Chromeleon 7.2 (Thermo Fisher Scientific, Waltham, MA). Prior to the experiment, *N*-glycans were removed from TrCel7A by overnight endo H treatment and separated with 10 kDa MWCO VivaSpin columns.

Kinetic characterization of *Tr*Cel7A before and after enzymatic *O*-deglycosylation

Intact (Figure 1B) and *O*-deglycosylated (*Nn*GH92 and *Aa*GH2) *Tr*Cel7A was subjected to a steady-state kinetic characterization on Avicel PH101 (Sigma-Aldrich, St. Louis, MO). This characterization was done in an identical setup as described elsewhere (Kołaczkowski et al. 2020). The substrate was washed in MiliQ water and then washed in 50 mM sodium acetate buffer pH 5.0. After the washes, it was stored in the same buffer in the presence of 5 mM sodium azide. A single experimental setting was investigated in which intact and deglycosylated *Tr*Cel7A were saturated with Avicel (Michaelis-Menten kinetics). In this experiment, 230 µL substrate aliquots from the Avicel stock (vigorously stirred) with various concentrations were pipetted into 96-well

plate (96F 26960 Thermo Scientific, Waltham, MA). The hydrolysis reaction was initiated by addition of 20 µL enzyme with a final concentration of 0.1 µM. The reactions were conducted for 1 h at 25 °C mixed at 1100 rpm and quenched using a centrifuge for 3 min at 2500 *g*. Then 60 µL supernatant was withdrawn from each reaction and mixed with 90 µL p-hydroxybenzoic acid (PAHBAH) to perform reducing sugar assay (Lever 1973; Sørensen et al. 2015). The obtained colored products were measured at 405 nm in plate reader (Spectra Max 3; Molecular Devices, Sunnyvale, CA, USA). The absorbance readouts were recalculated to reducing sugar end concentration using a standard curve prepared with cellobiose (0-1 mM). The reaction rate curves were fitted with Michaelis-Menten equation: $\nu_{ss}/E_0 = V_{max}S_0/(K_M + S_0)$; using Origin Pro (version 2019, OriginLab Corporation, Northampton, MA, USA). All reactions were performed in triplicate.

Funding

Funding provided by the Roskilde University, Novozymes A/S, Innovation Fund Denmark [Grant number: 5150-00020B], the Novo Nordisk Foundation [Grant number: NNF15OC0016606 and NNFSA170028392] and the Carlsberg Foundation. This research was also supported in part by National Institute of General Medical Sciences R24 grant number R24GM137782 at the Complex Carbohydrate Research Center.

Acknowledgement

Research Associate Clive Phipps Walter (Novozymes) is thanked for technical assistance with mass spectrometry analysis. We would like to thank assistant Technical Director, PhD, Christian

Heiss (Complex Carbohydrate Research Center) for technical support and discussion on the interpretation of the *O*-glycomics results.

Data Availability Statement

The data for a GH92 α-1,2-mannosidase from *Neobacillus novalis* (*Nn*GH92) was deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under accession number LR963497.1 (GenBank sequence ID). The data for GH2 β-galactofuranosidase from *Amesia atrobrunnea* (*Aa*GH2) was deposited in the GenBank at NCBI under accession number MW316479. All other data are included in the main article and supplementary materials.

Conflict of interest statement

The authors declare the following conflicts of interest: Christian I. Jørgensen, Nikolaj Spodsberg, Mary A. Stringer, Kristian B. R. M. Krogh, Kenneth Jensen work for Novozymes A/S, a major manufacturer of industrial enzymes.

Abbreviations

CBM, carbohydrate-binding module; CD, catalytic domain; endo H, endoglycosidase H; ESI, electrospray ionization; HPAEC-PAD, high-performance anion-exchange chromatography with pulsed amperometric detection; LC, liquid chromatography; MS, mass spectrometry; GlcNAc, *N*-acetylglucosamine; PNGase F, peptide-*N*-glycosidase F; SDS-PAGE, sodium dodecyl sulphate–polyacrylamide gel electrophoresis.

References

Adney WS, Jeoh T, Beckham GT, Chou YC, Baker JO, Michener W, Brunecky R, Himmel ME. 2009. Probing the role of *N*-linked glycans in the stability and activity of fungal cellobiohydrolases by mutational analysis. *Cellulose*. 16(4):699–709.

Amore A, Knott BC, Supekar NT, Shajahan A, Azadi P, Zhao P, Wells L, Linger JT, Hobdey SE, Wall TAV, et al. 2017. Distinct roles of *N*- and *O*-glycans in cellulase activity and stability. *Proc Natl Acad Sci*. 114(52):13667–13672.

Badino SF, Bathke JK, Sørensen TH, Windahl MS, Jensen K, Peters GHJ, Borch K, Westh P. 2017. The influence of different linker modifications on the catalytic activity and cellulose affinity of cellobiohydrolase Cel7A from *Hypocrea jecorina*. *Protein Eng Des Sel*. 30(7):495–501.

Balaguer E, Neusüss C. 2006. Glycoprotein characterization combining intact protein and glycan analysis by capillary electrophoresis-electrospray ionization-mass spectrometry. *Anal Chem*. 78(15):5384–5393.

Beckham GT, Dai Z, Matthews JF, Momany M, Payne CM, Adney WS, Baker SE, Himmel ME. 2012. Harnessing glycosylation to improve cellulase activity. *Curr Opin Biotechnol*. 23(3):338–345.

Borch K, Jensen K, Krogh K, McBrayer B, Westh P, Kari J, Olsen J, Sorensen T, Windahl M, Xu H. 2014. Cellobiohydrolase variants and polynucleotides encoding same. International Patent WO2014138672.

Briliūtė J, Urbanowicz PA, Luis AS, Baslé A, Paterson N, Rebello O, Hendel J, Ndeh DA, Lowe EC, Martens EC, et al. 2019. Complex *N*-glycan breakdown by gut *Bacteroides* involves an extensive enzymatic apparatus encoded by multiple co-regulated genetic loci. *Nat Microbiol*.

4(9):1571–1581.

Cherry JR, Fidantsef AL. 2003. Directed evolution of industrial enzymes: An update. *Curr Opin Biotechnol*. 14(4):438–443.

Cuskin F, Lowe EC, Temple MJ, Zhu Y, Cameron EA, Pudlo NA, Porter NT, Urs K, Thompson AJ, Cartmell A, et al. 2015. Human gut *Bacteroidetes* can utilize yeast mannan through a selfish mechanism. *Nature*. 517(7533):165–169.

Freeze HH, Kranz C. 2010. Endoglycosidase and glycoamidase release of *N*-linked glycans. *Curr Protoc Mol Biol*. 89(1):17.13A.1-17.13A.25.

Gerngross TU. 2004. Advances in the production of human therapeutic proteins in yeasts and filamentous fungi. *Nat Biotechnol*. 22(11):1409–1414.

Gomathinayagam S, Hamilton SR. 2014. In vitro enzymatic treatment to remove *O*-linked mannose from intact glycoproteins. *Appl Microbiol Biotechnol*. 98(6):2545–2554.

Goto M. 2007. Protein *O*-glycosylation in fungi: Diverse structures and multiple functions. *Biosci Biotechnol Biochem*. 71(6):1415–1427.

Gupta SK, Shukla P. 2018. Glycosylation control technologies for recombinant therapeutic proteins. *Appl Microbiol Biotechnol*. 102(24):10457–10468.

Harrison MJ, Nouwens AS, Jardine DR, Zachara NE, Gooley AA, Nevalainen H, Packer NH. 1998. Modified glycosylation of cellobiohydrolase I from a high cellulase- producing mutant strain of *Trichoderma reesei*. *Eur J Biochem*. 256(1):119–127.

Hopkins D, Gomathinayagam S, Hamilton SR. 2015. A practical approach for *O*-linked mannose removal: the use of recombinant lysosomal mannosidase. *Appl Microbiol Biotechnol*.

99(9):3913–3927.

Jensen K, Østergaard PR, Wilting R, Lassen SF. 2010. Identification and characterization of a bacterial glutamic peptidase. *BMC Biochem*. 11:47.

Kołaczkowski BM, Schaller KS, Sørensen TH, Peters GHJ, Jensen K, Krogh KBRM, Westh P. 2020. Removal of *N*-linked glycans in cellobiohydrolase Cel7A from *Trichoderma reesei* reveals higher activity and binding affinity on crystalline cellulose. *Biotechnol Biofuels*. 13(1):136.

Komachi Y, Hatakeyama S, Motomatsu H, Futagami T, Kizjakina K, Sobrado P, Ekino K, Takegawa K, Goto M, Nomura Y, et al. 2013. gfsA encodes a novel galactofuranosyltransferase involved in biosynthesis of galactofuranose antigen of *O*-glycan in *Aspergillus nidulans* and *Aspergillus fumigatus*. *Mol Microbiol*. 90(5):1054–1073.

Lever M. 1973. Calorimetric Determination and Acid Fluorometric with Hydrazide. *Anal Biochem*. 281:274–281.

Li Y, Li R, Yu H, Sheng X, Wang J, Fisher AJ, Chen X. 2020. *Enterococcus faecalis* α1–2-mannosidase (EfMan-I): an efficient catalyst for glycoprotein *N*-glycan modification. *FEBS Lett*. 594(3):439–451.

Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. 2014. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res*. 42(D1):490–495.

Matsunaga E, Higuchi Y, Mori K, Yairo N, Oka T, Shinozuka S, Tashiro K, Izumi M, Kuhara S, Takegawa K. 2015. Identification and characterization of a novel galactofuranose-specific β-D-galactofuranosidase from *Streptomyces* species. *PLoS One*. 10(9):1–16.

Miletti LC, Marino C, Mariño K, De Lederkremer RM, Colli W, Alves MJM. 1999. Immobilized

4-aminophenyl 1-thio-β-d-galactofuranoside as a matrix for affinity purification of an exo-β-d-galactofuranosidase. *Carbohydr Res*. 320(3–4):176–182.

Neelamegham S, Aoki-Kinoshita K, Bolton E, Frank M, Lisacek F, Lütteke T, O'Boyle N, Packer NH, Stanley P, Toukach P, et al. 2019. Updates to the Symbol Nomenclature for Glycans guidelines. *Glycobiology*. 29(9):620–624.

Orlean P. 2012. Architecture and biosynthesis of the *Saccharomyces cerevisiae* cell wall. *Genetics*. 192(3):775–818.

Payne CM, Resch MG, Chen L, Crowley MF, Himmel ME, Taylor LE, Sandgren M, Ståhlberg J, Stals I, Tan Z, et al. 2013. Glycosylated linkers in multimodular lignocellulose-degrading enzymes dynamically bind to cellulose. *Proc Natl Acad Sci U S A*. 110(36):14646–51.

Reily C, Stewart TJ, Renfrow MB, Novak J. 2019. Glycosylation in health and disease. *Nat Rev Nephrol*. 15(6):346–366.

Robb M, Hobbs JK, Woodiga SA, Shapiro-Ward S, Suits MDL, McGregor N, Brumer H, Yesilkaya H, King SJ, Boraston AB. 2017. Molecular Characterization of *N*-glycan Degradation and Transport in *Streptococcus pneumoniae* and Its Contribution to Virulence. Mitchell TJ, editor. *PLOS Pathog*. 13(1):e1006090.

Rubio MV, Terrasan CRF, Contesini FJ, Zubieta MP, Gerhardt JA, Oliveira LC, De Souza Schmidt Gonçalves AE, Almeida F, Smith BJ, De Souza GHMF, et al. 2019. Redesigning *N*-glycosylation sites in a GH3 β-xylosidase improves the enzymatic efficiency. *Biotechnol Biofuels*. 12(1):1–14.

Shajahan A, Heiss C, Ishihara M, Azadi P. 2017. Glycomic and glycoproteomic analysis of glycoproteins—a tutorial. *Anal Bioanal Chem*. 409(19):4483–4505.

Shajahan A, Supekar NT, Chapla D, Heiss C, Moremen KW, Azadi P. 2020. Simplifying Glycan Profiling through a High-Throughput Micropermethylation Strategy. *SLAS Technol Transl Life Sci Innov*. 25(4):367–379.

Shajahan A, Supekar NT, Heiss C, Ishihara M, Azadi P. 2017. Tool for Rapid Analysis of Glycopeptide by Permethylation via One-Pot Site Mapping and Glycan Analysis. *Anal Chem*. 89(20):10734–10743.

Solá RJ, Griebenow K. 2009. Effects of glycosylation on the stability of protein pharmaceuticals. *J Pharm Sci*. 98(4):1223–1245.

Sørensen TH, Cruys-Bagger N, Windahl MS, Badino SF, Borch K, Westh P. 2015. Temperature effects on kinetic parameters and substrate affinity of Cel7A cellobiohydrolases. *J Biol Chem*. 290(36):22193–22202.

Tefsen B, Lagendijk EL, Park J, Akeroyd M, Schachtschabel D, Winkler R, Die I van, Ram AFJ. 2012. Fungal α-arabinofuranosidases of glycosyl hydrolase families 51 and 54 show a dual arabinofuranosyl- and galactofuranosyl-hydrolyzing activity. *Biol Chem*. 393(8):767–775.

Varki A, Cummings RD, Aebi M, Packer NH, Seeberger PH, Esko JD, Stanley P, Hart G, Darvill A, Kinoshita T, et al. 2015. Symbol nomenclature for graphical representations of glycans. *Glycobiology*. 25(12):1323–1324.

Wallis GLF, Hemming FW, Peberdy JF. 2001. An extracellular β-galactofuranosidase from *Aspergillus niger* and its use as a tool for glycoconjugate analysis. *Biochim Biophys Acta - Gen Subj*. 1525(1–2):19–28.

Zhu Y, Suits MDL, Thompson AJ, Chavan S, Dinev Z, Dumon C, Smith N, Moremen KW, Xiang Y, Siriwardena A, et al. 2010. Mechanistic insights into a Ca^{2+}-dependent family of α-

mannosidases in a human gut symbiont. *Nat Chem Biol*. 6(2):125–132.

Figure legends

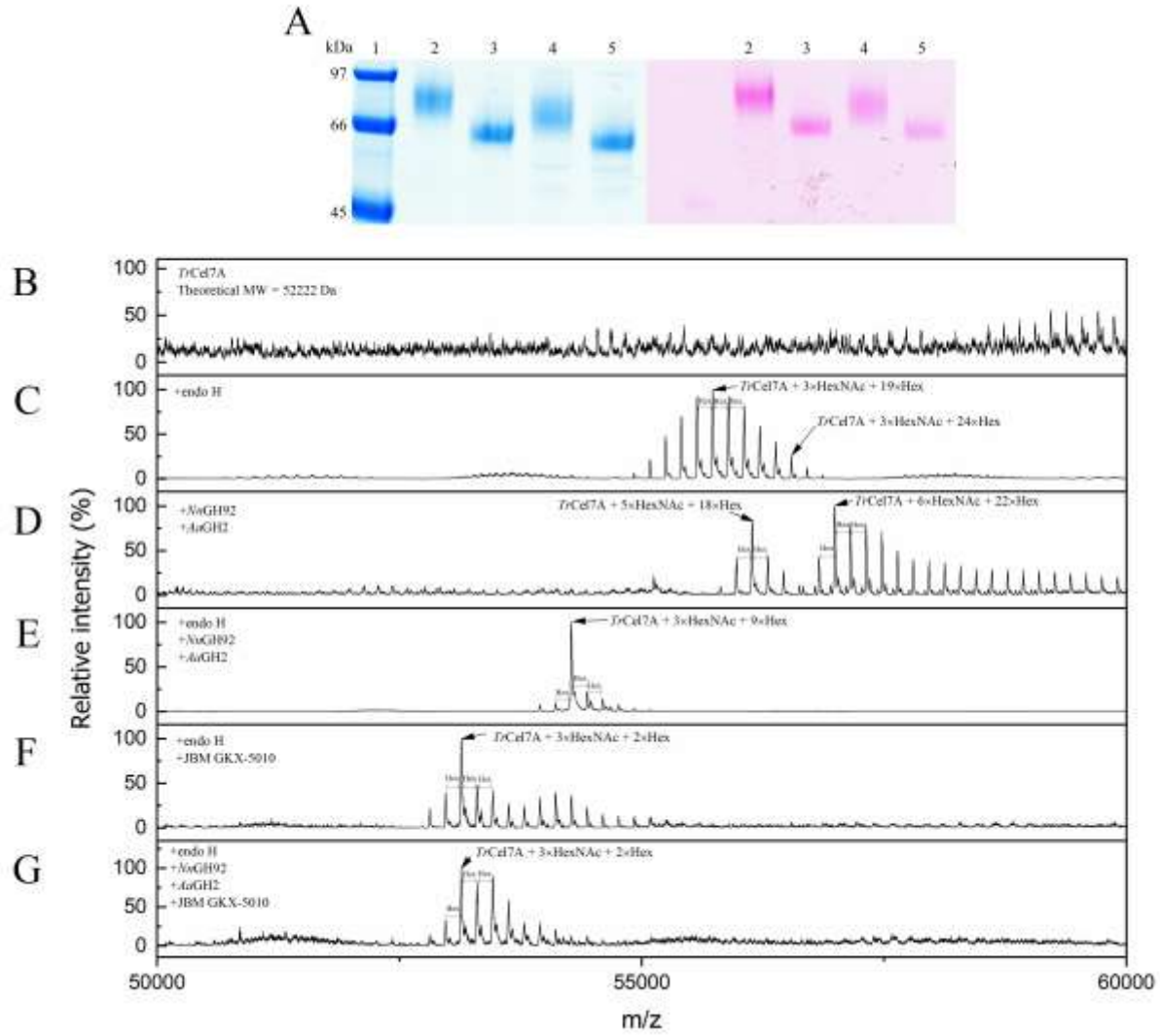

Figure 1. Evaluation of *Nn*GH92, *Aa*GH2 and JBM GKX-5010 deglycosylation effectiveness of *Tr*Cel7A WT.

(A) SDS-PAGE analysis: Coomassie blue staining (left) and periodic acid-Schiff staining method (right). Lane 1: Marker LMW (GE Healthcare) molecular weight standard; lane 2: *Tr*Cel7A; lane 3: +endo H; lane 4: +*Nn*GH92 and *Aa*GH2; lane 5: +endo H, *Nn*GH92 and *Aa*GH2. (B-F) LC-MS intact protein mass spectra of *Tr*Cel7A; (B) untreated; (C) +endo H; (D) +*Nn*GH92 and *Aa*GH2; (E) +*Nn*GH92, *Aa*GH2 and endo H; (F) JBM GKX-5010; (G) +endo H, *Nn*GH92, *Aa*GH2 and JBM GKX-5010. The peaks indicated with an arrow match the mass of *Tr*Cel7A + 3×HexNAc and the corresponding number of Hex. In the mass spectra (B-G) the adjacent peaks correspond to the mass difference of one hexose (162 Da). Theoretical molecular weight (MW) of *Tr*Cel7A was calculated based on the amino acid sequence without signal peptide from P62694 (UniProt entry).

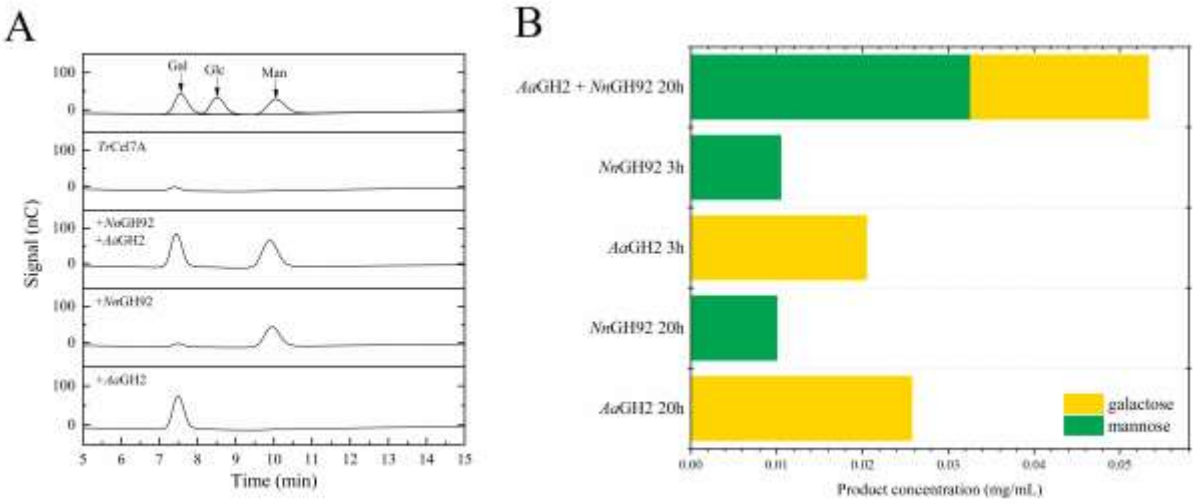

Figure 2. Product profile after enzymatic hydrolysis of endo H treated *Tr*Cel7A linker *O*-glycans with *Nn*GH92 and *Aa*GH2 measured by HPAEC-PAD Dionex PA1.

(A) HPAEC-PAD traces of 3 h long incubations with *Nn*GH92 and/or *Aa*GH2. The top panel presents monosaccharide standards, including galactose (Gal), glucose (Glc) and mannose (Man). *Tr*Cel7A panel correspond to a negative control. The subsequent panels represent 3 h enzymatic treatments of *Tr*Cel7A by the mixture of *Nn*GH92 and *Aa*GH2, and either *Nn*GH92 or *Aa*GH2. (B) *Tr*Cel7A was incubated with either *Nn*GH92 or *Aa*GH2 for 3 h and 20 h and the released mannose (green) and galactose (yellow) was quantified. The combined effect of *Nn*GH92 and *Aa*GH2 on mannose and galactose release from *Tr*Cel7A was evaluated by sequential addition of the enzymes. This was done by initially running a 3 h incubation with *Aa*GH2, followed by an extended incubation (17 h) in presence of both *Aa*GH2 and *Nn*GH92.

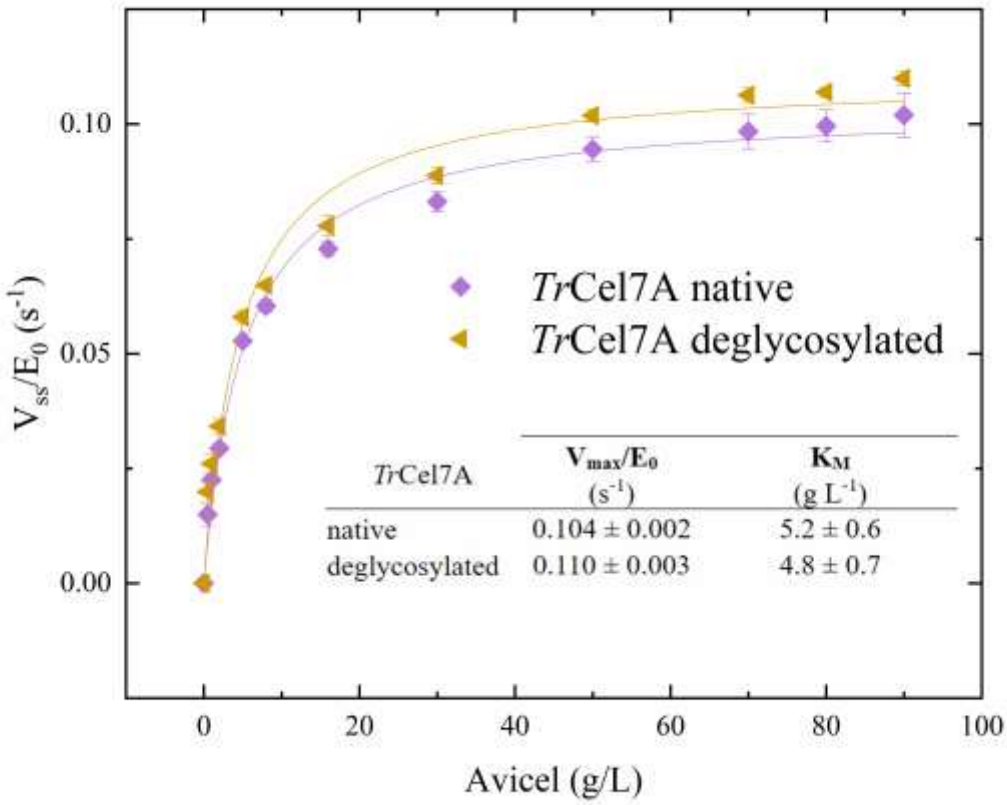

Figure 3. Steady-state kinetic analysis of *Tr*Cel7A with two different glycosylation patterns: native (Figure 1B) and *O*-deglycosylated (*Nn*GH92 and *Aa*GH2).

The analysis was done at 25°C on the cellulosic substrate Avicel. *Tr*Cel7A was enzymatically deglycosylated by *Nn*GH92 and *Aa*GH2. Error bars represent standard deviations from triplicate measurements.

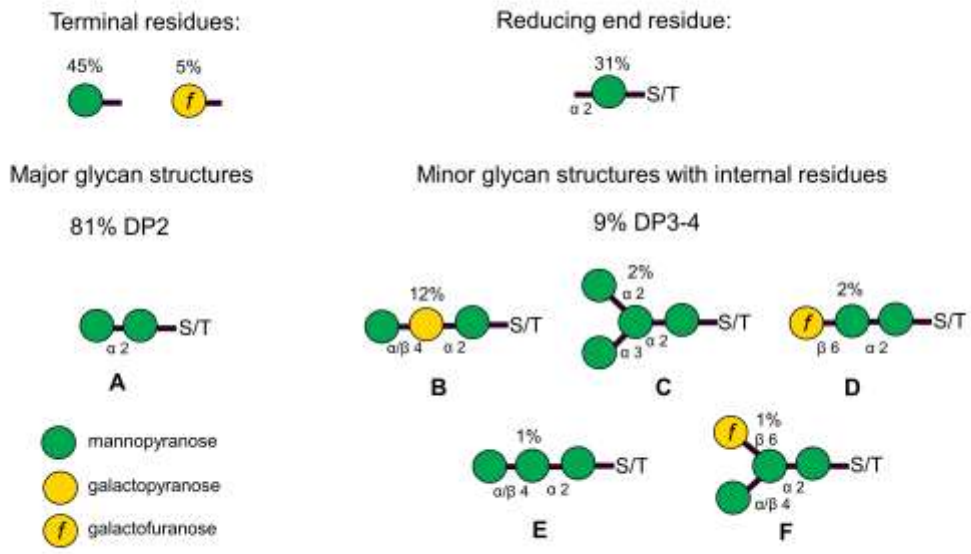

Figure 4. Possible *O*-glycan structures from *Tr*Cel7A expressed in *A. oryzae*.

The presented structures are based on the linkage analysis (Figure S1, Table SI) and the exo-glycosidase treatment. Only glycan structures having degrees of polymerization (DP) of 1-4 are demonstrated. The percentage values above the residues indicate their relative abundance (Table SI). The glycans were represented by following the SNFG guideline.(Varki et al. 2015; Neelamegham et al. 2019)

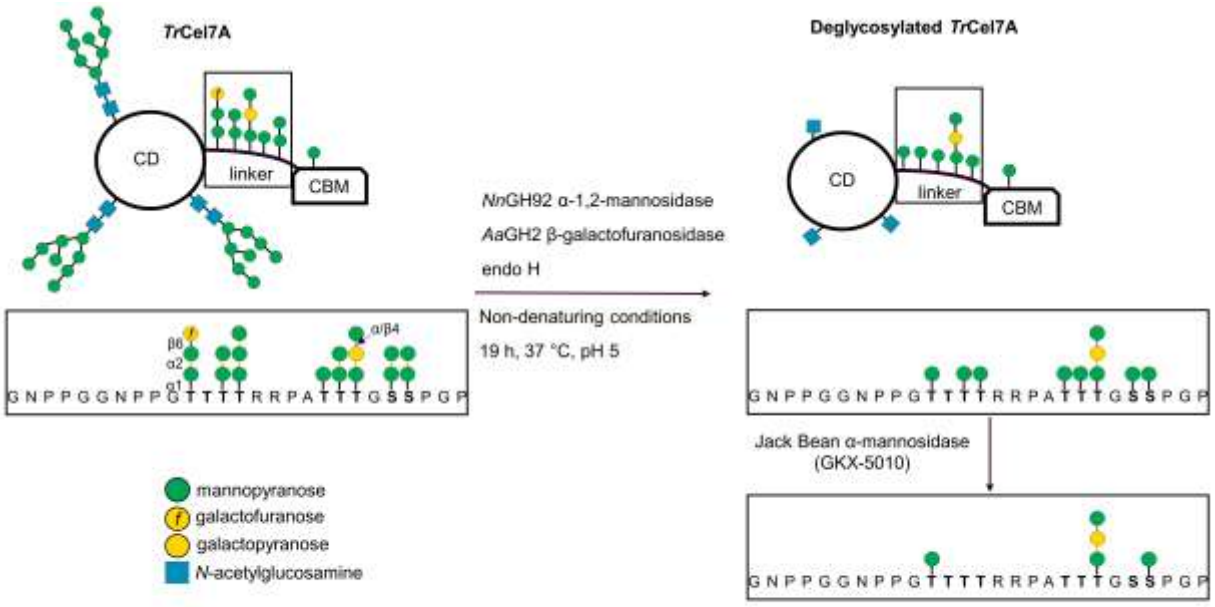

Figure 5. Scheme of the proposed model of enzymatic *O*-deglycosylation of *Tr*Cel7A WT expressed in *A. oryzae*.

The enzyme architecture consists of the catalytic domain (CD) linked to carbohydrate-binding module (CBM) through heavily *O*-glycosylated linker. The *O*-glycosylation profile demonstrates a simplified 'snapshot' of glycoforms present in the whole *O*-glycan population. *O*-linked glycans present on the *Tr*Cel7A are trimmed by *Nn*GH92 α-1,2-mannosidase and *Aa*GH2 β-galactofuranosidase. *N*-linked glycans present in the CD of *Tr*Cel7A can be removed by the action of endo H. The removal of the mannose residues directly attached to the Ser/Thr residues was achieved by the application of jack bean α-mannosidase (GKX-5010). The lines between the circles colored in black, purple, yellow and grey correspond to the glycosidic bonds with configuration of α1, α2, β6 and α/β4. The glycans were represented by following the SNFG guideline.(Varki et al. 2015; Neelamegham et al. 2019)