Biochemical characterization of a glycoside hydrolase family 43 -D-galactofuranosidase from the fungus Aspergillus niger

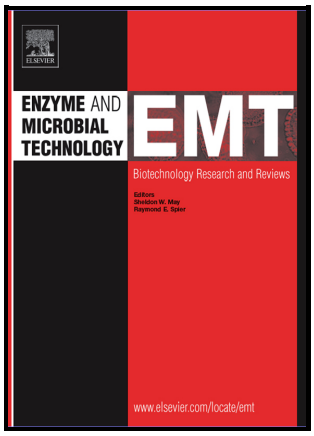
Bulmer, Gregory S.; Yuen, Fang Wei; Begum, Naimah; Jones, Bethan S.; Flitsch, Sabine L.; van Munster, Jolanda M.

Published in:
Enzyme and Microbial Technology

DOI:
10.1016/j.enzmictec.2022.110170

First published: 05/12/2022

Document Version
Version created as part of publication process; publisher's layout; not normally made publicly available



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

Gregory S. Bulmer, Fang Wei Yuen, Naimah Begum, Bethan S. Jones, Sabine L. Flitsch, Jolanda M. van Munster

PII: S0141-0229(22)00189-2

DOI: https://doi.org/10.1016/j.enzmictec.2022.110170

Reference: EMT110170

To appear in: *Enzyme and Microbial Technology*

Received date: 19 July 2022
Revised date: 27 November 2022
Accepted date: 30 November 2022

Please cite this article as: Gregory S. Bulmer, Fang Wei Yuen, Naimah Begum, Bethan S. Jones, Sabine L. Flitsch and Jolanda M. van Munster, Biochemical characterization of a glycoside hydrolase family 43 β-D-galactofuranosidase from the fungus *Aspergillus niger*XynD is a galactofuranosidase, *Enzyme and Microbial Technology,* (2022) doi:https://doi.org/10.1016/j.enzmictec.2022.110170

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

Running title: XynD is a galactofuranosidase

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

[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: Manchester Institute of Biotechnology (MIB) & School of Natural Sciences, The University of Manchester, 131 Princess Street, Manchester, M1 7DN, United Kingdom. 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 characterization of a galactofuranosidase from the industrially important fungus *Aspergillus niger*. Analysis of glycoside hydrolase family 43 subfamily 34 (GH43_34) members via conserved unique peptide patterns and phylogeny, revealed the occurrence of distinct clusters and, by comparison with specificities of characterized 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 produced 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 µM min^{-1}. The characterization of this first fungal GH43 galactofuranosidase offers further molecular insight into the degradation of Gal*f*-containing structures.

Keywords:

Aspergillus, galactofuranosidase, polysaccharide, galactofuranose, glycobiology, recombinant protein expression, XynD

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* (Fig. 1C)[24,25].

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, which have activity on both 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 characterization. *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 uncharacterized. 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 carbonsource, via degradation by Gal*f*ase, during carbon limitation to produce a source of galactose[32].

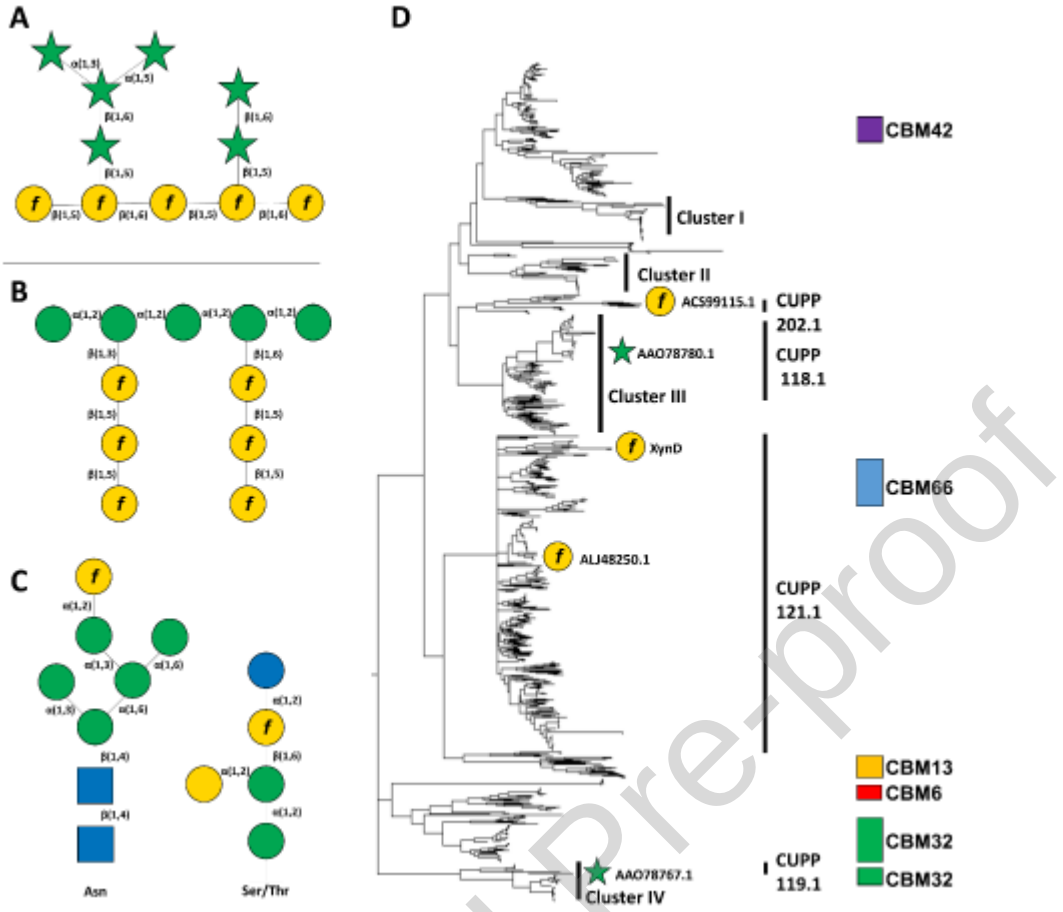


Figure 1: The presence of Galf throughout the fungal and bacterial kingdoms is not uniform and contains a variety of structures of both Galf backbones and Galf side chains, for example. **A** Structure of M. tuberculosis arabinogalactan fragment, **B** A. fumigatus galactomannan fragment, **C** N- and O-linked glycans. **D** Phylogenetic tree of GH43_34 family proteins, annotated with distinct phylogenetic clades and CUPP groups. Known activities are demonstrated as follows: β-galactofuranosidase (yellow circle), α-arabinofuranosidase (green star). Clades where all sequences are connected to a carbohydrate binding module are indicated by a colored bar labelled with the corresponding CBM family.

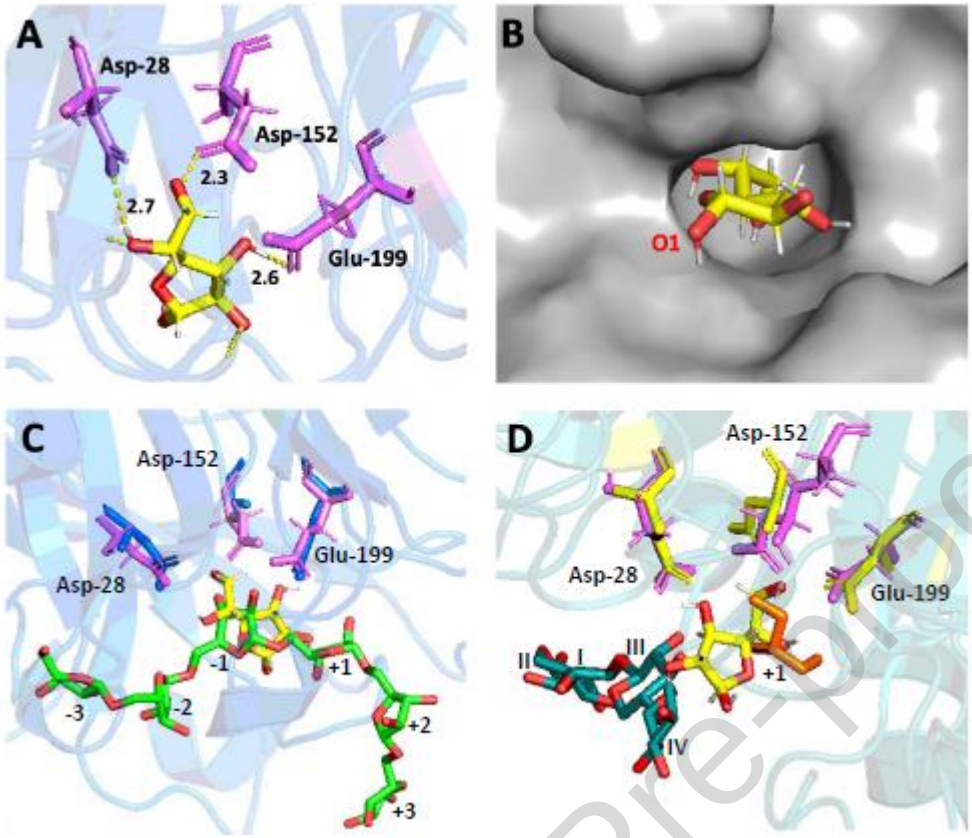


Figure 2: A Molecular docking (AutoDock VINA as implemented in YASARA) of galactofuranose (yellow) in a model of A. niger XynD generated using PDB 3QZ4 as a template. Catalytic residues Asp-28, Asp-152, Glu-199 are displayed in lilac with bonding lengths indicated in Å B XynD surface highlighting restricted binding pocket C GH43 C. japonicus α-L-arabinase Abf43A (GH43_5, PDB 1GYE) depicted with catalytic triad (blue) and arabinohexaose ligand (green), binding sites as indicated, overlaid with catalytic triad of XynD (liliac, catalytic residues indicated) and Galf ligand (yellow) D B. subtilis arabinoxylan arabinofuranohydrolase BsAXH-m2,3 (GH43_16, PDB 3C7G) depicted with catalytic triad (yellow), glycerol (orange) and xylotetraose ligand (grey), binding sites as indicated, overlaid with XynD (liliac, catalytic residues indicated) and Galf ligand (yellow).

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 4 biochemically characterized 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 in other enzyme families[28,33,34] ,and Ara*f* and Gal*f* are structurally similar, we hypothesized that Ara*f*ases and Gal*f*ases may share structurally similar active sites, and that therefore fungal Gal*f*ases were potentially harboured in GH families that contain 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 produced during *A. niger* interaction with plant biomass, however the biochemical function of XynD had yet to be elucidated. We express and characterize 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

Exploration of GH43_34 sequence space

We aimed to elucidate the role and biochemical activity of the GH43_34 enzyme XynD 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 characterized 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 (bootstrap values > 70) clusters (cluster I-IV) within the family (Fig. 1D). All 4 highly conserved clusters contain only bacterial members, with no fungal members (See Fig. S1 for bootstrap values and fungal member locations). For those organisms containing multiple GH43_34 members, different enzymes can be located in different clusters. Cluster III contains a characterised arabinofuranosidase (AAO78780.1) whilst cluster IV contains another biochemically characterized arabinofuranosidase (AA078767.1). The previously biochemically characterised galactofurosidases of GH43_34 (ACS99115.1 and ALJ48250.1) lie outside of these highly conserved clusters, where ALJ48250.1 seems to group with XynD but this is supported by very minimal (52 %) bootstrap values.

We also assessed clustering of the GH43_34 subfamily members via the conserved unique peptide patterns (CUPP) method[41,42]. This grouping of CAZymes enables the functional prediction of uncharacterised enzymes if characterised members exist in the CUPP group. The characterised Ara*f*ases AAO78767.1 and AAO78780.1 were assigned to CUPP groups 119.1 and 118.1, respectively. The characterised Gal*f*ase ACS99115.1 was assigned to CUPP group 202.1. Both XynD and characterised Gal*f*ase ALJ48250.1 were assigned to CUPP group 121.1. CUPP analysis thus supports the phylogenettic division of the four publicly available characterized GH43_34 members into four distinct groups and places XynD in the same group as ALJ48250.1, further suggesting that XynD may be a galactofuranosidase. CUPP groups mapped onto the clusters and groupings of the phylogenetic tree (Fig 1D), further strenghtening the analysis. Based on this, we hypothesize that *A. niger* XynD (An11g03120, GenBank: CAK40644.1) may have galactofuranosidase activity.

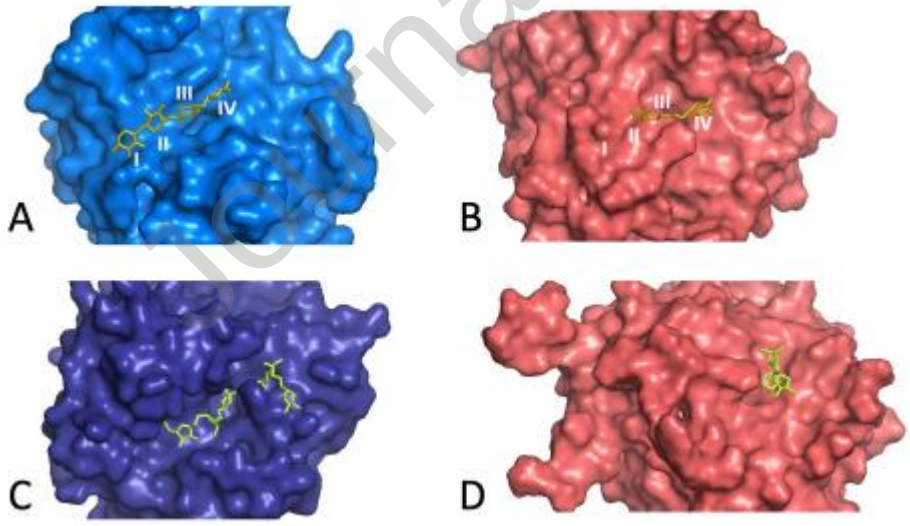


Figure 3 A Xylotetraose (X4) binding to B. subtilis arabinoxylan arabinofuranohydrolase BsAXH-m2,3 (PDB 3C7G), binding sites indicated **B** Superimposition of XynD over 3C7G with X4 ligand, **C** Arabinohexaose (A6) binding to C. japonicus α-L-arabinase Abf43A (PDB 1GYE) with binding sites indicated, **D** Superimposition of XynD over 1GYE with A6 ligand.

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 using YASARA software. Of the crystallised GH43 subfamily 34 members, the structure of GH43_34 arabinofuranosidase from *Bacteroides thetaiotaomicron* BT_3675 (PDB 3QZ4)[43] gave the greatest percentage coverage for modelling (89%), with 37% sequence identity. GH43 family enzymes display a five-bladed β-propeller fold[44], which was also predicted in the XynD model. No kinetic or substrate binding data is available for 3QZ4. Therefore, using this model, we applied molecular docking simulations to gain insights into the binding of potential substrates in the active site. Docking studies showed energetically favourable binding of galactofuranose to the catalytic domain of XynD with interaction with five residues: 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 other GH43 family members[44,45]. 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 that lacks space for further substrate binding subsites (Fig. 2B). Confirming our analysis, protein modelling and substrate docking using a model obtained via AlphaFold gave similar results.

Comparison of the XynD model with the *Cellvibrio japonicus* α-L-arabinanase Arb43A (GH43_5) complexed with arabinohexaose (1GYE) (Fig. 2C) and *Bacillus subtilis* arabinoxylan arabinofuranohydrolase BsAXH-m2,3 (GH43_16) in complex with xylotetraose (3C7G) (Fig. 2D), suggested the overall fold was conserved and position of the catalytic triad similar. XynD residues Trp-219 and Arg-286, who in the XynD model appear to interact with the Gal*f* substrate, were identified via comparison of sequence alignments as equivalent to BsAXH-m2,3 Phe-244 and Arg-321, respectively. In BsAXH-m2,3 these residues are involved in hydrophobic stacking at the II subsite and +1 subsite stability[46], respectively. Trp-219 is found exclusively in sequences from GH43_34 phylogenetic clusters that contain Gal*f*ases (Fig. S2), whilst the majority of those in the Ara*f*ase clusters contain a Thr-219. This suggests a role played by the residue in determining substrate specificity.

Analysis of the protein surface and accessibility of subsites identified in Arb43A and BsAXH-m2,3 further confirms potential differences in substrate specificity of XynD. When superimposed against 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).

Cloning, expression and purification of XynD

To confirm the activity of XynD experimentally, we heterologous expressed the gene. The predicted signal peptide (residues 1-24) and presence of four putative N-glycosylation sites suggests that XynD is a secreted protein. We therefore attempted expression in *Pichia pastoris*, based on a codon optimised cDNA sequence and the pPICZα expression system, but initial trials 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 predicted 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 13.8 µM 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-β-lactose. 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 a commercially available *Aspergillus* galactomannan ELISA assay (Platelia 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

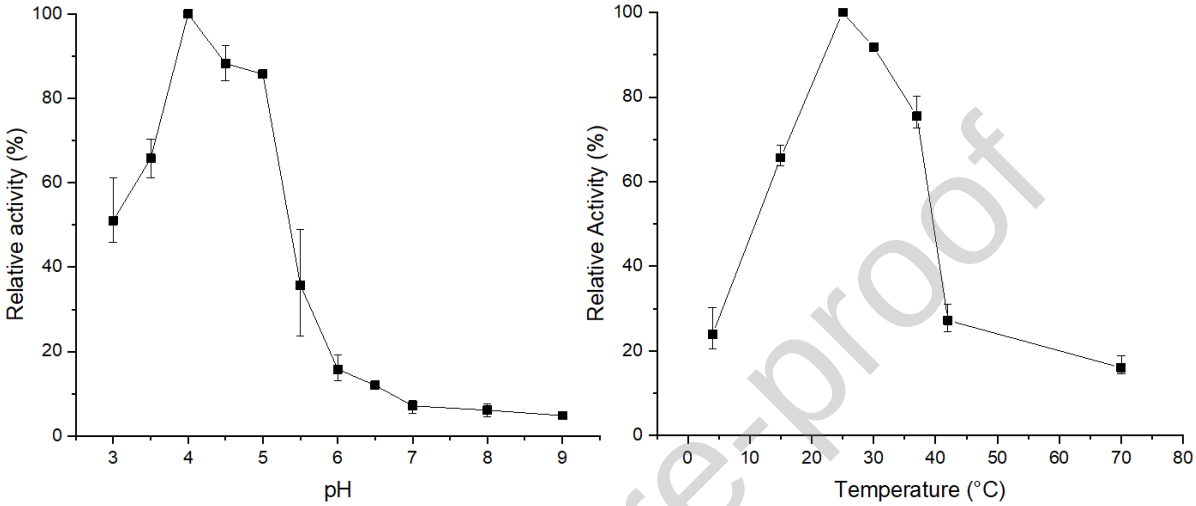


Figure 4: Optimisation of reaction conditions for pNP-Gal*f* hydrolysis by XynD **A** Relative activity at pH 3-9 **B** Relative activity at temperatures 4-70 °C

detected against the β-(1,5)-linked oligosaccharides 4MU-Galf_3, 4MU-Galf_4, 4MU-Galf_5, 4MU-Galf_6, a gift of T. Oka[47]. 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 T.L. Lowary[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[48] 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[48] 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, indicating that XynD is not active on these Gal*f* N-glycans.

Biochemical characterization

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).

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 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.

XynD has previously been annotated as a putative xylosidase in the literature and our identification of its galactofuranosidase activity highlights how detailed sequence analysis combined with biochemical characterization is critical to understand the roles of glycoside hydrolases. CUPP analysis combined with phylogenetic analysis of the subfamily followed by mapping of biochemically characterized members, demonstrated how GH43_34 is grouped into multiple clusters containing characterized members with arabinofuranosidase activity. AAO78780.1 (cluster III) acts on the α1,3-Araf linkage of single α1,3-Araf-substituted gum arabic arabinogalactan protein[38], whilst AAO78767.1 (cluster IV) acts on α1,3-Araf linkage present in grape rhamnogalacturonan-II[39]. This could therefore suggest potential roles of cluster III and cluster IV members as enzymes active against glycoproteins and complex polysaccharides, respectively. Association of a subset of grouped members with CBM42, which are commonly associated with GH54 arabinofuranosidases and bind arabinofuranose present in arabinoxylan[49], may suggest arabinofuranosidase activity for this group. The grouping of bacterial members with galactofuranosidase activity and XynD, our structural modelling, substrate docking and biochemical characterization confirmed our hypothesis that as a fungal member of this group, XynD, had galactofuranose activity.

Previously studies of GH family 43 reveal only the overall fold and the catalytic residues (Asp, Glu, and its pK_a-modulating Asp) are well conserved[44], and our molecular docking studies suggests a binding pose of galactofuranose in XynD that is stabilized to these from local Asp, Glu, Trp and Arg residues. XynD appears to have only a small pocket for substrate binding, which explains the lack of catalysis demonstrated against large oligo- and polysaccharides. Detailed interpretation of the model and substrate docking should however be performed with due care, the low sequence identity between template and target limits the reliability of the model.

Our sequence alignments highlighted how an amino acid 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 in our model positioned to affect sugar binding and we hypothesised that mutation of this site may switch enzyme activity from a galactofuranosidase to 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 produced and biochemically characterized. We showed XynD demonstrates hydrolytic activity against pNP-Gal*f.* Interestingly, despite the substantial structural similarities between Gal*f* and Ara*f,* XynD does not show activity against pNP-Ara*f* and is thus thereby differs from other characterized *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. Furthermore, the unidentified Gal*f*ase also shows activity against glycoproteins (GAM-1) whilst xynD did not. The major extracellular β-galactofuranose from *A. niger* is therefore not XynD, meaning unidentified β-galactofuranoses of the fungus remain.

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[30,34,50–52]. Optimal temperature of characterized galactofuranosidases varies, for example, the Gal*f*-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[51,53]. Since XynD was produced 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 in literature. Gene co-expressing networks, based on 155 transcriptomics experiments[54] and available via FungiDB, highlighted expression of *xynD* is positively correlated with 110 genes, and gene ontology (GO) enrichment analysis using REViGO[55] (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*[56] and other genes undergoing such repression[57]. Whilst *xynD* is not expressed during growth on maltose in the fungal wild-type[57], mutants with an inactivated *amyR* (encoding a regulator of starch degradation induced by maltose[58]) display upregulation of *xynD* transcription[57]. Similarly, *xynD* expression on arabinose is low in the wild-type, but an Δ*araR* disruption mutant showed upregulation of *xynD* transcription[57]. In the Δ*amyR*[59] and Δ*araR*[60] 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 pattern is consistent with a response to carbon limitation and concurs with the expression response to starvation seen with 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[61] then subsequently aligned and trimmed to just the catalytic domain using BioEdit[62]. Evolutionary analyses were conducted in MEGA X[63]. The evolutionary history was inferred by using the Maximum Likelihood method and JTT matrix-based model[64]. 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.

Conserved unique peptide patterns (CUPP)

CUPP groups in GH43_34 were downloaded the CUPP-CAZyme library groups in the publically available CUPP database[42] (accessed 22/11/2022). Characterised GH43_34 members were compared with proteins of unknown functions in the same CUPP group, and CUPP groups correlated with placement of sequences in the phylogenetic trees.

Computational docking studies

Enzyme modelling and molecular docking simulation were performed using the YASARA software from YASARA Biosciences GmbH. As there is no PDB structure of XynD a model was produced via YASARA 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. In addition, a second XynD model was produced using AlphaFold Colab notebook using AlphaFold model parameters[65], this model was subsequently used for substrate docking via the same workflow.

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: MSHPQQNLLATTTSNTKAGNPVFPGWYADPEARLFNAQYWIYPTYSADYSEQTFFDAFSSPDLLTWTKHPTILNITNIPWSTNRAAWAPS VGRKLRSSANAEEEYDYFMYFSVGDGTGIGVAKSTTGKPEGPYEDVLGEPLVNGTVYGAEAIDAQIFQDDDGRNWLYFGGWSHAVVVEL GEDMISLKGDYLEITPEGYVEGPWMLKRNGIYYYMFSVGGWGDNSYGVSYVTADSPTGPFSSTPKKILQGNDAVGTSTGHNSVFTPDGQD YYIVYHRRYVNDTARDHRVTCIDRMYFNEAGEILPVNITLEGVEGRTLS(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}, equivalent to 13.8 µM) 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.L. Lowary, The Institute of Biological Chemistry, Academia Sinica, Taiwan) in 50 mM citric acid-phosphate buffer pH 5. 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 and J. Park (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[65]. 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, stability and kinetic parameters

To identify the optimal pH for activity, 1 mM pNP-β-Gal*f* was 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-Gal*f* under standard condition reactions at pH 5. Kinetic parameters were assessed by incubation of XynD in the optimised conditions with substrate concentrations ranging from 0.125-10 mM, OD measurements were taken in triplicate every 10 minutes for 60 minutes in total (Fig. S9)

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).

Data availability

All data are contained within the article and the supporting information.

Acknowledgements

We thank Dr A. Ram and J. Park of The University of Leiden, the Netherlands, for providing us with *A. niger* galactomannan, Dr T.L Lowary of The Institute of Biological Chemistry, Academia Sinica, Taiwan 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. Marino, C., Rinflerch, A. & De Lederkremer, R. M. Galactofuranose antigens, a target for diagnosis of fungal infections in humans. *Futur. Sci. OA* **3**, (2017).

2. Peltier, P., Euzen, R., Daniellou, R., Nugier-Chauvin, C. & Ferrières, V. Recent knowledge and innovations related to hexofuranosides: structure, synthesis and applications. *Carbohydrate Research* vol. 343 1897–1923 (2008).

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

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

5. Heiss, C. *et al.* Unusual galactofuranose modification of a capsule polysaccharide in the pathogenic yeast Cryptococcus neoformans. *J. Biol. Chem.* **288**, 10994–11003 (2013).

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

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

8. Morelle, W. *et al.* Galactomannoproteins of Aspergillus fumigatus. *Eukaryot. Cell* **4**, 1308–1316 (2005).

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

10. Gsaller, F. *et al.* Sterol Biosynthesis and Azole Tolerance Is Governed by the Opposing Actions of SrbA and the CCAAT Binding Complex. *PLoS Pathog.* **12**, (2016).

11. Xue, X. *et al.* Synthetic polyprenol-pyrophosphate linked oligosaccharides are efficient substrates for mycobacterial galactan biosynthetic enzymes. *Org. Biomol. Chem.* **16**, 1939–1957 (2018).

12. Angala, S. K., Belardinelli, J. M., Huc-Claustre, E., Wheat, W. H. & Jackson, M. The cell envelope glycoconjugates of Mycobacterium tuberculosis. *Critical Reviews in Biochemistry and Molecular Biology* vol. 49 361–399 (2014).

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

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

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

16. Matveev, A. L. *et al.* Novel mouse monoclonal antibodies specifically recognize Aspergillus fumigatus galactomannan. *PLoS One* **13**, (2018).

17. Fontaine, T. *et al.* Molecular organization of the alkali-insoluble fraction of Aspergillus fumigatus cell wall. *J. Biol. Chem.* **275**, 27594–27607 (2000).

18. Latge, J. P. *et al.* Chemical and immunological characterization of the extracellular galactomannan of Aspergillus fumigatus. *Infect. Immun.* **62**, 5424–5433 (1994).

19. D'Haese, J. *et al.* Detection of galactomannan in bronchoalveolar lavage fluid samples of patients at risk for invasive pulmonary aspergillosis: Analytical and clinical validity. *J. Clin. Microbiol.* **50**, 1258–1263 (2012).

20. Muszkieta, L. *et al.* 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).

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

22. Clarke, B. R. *et al.* 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**, 5411–5418 (1995).

23. Seničar, M. *et al.* Galactofuranose-related enzymes: Challenges and hopes. *Int. J. Mol. Sci.* **21**, (2020).

24. Wang, L. *et al.* Mapping N-linked glycosylation sites in the secretome and whole cells of aspergillus niger using hydrazide chemistry and mass spectrometry. *J. Proteome Res.* **11**, 143–156 (2012).

25. Goto, M. Protein O-glycosylation in fungi: Diverse structures and multiple functions. *Bioscience, Biotechnology and Biochemistry* vol. 71 1415–1427 (2007).

26. Boecker, S. *et al.* Aspergillus niger is a superior expression host for the production of bioactive fungal cyclodepsipeptides. *Fungal Biol. Biotechnol.* **5**, (2018).

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

28. Matsunaga, E. *et al.* Identification and characterization of a novel galactofuranose-specific β-D-galactofuranosidase from Streptomyces species. *PLoS One* **10**, (2015).

29. Matsunaga, E. *et al.* Characterization of a PA14 domain-containing galactofuranose-specific β-D-galactofuranosidase from Streptomyces sp. *Biosci. Biotechnol. Biochem.* **81**, 1314–1319 (2017).

30. Shen, L. *et al.* The endogenous galactofuranosidase GlfH1 hydrolyzes mycobacterial arabinogalactan. *J. Biol. Chem.* **295**, 5110–5123 (2020).

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

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

33. Tefsen, B. *et al.* Fungalα-arabinofuranosidases of glycosyl hydrolase families 51 and 54 show a dual arabinofuranosyl- and galactofuranosyl-hydrolyzing activity. *Biol. Chem.* **393**, 767–775 (2012).

34. Matsunaga, E. *et al.* Identification and characterization of β-D-galactofuranosidases from Aspergillus nidulans and Aspergillus fumigatus. *J. Biosci. Bioeng.* **131**, 1–7 (2021).

35. Wu, M. *et al.* Genetic determinants of in vivo fitness and diet responsiveness in multiple human gut Bacteroides. *Science (80-.).* **350**, (2015).

36. Mahowald, M. A. *et al.* Characterizing a model human gut microbiota composed of members of its two dominant bacterial phyla. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 5859–5864 (2009).

37. Helbert, W. *et al.* Discovery of novel carbohydrate-active enzymes through the rational exploration of the protein sequences space. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 6063–6068 (2019).

38. Cartmell, A. *et al.* A surface endogalactanase in Bacteroides thetaiotaomicron confers keystone status for arabinogalactan degradation. *Nat. Microbiol.* **3**, 1314–1326 (2018).

39. Ndeh, D. *et al.* Complex pectin metabolism by gut bacteria reveals novel catalytic functions. *Nature* **544**, 65–70 (2017).

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

41. Barrett, K. & Lange, L. Peptide-based functional annotation of carbohydrate-active enzymes by conserved unique peptide patterns (CUPP). *Biotechnol. Biofuels* **12**, 1–21 (2019).

42. Barrett, K., Hunt, C. J., Lange, L. & Meyer, A. S. Conserved unique peptide patterns (CUPP) online platform: peptide-based functional annotation of carbohydrate active enzymes. *Nucleic Acids Res.* **48**, W110–W115 (2020).

43. (JCSG), J. C. for S. G. 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) doi:10.2210/pdb3QZ4/pdb.

44. Nurizzo, D. *et al.* Cellvibrio japonicus α-l-arabinanase 43a has a novel five-blade β-propeller fold. *Nat. Struct. Biol.* **9**, 665–668 (2002).

45. Brüx, C. *et al.* 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**, 97–109 (2006).

46. Vandermarliere, E. *et al.* 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**, 39–47 (2009).

47. Chihara, Y. *et al.* Biosynthesis of β-(1,5)-Galactofuranosyl Chains of Fungal- Type and. *mSphere* **5**, 1–19 (2020).

48. 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.* **256**, 149–158 (1994).

49. Miyanaga, A. *et al.* The family 42 carbohydrate-binding module of family 54 α-L- arabinofuranosidase specifically binds the arabinofuranose side chain of hemicellulose. *Biochem. J.* **399**, 503–511 (2006).

50. Matsunaga, E. *et al.* Characterization of a PA14 domain-containing galactofuranose-specific β- <scp>d</scp> -galactofuranosidase from *Streptomyces* sp. *Biosci. Biotechnol. Biochem.* **81**, 1314–1319 (2017).

51. Seničar, M. *et al.* Galactofuranosidase from JHA 19 Streptomyces sp.: subcloning and biochemical characterization. *Carbohydr. Res.* **480**, 35–41 (2019).

52. Matsunaga, E. *et al.* Identification and Characterization of a Novel Galactofuranose-Specific β-D-Galactofuranosidase from Streptomyces Species. *PLoS One* **10**, e0137230 (2015).

53. Matsunaga, E. *et al.* Characterization of a PA14 domain-containing galactofuranose-specific β-D-galactofuranosidase from Streptomyces sp. *Biosci. Biotechnol. Biochem.* **81**, 1314–1319 (2017).

54. Schäpe, P. *et al.* Updating genome annotation for the microbial cell factory Aspergillus Niger using gene co-expression networks. *Nucleic Acids Res.* **47**, 559–569 (2019).

55. Supek, F., Bošnjak, M., Škunca, N. & Šmuc, T. Revigo summarizes and visualizes long lists of gene ontology terms. *PLoS One* **6**, (2011).

56. Coutinho, P. M. *et al.* Post-genomic insights into the plant polysaccharide degradation potential of Aspergillus nidulans and comparison to Aspergillus niger and Aspergillus oryzae. *Fungal Genet. Biol.* **46 Suppl 1**, S161–S169 (2009).

57. Gruben, B. S. *et al.* Expression-based clustering of CAZyme-encoding genes of Aspergillus niger. *BMC Genomics* **18**, (2017).

58. Vankuyk, P. A., Benen, J. A. E., Wösten, H. A. B., Visser, J. & De Vries, R. P. 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**, 285–293 (2012).

59. Zhang, H. *et al.* 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).

60. Kowalczyk, J. E. *et al.* Combinatorial control of gene expression in Aspergillus Niger grown on sugar beet pectin. *Sci. Rep.* **7**, (2017).

61. El-Gebali, S. *et al.* The Pfam protein families database in 2019. *Nucleic Acids Res.* **47**, D427–D432 (2019).

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

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

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

65. Park, J. *et al.* Identification of the UDP-glucose-4-epimerase required for galactofuranose biosynthesis and galactose metabolism in A. niger. *Fungal Biol. Biotechnol. 2014 11* **1**, 1–12 (2014).

CrediT author statement

Gregory Bulmer; conceptualization, Investigation, methodology, formal analysis, visualization, validation,writing-original draft. Fang Wei Yuen; investigation, formal analysis. Naimah Begum; investigation. Bethan Jones; methodology, validation. Sabine Flitsch; methodology, writing-reviewing and editing. Jolanda van Munster; Conceptualization, Supervision, project administration, validation,funding acquisition, writing-reviewing and editing.

Highlights

- Sequence analysis suggests *Aspergillus niger* XynD is a galactofuranosidase
- Molecular docking indicates galactofuranose as substrate for XynD
- Biochemical characterisation of XynD confirms Galfase activity
- A broad range of possible natural substrates is assessed