The latest version is at http://www.jbc.org/cgi/doi/10.1074/jbc.RA119.008511

Molecular mechanism of *Aspergillus fumigatus* biofilm disruption by fungal and bacterial glycoside hydrolases

François Le Mauff[1,2,3*], Natalie C Bamford[4,5*], Noor Alnabelseya[4,5#], Yongzhen Zhang[6], Perrin Baker[4†], Howard Robinson[7], Jeroen D C Codée[6], P. Lynne Howell[4,5,$], Donald C. Sheppard[1,2,3$]

[1]Department of Microbiology and Immunology, Faculty of Medicine, McGill University, Montreal, Quebec, Canada; [2]Infectious Disease and Immunity in Global Health, Research Institute of McGill University Health Center, Montreal, Quebec, Canada; [3]McGill Interdisciplinary Initiative in Infection and Immunity, Montreal, Quebec, Canada; [4]Program in Molecular Medicine, The Hospital for Sick Children, Toronto, Ontario, Canada; [5]Department of Biochemistry, Faculty of Medicine, University of Toronto, Toronto, Ontario, Canada; [5]Leiden Institute of Chemistry, Leiden University, Leiden, The Netherlands; [6]Photon Science Division, Brookhaven National Laboratory, 743 Brookhaven Avenue, Upton, New York, 11973-5000, USA

Running title: Molecular mechanism of glycoside hydrolases $Sph3_h$ and $PelA_h$

[*] contributed equally to this work

Present addresses: [#] Diassess Inc.,1412 62nd street, Emeryville, CA, USA 94608, [†] Covalon Technologies Ltd., 660 Tech Ave #5, Mississauga, ON, Canada, L4W 5S7

[$]To whom correspondence should be addressed. Tel: 416-813-5378: Email: howell@sickkids.ca or Tel.: 514-934-1934, Ext. 36104: E-mail: donald.sheppard@mcgill.ca

Keywords: Glycoside hydrolases, Biofilm, Aspergillus fumigatus, extracellular matrix, protein structure, virulence factor, lung infection, substrate specificity, galactosaminogalactan

ABSTRACT

During infection, the fungal pathogen *Aspergillus fumigatus* forms biofilms that enhance its resistance to antimicrobials and host defenses. An integral component of the biofilm matrix is galactosaminogalactan (GAG), a cationic polymer of α-1,4-linked galactose and partially deacetylated N-acetylgalactosamine (GalNAc). Recent studies have shown that recombinant hydrolase domains from Sph3, an *A. fumigatus* glycoside hydrolase involved in GAG synthesis, and PelA, a multi-functional protein from *Pseudomonas aeruginosa* involved in Pel polysaccharide biosynthesis, can degrade GAG, disrupt *A. fumigatus* biofilms, and attenuate fungal virulence in a mouse model of invasive aspergillosis. The molecular mechanisms by which these enzymes disrupt biofilms have not been defined. We hypothesized that the hydrolase domains of Sph3 and PelA ($Sph3_h$ and $PelA_h$, respectively) share structural and functional similarities given their ability to degrade GAG and disrupt *A. fumigatus* biofilms. MALDI-TOF enzymatic fingerprinting and NMR experiments revealed that both proteins are retaining endo-α-

1,4-*N*-acetylgalactosaminidases with a minimal substrate size of seven residues. The crystal structure of $PelA_h$ was solved to 1.54 Å and structure alignment to $Sph3_h$ revealed that the enzymes share similar catalytic site residues. However, differences in the substrate binding clefts result in distinct enzyme-substrate interactions. $PelA_h$ hydrolyzed partially deacetylated substrates better than $Sph3_h$, a finding that agrees well with $PelA_h$'s highly electronegative binding cleft versus the neutral surface present in $Sph3_h$. Our insight into $PelA_h$'s structure and function necessitate the creation of a new glycoside hydrolase family, GH166, whose structural and mechanistic features, along with those of GH135 (Sph3), are reported here.

INTRODUCTION

Aspergillus fumigatus is a ubiquitous filamentous fungus that causes invasive pulmonary infections in patients who are immunosuppressed due to cytotoxic chemotherapy, organ and stem cell transplantation, or biological therapies such as tumor necrosis factor inhibitors (1-3). During pulmonary infection, filamentous hyphae of *A. fumigatus* grow within biofilms: multicellular communities of organisms embedded in a self-produced extracellular matrix (4). This biofilm extracellular matrix plays several roles in the pathogenesis of invasive aspergillosis including mediating the adherence of hyphae to host tissues and enhancing resistance to antifungal drugs and host immune defenses (5-8).

Solid-state NMR spectroscopy analyses have revealed that polysaccharides are the most abundant component of the *A. fumigatus* extracellular matrix (9). Immunohistochemical and electron microscopy studies of *Aspergillus* biofilms have identified α-1,3-glucan, galactomannan and galactosaminogalactan (GAG) as key polysaccharide components of the matrix (10). GAG is a linear cationic polymer of α-1,4-D-galactose and partially deacetylated α-1,4-*N*-acetyl-D-galactosamine (GalNAc) that is produced by activity of the products of a five-gene cluster located on the chromosome 3 (11). GAG synthesis is thought to be initiated by the synthesis of UDP-GalNAc and UDP-Gal by the glucose-4-epimerase Uge3 (12). These sugars are then linked and exported through the action of the predicted glycosyltransferase Gtb3, and the resulting polymer is partially deacetylated in the extracellular space by secreted Agd3 (13). Cleavage of the emerging polymer is thought to be mediated by two glycoside hydrolases encoded within this gene cluster, Sph3 and Ega3 (11). GAG, plays a central role in biofilm formation and virulence (14). GAG-deficient strains are markedly impaired in their ability to form biofilms and adhere to pulmonary epithelial cells *in vitro* (15). Hyphal-associated GAG enhances resistance to neutrophil extracellular traps (16) and conceals pathogen-associated molecular patterns such as ß-glucan from immune detection (17). Secreted GAG also promotes infection by inducing neutrophil apoptosis and the production of the anti-inflammatory cytokine IL-1 receptor antagonist (18). Consistent with these observations, GAG-deficient strains of *A. fumigatus* exhibit attenuated virulence in mouse models of invasive aspergillosis (17).

The importance of GAG in biofilm formation and virulence suggest that this exopolysaccharide is a promising therapeutic target. We recently reported that the recombinant glycoside hydrolase domains from two microbial proteins, *A. fumigatus* Sph3 and *Pseudomonas aeruginosa* PelA, degrade GAG and disrupt *A. fumigatus* biofilms (19). PelA is a multidomain protein with both deacetylase and glycoside hydrolase activity that is required for production of the *P. aeruginosa* Pel polysaccharide (20-22). In addition to its degrading *A. fumigatus* biofilms, the recombinant hydrolase domain of PelA ($PelA_h$) also disrupts Pel-dependent biofilms of *P. aeruginosa* (21). While, the structure and anomeric configuration of the Pel polysaccharide remains to be determined, it has been found to contain partially deacetylated 1,4-linked GalNAc and *N*-acetylglucosamine (GlcNAc) in a ratio of 5:1

(20,23). The presence of GalNAc-rich regions in both the GAG and Pel polysaccharides suggests that $PelA_h$ may also be specific for GalNAc.

Structural and functional characterization of glycoside hydrolase domain of Sph3 ($Sph3_h$) revealed that this enzyme is essential for GAG biosynthesis and belongs to the glycoside hydrolase (GH) family 135 (24). $Sph3_h$ has an $(\beta/\alpha)_8$ fold with a shallow conserved active site groove. Co-crystallization of $Sph3_h$ with GalNAc and mutagenesis of residues within the active site groove suggest that this enzyme hydrolyses GAG within GalNAc-rich regions of the polymer (24).

Less is known about structure and function of the glycoside hydrolase domain of PelA, which is predicted by Phyre[2] to contain a (β/α)-barrel fold (20) and by the conserved domain database to belong to the GH114 superfamily (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). However, a BLAST search against all CAZy, GH114 members, and a search of the sequence against all Hidden Markov Models of CAZy families, failed to find any significant sequence identity with classified glycoside hydrolases, suggesting that $PelA_h$ belongs to a new CAZy family.

Herein, we characterize the molecular mechanisms underlying the cross-kingdom GAG activity of the glycoside hydrolases, $Sph3_h$ and $PelA_h$. Mass spectrometry enzymatic fingerprinting and NMR studies revealed that both enzymes share retaining endo-α-1,4-*N*-acetylgalactosaminidase activity and require a minimal substrate length of seven GalNAc residues. The substrate interacts with the binding cleft differently in the two enzymes leading $Sph3_h$ to cleave substrates closer to the non-reducing end while $PelA_h$ cleaves proximal to the reducing end of oligosaccharides. The enzymes also differ in their ability to cleave partially deacetylated substrates. Although neither enzyme is active against fully deacetylated oligosaccharides, $PelA_h$ has a higher propensity to cleave GalNAc linkages within regions of partially deacetylated GAG. The structure of $PelA_h$ was determined, and comparison with $Sph3_h$, revealed a high degree of structural similarity within the catalytic site of these enzymes. The presence of a deeper, more electronegative groove in $PelA_h$ likely underlies its ability to bind and cleave cationic partially deacetylated substrates more effectively than $Sph3_h$.

RESULTS:

$Sph3_h$ and $PelA_h$ are α-1,4-N-acetylgalactosaminidases

To elucidate the molecular mechanisms by which $Sph3_h$ and $PelA_h$ mediate biofilm disruption, these enzymes were incubated with pre-grown *A. fumigatus* biofilms and the profile of oligosaccharides that were released was analyzed by Matrix Assisted Laser Desorption and Ionization – Time of Flight Mass Spectrometry (MALDI-TOF MS) enzymatic fingerprinting. Consistent with our previous studies, nanomolar concentrations of $Sph3_h$ or $PelA_h$ disrupted *A. fumigatus* biofilms (Fig. 1A and 1B) (19). Analysis of the oligosaccharides released by treatment with either enzyme revealed similar MS fingerprints, with spectra displaying ions of m/z ratio corresponding to *N*-acetylhexosamine (HexNAc) oligomers. Treatment of biofilms with 100 nM $Sph3_h$ produced ions with a m/z ratio from 1056.5684 to 2478.2696 with a repetitive occurrence of 203.1002 +/- 0.0059 matching with the HexNAc m/z ratio (25). This profile is consistent with the release of HexNAc oligomers ranging in size from pentamers to dodecamers (12-mers) (Fig. 1A). Treatment of fungal biofilms with 100 nM $PelA_h$ generated a similar spectrum of ions ranging from m/z 1056.6345 to 3290.8078, with an interval m/z difference of 203.1107 +/- 0.0053 between ions, suggesting that $PelA_h$ treatment released HexNAc oligosaccharides ranging in size from pentamers to hexadecamer (16-mers) (Fig. 1B). MS-MS fragmentation analysis confirmed these ions were composed uniquely of m/z ratio 203.10 units, suggesting that these oligosaccharides are HexNAc homopolymers (Fig. 1C). Gas chromatography

coupled to mass spectrometry (GC-MS) analysis of the monosaccharide composition of these oligosaccharides confirmed they are composed solely of GalNAc, supporting the hypothesis that these enzymes cleave regions of homo-GalNAc within GAG (Fig. 1D).

To further confirm the specificity of $Sph3_h$ and $PelA_h$ for the α-1,4-GalNAc linkages within GAG, the ability of these enzymes to cleave synthetic oligo-α-1,4-galactose, and oligo-α-1,4-galactosamine were tested. Neither $Sph3_h$ or $PelA_h$ exhibited activity against these components of GAG, even at concentrations as high as 5 μM (Fig. 2A and B). Similarly, neither enzyme was able to degrade chitin (β-1,4 linked N-acetylglucosamine), the only other major hexosamine-containing polysaccharide within the fungal cell wall (Fig. 2C). These data suggest that biofilm disruption by $Sph3_h$ and $PelA_h$ is a consequence of cleavage of α-1,4-GalNAc homopolymeric regions within GAG.

$Sph3_h$ and $PelA_h$ are α-1,4-GalNAc endo-acting hydrolases with a minimal substrate length of 7 GalNAc units.

To confirm the α-1,4-*N*-acetylgalactosaminidase activity of $Sph3_h$ and $PelA_h$, MALDI-TOF MS enzymatic fingerprinting was performed using a purified fraction of α-1,4-GalNAc oligosaccharides obtained by partial $Sph3_h$ digestion of *A. fumigatus* biofilms (Fig 2D). Treatment of these predominately decameric to dodecameric oligosaccharides with either 1 μM $Sph3_h$ (Fig 2E) or 1 μM $PelA_h$ (Fig 2F) for 1 hour resulted in a shift of the spectra consistent with final product sizes of penta-, hexa- and heptamers, confirming the ability of these enzymes to cleave α-1,4-GalNAc linkages.

To determine the minimum substrate size that can be cleaved by $Sph3_h$ and $PelA_h$, α-1,4-GalNAc hexamers and heptamers purified from partial $Sph3_h$ digestion of *A. fumigatus* biofilms were treated with each enzyme and the degradation products analyzed by MALDI-TOF MS fingerprinting. GalNAc heptasaccharides but not hexasaccharides were rapidly hydrolyzed by both enzymes (Fig. 3A-D). Hydrolysis of GalNAc heptasaccharides by $Sph3_h$ resulted in the accumulation of pentasaccharides (Fig. 3B), while $PelA_h$ hydrolysis produced predominantly both tetra- and pentasaccharides (Fig. 3D) suggesting that these enzymes function as endo-acting glycoside hydrolases. The ability of $PelA_h$ to degrade the heptamers into two sets of products suggests some flexibility in the positioning of the oligosaccharide in the $PelA_h$ binding site.

$Sph3_h$ and $PelA_h$ are retaining endoglycoside hydrolases.

To investigate the molecular mechanism of the two glycoside hydrolases, the stereochemical outcome of oligosaccharide cleavage by these enzymes was evaluated by ^{1}H-NMR spectroscopy using synthetic GalNAc octamers as a substrate. Spectral analysis of the pre-reaction substrate demonstrated the presence of four doublets with coupling constants between 3.5 and 4.0 Hz in the anomeric region of the spectra, characteristic of α-glycosidic linkages. The addition of $Sph3_h$ or $PelA_h$ led to the appearance of a new peak at 5.29 ppm with a coupling constant of 4.0 Hz characteristic of an α-anomer (Fig 4A and B). Acquisition of 1D spectra at 25 °C and 37 °C (Fig 4A and 4B), and of 2D COSY ^{1}H-^{1}H spectra (Supplemental Fig. 1) after 24 h of reaction revealed the appearance of a β-anomer signal at 4.73 ppm on the 1D spectra consistent with secondary mutarotation at the new reducing end. The addition of purified monomeric GalNAc to samples at the end of the experiment resulted in the appearance of anomeric signals at 4.66 and 5.25 ppm. These signals were distinct from the experimental products, demonstrating that neither $Sph3_h$ nor $PelA_h$ released monosaccharides from GalNAc octamers (supplemental Figure 2). Collectively, these data suggest that $Sph3_h$ and $PelA_h$ are retaining endoglycosidases.

$PelA_h$ has a (β/α) barrel with a deep substrate binding groove.

To shed light on the mechanisms by which $PelA_h$ and $Sph3_h$ hydrolyse α-1,4-GalNAc polymers, and to complement the available structure of $Sph3_h$, we pursued the structure determination of $PelA_h$. Crystals of $PelA_h$ formed readily in a variety of different crystallization conditions and the structure was subsequently determined using selenomethionine incorporation and single-wavelength anomalous diffraction method to 1.54 Å and refined to a final R_{work} and R_{free} of 16.9 % and 19.2 %, respectively. Unambiguous electron density allowed the modeling of residues 48 to 300. No interpretable density was observed for residues 47, 301-303, or the hexa-histidine purification tag. The structure of $PelA_h$ revealed a β_8/α_7-barrel fold with a C-terminal region lacking secondary structure (Fig. 5A). This region, residues 290-300, packs against the core β-barrel and makes multiple contacts to neighboring helices suggesting that the lack of helical structure is not an artifact of the construct design. Furthermore, secondary structure prediction using Jpred4 (26) predicts that residues 297-310 form a β-strand supporting the lack of an eight helix.

In addition to the core (β/α)-fold, $PelA_h$ contains two structural insertions. The first is a 38 amino acid loop between β3 and α3 containing a two strand anti-parallel β-sheet and two α-helices (Loop3 in Fig 5A). The second is a β-hairpin loop (βHL) between β6 and α6. These insertions contribute to the creation of a 15 Å deep, ~36 Å long, active site groove on the "top" face of the barrel at the C-termini of the β-strands (Fig. 5A). This groove is composed of several aromatics that could potentially participate in polysaccharide-protein interactions (Fig. 5B).

When the structure of $PelA_h$ was submitted to the tertiary structure similarity server, DALI (27), the most similar structures found were: (i) a hypothetical protein TM1410 from *Thermatoga maritima* which has sequence similarity to GH114 family members (PDB 2AAM, 2.5 Å root-mean-square deviation RMSD over 203 residues); (ii) a cycloisomaltooligosaccharide glucotransferase from GH66 (PDB 3WNK, 2.9 Å RMSD over 165 residues); (iii) Cwp19, a peptidoglycan hydrolase reported as a GH-like 10 family member (PDB 5OQ2, 2.9 Å rmsd over 193 residues), and (iv) a dextranase from GH66 (PDB 3VMN, 2.9 RMSD over 155 residues). TM1410 has the highest similarity in structure and highest sequence identity to $PelA_h$ (15.7%). TM1410 also contains an insertion after β3, consisting of a three-stranded β-sheet and a small α-helix ($Loop3^{TM1410}$). $Loop3^{TM1410}$ folds further over the putative active site than the equivalent Loop3 in $PelA_h$ (Fig. 5C). Electron density corresponding to a ring containing ligand, as well as multiple glycerol molecules, was found in the deep groove of the TM1410 structure. It is possible that $Loop3^{TM1410}$ has some flexibility and the presence of a ligand in TM1410 causes the loop to cap the groove. $Loop3^{PelA}$ and the βHL have the highest B-factors of the $PelA_h$ structure, reaching 134 $Å^2$ in Loop3 as compared to an average protein B-factor of 31.6 $Å^2$, suggesting conformational heterogeneity in the crystal, further supporting the flexibility of these regions (Fig. 5D).

$PelA_h$ is most closely related to GH114 family as previously noted. According to the automated carbohydrate active enzymes annotation server dbCAN2, $PelA_h$ has very low overlap with the GH114 HMM profile and has insufficient sequence identity to be assigned to this family (28,29). These results suggest that $PelA_h$ represents a new GH family related to GH114. Our experimental evidence that $PelA_h$ exhibits glycoside hydrolase activity allows for the creation of GH166 family. Although no GH114 structure has been solved to date, the similarity in predicted structure suggests that GH166 and GH114 may create a new GH clan (30).

Structural comparison of GAG hydrolyzing $PelA_h$ and $Sph3_h$ reveals architectural differences but conserved active site residues.

A superimposition of $PelA_h$ and $Sph3_h$ (RMSD of 3.05 Å over 167 Cα atoms) aligned the active site

of $Sph3_h$ (24) with that of $PelA_h$ (19) and revealed that the insertions in the $PelA_h$ structure are unique to this protein (Fig. 6A). $Sph3_h$ has a much shallower binding cleft compared to $PelA_h$ (Fig. 6B). Active sites situated in clefts and tunnels suggest endo-activity and correlates with increased processivity (31-33). Shallow substrate grooves correlate with low processivity but have been found in enzymes that cleave crystalline, recalcitrant substrates (31-33). The deeper cleft found in $PelA_h$ is indicative of a more processive activity as compared to $Sph3_h$.

Despite differences in cleft architecture, examination of the putative active sites (RMSD 0.25 Å over active site motif backbone atoms) revealed a high degree of conservation, with residues D166, N202 and E222 of $Sph3_h$ that interact with GalNAc in our co-crystal structure (24), superimposing with residues D160, N199 and E218 of $PelA_h$ (Fig. 5CD). $Sph3_h$ activity was previously shown to be dependent on D166 as conservative mutation of this residue abolished $Sph3_h$ degradation of GAG (24). Mutation of E222 to alanine also abolished $Sph3_h$ activity but replacement with glutamine had a lesser affect. Biofilm disruption by $PelA_h$ was shown to involve E218, which aligns to E222 of $Sph3_h$, suggesting this residue is involved in catalysis (19). The similarity of the residues within the catalytic site of $Sph3_h$ and $PelA_h$ suggests that they share a common catalytic mechanism. The distance between carboxyl side chains of the catalytic residues D166 and E222 in $Sph3_h$ and the homologous residues D160 and E218 in $PelA_h$ are 5.2 – 6.7 Å and 4.8 – 6.8 Å, respectively. This distance is consistent with a retaining (~ 5.5 Å) rather than inverting mechanism of cleavage (~ 10 Å) (32).

$Sph3_h$ and $PelA_h$ have different distribution of substrate binding subsites.

To determine if $Sph3_h$ and $PelA_h$ cleave GalNAc oligosaccharides closer to the reducing or non-reducing end, an oligosaccharide preparation enriched in $(\alpha\text{-}1,4\text{-}GalNAc)_9$ was reduced by sodium borohydride treatment, conferring an additional m/z ratio of 2 to the reducing end. These reduced oligosaccharides were then incubated with each enzyme and the resulting products analyzed by MALDI-TOF MS. Both enzymes produced multiple oligosaccharides products ranging in length from tetramers to heptamers. However, oligosaccharides produced by $Sph3_h$ treatment were consistent with cleavage of $(\alpha\text{-}1,4\text{-}GalNAc)_9$ near the reducing end of the substrate, while $PelA_h$ treatment released reduced $GalNAc_{4\text{-}6}$ consistent with cleavage near the non-reducing end (Fig 7).

Mapping of sequence conservation based on alignments of $Sph3_h$ and $PelA_h$ to respective homologous proteins revealed high degrees of sequence conservation in the active site groove of each enzymes (Fig 7C and D). However, closer examination of the patterns of conservation shows differences between the two enzymes (Fig. 7C and D). For $Sph3_h$, surface conservation extends further on the non-reducing side of +1/-1 cleavage site suggesting approximately five conserved substrate binding subsites (-5 to -1, Fig 7C). This observation is consistent with our data that shows that this enzyme produces a minimum length pentasaccharide from the non-reducing end (Fig 7A). In contrast, $PelA_h$ has surface conservation that extends on the reducing end side of cleavage and at least five subsites could be mapped (+1 to +5, Fig 7B). Surface residue conservation thus correlates well with the results of the $GalNAc_9$ hydrolysis experiments.

$PelA_h$ cleaves deacetylated-rich regions within GAG

Comparison of the structures of $Sph3_h$ and $PelA_h$ revealed that the $PelA_h$ active site groove is more electronegative than that of $Sph3_h$ (Fig. 8A), suggesting that $PelA_h$ may be able to bind and potentially cleave cationic oligosaccharides, such as partially deacetylated oligo-GalNAc. Although neither enzyme was able to cleave α-1,4 GalN homopolymers, it is possible that these enzymes may be able to cleave α-1,4 GalNAc linkages

within GalN-rich regions of GAG. Although GalN-containing oligosaccharides were not detected during the biofilm disruption assay (Fig 1), these cationic degradation products could have remained adherent to the negatively charged hyphal cell wall. To test this hypothesis, a cell-free GAG degradation assay was performed using purified secreted GAG. $Sph3_h$ treatment of secreted GAG resulted in the release of predominately homo-GalNAc oligosaccharides (86%) with a smaller amount of mono-deacetylated GalNAc oligosaccharides (14%, Fig. 8B). In contrast, treatment of secreted GAG with $PelA_h$ produced almost exclusively partially deacetylated GalNAc oligosaccharides (98%, Fig. 8B) suggesting that $PelA_h$ preferentially binds and degrades GalN-containing regions of GAG. Consistent with our prior results, no galactose homo- or heteropolymers, nor GalN homopolymers were found, confirming that $Sph3_h$ and $PelA_h$ specifically cleave α-1,4 GalNAc linkages (Fig. 2B).

DISCUSSION:

We previously described the structure of $Sph3_h$ and found that this enzyme constituted the first member of a new glycoside hydrolase family, GH135, with activity against GAG and *Aspergillus* biofilms (24). More recently, the glycoside hydrolase domain of PelA from *P. aeruginosa* has also been reported to disrupt *A. fumigatus* biofilms (19). However, the molecular mechanism and specificity these enzymes had not been elucidated.

Herein, the mechanism by which these enzymes disrupt *A. fumigatus* biofilms was studied using a combination of structural biology, and mass and NMR spectroscopy. These studies revealed that $Sph3_h$ and $PelA_h$ share a conserved active site and cleave α-1,4-GalNAc glycosidic linkages using a retaining enzyme mechanism. Despite these similarities of specificity and mechanism, the interaction between the enzymes and the substrate differ. The $Sph3_h$ has a neutral electrostatic cleft in contrast to the deep electronegative groove of the $PelA_h$. These observations correlate with the nature of the products observed after cleavage of secreted GAG. The electronegative $PelA_h$ was able to bind and release cationic GalN-containing oligosaccharide more efficiently, while the neutral $Sph3_h$ released mostly neutral GalNAc oligosaccharides. Furthermore, despite interacting with the same substrate, the two glycoside hydrolases have extended conserved substrate binding sites on opposites sides of the catalytic -1/+1 site. This difference leads to $Sph3_h$ cleaving closer to the reducing end, while $PelA_h$ cleaves nearer to the non-reducing end of oligosaccharides.

$Sph3_h$ is structurally similar to GH27 family members which rely on two acidic residues to cleave the glycosidic bond (24). GH27 catalytic residues align structurally with D166 and E222 in $Sph3_h$. Previously, mutagenesis studies on $Sph3_h$ identified D166 as essential for catalysis (24). E222 was not essential *in vitro* as mutation to glutamine retained some activity (24). Replacement of the catalytic acid/base residue with glutamine does not abolish activity in some retaining hydrolases but slows the rate of reaction (34,35). Mutation of the acid/base glutamate in the *Sulfolobus solfataricus* β-glycosidase reduced activity 10- to 60-fold depending on the substrate (36). In $Sph3_h$, E222 is coordinated by Y88 and N202, and these residues may aid in activation of the glutamine in the E222Q mutant. $PelA_h$ has equivalent catalytic residues that align structurally to those of $Sph3_h$. NMR analysis found that both enzymes utilize a retaining mechanism. Analysis of the structure of $Sph3_h$ and $PelA_h$ revealed that the distances between the carboxyl side chains of the catalytic residues E222 and D166 in $Sph3_h$ and E218 and D160 in $PelA_h$ correspond to the average distance between the catalytic residues of retaining hydrolases (32). Thus, the structures of $Sph3_h$ and $PelA_h$ and the NMR results are consistent with these enzymes utilizing a retaining mechanism (Fig. 4).

Our findings demonstrate that $PelA_h$ and $Sph3_h$ share a similar (β/α) barrel fold with central grooves containing a highly conserved active sites,

underlying their endo-α-1,4-*N*-acetyl-galactosaminidase activity. The high identity between the catalytic residues of the active site suggests that the difference in the ability of these enzymes to hydrolyze oligosaccharides containing deacetylated residues is a reflection of differences in polysaccharide binding affinity. This hypothesis is supported by the fact that, unlike $Sph3_h$, the binding groove of $PelA_h$ is highly electronegative, suggesting an affinity for cationic substrates as seen in the soluble GAG degradation assay. PelA is a multidomain protein with both hydrolase and carbohydrate deacetylase activity. The deacetylase activity has been shown to be required for *in vivo* Pel dependent biofilm formation (20). Our results suggest that the PelA deacetylase domain likely acts on Pel first, rendering the polysaccharide cationic prior to hydrolysis of the polymer. Similar findings have been reported for the poly-β1-6-N-acetyl-D-glucosamine (PNAG) modifying enzyme, PgaB, which contains both a deacetylase domain and a GH153 hydrolase domain (37). Studies of PgaB demonstrated that the GH153 domain requires a specific pattern of partially deacetylated PNAG as a substrate (37). $PelA_h$ appears to be more promiscuous in its substrate specificity, as the enzyme can hydrolyze pure acetylated oligomers.

Another difference between $Sph3_h$ and $PelA_h$ was revealed by their activity on reduced GalNAc nonamers (Fig 7). $Sph3_h$ hydrolyzed oligosaccharides proximal to the reducing end while $PelA_h$ cleaved near the non-reducing end. This observation suggests differences in the roles of the substrate binding residues on either side of the catalytic site in each enzyme. This hypothesis is substantiated by the differences in surface residue conservation in the putative binding grooves of $Sph3_h$ and $PelA_h$. $Sph3_h$ contains a longer stretch of conserved residues on the non-reducing side of the -1/+1 cleavage site. The conserved residues in $Sph3_h$ span the -5 to +2 binding subsites. In contrast, the conserved surface of $PelA_h$ aligns with -2 to +5 subsites.

Previously, we reported that $PelA_h$ was able to disrupt both GAG- and Pel-dependent biofilms. In contrast, $Sph3_h$ was only active only against GAG-dependent fungal biofilms, and was unable to cleave Pel-dependent biofilms despite being able to bind Pel polysaccharide(19,24). The results of our structure function studies reveal important differences in the substrate specificity and structure of these enzymes that may explain these observations. $PelA_h$ was found to have a deeper, more electronegative substrate binding groove and exhibit preferential activity against partially deacetylated substrates compared with $Sph3_h$. While $Sph3_h$ treatment of soluble GAG releases predominately acetylated oligomers, treatment with $PelA_h$ releases 98% partially deacetylated products. Although the detailed structure of Pel has not been determined, compositional studies suggest it contains both GlcNAc and GalNAc (23). The degree of deacetylation, and the identity of the sugars deacetylated in the Pel polysaccharide have not been determined. Our results suggest the inactivity of Sph3h against Pel may reflect the fact that Pel is more extensively deacetylated than GAG. Alternately, while the activity of PelA against α-linked N-acetyl galactosamine suggests that these α-linkages are present within Pel, it is possible that the presence of GlcNAc sugars within the Pel polymer may interfere with $Sph3_h$ ability to cleave the Pel polymer. Distinguishing among these possibilities, and confirming the presence of α-linked N-acetyl galactosamine in Pel will require a complete structural analysis of this polysaccharide.The results of these studies assign enzymatic function to two new glycoside hydrolases. $Sph3_h$ is the only GH135 family member that has been functionally characterized. While the structure of $Sph3_h$ had been previously determined (24), the current study expands our understanding of the activity of this enzyme and demonstrates that $Sph3_h$ functions as a retaining endo-α-1,4-N-acetylgalactosaminidase with specificity for fungal GAG. We also found that the bacterial enzyme $PelA_h$ is a retaining endo-α-1,4-N-

acetylgalactosaminidase. As the sequence identity between $PelA_h$ and current glycoside hydrolase families is low, our structural and functional characterization of this protein has enables us to classify $PelA_h$ as the first member of a new family, GH166.

MATERIAL AND METHODS:

Protein expression and purification of $PelA_h$ for structural studies - $PelA_h$ encompassing residues 47-303 of the mature $PelA_h$ protein was expressed and purified as previously described (21). Selenomethionine (Se-Met) labeled protein was produced as previously described (38) with B834 Met- E. coli cells (Novagen) and purified as described for the native protein.

Biofilm disruption assay - 10^4 A. fumigatus conidia were grown in Brian media in polystyrene, 96-well plates non-tissue culture treated for 21 h at 37 °C and then treated with the indicated concentration of glycoside hydrolase in 1X PBS for 1 h at room temperature under gentle agitation. Biofilms were then washed, stained with 0.1% (w/v) crystal violet and destained with 100% ethanol for 10 min. The optical density of the destain solution was measured at 600 nm.

MALDI-TOF MS enzymatic fingerprint - Products of enzymatic digestions were diluted in 0.2% Trifluoroacetic acid (TFA) before being spotted on the MALDI-TOF plate in a ratio 1:1 (v:v) with 5 mg/ml dihydroxybenzoic acid (DHB) matrix reconstituted in acetonitrile ACN - 0.2% TFA (70:30, v:v). Spectra were recorded on a Bruker UltrafleXtreme in positive reflector mode and represent an accumulation of 5000 laser shots. MALDI-TOF MS/MS experiments were performed using the same mass spectrometer.

Gas chromatography-Mass spectrometry monosaccharide composition - Oligosaccharides enzymatically extracted from biofilm were hydrolysed with 6M hydrochloric acid (HCl) for 4 h at 100 °C. After drying, samples were derivatized and analyzed as previously described (39). Briefly, samples were then converted in methyl glycosides by heating in 1 M methanol-HCl (Supelco) for 16 h at 80°C. Samples were dried, washed twice with methanol prior re-N-acetylating hexosamine residues. Re-N-acetylation was performed by incubation with a mix of methanol : pyridine : anhydride acetic (10 : 2 : 3) for 1h at room temperature. Samples were then treated with hexamethyldisilazane : trimethylchlorosilane : pyridine solution (3: 1: 9, Supelco) for 20 min at 80 °C. The resulting trimethylsilyl methyl glycosides were dried, resuspended in 1 mL of cyclohexane and injected in the Trace1300 GC-MS system equipped with a CP-Sil5-CB capillary column (Agilent Technologies). Elution was performed with the following temperature gradient: 120 °C to 160 °C at a rate of 10 °C/min, 160 °C to 220 °C at a rate of 1.5 °C/min, 220 °C to 280 °C at a rate of 20 °C/min. Identification and quantification of each monosaccharides were carried out using standards and response factors determined for each monosaccharide.

Production of oligosaccharide of chitin, α-1,4-Gal and α-1,4-GalN and specificity study - Oligosaccharides of Chitin/chitosan were produced by acidic partial hydrolysis of chitin from shrimp shell (Sigma). Briefly, chitin was incubated in 0.1 M HCl for 2 h at 100 °C. Solubilized oligosaccharides were then purified on a Hypersep Hypercarb SPE cartridge (Thermofisher) conditioned as per manufacturer instructions. After loading the sample, the cartridge was washed with water, 5% (v/v) ACN, and oligosaccharides were eluted with 50% (v/v) ACN.

The α-1,4-Gal and α-1,4-GalN oligosaccharides were chemically synthetized based on the use of di-tert-butylsilylene (DTBS) group-protected building blocks. These extremely powerful cis-galactosylating agents only produced α-isomer products. After glycosylation reactions, the DTBS

group was removed with HF.pyridine. The resulting free 6-hydroxyl group was protected with benzoyl group selectively to afford the acceptor. The final compounds were obtained after deprotection by saponification, debenzylation, azide reduction and acetylation of amine groups.

All oligosaccharides were incubated with 5 µM $Sph3_h$ or $PelA_h$ for 1 h at room temperature and analyzed by MALDI-TOF MS. Controls of enzyme functionality were performed using as substrate α-1,4-GalNAc oligosaccharides produced as described below.

Enzymatic mechanism determination by 1H-NMR – Proton NMR spectra were recorded on an AVANCE III HD 600 NMR spectrometer (Ascend™ 600 magnet - Bruker Biospin Ltd.) operating at a frequency of 600.17 MHz for 1H and equipped with a quadruple resonance CryoProbe (CPQCI 1H-31P/13C/15N) and a SampleJETTM autosampler. For each sample, 0.45 mg of α-1,4-GalNAc octamer was resuspended in 160 µL 0.1X PBS in D2O containing 0.5 mM trimethylsilylpropanoic acid (TSP) and transferred in a 3 mm NMR tube. The 1H NMR spectra were continuously acquired at 25 °C for 4 h with lock and shim performed on every 10 experiments. A new spectrum at 25°C was recorded after 24h to observe the mutarotation and confirmation of this event was validated with the acquisition of ^{1}H NMR spectra at 37 °C and a 2D ^{1}H-^{1}H Correlation Spectroscopy (COSY) spectra at 25 °C. The ^{1}H spectra were acquired using the pulse sequence *noesypr1d* (Bruker Biospin Ltd) in order to achieve good suppression of the water signal. Each ^{1}H spectrum was acquired with 32 scans, a ^{1}H 90º pulse length of 7.8 µs, a mixing time of 10 ms, a spectral width of 12 kHz, a recycle delay of 4 s for a total of 66K data points. The 2D ^{1}H-^{1}H COSY spectra were acquired using 16 scans with a ^{1}H 90º pulse length of 8 µs, a spectral width of 3 kHz in both dimensions, a repetition delay of 1.8 s for a total of 2048 data points in F2 and 128 increments in F1. All spectra were processed using TOPSPIN software (version 3.5 pl 7, Bruker Biospin Ltd).

Production of pure α-1,4-GalNAc oligosaccharides – A. fumigatus biofilm were incubated with 5 nM $Sph3_h$ for 1 h at room temperature, solubilized oligosaccharides were then further purified on a Sep-pak C18 cartridge. In brief, cartridges were conditioned using absolute ethanol followed by water. Samples were then loaded onto the cartridge before washing and eluting using a 0.25% (v/v) step gradient of methanol from 0 to 4% (v/v) followed by a 1% (v/v) step gradient of ACN from 1 to 4%.

Reduction of oligosaccharides - Reduction of the α-1,4-GalNAc 9-mer was performed resuspending the oligosaccharides in 1 M ammonium hydroxide containing 10 mg/ml sodium borohydride and incubating overnight at room temperature. Reaction was quenched adding dropwise 30% (v/v) acetic acid and samples were purified using a Hypersep Hypercarb SPE cartridge (Thermofisher) as per chitin oligosaccharide purification.

Crystallization, data collection, and structure solution - Purified $PelA_h$ was concentrated to ~20 mg/mL and crystallization trials were performed using MCSG1-4 sparse-matrix screens (Microlytic) in 48-well hanging-drop VDX plates (Hampton Research) using a 2 µL drop with a 1:1 protein:precipitant ratio at 20 °C. Initial crystallization hits were obtained in several conditions. Diffraction quality native $PelA_h$ crystals were grown using 0.1 M Bis-Tris pH 7.5, 25% (w/v) PEG MME 5000 at a 1:2 ratio of protein to crystallization solution at 20 °C. Se-Met-labeled $PelA_h$ was crystallized in a similar condition with 26% (w/v) PEG MME 5000. Both crystals were cryoprotected for 10 s in mother liquor supplemented with 15% (v/v) ethylene glycol prior to vitrification in liquid nitrogen.

Diffraction data were collected at -173°C with wavelengths of 0.9791 and 1.075 Å on beamline X29, National Synchrotron Light Source (NSLS) (**Table 1**) for the Se-Met labeled and native crystal,

respectively. A high redundancy dataset was generated for the Se-Met labeled $PelA_h$ by collecting 90 images with 2° oscillation at 90% beam attenuation and with an exposure time of 0.3 s/image and 360 images with 1° oscillation with 50% beam attenuation with an exposure time of 0.4 s/image on an ADSC Quantum-315 detector with a 260 mm crystal-to-detector distance. The native $PelA_h$ dataset was collected using the same strategy as described above for the Se-Met crystal but with a 180 mm crystal-to-detector distance. Autosol (40) was used to determine initial phases and generate a density-modified map. The resulting electron density map was of high quality and enabled PHENIX AutoBuild to build >95 % of the protein. The remaining residues were built manually in COOT (41,42) and the structure refined using PHENIX.REFINE (43). Translation/Libration/Screw (TLS) groups were added to the refinement in PHENIX through the use of the TLSMD server (44,45).

All structure figures were generated using the PyMOL molecular graphics system (DeLano Scientific) (46), or Chimera (47) for electrostatics using APBS (48). Structural similarity to deposited structures in the protein data bank (PDB) was determined using DALI and structure alignment was performed in COOT (27,41). Amino acid conservation was calculated using the Consurf server (49) aligned to 113 proteins using the default settings. Programs used for crystallographic data processing and analysis were accessed through SBGrid (50).

Secreted GAG purification and digest - Secreted purified GAG was prepared as previously reported (13). Briefly, culture supernatant of a 3 days old Af293 culture was filtered on Miracloth prior to be ethanol precipitated. Precipitate was then successively washed with 70% (v/v) ethanol twice, 150 mM NaCl and water. The remaining gel was then freeze dried. The dried purified GAG was incubated with the 1 µM glycoside hydrolase for 1 h in 0.1X PBS. The released soluble oligosaccharides were then analyzed using the MALDI-TOF MS enzymatic fingerprint technique.

ACKNOWLEDGMENTS:

We would like to thank Dr Todd Lowary from the Department of Chemistry, University of Alberta, for its assistance in the design and interpretation of the NMR spectra. We would like to thank the Drug Discovery platform of the Research Institute of the McGill Health Center for the use of the MALDI-TOF mass spectrometer and NMR equipment.

AUTHOR CONTRIBUTIONS:

F. Le Mauff: Main author of the manuscript, and conceived, ran, interpreted all the mass spectrometry and NMR experiments and data.

N. Bamford: Main author of the manuscript, participated to the elucidation and analysis of $PelA_h$ structure and generated all associated figures.

N. Alnabelseya, P. Baker, H. Robinson: participated to the structure determintaion of $PelA_h$.

Y. Zhang and J. Codee: Synthesized the short synthetic oligosaccharides and reviewed the manuscript.

L. Howell and D. Sheppard: Conceived the project, designed the experiments and wrote the manuscript.

FOOTNOTES – Research described in this paper is supported by operating grants from the *Canadian Institutes of Health Research* (CIHR) (#81361 to P.L.H. and D.C.S, FDN-154327 and #43998 to P.L.H, and

123306 and FDN-159902 to D.C.S.) and Cystic Fibrosis Canada (D.C.S. and P.L.H. #558692). N.C.B has been supported in part by a graduate scholarship from the *Natural Sciences and Engineering Research Council of Canada* (NSERC) including Vanier and CGS-M, Mary H. Beatty, and Dr. James A. and Connie P. Dickson Scholarships from the University of Toronto, Cystic Fibrosis Canada, and The Hospital for Sick Children. P.L.H. is the recipient of a Tier I Canada Research Chair. D.C.S is supported by a Chercheur-Boursier Award from the *Fonds de Recherche Quebec Santé* (FRQS) #34961. Beam Line X29 at the National Synchrotron Light Source is supported by the United States Department of Energy Office and by the National Center for Research Resources Grant P41RR012408 and the National Institute of General Medical Sciences Grant P41GM103473, from the National Institutes of Health. The authors gratefully acknowledge support from the European Research Council (ERC), grant awarded to JDCC, grant number ERC-CoG-726072-GLYCONTROL.

The atomic coordinates and structure factors for the $PelA_h$ structure (code 5TCB) have been deposited in the Protein Data Bank (http://www.rcsb.org/).

The authors declare that they have no conflicts of interest with the contents of this article.

REFERENCES:

1. Latgé, J.-P. (2001) The pathobiology of Aspergillus fumigatus. *Trends in microbiology* **9**, 382-389
2. Taccone, F. S., Van den Abeele, A. M., Bulpa, P., Misset, B., Meersseman, W., Cardoso, T., Paiva, J. A., Blasco-Navalpotro, M., De Laere, E., Dimopoulos, G., Rello, J., Vogelaers, D., and Blot, S. I. (2015) Epidemiology of invasive aspergillosis in critically ill patients: clinical presentation, underlying conditions, and outcomes. *Critical care (London, England)* **19**, 7
3. Mehrad, B., Strieter, R. M., and Standiford, T. J. (1999) Role of TNF-α in Pulmonary Host Defense in Murine Invasive Aspergillosis. *The Journal of Immunology* **162**, 1633-1640
4. Müller, F.-M. C., Seidler, M., and Beauvais, A. (2011) Aspergillus fumigatus biofilms in the clinical setting. *Medical mycology* **49**, S96-S100
5. Pierce, C. G., Srinivasan, A., Uppuluri, P., Ramasubramanian, A. K., and López-Ribot, J. L. (2013) Antifungal therapy with an emphasis on biofilms. *Current opinion in pharmacology* **13**, 726-730
6. Seidler, M. J., Salvenmoser, S., and Müller, F.-M. C. (2008) Aspergillus fumigatus forms biofilms with reduced antifungal drug susceptibility on bronchial epithelial cells. *Antimicrobial agents and chemotherapy* **52**, 4130-4136
7. Kernien, J. F., Snarr, B. D., Sheppard, D. C., and Nett, J. E. (2017) The interface between Fungal Biofilms and innate immunity. *Frontiers in immunology* **8**, 1968
8. Ramage, G., Rajendran, R., Sherry, L., and Williams, C. (2012) Fungal biofilm resistance. *International journal of microbiology* **2012**
9. Reichhardt, C., Ferreira, J. A., Joubert, L.-M., Clemons, K. V., Stevens, D. A., and Cegelski, L. (2015) Analysis of the Aspergillus fumigatus biofilm extracellular matrix by solid-state nuclear magnetic resonance spectroscopy. *Eukaryotic cell* **14**, 1064-1072

10. Loussert, C., Schmitt, C., Prevost, M. C., Balloy, V., Fadel, E., Philippe, B., Kauffmann-Lacroix, C., Latgé, J. P., and Beauvais, A. (2010) In vivo biofilm composition of Aspergillus fumigatus. *Cellular microbiology* **12**, 405-410
11. Sheppard, D. C., and Howell, P. L. (2016) Biofilm exopolysaccharides of pathogenic fungi: lessons from bacteria. *Journal of Biological Chemistry* **291**, 12529-12537
12. Lee, M. J., Gravelat, F. N., Cerone, R. P., Baptista, S. D., Campoli, P. V., Choe, S. I., Kravtsov, I., Vinogradov, E., Creuzenet, C., Liu, H., Berghuis, A. M., Latge, J. P., Filler, S. G., Fontaine, T., and Sheppard, D. C. (2014) Overlapping and distinct roles of Aspergillus fumigatus UDP-glucose 4-epimerases in galactose metabolism and the synthesis of galactose-containing cell wall polysaccharides. *J Biol Chem* **289**, 1243-1256
13. Lee, M. J., Geller, A. M., Bamford, N. C., Liu, H., Gravelat, F. N., Snarr, B. D., Le Mauff, F., Chabot, J., Ralph, B., and Ostapska, H. (2016) Deacetylation of fungal exopolysaccharide mediates adhesion and biofilm formation. *MBio* **7**, e00252-00216
14. Briard, B., Muszkieta, L., Latgé, J.-P., and Fontaine, T. (2016) Galactosaminogalactan of Aspergillus fumigatus, a bioactive fungal polymer. *Mycologia* **108**, 572-580
15. Sheppard, D. C. (2011) Molecular mechanism of Aspergillus fumigatus adherence to host constituents. *Curr Opin Microbiol* **14**, 375-379
16. Lee, M. J., Liu, H., Barker, B. M., Snarr, B. D., Gravelat, F. N., Al Abdallah, Q., Gavino, C., Baistrocchi, S. R., Ostapska, H., and Xiao, T. (2015) The fungal exopolysaccharide galactosaminogalactan mediates virulence by enhancing resistance to neutrophil extracellular traps. *PLoS pathogens* **11**, e1005187
17. Gravelat, F. N., Beauvais, A., Liu, H., Lee, M. J., Snarr, B. D., Chen, D., Xu, W., Kravtsov, I., Hoareau, C. M., and Vanier, G. (2013) Aspergillus galactosaminogalactan mediates adherence to host constituents and conceals hyphal β-glucan from the immune system. *PLoS pathogens* **9**, e1003575
18. van de Veerdonk, F. L., Gresnigt, M. S., Romani, L., Netea, M. G., and Latgé, J.-P. (2017) Aspergillus fumigatus morphology and dynamic host interactions. *Nature Reviews Microbiology* **15**, 661
19. Snarr, B. D., Baker, P., Bamford, N. C., Sato, Y., Liu, H., Lehoux, M., Gravelat, F. N., Ostapska, H., Baistrocchi, S. R., and Cerone, R. P. (2017) Microbial glycoside hydrolases as antibiofilm agents with cross-kingdom activity. *Proceedings of the National Academy of Sciences* **114**, 7124-7129
20. Colvin, K. M., Alnabelseya, N., Baker, P., Whitney, J. C., Howell, P. L., and Parsek, M. R. (2013) PelA deacetylase activity is required for Pel polysaccharide synthesis in Pseudomonas aeruginosa. *Journal of bacteriology* **195**, 2329-2339
21. Baker, P., Hill, P. J., Snarr, B. D., Alnabelseya, N., Pestrak, M. J., Lee, M. J., Jennings, L. K., Tam, J., Melnyk, R. A., and Parsek, M. R. (2016) Exopolysaccharide biosynthetic glycoside hydrolases can be utilized to disrupt and prevent Pseudomonas aeruginosa biofilms. *Science advances* **2**, e1501632
22. Franklin, M. J., Nivens, D. E., Weadge, J. T., and Howell, P. L. (2011) Biosynthesis of the Pseudomonas aeruginosa Extracellular Polysaccharides, Alginate, Pel, and Psl. *Front Microbiol* **2**, 167
23. Jennings, L. K., Storek, K. M., Ledvina, H. E., Coulon, C., Marmont, L. S., Sadovskaya, I., Secor, P. R., Tseng, B. S., Scian, M., and Filloux, A. (2015) Pel is a cationic exopolysaccharide that cross-links extracellular DNA in the Pseudomonas aeruginosa biofilm matrix. *Proceedings of the National Academy of Sciences* **112**, 11353-11358

24. Bamford, N. C., Snarr, B. D., Gravelat, F. N., Little, D. J., Lee, M. J., Zacharias, C. A., Chabot, J. C., Geller, A. M., Baptista, S. D., and Baker, P. (2015) Sph3 is a glycoside hydrolase required for the biosynthesis of galactosaminogalactan in Aspergillus fumigatus. *Journal of Biological Chemistry* **290**, 27438-27450
25. Yang, S., Wang, M., Chen, L., Yin, B., Song, G., Turko, I. V., Phinney, K. W., Betenbaugh, M. J., Zhang, H., and Li, S. (2015) QUANTITY: an isobaric tag for quantitative glycomics. *Scientific reports* **5**, 17585
26. Drozdetskiy, A., Cole, C., Procter, J., and Barton, G. J. (2015) JPred4: a protein secondary structure prediction server. *Nucleic Acids Res* **43**, W389-394
27. Holm, L., and Laakso, L. M. J. N. a. r. (2016) Dali server update. **44**, W351-W355
28. Yin, Y., Mao, X., Yang, J., Chen, X., Mao, F., and Xu, Y. (2012) dbCAN: a web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res* **40**, W445-451
29. Zhang, H., Yohe, T., Huang, L., Entwistle, S., Wu, P., Yang, Z., Busk, P. K., Xu, Y., and Yin, Y. (2018) dbCAN2: a meta server for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res* **46**, W95-W101
30. Henrissat, B., and Bairoch, A. (1996) Updating the sequence-based classification of glycosyl hydrolases. *Biochemical Journal* **316**, 695
31. Payne, C. M., Baban, J., Horn, S. J., Backe, P. H., Arvai, A. S., Dalhus, B., Bjørås, M., Eijsink, V. G., Sørlie, M., and Beckham, G. T. (2012) Hallmarks of processivity in glycoside hydrolases from crystallographic and computational studies of the Serratia marcescens chitinases. *Journal of Biological Chemistry*, jbc. M112. 402149
32. Davies, G., and Henrissat, B. (1995) Structures and mechanisms of glycosyl hydrolases. *Structure* **3**, 853-859
33. Kleywegt, G. J., Zou, J.-Y., Divne, C., Davies, G. J., Sinning, I., Ståhlberg, J., Reinikainen, T., Srisodsuk, M., Teeri, T. T., and Jones, T. A. (1997) The crystal structure of the catalytic core domain of endoglucanase I from Trichoderma reesei at 3.6 Å resolution, and a comparison with related enzymes1. *Journal of molecular biology* **272**, 383-397
34. Cecchini, D. A., Fauré, R., Laville, E., and Potocki-Veronese, G. J. F. l. (2015) Biochemical identification of the catalytic residues of a glycoside hydrolase family 120 β-xylosidase, involved in xylooligosaccharide metabolisation by gut bacteria. **589**, 3098-3106
35. Zechel, D. L., and Withers, S. G. J. A. o. c. r. (2000) Glycosidase mechanisms: anatomy of a finely tuned catalyst. **33**, 11-18
36. Moracci, M., Capalbo, L., Ciaramella, M., Rossi, M. J. P. E., Design, and Selection. (1996) Identification of two glutamic acid residues essential for catalysis in the β-glycosidase from the thermoacidophilic archaeon Sulfolobus solfataricus. **9**, 1191-1195
37. Little, D. J., Pfoh, R., Le Mauff, F., Bamford, N. C., Notte, C., Baker, P., Guragain, M., Robinson, H., Pier, G. B., and Nitz, M. (2018) PgaB orthologues contain a glycoside hydrolase domain that cleaves deacetylated poly-β (1, 6)-N-acetylglucosamine and can disrupt bacterial biofilms. *PLoS pathogens* **14**, e1006998
38. Lee, J. E., Cornell, K. A., Riscoe, M. K., and Howell, P. L. (2001) Structure of E. coli 5′-methylthioadenosine/S-adenosylhomocysteine nucleosidase reveals similarity to the purine nucleoside phosphorylases. *Structure* **9**, 941-953
39. Le Mauff, F., Loutelier-Bourhis, C., Bardor, M., Berard, C., Doucet, A., D'Aoust, M. A., Vezina, L. P., Driouich, A., Couture, M. M., and Lerouge, P. (2017) Cell wall biochemical alterations during Agrobacterium-mediated expression of haemagglutinin-based influenza virus-like vaccine particles in tobacco. *Plant Biotechnol J* **15**, 285-296

40. Terwilliger, T. C., and Berendzen, J. (1999) Automated MAD and MIR structure solution. *Acta Crystallographica Section D: Biological Crystallography* **55**, 849-861
41. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D: Biological Crystallography* **60**, 2126-2132
42. Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., and Grosse-Kunstleve, R. W. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D: Biological Crystallography* **66**, 213-221
43. Afonine, P. V., Mustyakimov, M., Grosse-Kunstleve, R. W., Moriarty, N. W., Langan, P., and Adams, P. D. (2010) Joint X-ray and neutron refinement with phenix. refine. *Acta Crystallographica Section D: Biological Crystallography* **66**, 1153-1163
44. Painter, J., and Merritt, E. A. (2006) Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallographica Section D* **62**, 439-450
45. Painter, J., and Merritt, E. A. (2006) TLSMD web server for the generation of multi-group TLS models. *Journal of Applied Crystallography* **39**, 109-111
46. DeLano, W. L. (2002) The PyMOL molecular graphics system. *http://www. pymol. org*
47. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera—a visualization system for exploratory research and analysis. *Journal of computational chemistry* **25**, 1605-1612
48. Baker, N. A., Sept, D., Joseph, S., Holst, M. J., and McCammon, J. A. (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proceedings of the National Academy of Sciences* **98**, 10037-10041
49. Ashkenazy, H., Erez, E., Martz, E., Pupko, T., and Ben-Tal, N. (2010) ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic acids research* **38**, W529-W533
50. Morin, A., Eisenbraun, B., Key, J., Sanschagrin, P. C., Timony, M. A., Ottaviano, M., and Sliz, P. (2013) Cutting edge: Collaboration gets the most out of software. *elife* **2**, e01456
51. Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66**, 12-21
52. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221

ABBREVIATIONS:

GAG: Galactosaminogalactan

GH: Glycoside Hydrolase

CAZy : Carbohydrate Active Enzyme

BLAST: Basic Local Alignment Search Tool

HexNAc: N-Acetylhexosamine

RMSD: Root Mean Square Deviation

PNAG: Poly-β-1,6-N-acetyl-D-glucosamine

DHB : DiHydrohybenzoic acid

ACN : Acetonitrile

HCl : Hydrolchloric acid

DTBS : Ditert-butylsilylene

TSP : Trimethylsilylopropanoic acid

SPE : Solid Phase Extraction

NSLS : National Synchroton Light Source

TABLE 1: Summary of data collection and refinement statistics.

(Values in parentheses correspond to the highest resolution shell).

	SeMet $PelA_h$	**$PelA_h$**
Data collection		
Beamline	NSLS X29	NSLS X29
Wavelength (Å)	0.979	1.075
Space group	$P2_12_12$	$P2_12_12$
Cell dimensions		
a, *b*, *c* (Å)	61.3, 85.2, 47.2	65.0, 84.0, 47.2
α, β, γ (°)	90, 90, 90	90, 90, 90
Resolution (Å)	50.0 – 1.90 (1.97-1.90)	50.00 – 1.54 (1.58-1.54)
No. reflections	304767	665419
No. of unique reflections	19863	39193
$I / \sigma I$	21.7 (4.5)	18.6 (6.6)
Completeness (%)	99.8 (99.7)	99.9 (99.5)
R_{merge} (%)	12.2 (62.5)	18.1 (44.3)
Refinement		
R_{work} / R_{free}		16.9 / 19.2
No. of atoms		
Protein		1935
Water		193
Average B-factors (Å^2)		
Protein		31.6
Water		38.3
RMS deviations		
Bond lengths (Å)		0.008
Bond angles (°)		0.94
Ramachandran plot		
Total favored (%)		96.8
Total allowed (%)		100
Coordinate error (Å)		0.13
PDB code		**5TCB**

[1]$R_{merge} = \sum\sum | I(k) - \langle I \rangle | / \sum I(k)$ where $I(k)$ and $\langle I \rangle$ represent the diffraction intensity values of the individual measurements and the corresponding mean values. The summation is over all unique measurements.

[2]$R_{work} = \sum ||F_{obs}| - k|F_{calc}|| / |F_{obs}|$ where F_{obs} and F_{calc} are the observed and calculated structure factors, respectively.

[2]R_{free} is the sum extended over a subset of reflections (5%) excluded from all stages of the refinement.

[3]As calculated using MolProbity (51).

[4]Maximum-Likelihood Based Coordinate Error, as determined by PHENIX (52).

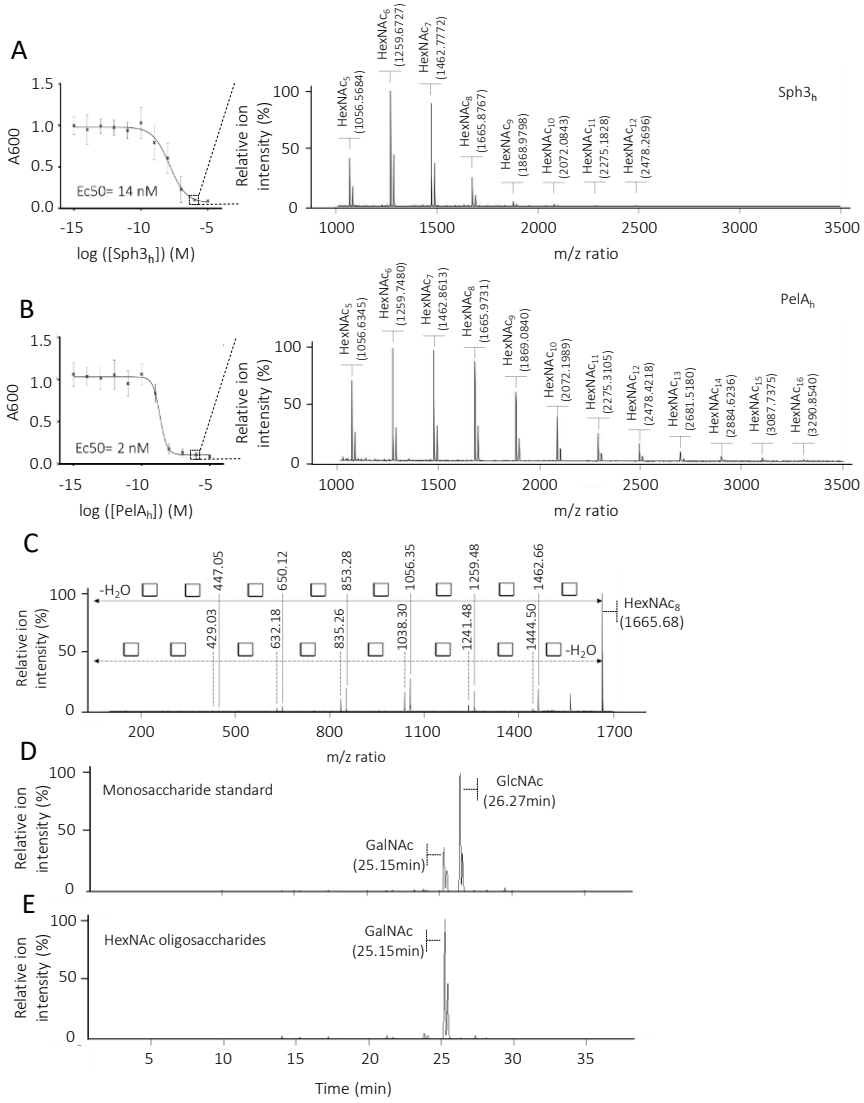

Figure 1. *A. fumigatus* biofilm degradation by $Sph3_h$ and $PelA_h$ release α-1,4-GalNAc oligosaccharides. (A) Crystal violet staining of pre-grown *A. fumigatus* biofilm treated with the indicated concentration of $Sph3_h$ and the associated MALDI-TOF MS spectra of oligosaccharides released from the biofilm treated with 100nM $Sph3_h$ (circled concentration) **(B)** Crystal violet staining of pre-grown *A. fumigatus* biofilms treated with the indicated concentration of $PelA_h$ and the associated MALDI-TOF MS spectra of oligosaccharides released from the biofilm treated with 100nM $PelA_h$ (circled concentration). **(C)** MS-MS fragmentation spectra of m/z=1665.7 ion issued of the $Sph3_h$ treatment and matching with $HexNAc_8$ oligosacaccharide. This structure is a homopolymer of units with an average m/z of 203.1086, corresponding to the m/z ratio of a N-acetylhexosamine (white square), consistent with a $HexNAc_8$ oligosaccharide. **(D)** GC-MS extracted ion chromatogram m/z=173.0 of HexNAc monosaccharide standards mix. **(E)** GC-MS extracted ion chromatogram m/z=173.0 of the biofilm released HexNAc oligosaccharides.

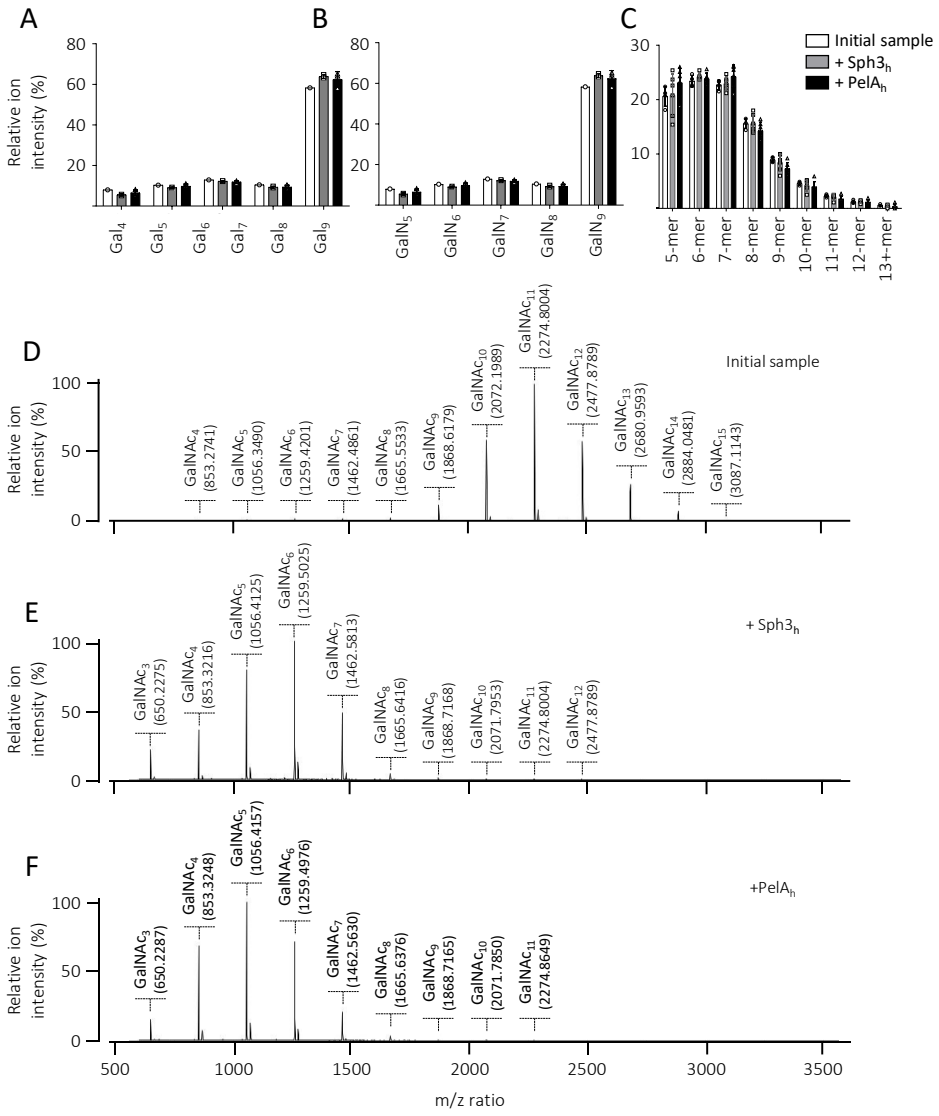

Figure 2. $Sph3_h$ and $PelA_h$ are specific for α-1,4-GalNAc. MALDI-TOF MS analysis of the products released by the incubation of $Sph3_h$ and $PelA_h$ with **(A)** a fraction enriched in α-1,4-galactose 9-mers, and **(B)** a fraction enriched in α-1,4-galactosamine 9-mers, **(C)** chemically partially hydrolysed chitin. Initial sample is represented in white, following $Sph3_h$ treatment in grey and $PelA_h$ treatment in black. MALDI-TOF MS spectra of **(D)** a sample enriched in α-1,4-GalNAc 4-mer to 15-mer. MALDI-TOF MS spectra of the enzymatic products released by **(E)** $Sph3_h$ and **(F)** $PelA_h$ from sample presented in D.

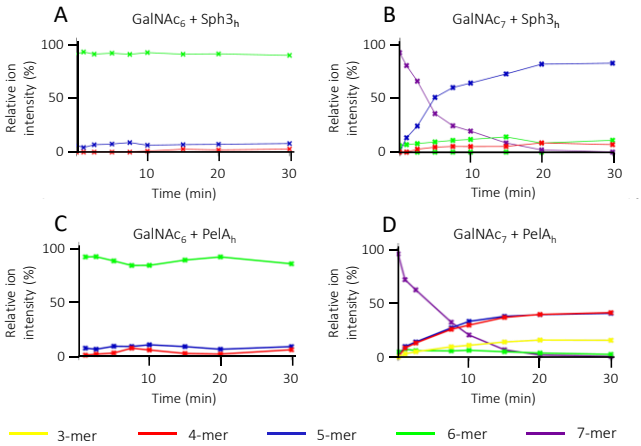

Figure 3. $Sph3_h$ and $PelA_h$ are endo-N-acetylgalactosaminidases with a minimum substrate size of seven residues. Degradation time course of pure oligosaccharides monitored over 30 minutes by MALDI-TOF MS. $Sph3_h$ degradation kinetic of **(A)** α-1,4-GalNAc 6-mers and **(B)** α-1,4-GalNAc 7-mers. $PelA_h$ degradation kinetic of **(C)** α-1,4-GalNAc 6-mers and **(D)** α-1,4-GalNAc 7-mers. Relative proportion of each ions was calculated and reported here under the color purple for 7-mer; green for 6-mer, blue for 5-mer, red for 4-mer and yellow for 3-mer.

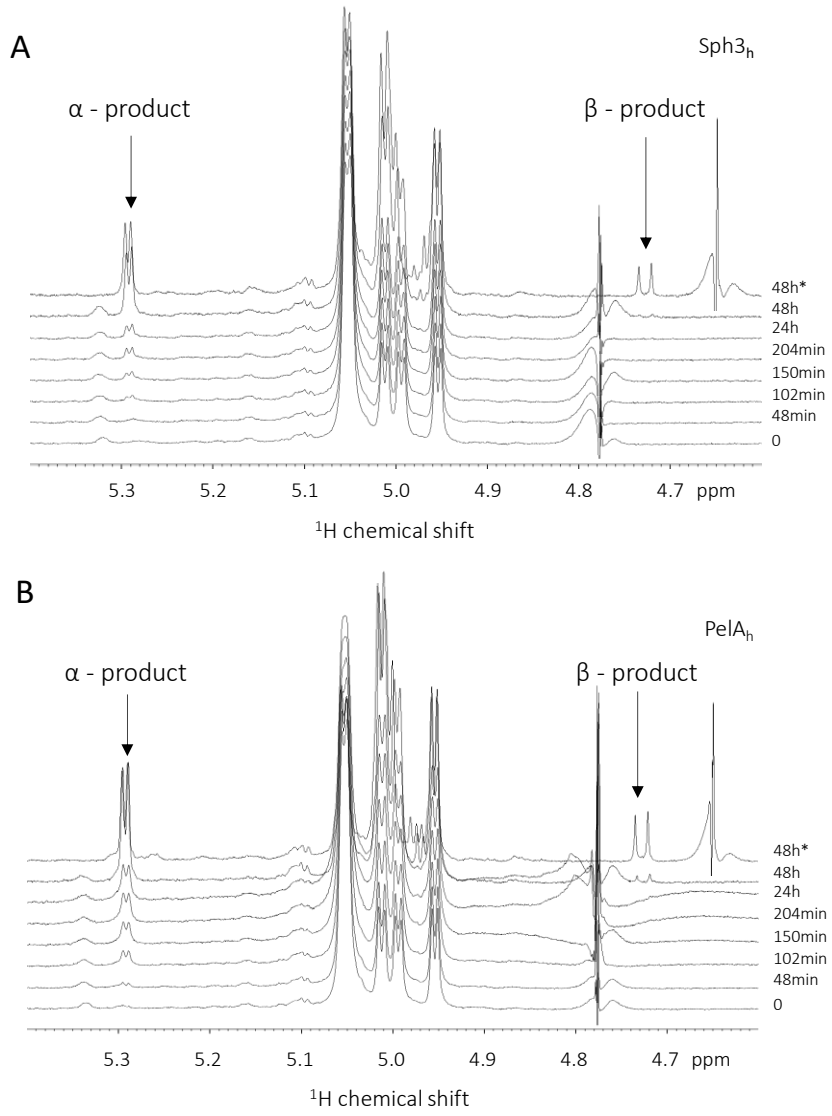

Figure 4. $Sph3_h$ and $PelA_h$ are acting as retaining glycoside hydrolases. Degradation time course of α-1,4-GalNAc octamers by **(A)** $Sph3_h$ and **(B)** $PelA_h$ monitored by NMR spectroscopy highlighting the region between 4.6 to 5.4 ppm. All spectra were recorded at 25°C except the final point at 48h, which was also acquired at 37 °C as indicated by the *.

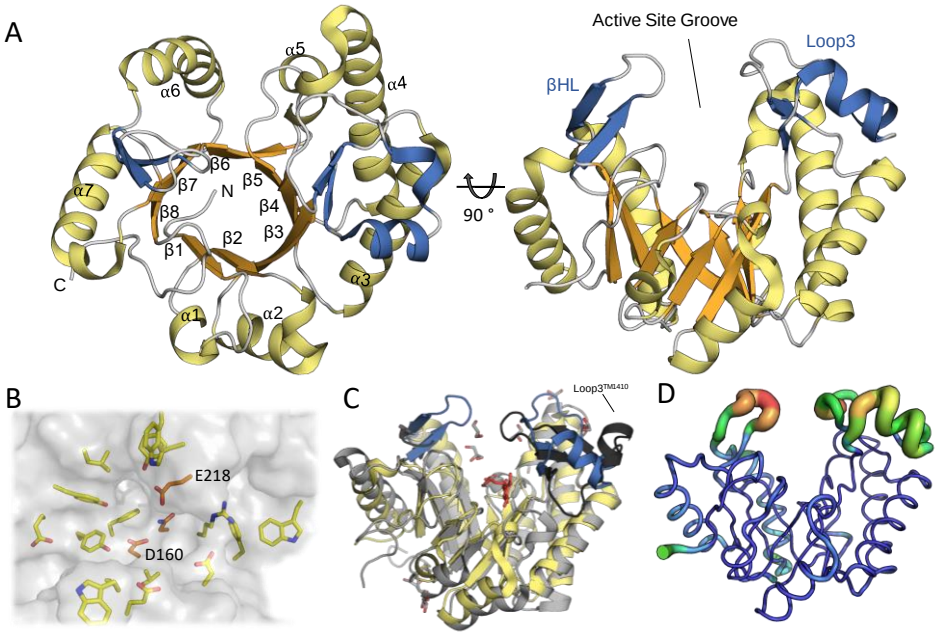

Figure 5. The structure of $PelA_h$ reveals a (β/α)-barrel. (A) Tertiary structure of $PelA_h$ with the (β/α) fold coloured in orange and yellow, respectively. The β-hairpin (βHL) and the insertion after β3 (Loop3) are coloured in blue. **(B)** The active site groove is shown with a transparent surface allowing for visualization of the conserved residues residing in the cleft based on Consurf analysis. In orange are the three residues identified as highly conserved in Sph3 homologues. **(C)** Cα alignment of $PelA_h$ (yellow and blue) to hypothetical protein TM1410 (PDB 2AAM, grey and black) shows similarity in tertiary structure topology. The insertions of $PelA_h$ are coloured as in panel (A) showing that TM1410 also contains these additions including a loop after β3 ($Loop3^{TM1410}$, black). An unknown ligand (red) and glycerol (grey) were found in the groove of the TM1410 structure. (D) Visualization of $PelA_h$ structure B-factors, coloured blue to red for relative low to high values.

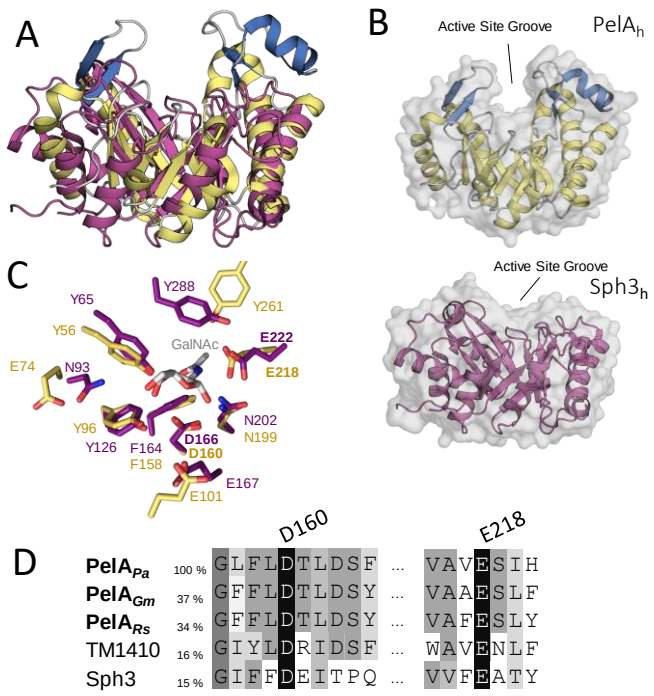

Figure 6. $PelA_h$ and $Sph3_h$ differ in their substrate binding cleft architecture but share catalytic motifs. (A) Tertiary structure alignment of $Sph3_h$ (PDB 5D5G, purple) with $PelA_h$. **(B)** Transparent surface representation of $PelA_h$ (yellow and blue) and $Sph3_h$ (purple) in the same orientation shows the relative depths of the active site groove. **(C)** Alignment of the active site residues of $PelA_h$ (yellow) and $Sph3_h$ (purple) based on the $Sph3_h$ active site motifs shows high identity between the hydrolases around the GalNAc (grey) binding site of $Sph3_h$ (PDB 5D6T). **(D)** Primary sequence alignment of *P. aeruginosa* $PelA_h$ ($PelA_{Pa}$) with homologues from *Geobacter metallireducens* ($PelA_{Gm}$), *Ralstonia solanacearum* ($PelA_{Rs}$) as well as TM1410 and Sph3 (*Aspergillus clavatus*) done by MUSCLE. Sequence identity to $PelA_h$ is listed based on MUSCLE alignment for the two homologues and Sph3. Sequence identity to TM1410 is based on structural alignment.

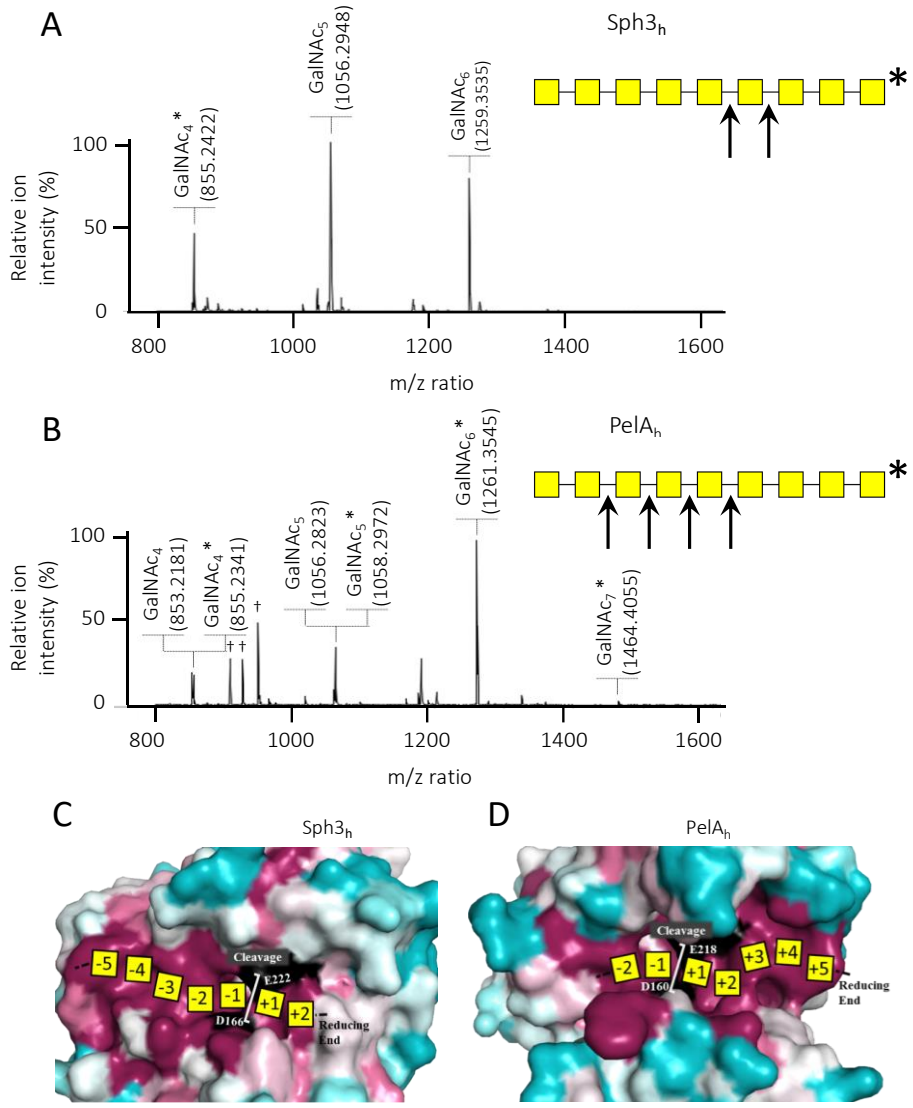

Figure 7. Sph3$_h$ and PelA$_h$ do not hydrolyze the oligosaccharide at the same location. (A) MALDI-TOF MS spectra of the enzymatic product of Sph3$_h$ and **(B)** PelA$_h$ cleavage of reduced α-1,4-GalNAc 9-mers. and their associated schematic view of the cleavage sites. Yellow squares represent GalNAc, arrows indicate enzymatic cleavage sites, * indicates the reducing end of the oligosaccharides. † indicates matrix ion signals**. (C)** Surface representation coloured by conservations (conserved in magenta and variable in teal, catalytic residues in black) showing a proposed map of a heptamer substrate and product subsites on Sph3$_h$ and **(D)** PelA$_h$.

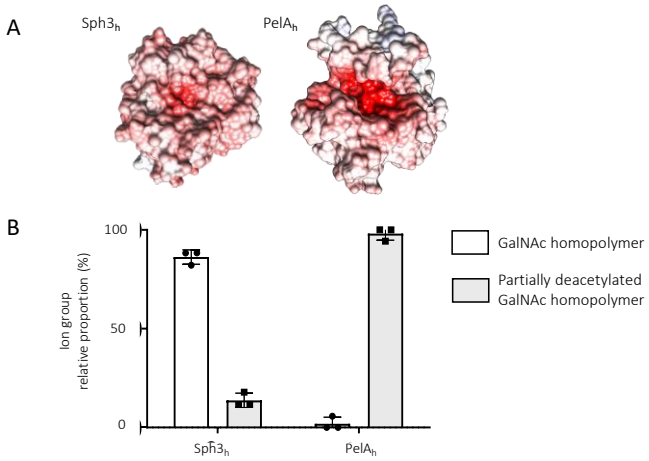

Figure 8. Difference in electrostatic charge surface predict $PelA_h$ is preferentially able to cleave oligosaccharides containing deacetylated GalNAc. (A) Electrostatic surface representation of $Sph3_h$ and $PelA_h$ generated using APBS in Chimera. Quantitative electrostatics are colored from red (-15kT) to blue (+15kT). (B) Relative proportions of oligosaccharides obtained from digestion of purified, secreted GAG degradation products with 1μM of $Sph3_h$ or $PelA_h$. Oligosaccharide products were detected by MALDI-TOF MS. Ions were categorized according to their composition as indicated in the legend. No galactose-homopolymers were detected.

Molecular mechanism of *Aspergillus fumigatus* biofilm disruption by fungal and bacterial glycoside hydrolases

François Le Mauff, Natalie C. Bamford, Noor Alnabelseya, Yongzhen Zhang, Perrin Baker, Howard Robinson, Jeroen D.C. Codée, P. Lynne Howell and Donald C. Sheppard

J. Biol. Chem. published online June 5, 2019 originally published online June 5, 2019
