Article

Elucidating Sequence and Structural Determinants of Carbohydrate Esterases for Complete Deacetylation of Substituted Xylans

Leena Penttinen 1, Vera Kouhi 2, Régis Fauré 3, Tatiana Skarina 4, Peter Stogios 4, Emma Master 4,5 and Edita Jurak 6,*

1 Department of Chemistry, University of Eastern Finland, 80130 Joensuu, Finland; leena.penttinen@uef.fi
2 Department of Pharmacy, University of Helsinki, 00790 Helsinki, Finland; vera.kouhi@helsinki.fi
3 Toulouse Biotechnology Institute, Bio & Chemical Engineering, Université de Toulouse, Centre National de la Recherche Scientifique, Institut National Des Sciences Appliquées, 31077 Toulouse, France; regis.fautre@insa-toulouse.fr
4 Department of Chemical Engineering and Applied Chemistry, University of Toronto, Toronto, ON M5S 3E5, Canada; tanskarina@gmail.com (T.S.); p.stogios@utoronto.ca (P.S.); emma.master@utoronto.ca (E.M.)
5 Department of Bioproducts and Biosystems, Aalto University, 02150 Espoo, Finland
6 Department of Bioproduct Engineering, University of Groningen, 9747 Groningen, The Netherlands
* Correspondence: e.jurak@rug.nl

Abstract: Acetylated glucuronoxylan is one of the most common types of hemicellulose in nature. The structure is formed by a β-(1→4)-linked D-xylopyranosyl (Xylp) backbone that can be substituted with an acetyl group at O-2 and O-3 positions, and α-(1→2)-linked 4-O-methylglucopyranosylnurononic acid (MeGlcA). Acetyl xylan esterases (AcXE) that target mono- or doubly acetylated Xylp are well characterized; however, the previously studied AcXE from Flavobacterium johnsoniae (FjoAcXE) was the first to remove the acetyl group from 2-O-MeGlcpA-3-O-acetyl-substituted Xylp units, yet structural characteristics of these enzymes remain unspecified. Here, six homologs of FjoAcXE were produced and three crystal structures of the enzymes were solved. Two of them are complex structures, one with bound MeGlcA and another with acetate. All homologs were confirmed to release acetate from 2-O-MeGlcPA-3-O-acetyl-substituted xylan, and the crystal structures point to key structural elements that might serve as defining features of this unclassified carbohydrate esterase family. Enzymes comprised two domains: N-terminal CBM domain and a C-terminal SGNH domain. In FjoAcXE and all studied homologs, the sequence motif around the catalytic serine is Gly-Asn-Ser-Ile (GNSI), which differs from other SGNH hydrolases. Binding by the MeGlcA-Xylp ligand is directed by positively charged and highly conserved residues at the interface of the CBM and SGNH domains of the enzyme.

Keywords: acetyl xylan esterase; carbohydrate esterase; xylan; SGNH hydrolase

1. Introduction

Hemicelluloses represent the second most abundant polysaccharide in plant fiber (i.e., lignocellulose), after cellulose. The composition and structure of hemicelluloses vary depending on botanical source. 4-O-Methyl-glucuronoxylans (GXs) comprise approximately 20% of wood fiber from deciduous trees [1,2] and are characterized by a β-(1→4)-linked D-xylopyranosyl (Xylp) backbone that can be acetylated and/or substituted by α-(1→2)-linked 4-O-methyl/D-glucopyranosylnurononic acid (MeGlcA) at a reported frequency of 1 MeGlcA for every 10–15 Xylp, and acetyl units at C-2 and/or C-3 positions [2–4]. The presence of such side groups can impact xylan properties, including water solubility, rheology, and nutrient value, and can hinder the digestibility of xylan and the access of xylan-degrading enzymes [5–9]. Accordingly, enzymes that selectively remove
xylan side groups are important for tailoring xylan structures and increasing the accessibility of main-chain xylanases that convert xylan polymers to oligomers or monomeric xylose. Whereas tailored xylan polymers can be processed into marketable materials, such as packaging films and hydrogels [6,10–12], further enzymatic deconstruction of xylan to xylo-oligosaccharides and xylose is needed for the production of prebiotics, low caloric sweeteners such as xylitol, and fermentation products [13,14].

Enzymes targeting side groups of GX have been classified into multiple carbohydrate-active enzyme families (CAZymes) [15]. These include α-glucuronidases from glycoside hydrolase (GH) families GH67 and GH115, and acetyl xylan esterases (AcXEs) that have been primarily reported in carbohydrate esterase (CE) families CE1, CE4–CE6, and CE16 [15]. Whereas α-glucuronidases from family GH115 hydrolyze the α-(1→2)-linkage between MeGlcA and mono-substituted Xylp at internal and end positions of GX [16–20], GH67 activity is restricted to such linkages at the non-reducing end of xylans [21–23]. CEs acting on GX include those that act on O-2 and O-3 monoacetylated Xylp (CE2, CE3, CE4), and others with preferential activity towards 2,3-di-O-acetylated Xylp (CE1, CE5, CE6) [24]. Certain family CE16 enzymes were previously shown to act on non-reducing MeGlcA-(1→2)-3-O-acetyl-Xylp positions [24,25], possibly as a result of acetyl group migration from O-3 to O-4 on the non-reducing Xylp [26]. In 2018, we identified an unclassified carbohydrate esterase from Flavobacterium johnsoniae (FjoAcXE) with the ability to deacetylate MeGlcA-(1→2)-3-O-acetyl-Xylp subunits at internal positions of the xylan backbone [27]. This finding was subsequently reported for a second acetyl xylan esterase from Flavobacterium johnsoniae that shares over 98% identity with FjoAcXE, but from a different F. johnsoniae strain [28].

Prior to its characterization, FjoAcXE was annotated as a protein of unknown function (PUF) belonging to the polysaccharide utilization locus (PUL) PUL20 of F. johnsoniae UW 101, which comprises CAZymes predicted to act on xylans including sequences belonging to families GH43, GH115, CE1 and CE6 (http://www.cazy.org/PULDB/, accessed on 13 February 2022) [29]. The FjoAcXE sequence was predicted to include a signal sequence for secretion and a SGNH hydrolase-type esterase domain belonging to GDSL-like lipase/acylhydrolase-type domain (Pfam domain PF12472). Structural models of FjoAcXE showed an α/β/α-fold that is characteristic to SGNH hydrolases and an N-terminal CBM-like structure. Different to most similar esterases from families CE2 and CE12, three loops surrounding the predicted catalytic site of FjoAcXE adopted unique conformations predicted to play a role in substrate preference and substrate binding. The found activity completes the deacetylation of xylan and was shown to boost α-glucuronidase activity by offering α-glucuronidase free access to xylan unit [27].

To expand our knowledge of MeGlcA-(1→2)-3-O-acetyl-Xylp acting CE enzymes and elucidate their structural characteristics, we selected six FjoAcXE homologous that shared 47–98% sequence identity and produced them for structural and biochemical characterization.

2. Results and Discussion

2.1. Candidate Selection and Recombinant Protein Production

Homologous sequences of FjoAcXE were retrieved through BLASTp search against GenBank and six sequences comprising Lipase_GDSL_2 (PF13472) and SGNH_hydro domains were selected for further studies (Table S1). Alignment of the six selected sequences with FjoAcXE (Figure 1) revealed that all homologs possess Ser, Gly, Asn, and His (SGNH) residues that are the signature feature of SGNH-type hydrolases, along with the conserved Ser, Asp, and His residues predicted to form the catalytic center. The six selected sequences were functionally expressed in E. coli and purified using affinity chromatography (Figure S1).
2.2. General Biochemical Properties

The recombinantly produced proteins showed no or negligible activity on all tested pNP-glycosides and selected commercially available polysaccharides (Table S2). Instead, clear activity was observed for all targets on pNP-acetate (pNPA) and 4-MUA (Table 1). Given the pH stability of 4-MUA [30], this substrate was used to evaluate the pH optimum of each enzyme. In all cases, the pH optimum was between pH 6.0 and pH 7.5 (Table S3), which is slightly lower than the pH optimum reported for FjoAcXE (pH 8.0, 25). With the exception of AimAcXE, all enzymes retained over 50% activity after 30 min at 55 °C (Table S4) and thus displayed higher temperature stability than FjoAcXE.

Table 1. Percentage of hydrolyzed substrate by the studied acetylxylan esterases in 2 h using 1 mM substrate and 40 °C temperature (n = 3).

| Substrate | pH | FspAcXE | FspF52AcXE | CspAcXE | PbeAcXE | FreAcXE | AimAcXE |
|-----------|----|---------|------------|---------|---------|---------|---------|
| pNPA      | 5.5| 23      | 27         | 52      | 42      | 47      | 24      |
|           | 7  | 47      | 47         | 57      | 55      | 50      | 53      |
| 4-MUA     | 5.5| 35      | 32         | 33      | 31      | 25      | 25      |
|           | 7  | 45      | 42         | 48      | 47      | 40      | 54      |

2.3. Activity on Acetylated Oligosaccharides

Based on the general biochemical properties of the targeted enzymes, activity studies using acetylated carbohydrates were performed at pH 7.0 and 30 °C. The per-acetylated xylooligosaccharides display acetyl substitutions at the reducing anomeric O-1 position, non-reducing O-4 position as well as all O-2 and O-3 positions of the Xylp subunits. Given that the theoretical maximal yield of acetic acid was determined to be 3 µg/mL, approximately 60% of total acetyl groups were released by all selected enzymes after 20 h, including FjoAcXE (Figure 2A). This apparent limit to the extent of deacetylation likely reflects the regioselectivity of the enzymes on the synthetic per-acetylated xylooligosaccharides. Moreover, these results confirmed that, like FjoAcXE, the enzymes reported for the first time herein are carbohydrate esterases.
Figure 2. Acetic acid release from (A) per-acetylated xylooligosaccharide (DP 4–7) using 2 µg of enzyme, and (B) from CF-XOS using 1 µg enzyme. Reactions (60 µL in Subfigure (A) and 50 µL in Subfigure (B)) were performed at 50 mM HEPES and pH 7.0, and were incubated for 20 min, 4 h, and 20 h. Acetic acid release was detected with acetic acid kit (K-ACETRM, Megazyme). Error bars represent standard deviation (n = 3).

All selected enzymes also deacetylated xylooligosaccharides from corn fiber (CF-XOS) fraction (Figure 2B), albeit to a lesser extent than per-acetylated xylooligosaccharides. The theoretical maximal yield of acetic acid was calculated to be 1.2 µg/mL; therefore, less than 20% of total acetyl groups were released by all enzymes, including FjoAcXE after 20 h. Previously, it was shown that FjoAcXE is able to target 2-O-acetyl-Xylp where the same Xylp is substituted at the O-3 position with L-arabinosyl residues [27]. The CF-XOS used is expected to be prevalently feruloylated [31]. Given that FjoAcXE was unable to release acetyl groups that neighbor feruloylated side groups, it can be speculated that the remaining acetyl groups belong to that CF-XOS.

2.4. Complementation of α-Glucuronidase Activity

AcXE activity on MeGlcPα-(1→2)-3-O-acetyl-Xylp was evaluated by testing the synergistic effect of the AcXEs with GH115 α-glucuronidase from Amphibacillus xylanus (AxyAgu115A) on acetylated glucuronoxylan. Approximately ten-times higher MeGlcPα was released by α-glucuronidase after adding AcXE to the reaction (Figure 3). The result clearly shows that the studied AcXEs deacetylate MeGlcPα-(1→2)-3-O-acetyl-Xylp of glucuronoxylan, and possibly to some extent 2-O-acetyl-Xyl from neighboring xylopyranosyl residues. This catalytic specificity is only found in the previously reported FjoAcXE and the six homologs of this study. The apparent uniqueness of this activity motivated structural characterization of the studied CEs to uncover features that promote binding of the MeGlcPα-(1→2)-3-O-acetyl-Xylp substrate.
We attempted to crystallize FjoAcXE and its homologs FspAcXE, FspF52AcXE, CspAcXE, PbeAcXE, FreAcXE and AimAcXE. We successfully solved the structure of the CBM-like domain of FjoAcXE (to resolution 2.45 Å) and multiple structures of the full-length CspAcXE and PbeAcXE enzymes. Based on the protein sequence of FjoAcXE, the CBM-like domains do not clearly group with the existing CBM families. The crystal structure of CspAcXE was solved in the native state (1.30 Å). Three crystal structures of AcXE from Prolixibacter bellariivorans (PbeAcXE) were also solved, including a native form (1.35 Å) and two complex structures, one with MeGlcP-A-Xylp (to 1.26 Å) and another with acetate (1.13 Å). Full x-ray crystallographic statistics can be found in Table S5.

The sequence identity between the CspAcXE and PbeAcXE is 52.4%. Both comprise an N-terminal CBM-like domain consisting of three antiparallel β-sheets, and a C-terminal SGNH domain where the central β-sheet comprises five parallel β-strands flanked on both sides with α-helices (Figure 4). In both cases, the CBM-like and catalytic SGNH domains are tightly packed and connected through a loop of six to seven amino acids. The residues forming the CBM-like domains in PbeAcXE and CspAcXE are 1–175 and 1–176, respectively, whereas residues forming the SGNH domain in PbeAcXE and CspAcXE are 180–385 and 181–386, respectively. The first seven N-terminal residues of PbeAcXE and the first 10 residues of CspAcXE did not have electron density and so were not included in the final structures. Moreover, residues 350 to 365 in CspAcXE, which form a loop next to the active site, had unclear or missing electron density and so were omitted from the solved structure; this could be the result of the absence of ligand to stabilize its conformation, or due to the absence of many crystal packing contacts in this region of the protein.
As indicated by a structural similarity search against PDB using Dali (Figure 5), the structures of \textit{PbeAcXE} and \textit{CspAcXE} are most similar to structures of GDSL hydrolases of unknown specificity and some with esterase/lipase activity, i.e., TesA from \textit{E. coli} and \\textit{Pseudomonas aeruginosa} [32] (Table S6). There was also a clear structural similarity with known acetyl xylan esterases, including \textit{AlAXEase} from \textit{Arcticibacterium luteifluviistationis} (PDB ID 7ddy), Ax2 from \textit{Geobacillus stearothermophilus} (PDB ID 4jko) [33], a CE3 enzyme from \textit{Talaromyces cellulolyticus} (PDB ID 5b5s) [34], and CE2 enzyme from \textit{Cellvibrio japonicus} (PDB ID 2waa) [35] and \textit{Butyrivibrio proteoclasticus} B316 (PDB ID 3u37) [36] (Figure 5).
The CBM-like domains of FjoAcXE, PbeAcXE and CspAcXE adopted a similar structure (Figure 4). The construct of FjoAcXE that successfully crystallized did not include the first 36 amino acids that comprise the CBM; residues 37–40 and residues 137–148 were also disordered in the structure. Due to the incomplete FjoAcXE CBM structure, further analysis will focus on the PbeAcXE and CspAcXE structures, although based on our analysis in subsequent sections, we expect that these regions would participate in the active site, i.e., Lys34 and Arg148 of FjoAcXE. The CBM from these proteins was most similar to a CBM1 domain from GH10 enzyme from Bacteroides intestinalis (PDB 4qpw) [37], CBM35 from a GH31 enzyme from Paenibacillus sp. 598K (PDB 5x7s) [38], and the Ig-like domain from a GH86 enzyme from Bacteroides uniformis (PDB 5ta1) [39] and the β-sandwich domain from a GH87 enzyme from Streptomyces thermodiastaticus (PDB 7c7d) [40] (Dali Z-scores between 6.4 and 6.9, RMSDs 3.3–4.2 Å over 94–102 Cα atoms) (Figure S2). Based on a more detailed comparison with the structure of the CBM1 in complex with xylotriose (Figure S2), the CBM from PbeAcXE and CspAcXE does not appear to have an equivalent ligand binding pocket due to the positioning of the loops 73–82 and 146–160; however, conformational changes in this area in the presence of the appropriate ligand may open a binding pocket.

2.6. Sequence Determinants of the New CE Family

The active site of PbeAcXE and CspAcXE is located at the C-terminal end of the central β-sheet of the SGNH domain, as is common for all α/β/α hydrolases [41]. The Ser-Asp-His catalytic triad is formed by Ser188, Asp363, and His366 in PbeAcXE, and Ser189, Asp363, and His366 in CspAcXE. In both cases, the active sites are on the surface of the enzymes and highly exposed to solvent, where the serine occurs in a loop following the β1-strand, and the aspartate and histidine are positioned on a long loop between the β5-strand and a short α-helix (Figure 4). In PbeAcXE, the oxyanion hole that stabilized the oxyanion intermediate during catalysis is formed by the main chain amides of Ser188 and Gly228, and the amine side group of Asn265 (Figure 6). In CspAcXE, these residues are Ser189, Gly229 and Asn266.

![Figure 6](image-url) The complex crystal structures of acetylxyan esterase from Prolixibacter bellariivorans (PbeAcXE) with methylated D-glucuronic acid linked to D-xylopyranose (MeGlcA-(1→2)-Xylp). (A) The polder omit electron density map for the ligand is shown at 3σ contour level in green. Surface charges are shown in blue (positive charge) and red (negative charge). (B) The amino acid residues at the active site that participate in ligand binding in PbeAcXE are labelled and shown in sticks. The hydrogen bonds between these residues and the ligand are marked as yellow dash lines. (C) Superposed complex structures of PbeAcXE with MeGlcA-(1→2)-Xylp and acetate. Acetate is shown in green. The catalytic Ser188 exists in the same conformation in (B,C).

The SGNH lipase/hydrolase-type proteins have four sequence blocks each containing a conserved catalytic or oxyanion hole residue: Serine in block I, glycine in block II, asparagine in block III and histidine in block V [42]. In SGNH hydrolases, especially lipases, the catalytic serine is often part of a GDSL sequence motif [42,43]. Interestingly, a GDSL motif is not found in AcXEs; however, a shorter GDS motif was found in an AcXE from Geobacillus stearothermophilus [44] and in an AcXE from Caldicellulosiruptor bescii [45]. In FjoAcXE and all studied homologs, the sequence motif around the catalytic serine is Gly-Asn-Ser-Ile (GNSI). The most common consensus sequence in α/β hydrolases is GXSXG, which defines the nucleophilic elbow of an active site [46,47]. Therefore, like the GDSL-hydrolases, the GNSI-AcXEs do not have the nucleophilic elbow; this is corroborated by the
crystal structures of \textit{PbeAcXE} and \textit{CspAcXE}, which show that the Asn in this motif (Asn187 and Asn188) are buried into the interior of the enzymes. Since the GNSI sequence motif has not been reported in any other carbohydrate esterase or lipases, this may be a signature feature of \textit{FjoAcXE}-like AcXEs.

2.7. Five Conserved Loops Shape the Active Site

The substrate-binding sites in \textit{PbeAcXE} and \textit{CspAcXE} are mainly shaped by five conserved loops, from which two belong to the CBM domain and three to the SGNH domain. The loops in \textit{PbeAcXE} are 16–35, 127–142, 185–203, 224–244, and 346–369; in \textit{CspAcXE}, the loops are 19–37, 127–144, 187–204, 227–242, and 349–368 (Figure 7). The 185–203/187–204 loop also contains one short \(\alpha\)-helix. Here, the loops are named loop1 to loop5 by their order of occurrence in the sequence. All five loops include conserved regions in \textit{FjoAcXE} and the studied homologs. Loop1 is long and highly conserved but contributes only one residue (Asn26 of \textit{PbeAcXE} and Asn28 of \textit{CspAcXE}) to the active site. The GNSIXDGRGS region in loop3 includes the GNSI motif sequence stretches that overlap with block I of the SGNH blocks (i.e., residues 186–195 in \textit{PbeAcXE} and 187–196 in \textit{CspAcXE}). Loop4 overlaps with block II of the SGNH blocks and comprises two conserved stretches: GIGG and GGLGP (226–229 and 236–240 in \textit{PbeAcXE}, and 227–230 and 235–239 in \textit{CspAcXE}). Lastly, loop5 comprises the DXLHP motif, which include the catalytic aspartate and histidine in block V of the SGNH blocks (Asp363 in both \textit{PbeAcXE} and \textit{CspAcXE}; this region was unmodelled in \textit{CspAcXE} due to disorder).

![Figure 7. Overlaid crystal structures of acetyl xylan esterases from \textit{Prolixibacter bellariivorans}, (gray) in complex with MeGlc\(\text{pA-(1\rightarrow2)-Xylp}\) ligand (yellow) and \textit{Chryseobacterium} sp. YR480 (light green). The five conserved loops in the AcXEs are shown in yellow, orange, red, brown and black. The catalytic serine is shown in sticks.](image)

2.8. Substrate Binding at the Active Site

Two complex structures of \textit{PbeAcXE} were solved, one with bound MeGlc\(\text{pA-(1\rightarrow2)-Xylp}\) and another with acetate. These two complex structures allowed for the examination of substrate binding in AcXEs that are active towards MeGlc\(\text{pA-(1\rightarrow2)-3-O-acetyl-Xylp}\). The MeGlc\(\text{pA-(1\rightarrow2)-Xylp}\) ligand is bound into a positively charged and highly conserved (Figure S3) environment formed by Arg25, Asn26, Arg139 and Arg193 (Figure 6). Specifically, the carboxylic acid group of MeGlc\(\text{pA}\) is bound to Arg139 in the bidentate way and to the catalytic Ser188; the O-3 and O-5 oxygen of the Xylp ring is also bound to Ser188.

In addition to coordination by Arg139, the hydroxyl groups at the C-2 and C-3 positions of MeGlc\(\text{pA}\) and O-5 of the pyranose ring are bound by Arg25, Asn26 and Arg193,
respectively. At the same time, the Xylp is further coordinated by Ser188, Arg193, and His366. Residues Arg25, Asn26 and Arg139 all belong to the CBM, whereas Ser188, Arg193 and His366 belong to the SGNH domain. This reveals that despite the catalytic site being localized to the SGNH domain, the CBM domain plays a role in binding MeGlcPA-(1→2)-3-O-acetyl-substituted xylan and enabling the unique activity of this type of AcXEs.

Interestingly, in the PbeAcXE·MeGlcPA-(1→2)-Xylp crystal we noticed weak additional density extending beyond the O-1 position of the Xylp unit (Figure S4). We attempted to model two additional Xylp units into this density but given the poor quality of the density they were not included into the final deposited structure.

The complex structure of PbeAcXE with acetate shows that acetate is bound to the oxyanion hole by Ser188, His366, Asn265 and Gly229, and the α-carbon is localized in a hydrophobic pocket formed by Ile189, Val264 and Leu365. Based on a superposition of the PbeAcXE-acetate complex with the PbeAcXE·MeGlcPA-(1→2)-Xylp complex, the location of the acetate corresponds to the O4 position of the bound Xylp; this is consistent with PbeAcXE’s recognition of the non-reducing end, i.e., MeGlcPA-(1→2)-4-O-acetyl-Xylp. The catalytic Ser188 was modeled in two conformations, suggesting flexibility during catalysis. All other residues are conserved over the studied AcXEs except Arg25, which can be either replaced by histidine, lysine or asparagine.

3. Materials and Methods

3.1. Materials

All polysaccharides were purchased from Megazyme (Ireland) with the exception of oat spelt xylan, which was purchased from Sigma-Aldrich (St. Louis, MO, USA) (Table S1). All para-nitrophenol (pNP) substrates and 4-methylumbelliferyl acetate (4-MUA) were purchased from Sigma (St. Louis, MO, USA) (Table S1). The per-acetylated xylo-oligosaccharide mixture comprised chemically acetylated xylooligosaccharides (per-acetylated) with degree of polymerization between 4 and 7 D-xylosyl units [48]. Acetylated glucurono-xylooligosaccharides (Ac-XOS) were isolated by steam extraction of milled chips from Eucalyptus wood [49] and was a kind gift from Prof. Maija Tenkanen (University of Helsinki, Finland). Acetylated and feruloylated AcFaXOS from corn fiber (CF-XOS) was isolated by mild acid hydrolysis and was received as a kind gift from Prof. Mirjam Kabel (University of Wageningen, the Netherlands). The characterization of the CF-XOS (fraction B) was previously described by Appeldoorn et al. [31]. All other chemicals were analytical grade and were ordered from Sigma-Aldrich or Fisher Scientific.

FjoAcXE and AxyAgu115A were prepared as previously described [16,27].

3.2. Candidate Selection

Six homologues that share sequence identity between 47 and 98% with FjoAcXE (accession number ABQ06890.1) were selected. The protein sequences are listed in Table S2, and included selections from Flavobacterium sp. 40S8 (WP_073411463.1), Flavobacterium sp. F52 (WP_008467835.1), Chryseobacterium sp. YR480 (CspAcXE, WP_047423182.1), Prolibacter bellariivorans (PbeAcXE, WP_027586161.1), Flavobacterium reichenbachii (FreAcXE, WP_035683315.1), and Alkaliflexus imshenetskii DSM 15055 (AimAcXE, WP_044117739.1).

3.3. Protein Expression and Purification

For biochemical studies, selected genes were synthesized and cloned into pET-29b + plasmid (C-terminal His-Tag) by Genscript (Piscataway, NJ, USA) and then transformed into E. coli BL21(DE3) for protein expression. Transformants were propagated at 30 °C in 2 L Erlenmeyer flasks containing 500 mL of Luria-Bertani (LB) Broth-Miller supplemented with 50 µg/mL kanamycin, 0.5 M sorbitol and 0.025 M glycine betaine. When cultures reached an OD600 of ~ 0.6, recombinant protein expression was induced overnight at 16 °C using 0.5 mM isopropyl β-D-1-thiogalactopyranoside. Cells were harvested by centrifugation at 3000 × g for 30 min at 4 °C. The pellet was then suspended in lysis buffer (20 mM HEPES pH 7.4, 500 mM NaCl, Pierce™ Protease Inhibitor Tablets according to the manufacturer’s
instructions (1 tablet/10 mL; Thermo Fisher Scientific Inc. Mississauga, ON, Canada) and the cells were disrupted using an Emulsiflex homogenizer (Avestin Inc, Ottawa, ON, Canada) for 10 min in pulsation mode with pressure set between 10,000 and 15,000 psi. Cell debris was removed by centrifugation at 19,000 × g for 45 min at 4 °C.

The recombinant protein was further purified from clarified lysates using NiNTA resin. Briefly, cell lysates were passed through NiNTA resin using either a Preppy™ 12-port vacuum system, gravity column, or FPLC system (ÄKTA purifier, Cytiva Life Sciences, Vancouver, BC, Canada) to permit gradient elution. In all cases, the binding buffer was 20 mM HEPES (pH 7.4) and 500 mM NaCl, the elution buffer was the binding buffer plus 500 mM imidazole, and elution was achieved using 5 to 100% elution buffer. The resulting fractions were analyzed by 12% SDS-PAGE; selected fractions were pooled and then exchanged to 20 mM HEPES pH 7.4 using 10kDa Vivaspin 20 column (Vivascience GmbH., Hannover, Germany). Protein concentrations were measured using the Pierce™ BCA Protein Assay Kit (Thermo Fisher Scientific Inc., Mississauga, ON, Canada), and purities were evaluated by SDS-PAGE.

For crystallography, the genes for FspAcXE, CspAcXE, PbeAcXE, FreAcXE, AimAcXE, FjoAcXE and a construct coding for the CBM domain of FjoAcXE (residues 37–186) plus two mutations to introduce additional methionine residues (Leu62Met and Val113Met) were subcloned into pMCSG53 (coding for N-terminal TEV protease site and His-tag) and expressed in Magic competent E. coli BL21(DE3). Cells were grown in ZYP-5052 auto-inducing complex medium [50] by incubating for a few hours at 37 °C followed by transferring to 20 °C for overnight growth, and proteins were purified using Ni-NTA affinity chromatography. Expression of selenomethionine-derivatized proteins was carried out in M9 minimal media following the manufacturer’s instructions (Shanghai Medicilon, Shanghai, China). Expression was induced with isopropyl β-D-1-thiogalactopyranoside (IPTG) at 20 °C when the optical density at 600 nm (OD600) reached 1.2. The overnight cell culture was then collected by centrifugation at 6000 × g. Cells were resuspended in binding buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 5 mM imidazole, and 5% (v/v) glycerol) and lysed with a sonicator, and cell debris was removed by centrifugation at 20,000 × g. The clarified cell lysate was loaded on a 4 mL Ni-NTA column (Qiagen GmbH, Hilden, Germany) pre-equilibrated with 20 mL of binding buffer. The resin was washed with wash buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 55 mM imidazole, and 5% glycerol (v/v)), and the proteins were eluted with elution buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 250 mM imidazole, and 5% (v/v) glycerol). The His-tagged proteins were then subjected to overnight TEV cleavage using 50 µg of TEV protease per mg of His-tagged protein and dialyzed overnight against the pre crystallization buffer (10 mM HEPES pH 7.5, 300 mM KCl). The His-tag and TEV were removed by running the protein over the Ni-NTA column. The tag-free proteins were concentrated using a BioMax concentrator (MilliporeSigma, Toronto, ON, Canada) and passed through a 0.2 µm Ultrafree-MC centrifugal filter (MilliporeSigma, Toronto, ON, Canada). The purity of the protein was analyzed by SDS-polyacrylamide gel electrophoresis (Figure S1).

3.4. Enzyme Activity Screens

Activity screens using 1 mM pNP-linked substrates (pNP-α-L-arabinofuranoside, pNP-β-D-mannopyranoside, pNP-β-D-xylopyranoside, pNP-α-D-galactopyranoside, pNP-α-L-fucopyranoside, pNP-β-D-galactopyranoside, pNP-α-D-glucopyranoside, pNP-acetate, pNP-β-D-glucopyranoside, pNP-β-D-cellobioside) and 1 mM 4-MUA were performed at 40 °C in 96-well plates at pH 5.5 (50 mM sodium acetate) and pH 7.0 (50 mM HEPES). The absorbance of reactions containing pNP-linked substrates was followed continuously at 410 nm for up to 24 h. After 2 h, reactions containing 1 mM 4-MUA were centrifuged at 3900 rpm and supernatants were transferred to a new 96-well plate prior to measuring absorbance at 354 nm. The corresponding standard curve was prepared using eight points between 0.5 and 2.5 mM of 4-MU.
Glycoside hydrolase activity was evaluated using 0.6 mg/mL (total reaction volume of 200 µL) of 16 different polysaccharides (Table S2), where reactions were performed for up to 24 h at 40 °C in 96-well plates and three pH values: pH 5.5 (50 mM sodium acetate), pH 7.0 (50 mM HEPES), and pH 8.5 (50 mM HEPES). Following incubation, glycoside hydrolase activity was measured using the PAHBAH assay [51]. Briefly, 10 µL of each reaction were transferred to 200 µL of PAHBAH reagent (4-hydrobenzoic acid hydrazide) and samples were incubated at 40 °C for 30 min with shaking prior to measuring absorbance at 405 nm.

For all screens, reactions were performed in triplicate and contained 10% w/w of enzyme to dry weight of substrate.

3.5. Optimum Reaction Conditions

The pH optimum of each purified protein was tested using 50 mM Universal buffer (67 mM NaCH$_3$COO, 67 mM H$_3$BO$_3$, 67 mM H$_3$PO$_4$, adjusted to pH values between 3.5 and 9.0 using NaOH) and 0.5 mM 4-MUA. Reactions (200 µL) were initiated by adding 0.02 µg of enzyme (0.03 µg for FspAcXE) and continued for 30 min at 40 °C. 4-MUA was selected to determine the pH optimum of each enzyme given the relative stability of this substrate under alkaline conditions [29]. For all reactions using 4-MUA, absorbance was measured at 354 nm and the reaction mixture without enzyme was used as a blank.

The temperature stability of each enzyme was tested against 0.5 mM 4-MUA in 50 mM Universal buffer adjusted to the respected optimum pH. The reaction was started by addition of 0.02 µg of protein (0.03 µg for FspAcXE), and each suspension was incubated for up to 24 h at 20, 40, 55 and 70 °C. The total reaction volume was 200 µL.

3.6. Activity on Acetylated Xylooligosaccharides

The activities of studied AcXEs on acetylated xylooligosaccharides were tested using highly branched acetylated and feruloylated CF-XOS (described in Appeldoorn et al., 2013 [30]) and a per-acetylated xylo-oligosaccharide mixture (DP 4–7). Acetic acid release was determined using the acetic acid kit (K-ACETRM, Megazyme). Reactions (50 µL final volume) containing CF-XOS comprised 1 µg enzyme, and 50 mM HEPES buffer (pH 7.0) and 0.5 µg of CF-XOS, and were incubated at 30 °C for 20 min and 4 h. Reactions (60 µL final volume) containing the per-acetylated xylo-oligosaccharide mixture comprised 2 µg enzyme, 50 mM HEPES buffer (pH 7.0) and 0.5 µg of per-acetylated xylo-oligosaccharide mixture, and were incubated at 30 °C for 20 min, 4 h and 20 h. The maximal acetic acid release from each substrate was determined using a modified method described by Teleman et al., 2002 [4]. Briefly, 1% (w/v) substrate was suspended in 50 µL 0.5 M NaOH for 30 min at 70 °C and 450 rpm. The reaction solutions were then neutralized, and the release of acetic acid was measured with acetic acid kit. Control reactions without enzyme were run in parallel and the amount of non-enzymatically released acetic acid was subtracted from the values of enzyme–substrate reaction mixtures.

3.7. Complementation Assay Using AcXE and α-Glucuronidase

Activity levels of the studied AcXE and AxyAgu115A (GH115) on acetylated glucuronoxylan were tested alone and in pairs. In total, 1% (w/v) substrate and 0.5 µg of enzyme were incubated for 20 min in 50 mM HEPES pH 7.0 at 30 °C. The total reaction volume was 25 µL. Acetic acid release was detected using Acetic acid kit and release of glucuronic acid using D-glucuronic/D-galacturonic acid kit (Megazyme).

3.8. Protein Crystallization, Data Collection, and Structure Determination

All crystals were grown at room temperature using the vapor diffusion sitting drop method. For the selenomethionine-derivatized native CspAcXE crystal, 20 mg/mL protein was mixed with 20 mM acetylated xylooligosaccharides, and then with reservoir solution 30% (w/v) Jeffamine ED-2001 pH 7.0. For the selenomethionine-derivatized PbeAcXE:McGlcPA-Xylp complex crystal, 20 mg/mL protein was mixed with 20 mM acetylated xylooligosaccharides, and then with reservoir solution 0.1 M Tris pH 6.5 and 20% (w/v) PEG 3350; for
the selenomethionine-derivatized PbeAcXE-acetate complex crystal, 20 mg/mL protein was mixed with 70 mM xylose, then with reservoir solution 0.1 M Tris pH 8.5, 0.2 M sodium acetate, and 30% (w/v) PEG4K. For the native PbeAcXE crystal, 20 mg/mL protein was mixed with 20 mM xylolbiose, and then reservoir solution 0.2 M magnesium chloride and 25% (w/v) PEG3350. For the selenomethionine-derivatized FjoAcXE CBM Leu62Met + Val113Met mutant, 15 mg/mL protein was mixed with reservoir solution 1 M ammonium sulfate, 0.6 M sodium citrate and 0.1 M calcium chloride. All crystals were cryoprotected with paratone oil.

Diffraction data at 100 K were collected at beamline 21-ID of the Life Sciences Collaborative Access Team at the Advanced Photon Source, Argonne National Laboratory. Diffraction data were processed using HKL3000 [52]. Structures were solved by single anomalous dispersion (SAD) phasing using Phenix.AutoSol [53]. All model building and refinement were performed using Phenix.refine and Coot [54] with full B-factor anisotropy for all structures but the FjoAcXE CBM, which was refined with TLS parameterization. Atomic coordinates have been deposited in the Protein Data Bank with accession codes 7TOH, 7TOI, 7TOJ and 7TOK.

4. Conclusions

Six FjoAcXE-like enzymes were biochemically characterized, and solved structures were obtained for two of these (PbeAcXE and CspAcXE), revealing structure-functional relationships for AcXEs with potential to fully deacetylate glucuronoxylans. FjoAcXE-like AcXEs comprise a novel CBM-like domain and an SGNH domain. CBM-like domains share structural similarity to several types of β-structures, including CBM1 and CBM35. However, the CBM-like domains of FjoAcXE-like AcXEs do not clearly group with existing CBM families. The complex structure of PbeAcXE with bound MeGlcA-Xylp reveals interactions between the enzyme and xylan side groups and the elements that line the substrate specificity. The substrate is bound at the positively charged interface between the SGNH and CBM-like domains consisting of conserved amino acid residues. The MeGlcA is bound by Arg139 in a bidentate way, and we suggest that fully conserved Arg139 is crucial for the substrate specificity of FjoAcXE-like enzymes. Other amino acid residues that bind to hydroxyl groups of MeGlcA are Arg25 and Asn26. The Xylp unit is bound to Ser188, Arg193 and His366. Two distinguishing features of the active site include the catalytic serine in FjoAcXE and its homologs that occur in a novel GNSI sequence motif, and the absence of the GXSXG consensus sequence of α/β hydrolases. Together, the biochemical and structural characterization of FjoAcXE and its homologs open the way to establishing a new carbohydrate esterase family.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/molecules27092655/s1. Figure S1: Studied acetyl xylan esterases (AcXE) on SDS-PAGE. Sample 1: FspAcXE, 2: FspF52AcXE, 3: CspAcXE, 4: PbeAcXE, 5: FreAcXE, 6: AimAcXE; Figure S2: A crystal structure of the CBM domain of PbeAcXE superimposed with a CBM1 domain from GH10 enzyme from Bacteroides intestinalis (PDB ID 4qpw), CBM35 from a GH31 enzyme from Paenibacillus sp. 598K (PDB ID 5x7s), and the Ig-like domain from a GH86 enzyme from Bacteroidales uniformis (PDB ID 5ta1) and the β-sandwich domain from a GH87 enzyme from Streptomyces thermoductiasticus (PDB ID 7c7d); Figure S3: Sequence alignment of the studied acetyl xylan esterases (AcXE) and AcXE from Flavobacterium johnsoniae UW101 (FjoAcXE); Figure S4: Surface representation of a complex structure of Prolixibacter bellariivorans acetyl xylan esterase with U4m2XX, MeGlcA-(1→2)-Xylp(1→4)-Xylp-(1→4)-Xylp. (A) The conserved residues near the active site are shown in red. (B) The catalytic SGNH domain is shown in gray and the CBM domain in pale blue. (C) Positive Fo-Fc electron density in the complex structure after removing the ligand is shown at 2σ contour level in green. (D) The same positive Fo-Fc electron density is shown at 1σ contour level; Table S1: The six selected FjoAcXE-homologues; Table S2: Substrates used in this study; Table S3: pH optimum of FjoAcXE homologs; Table S4: Temperature optimum of FjoAcXE homologs at 20, 40, 55 and 70 °C; Table S5: X-ray crystallographic statistics; Table S6: Structural orthologs of PheAcXE and CspAcXE from Dali search (full-length structures).
References

1. Ebringerová, A.; Kramár, A.; Rendoš, F.; Domanský, R. Die Stufenextraktion Der Hemicellulosen Aus Dem Holz Der Hagebuche (Carpinus betulus L.). Holzforschung 1967, 21, 74–77. [CrossRef]
2. Biely, P.; Singh, S.; Puchart, V. Towards Enzymatic Breakdown of Complex Plant Xylan Structures: State of the Art. Biotechnol. Adv. 2016, 34, 1260–1274. [CrossRef] [PubMed]
3. Teleman, A.; Lundqvist, J.; Tjerneld, F.; Stålbrand, H.; Dahlman, O. Characterization of Acetylated 4-O-Methylglucuronoxylan Isolated from Aspen Employing 1H and 13C NMR Spectroscopy. Carbohydr. Res. 2000, 329, 807–815. [CrossRef]
4. Teleman, A.; Tenkanen, M.; Jacobs, A.; Dahlman, O. Characterization of O-Acetyl-(4-O-Methylglucuronoxylan Isolated from Birch and Beech. Carbohydr. Res. 2002, 337, 373–377. [CrossRef]
5. Ebringerová, A.; Heinzle, T. Xylan and Xylan Derivatives—Biopolymers with Valuable Properties, 1. Naturally Occurring Xylans Structures, Isolation Procedures and Properties. Macromol. Rapid Commun. 2000, 21, 542–556. [CrossRef]
6. Ebringerová, A. Structural Diversity and Application Potential of Hemicelluloses. Macromol. Symp. 2005, 232, 1–12. [CrossRef]
7. Shrestha, U.R.; Smith, S.; Pingali, S.V.; Yang, H.; Zahran, M.; Breunig, L.; Wilson, L.A.; Kowali, M.; Kubicki, J.D.; et al. Arabinose Substitution Effect on Xylan Rigidity and Self-Aggregation. Cellulose 2016, 23, 2267–2278. [CrossRef]
8. Xiang, Z.; Jin, X.; Huang, C.; Li, L.; Wu, W.; Qi, H.; Nishiyama, Y. Water Cast Film Formability of Sugarcane Bagasse Xylans Favored by Side Groups. Cel lulose 2020, 27, 7307–7320. [CrossRef]
9. Tiwari, U.P.; Fleming, S.A.; Rasheed, M.S.A.; Jha, R.; Dilger, R.N. The Role of Oligosaccharides and Polysaccharides of Xylan and Mannan in Gut Health of Monogastric Animals. J. Nutr. Sci. 2020, 9, e21. [CrossRef]
10. Mikkonen, K.S.; Tenkanen, M. Sustainable Food-Packaging Materials Based on Future Biorefinery Products: Xylans and Mannans in Gut Health of Monogastric Animals. Trends Food Sci. Technol. 2012, 28, 90–102. [CrossRef]
11. Chen, G.-G.; Qi, X.-M.; Guan, Y.; Peng, F.; Yao, C.-L.; Sun, R.-C. High Strength Hemicellulose-Based Nanocomposite Film for Food Packaging Applications. ACS Sustain. Chem. Eng. 2016, 4, 1985–1993. [CrossRef]
12. Farhat, W.; Venditti, R.A.; Hubbe, M.; Taha, M.; Becquart, F.; Ayoub, A. A Review of Water-Resistant Hemicellulose-Based Materials: Processing and Applications. ChemSusChem 2017, 10, 305–323. [CrossRef] [PubMed]
13. Deutschmann, R.; Dekker, R.E.H. From Plant Biomass to Bio-Based Chemicals: Latest Developments in Xylan Research. Biotechnol. Adv. 2012, 30, 1627–1640. [CrossRef] [PubMed]
14. Vuong, T.V.; Master, E.R. Enzymatic Upgrading of Heteroxylans for Added-Value Chemicals and Polymers. Curr. Opin. Biotechnol. 2022, 73, 51–60. [CrossRef]
15. Lombard, V.; Golacosta Ramulu, H.; Drula, E.; Coutinho, P.M.; Henrissat, B. The Carbohydrate-Active Enzymes Database (CAzy) in 2013. Nucleic Acids Res. 2014, 42, D490–D495. [CrossRef]
16. Yan, R.; Wang, W.; Vuong, T.V.; Xi, Y.; Skarina, T.; Di Leo, R.; Gatenholm, P.; Toriz, G.; Tenkanen, M.; Stogios, P.J.; et al. Structural Characterization of the Family GH115 α-Glucuronidase from Amphibacillus Xylanus Yields Insight into Its Coordinated Action with α-Arabinofuranosidases. New Biotechnol. 2021, 62, 49–56. [CrossRef]
17. Yan, R.; Vuong, T.V.; Wang, W.; Master, E.R. Action of a GH115 α-Glucuronidase from Amphibacillus Xylanus at Alkaline Condition Promotes Release of 4-O-Methylglucopyranosyluronic Acid from Glucuronoxylan and Arabinoglucuronoxylan. Enzym. Microb. Technol. 2017, 104, 22–28. [CrossRef]
42. Upton, C.; Buckley, J.T. A New Family of Lipolytic Enzymes? Trends Biochem. Sci. 1995, 20, 178–179. [CrossRef]
43. Akoh, C.C.; Lee, G.-C.; Liaw, Y.-C.; Huang, T.-H.; Shaw, J.-F. GDSL Family of Serine Esterases/Lipases. Prog. Lipid Res. 2004, 43, 534–552. [CrossRef] [PubMed]
44. Alalouf, O.; Balazs, Y.; Volkinshtein, M.; Grimpel, Y.; Shoham, G.; Shoham, Y. A New Family of Carbohydrate Esterases Is Represented by a GDSL Hydrolase/Acetylxylan Esterase from Geobacillus Stearothermophilus. J. Biol. Chem. 2011, 286, 41993–42001. [CrossRef]
45. Soni, S.; Sathe, S.S.; Odaneth, A.A.; Lali, A.M.; Chandrayan, S.K. SGNH Hydrolase-Type Esterase Domain Containing Cbes-AcXE2: A Novel and Thermostable Acetyl Xylan Esterase from Caldicellulosiruptor Bescii. Extremophiles 2017, 21, 687–697. [CrossRef]
46. Polgár, L. The Catalytic Triad of Serine Peptidases. Cell. Mol. Life Sci. 2005, 62, 2161–2172. [CrossRef]
47. Brumlik, M.J.; Buckley, J.T. Identification of the Catalytic Triad of the Lipase/Acyltransferase from Aeromonas Hydrophila. J. Bacteriol. 1996, 178, 2060–2064. [CrossRef]
48. Hameleers, L.; Penttinen, L.; Ikonen, M.; Jaillot, L.; Fauré, R.; Terrapon, N.; Deuss, P.J.; Hakulinen, N.; Master, E.R.; Jurak, E. Polysaccharide Utilization Loci-Driven Enzyme Discovery Reveals BD-FAE: A Bifunctional Feruloyl and Acetyl Xylan Esterase Active on Complex Natural Xylans. Biotechnol. Biofuels 2021, 14, 127. [CrossRef]
49. Gullón, P.; González-Muñoz, M.J.; Domínguez, H.; Parajó, J.C. Membrane Processing of Liquors from Eucalyptus Globulus Autohydrolysis. J. Food Eng. 2008, 87, 257–265. [CrossRef]
50. Studier, F.W. Protein Production by Auto-Induction in High-Density Shaking Cultures. Protein Expr. Purif. 2005, 41, 207–224. [CrossRef]
51. Lever, M. Carbohydrate Determination with 4-Hydroxybenzoic Acid Hydrazide (PAHBAH): Effect of Bismuth on the Reaction. Anal. Biochem. 1977, 81, 21–27. [CrossRef]
52. Minor, W.; Cymborowski, M.; Otwinowski, Z.; Chruschcz, M. HKL -3000: The Integration of Data Reduction and Structure Solution—From Diffraction Images to an Initial Model in Minutes. Acta Crystallogr. Sect. D Biol. Crystallogr. 2006, 62, 859–866. [CrossRef] [PubMed]
53. Liebschner, D.; Afonine, P.V.; Baker, M.L.; Bunkóczi, G.; Chen, V.B.; Croll, T.I.; Hintze, B.; Hung, L.-W.; Jain, S.; McCoy, A.J.; et al. Macromolecular Structure Determination Using X-Rays, Neutrons and Electrons: Recent Developments in Phenix. Acta Crystallogr. Sect. D Struct. Biol. 2019, 75, 861–877. [CrossRef]
54. Casañal, A.; Lohkamp, B.; Emsley, P. Current Developments in Coot for Macromolecular Model Building of Electron Cryo-microscopy and Crystallographic Data. Protein Sci. 2020, 29, 1055–1064. [CrossRef] [PubMed]