STRUCTURAL BIOLOGY

ISSN 2059-7983

Protein structure of a glycoside hydrolase family 30, subfamily 12 endo-1,4-β-xylanase

Franz J. St John,[a]* Casey Crooks,[a] Michael Endres,[b] Lily Pakdaman,[b] Lisa Koch,[a] Loreen Bynum,[a] Nathaniel Kuch,[a] Andrzej Joachimiak[b,c] and Kemin Tan[b,c]*

[a]Institute for Microbial and Biochemical Technology, Forest Products Laboratory, USDA Forest Service, Madison, WI 53726, USA, [b]eBERlight and Structural Biology Center, X-ray Science Division, Advanced Photon Source, Argonne National Laboratory, Lemont, Illinois, USA, and [c]Consortium for Advanced Science and Engineering, University of Chicago, Chicago, Illinois, USA. *Correspondence e-mail: fjstjohn@gmail.com, franz.j.st.john@usda.gov, ktan@anl.gov

Received 9 December 2025
Accepted 26 February 2026

Edited by R. McKenna, University of Florida, USA

Keywords: glycoside hydrolase family 30; GH30; GH30_12; endo-1,4-β-xylanases; *Anaerobacterium chartisolvens*.

PDB reference: *Ach*Xyn30A, 9ywn

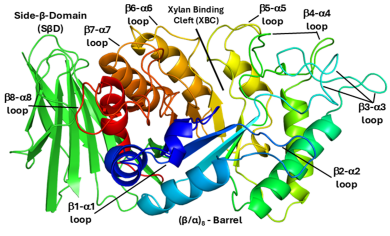
Supporting information: this article has supporting information at journals.iucr.org/d

We have determined the X-ray crystallographic protein structure of endo-1,4-β-xylanase (EX) A from *Anaerobacterium chartisolvens* (*Ach*Xyn30A), a homologue of the recent biochemically characterized glycoside hydrolase family 30, subfamily 12 (GH30_12) EX from *Acetivibrio clariflavus* (*Ac*Xyn30B). The N-terminal GH30 catalytic domains (CDs) of these two enzymes share approximately 63% amino-acid sequence identity and the full-length proteins each consist of the GH30_12 CD, a family 6 carbohydrate-binding module and a C-terminal dockerin domain. In this report, we offer additional support for the recent subfamily classification of these EXs and provide detailed X-ray crystallographic protein structure analysis of *Ach*Xyn30A, the first protein structure from this newly defined GH30 subfamily. We also provide comparative structural analysis using a generated *Ac*Xyn30B homology model as well as other GH30 subfamily enzymes. Additionally, we examine potential xylan-chain interactions informed by the protein structure. These characterized EXs further illustrate the diversity of xylan-degrading enzymes which have evolved within glycoside hydrolase family 30.

1. Introduction

Xylan is the most abundant hemicellulose found in land plants. This β-1,4-linked repeating xylose chain can be variably substituted with a variety of sugars and acetyl groups. The sugars primarily include α-1,2-linked 4-*O*-methylglucuronic acid (GlcA) and α-1,2- and α-1,3-linked arabinofuranose (Ara*f*; Timell, 1967). The variable occurrence, frequency and pattern of these substitutions along the xylan chain explain the large variety of xylan types which exist in nature. Xylan is variable not just between plant types but also between different tissues in the same plant (Kaprelyants *et al.*, 2019; Peña *et al.*, 2016; Jacobs *et al.*, 2001).

Lignocellulosic biomass is valued as a renewable resource. To make the most of this and of its natural abundance, utilization methods must be developed to achieve maximum efficiency towards defined products. Depending on the source, xylan constitutes between 8% and 33% of the dry weight of woody biomass. Softwoods, where the primary hemicellulose is a galactoglucomannan, contain lower amounts of xylan compared with the higher proportions found in hardwoods (Fengel & Wegener, 1979). During pretreatments of woody biomass, the hemicellulose component, of which xylan is a significant fraction, is readily hydrolyzed to monosaccharides, which can then undergo subsequent bioconversion into fuels and chemicals. Other applications for xylan include its use in material science (Nechita *et al.*, 2021; Mikkonen & Tenkanen, 2012), pharmaceuticals (Anderson & Perry, 2006; Petzold-



Welcke *et al.*, 2014; Ohbuchi *et al.*, 2010), nutraceuticals (Abik *et al.*, 2023; Granato *et al.*, 2022) and food products (Chakraborty & Bhowal, 2025). These latter applications will necessitate less severe pretreatment methods to obtain polymeric or large oligomeric xylan materials. Xylan-extraction processes and secondary xylan manipulation will require the use of xylan-active enzymes with well defined function.

The enzymes which degrade xylans either act to remove xylan main-chain appendages (*i.e.* GlcA or Ara*f*) and are commonly referred to as xylan-debranching enzymes, or directly cleave the xylan main chain. These latter, true xylanases include endo-functioning xylanases (endo-1,4-β-xylanases; EXs) and exo-functioning (terminal cleaving) xylosidases. EXs cleave the chain internally, away from the ends of the xylan chain, and xylosidases hydrolyze a single xylose from the nonreducing terminus of the xylan chain (St John *et al.*, 2024). While enzymes with xylosidase activity have been identified in numerous glycoside hydrolase (GH) enzyme families (Li *et al.*, 2023), the most common and best-studied EXs primarily derive from either the GH10 or GH11 enzyme families (Collins *et al.*, 2005; Chow *et al.*, 2022). These two families of EXs do not contain enzymes with other activities. More recently, the GH30 enzyme family was analyzed and divided into subfamilies on the basis of primary amino-acid sequence and enzyme-function determination. In this glycoside hydrolase family reassessment, a group of GH5 enzymes were reclassified as subfamilies in GH30 (St John *et al.*, 2010). GH30 enzymes within subfamily 8 (GH30_8) were shown to be functionally specific GlcA appendage-dependent EXs (glucuronoxylanases; GXs), which cleave the xylan main chain following recognition and coordination of the nearly ubiquitous GlcA xylan appendage (St John *et al.*, 2006, 2011; Vršanská *et al.*, 2007; Urbániková *et al.*, 2011; Freire *et al.*, 2016). Subsequent to this, several enzymes from GH30 subfamily 7 (GH30_7) were characterized and revealed a diversity of related xylan-hydrolysis functionalities like GH30_8 GXs, but also with significant additional activities such as EX activity (GX:EX; Biely *et al.*, 2014; Nakamichi, Fujii *et al.*, 2019, 2020) and xylobiohydrolase (XBH) activity (GX:XBH; Katsimpouras *et al.*, 2019; Nakamichi, Watanabe *et al.*, 2020; Nakamichi, Fouquet, Ito, Watanabe *et al.*, 2019; Šuchová *et al.*, 2020). Other activities include a primary XBH with secondary EX (XBH:EX) activity (Šuchová *et al.*, 2020) and reducing-end xylan exo-xylosidase (REX) activity with a secondary EX- or GX-like function (REX:EX) (Nakamichi *et al.*, 2023; Nakamichi, Fouquet, Ito, Matsushika *et al.*, 2019; Tenkanen *et al.*, 2013). Growing evidence indicates that GH30 enzymes have diverse functions in the enzymatic degradation of xylans (Puchart *et al.*, 2021).

When growing on cellulosic substrates, *Acetivibrio clariflavus* (basionyms *Clostridium clariflavum* and *Hungateiclostridium clariflavum*) expresses two GH30 enzymes. Both are among the top five most abundant enzymes in the *A. clariflavus* cellulosome (Artzi *et al.*, 2015). These enzymes have recently been assigned to newly established GH30 subfamilies. One of these enzymes, *Ac*Xyn30A, was characterized as being a xylobiohydrolase (XBH) and was assigned to GH30, subfamily 10 (GH30_10; Crooks *et al.*, 2021; Šuchová *et al.*, 2021; St John *et al.*, 2022); this functionally specific XBH is in contrast to the related GH30_7 dual-function glucuronoxylanase GX:XBH or XBH:EX enzymes referred to above. The second enzyme, *Ac*Xyn30B, was recently biochemically characterized and assigned to the newly established GH30_12 subfamily (Šuchová *et al.*, 2024). *Ac*Xyn30B acts as a non-specific endo-functioning xylanase, cleaving both highly substituted arabinoxylans and lesser substituted glucuronoxylans at similar rates.

Our laboratory originally pursued these new GH30 enzymes due to their apparently unique phylogenetic placement. For *Ac*Xyn30B, encoded by locus tag Clocl_2746 in the *A. clariflavus* genome, our initial studies indicated not only the likelihood of a new subfamily, but also a distinct placement within the GH30 family. To support the establishment of this new subfamily with more than a single characterized representative, we selected the homologous enzyme endo-1,4-β-xylanase A from *Anaerobacterium chartisolvens* (*Ach*Xyn30A; Horino *et al.*, 2014) for further study.

In this report, we present the first crystallographic structure of a GH30_12 EX. *Ach*Xyn30A consists of the typical dual-domain architecture of a GH30 $(\beta/\alpha)_8 + \beta$ fold (St John *et al.*, 2010) [also referred to in full as a GH30 catalytic domain (CD) or in part as a $(\beta/\alpha)_8$-barrel and side-β-domain, SβD). We provide a detailed structural characterization of *Ach*Xyn30A and perform comparisons with other GH30 xylan-active enzymes. To complement the previously reported biochemical information, we compare the *Ach*Xyn30A structure with a homology model of *Ac*Xyn30B and attempt to reconcile the reported biochemical activities with structural features. These structural and functional insights into these GH30_12 enzymes will help to determine how diverse β-1,4-xylan-hydrolyzing functions may be utilized in the process of biomass bioconversion.

2. Materials and methods

2.1. Primary amino-acid sequence analysis

The amino-acid sequence of locus tag DFR58_1374 (UniProt Accession A0A369AIV8) from the *A. chartisolvens* genome was utilized for this research. The Carbohydrate-Active enZYmes Database (CAZY; Drula *et al.*, 2022), UniProt (UniProt Consortium, 2025) and Entrez databases (Schuler *et al.*, 1996) were utilized for sequence collection. *BLASTp* was performed to collect homologous sequences (Altschul *et al.*, 1990). Defined sequence sets (a large proportion of characterized proteins), expanded sequence sets (a mixture of characterized and uncharacterized) and large discovery sets (sequence collection down to 30% identity) of *Ach*Xyn30A and representative members of other GH30 xylan-active enzyme subfamilies were utilized, as well as subsets of these collections weighting subfamilies either equally or in the same proportion as the discovery set. Sets were processed by trimming off secretion signal sequences (Almagro Armenteros *et al.*, 2019) and other non-GH30

domains such as carbohydrate-binding and dockerin domains identified using the Conserved Domain Database (Wang *et al.*, 2023). Alignments were performed utilizing the *MAFFT* algorithm (Katoh *et al.*, 2002) and bootstrapped (100×) maximum-likelihood trees were built using the *Lasergene* software (DNAStar, Madison, Wisconsin, USA).

2.2. Cloning, expression and purification of *Ach*Xyn30A

The translated genome sequence encoding *Ach*Xyn30A was trimmed to consist of only the amino-acid sequence of the GH30 CD. The signal-peptide cleavage site was predicted using *SignalP* 5.0 (Almagro Armenteros *et al.*, 2019) and additional C-terminal protein domains were trimmed based on homologous GH30 protein structures and *in silico* domain-prediction tools (Wang *et al.*, 2023). A codon-optimized nucleotide sequence for protein expression in *Escherichia coli* was ordered through Integrated DNA Technologies (IDT; Coralville, Iowa, USA) and was cloned using ligation-independent cloning into the pMCSG53 vector (Eschenfeldt *et al.*, 2013) with in-frame fusion to a N-terminal His-tag/TEV protease cleavage recognition site. The prepared expression plasmid construct was nucleotide-sequence verified to encode the amino-acid sequence MHHHHHHSSGVDLGT**ENLYFQS** NAMAYSVSVDINT.....AQSVVTVALE, with the N-terminal starting Met, His-tag, TEV recognition and cut site in bold and the cloning artifact sequence underlined. This expression construct was provided to the Structural Biology Center of the Advanced Photon Source at Argonne National Laboratory (Lemont, Illinois, USA) for high-throughput protein expression and structure determination (Tan *et al.*, 2017; Fan *et al.*, 2015; Eschenfeldt *et al.*, 2013). The theoretical molecular weight of the purified, TEV-processed *Ach*Xyn30A protein for crystallization is 47.2 kDa. Efforts to express this construct for protein biochemical studies using in-house auto-induction methods failed to yield significant protein.

For biochemical studies, an alternative construct was prepared using the same approach but rather included an in-frame C-terminal fusion with the pET-28-encoded His-tag. The optimized nucleic acid construct was sequence-verified to translate into the amino-acid sequence MAYSVSVDIN..... AQSVVTVALEHHHHHH. This expression product has a theoretical molecular weight of 47.8 kDa due primarily to the addition of the uncleavable affinity tag. This pET-28 construct was expressed in *E. coli* at 17°C for 36 h using autoinduction as originally described by Studier (2005). Cell pellets from 125 ml expression cultures received an EDTA-free Pierce protease-inhibitor Mini Tablet (Thermo Scientific, Rockford, Illinois, USA) and were solubilized in B-PER bacterial protein-extraction reagent (Thermo Scientific) according to the product instructions. Following a 15 min incubation at room temperature, the lysates were centrifuged at 31 000*g* at 15°C for 20 min. The supernatant was 0.22 µm filtered and the volume was amended using 5 *M* NaCl to contain 500 m*M* NaCl. This cell-free extract was subjected to immobilized metal-affinity chromatography (IMAC) using a 1 ml HisTrap HP column (Cytiva, Marlborough, Massachusetts; Crooks *et al.*, 2021). The *Ach*Xyn30A peak was desalted and further purified using a Mono Q column. Chromatographic purification was performed using a Bio-Rad BioLogic DuoFlow chromatography system (Hercules, California, USA). The resulting *Ach*Xyn30A protein peak was concentrated and buffer-exchanged to 1.8 mg ml^{-1} in 20 m*M* Tris pH 7.5 buffer. SDS–PAGE (Laemmli, 1970) using a BioRad Mini-PROTEAN Tetra Cell with 4–20% acrylamide gradient TGX gels was performed to confirm purity (Supplementary Fig. S1). This protein preparation was filter-sterilized, aliquoted and flash-frozen in liquid nitrogen for storage at −80°C. In all aspects of this research report, *Ach*Xyn30A consists of only the GH30 catalytic domain (CD), having two natively encoded C-terminal, noncatalytic domains removed.

2.3. Biochemical assay

Xylanase activity was measured using Nelson's test to quantify an increase in reducing-terminal sugar concentration (Nelson, 1944) as described previously (St John *et al.*, 2006, 2018). Reactions consisted of beechwood glucuronoxylan (BeWX) or wheat arabinoxylan (WAX) at 5 mg ml^{-1} in 50 m*M* sodium acetate buffer pH 5.5 at 40°C and were initiated by the addition of *Ach*Xyn30A to a final enzyme concentration of 2 µg ml^{-1}.

2.4. Crystallization, data collection and refinement of crystallographic data

Crystallization screening was performed by the sitting-drop vapor-diffusion method using a Mosquito Nanolitre Liquid Handler (TTP LabTech, Concord, California, USA) crystallization robot with CrystalQuick 96-well crystallization plates (Greiner, Monroe, California, USA). Protein concentrated to 11.37 mg ml^{-1} was used for crystal screening. For each crystallization condition, 0.4 µl protein solution and 0.4 µl precipitant solution were mixed and the mixture was equilibrated against 140 µl crystallization solution in each reservoir well. Microlytic crystallization screens MCSG-1–MCSG-4 (Anatrace Products, Maumee, Ohio, USA) were used and the plates were incubated at 16°C. Crystals appeared under multiple conditions, including 0.17 *M* ammonium acetate, 0.085 *M* sodium acetate–HCl pH 4.6, 25.5%(*w*/*v*) PEG 4000, 15%(*v*/*v*) glycerol, in which the best diffraction-quality crystals were obtained. Crystals were harvested and cryoprotected in the mother liquor supplemented with 25%(*v*/*v*) glycerol and then flash-cooled in liquid nitrogen.

Crystals were screened and diffraction data were collected at 100 K on the 19-ID (NYX) beamline of National Synchrotron Light Source II at Brookhaven National Laboratory. The intensities were integrated, scaled and merged with the *HKL*-3000 program suite (Minor *et al.*, 2006; Table 1). The structure factors were phased using molecular replacement (*MOLREP*; Vagin & Teplyakov, 2010) with a search model generated using *AlphaFold* (Jumper *et al.*, 2021). The final model was developed following iterative rounds of model rebuilding using *Coot* (Emsley *et al.*, 2010) and restrained refinement in *REFMAC* (Murshudov *et al.*, 2011). Differences between the

protomers in the asymmetric unit were investigated using the *LSQ Superpose* alignment tool in *Coot*. The final structure model was validated by analysis using *MolProbity* (Chen *et al.*, 2010; Table 1). The validated structure was deposited in the Protein Data Bank and assigned PDB code 9ywn. No crystals of *Ach*Xyn30A with xylooligosaccharides bound were obtained. Figures in this report were generated using *PyMOL* (DeLano, 2002). Structural alignments for comparative analysis between GH30 xylan-active subfamilies were performed using the pair fitting function in *PyMOL*. Specific atom positions from a set of catalytic-center proximal conserved amino acids were chosen for this purpose (Supplementary Fig. S5 and Supplementary Table S1).

Using the *Ach*Xyn30A crystallographic protein structure model as a template, *SWISS-MODEL* (Waterhouse *et al.*, 2018) was utilized to generate a homology model of the biochemically characterized GH30_12 enzyme *Ac*Xyn30B. The CDs of these enzymes share 62.5% homology, with almost perfect conservation in the xylan-binding cleft (XBC), so it is expected that the *Ac*Xyn30B model is a reliable prediction.

Table 1
Summary of crystallographic data for *Ach*Xyn30A.

Values in parentheses are for the highest resolution shell.

Data-collection statistics	
Space group	$P2_1$
a, b, c (Å)	89.3, 133.4, 137.7
α, β, γ (°)	90, 90, 90
Molecular weight† (Da)	47200
No. of amino acids†	426
Molecules in asymmetric unit	6
Wavelength (Å)	0.9857
Resolution (Å)	49.0–1.92 (1.95–1.92)
No. of unique reflections	239756
R_{meas} (%)	13.1 (116)
$R_{p.i.m.}$ (%)	7.2 (65)
$CC_{1/2}$	0.994 (0.403)
Completeness (%)	99.6 (96.2)
Multiplicity	3.3 (2.9)
$\langle I/\sigma(I)\rangle$	9.4 (0.89)
Solvent content (%)	56.68
Phasing	
Resolution range (Å)	49.0–3.00
Correlation coefficient‡ (%)	57.5
Refinement	
Resolution range (Å)	48.9–1.92
No. of reflections	227733
Completeness (%)	99.0
R_{work}/R_{free} (%)	16.9/20.5
No. of atoms (protein + heteroatoms)	22256
R.m.s.d., bond lengths (Å)	0.007
R.m.s.d., bond angles (°)	1.597
B factors (Å^2)	
Main chain	26.1
Side chain + H_2O	31.0
Wilson *B* factor (Å^2)	32.3
MolProbity validation	
Ramachandran outliers (%)	0.20
Ramachandran favored (%)	96.56
Rotamer outliers (%)	1.26
Clashscore	3.46
MolProbity score	1.44
PDB code	9ywn

† Not including the N-terminal His-tag and TEV protease recognition and cleavage sites. ‡ Molecular replacement: *AlphaFold*-generated model.

3. Results and discussion

3.1. Primary amino-acid sequence analysis and phylogeny

The first characterized GH30_12 subfamily member was identified in *A. clariflavus* (basonym *Clostridium clariflavum*; Artzi *et al.*, 2015; Šuchová *et al.*, 2024). This enzyme, *Ac*Xyn30B, was biochemically characterized by Šuchová and coworkers and established as the defining member of this new GH30 subfamily. Independent studies in our laboratory support these findings. While the conclusions reached are broadly consistent with those of Šuchová *et al.* (2024), our analysis presented in this current work offers further refinement of the distinctions between GH30 subfamilies. As detailed in Fig. 1, the original GH30 subfamily classification scheme (St John *et al.*, 2010) delineated the GH30 family into two groups based on the number of β-sheets of the dual-hinged side-β-domain (SβD) which derive from the N-terminus and the C-terminus. Group 1 enzymes, which include the originally classified GH30 members (Henrissat & Davies, 1997), have the first three β-strands encoded in the N-terminal region and the remaining six in the C-terminal region. In contrast, group 2 enzymes, which were more recently transferred into the GH30 family (St John *et al.*, 2010), encode only the first β-strand of the SβD in the N-terminal region and the remaining eight in the C-terminal region (Fig. 1). Therefore, *Ach*Xyn30A is empirically, by its structure, determined to be a group 2 GH30 enzyme.

However, in the enzyme-classification studies presented here, GH30_12 enzymes group more frequently, and with higher bootstrap values, with group 1 GH30 enzymes than with group 2. To confirm this, numerous types of amino-acid sequence data sets (see Section 2.1 above) were utilized. Fig. 2(*a*) illustrates a representative phylogeny of an equally weighted expanded set of 72 sequences. These findings indicate that detailed sequence analysis does not effectively partition group 1 and group 2 enzymes as in the past (St John *et al.*, 2010). This is likely due to the significantly increased diversity of available database sequences. Small variations in the aligned sequence dataset can markedly affect phylogenetic outcomes, emphasizing the importance of structural validation for GH30 classification. Notably, GH30_8 GX enzymes also now frequently group within group 1 in phylogenetic trees despite having a confirmed group 2 protein fold (Fig. 2*a*).

3.2. Function verification of *Ach*Xyn30A

Šuchová *et al.* (2024) recently published the first biochemical characterization of GH30_12 enzymes, demonstrating that the enzyme *Ac*Xyn30B from *A. clariflavus* exhibits similar activity on different xylan types. This indicates that it is a broadly functional EX. To evaluate the xylanase activity of *Ach*Xyn30A, we performed progress curve assays using WAX and BeWX as substrates (Fig. 2*b*). Activities were monitored by following the formation of reducing-end xylose at 40°C over time. The results showed that *Ach*Xyn30A showed a higher rate of hydrolysis for the more heavily substituted WAX compared with BeWX. The specific activity of

*Ach*Xyn30A from the progress curves was 3.32 ± 0.52 U mg^{-1} for WAX and 2.28 ± 0.17 U mg^{-1} for BeWX. These rate values are similar to those reported for *Ac*Xyn30B, indicating some similarity between these two enzymes (Šuchová *et al.*, 2024). Optimum conditions for 1–2 h reactions were determined to be 50°C at pH 5.5.

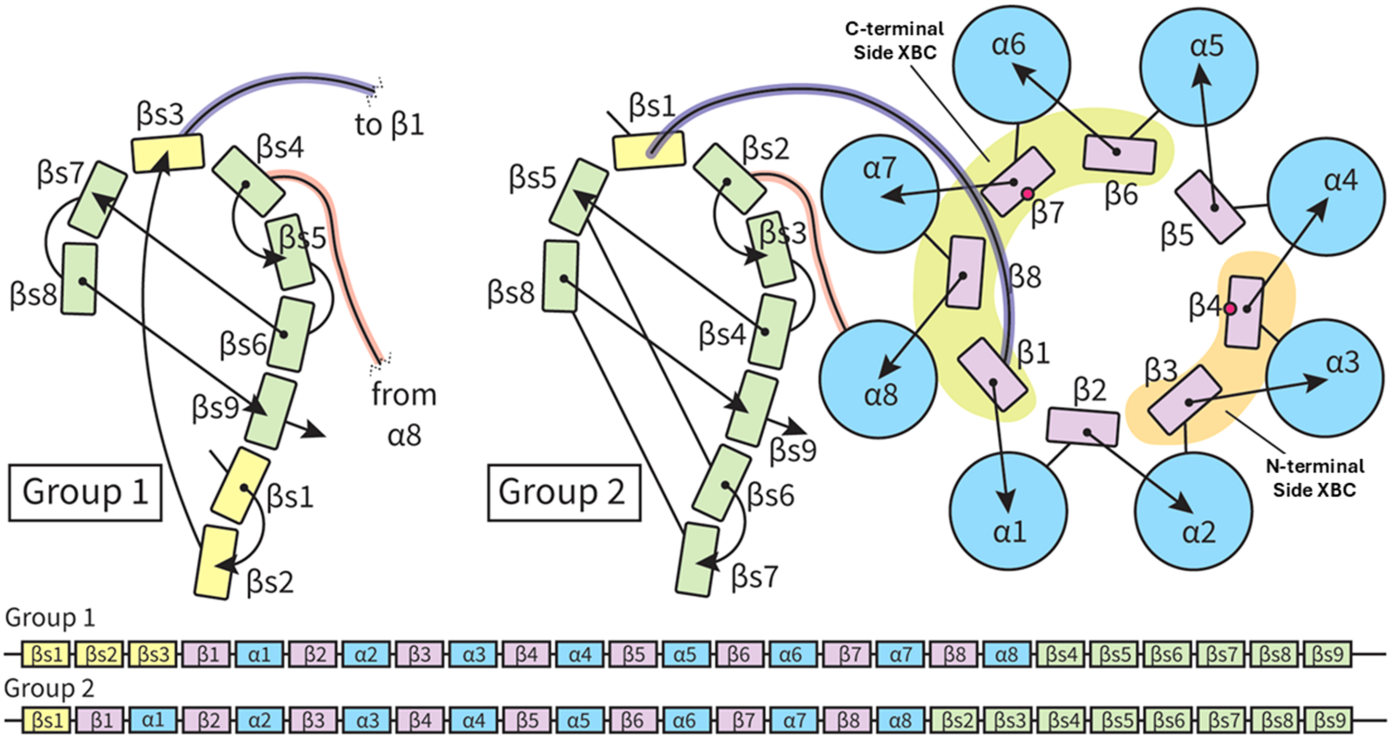
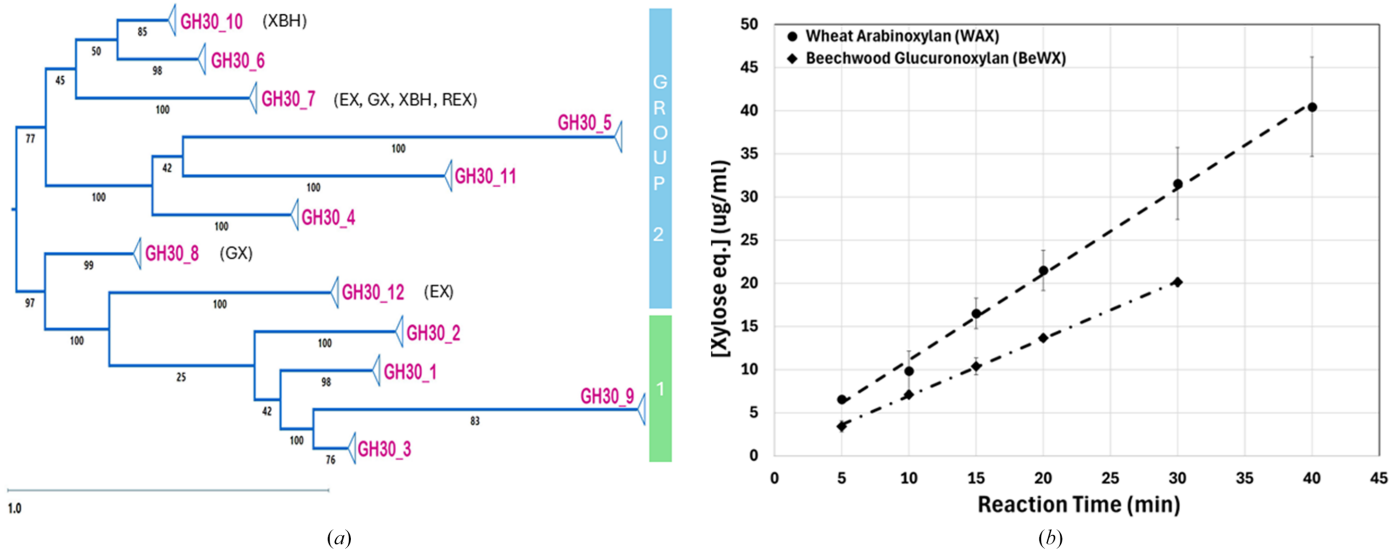


Figure 1
GH30 group 1 and group 2. Secondary-structure map of GH30 enzymes detailing the differences between GH30 group 1 and group 2 enzymes and providing the secondary-structure labels as used in this and related work. Also represented are the N-terminal linker (blue line) and the C-terminal linker (orange line) and the approximate location of the N-terminal and C-terminal sides of the xylan-binding cleft (XBC). The two red dots represent the locations of two catalytic residues, Glu143 after *β*4 and the acid/base Glu256 at the end of *β*7.



(*a*) (*b*)

Figure 2
Sequence analysis and activity. (*a*) Maximum-likelihood phylogenetic tree analysis reveals how GH30_12 enzymes (*Ach*Xyn30A, *Ac*Xyn30B and homologs) branch relative to other GH30 enzyme subfamilies. Our findings show that GH30_12 enzymes establish a separate branch among other GH30 enzymes and that this branch is (often) more closely related to GH30 group 1 enzymes (clustering with GH30_1, GH30_2 and GH30_3). Xylan-active subfamilies are additionally labeled with abbreviations of known xylan activities. (*b*) The xylanase activity of the purified enzyme was confirmed to generate a linear increase in reducing ends when incubated with wheat arabinoxylan (WAX) or with beechwood xylan (BeWX).

3.3. *Ach*Xyn30A protein structure

One asymmetric unit contains six *Ach*Xyn30A protomers and each protomer consists of the expected GH30 $(\beta/\alpha)_8 + \beta$ two-domain protein fold. These protomers are packed in a pseudo-*D*3-like symmetry with a dimer-of-trimers- or trimer-of-dimers-like assembly, which is suggestive of potential twofold, threefold or even sixfold biological symmetry for the intact enzyme molecule (Supplementary Figs. S2–S4). Use of the *PISA* server for the detection of biological assemblies did not find any potentially stable quaternary arrangements in solution (Krissinel & Henrick, 2007), largely due to the small interface areas between neighboring molecules within the hexametric assembly. This agrees with the result of high-performance gel-filtration chromatography molecular sizing of *Ach*Xyn30A performed during purification (data not shown). The six protomers are nearly identical based on pairwise structural alignment, with r.m.s.d. values ranging from 0.216 to 0.332 Å. Therefore, in the following presentation and discussion of the *Ach*Xyn30A protomer, only chain *A* (PDB entry 9ywn chain *A*) will be used as an example unless specified otherwise. The *Ach*Xyn30A protomer represents a canonical GH30 protein fold (Figs. 1 and 3). Fig. 3(*a*) shows the spatial relationship between the two domains of this CD and indicates the relative location of the XBC in which the xylan chain positions for catalysis. Fig. 3(*b*), with a 90° forward rotation, shows a birds-eye view with the catalytic nucleophile and acid–base catalyst highlighted in the central region of the XBC. This figure (and the alternative representation in Fig. 1) also indicates the two surfaces or sides which guide and position the xylan chain, thus forming the XBC. Also indicated in Fig. 3(*b*) are the glycone and aglycone regions of the XBC. These are the regions of the XBC which contain the catalytic enzyme–substrate complex and the side which contains the leaving group, respectively. A defining feature of this type of protein

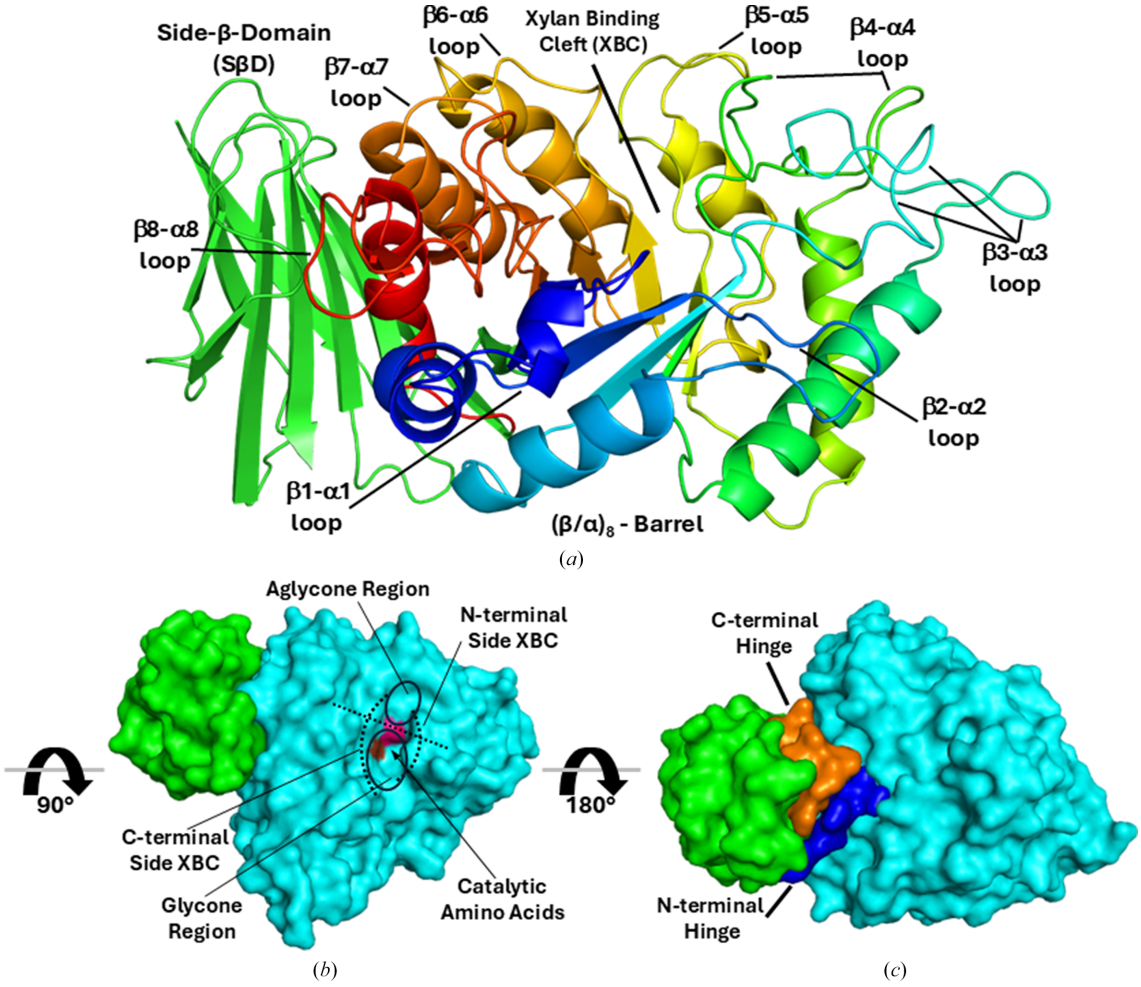


Figure 3
Protein structure of *Ach*Xyn30A at 1.92 Å resolution. (*a*) Cartoon view of the secondary-structure elements of *Ach*Xyn30A, showing it to be comprised of a $(\beta/\alpha)_8$-barrel domain (rainbow) and a side-β-domain (SβD, green) which compose the full GH30 $(\beta/\alpha)_8 + \beta$ dual-domain protein fold. The xylan-binding cleft (XBC) is located on the top (indicated), catalytically active face of the barrel domain. Each β–α loop on the $(\beta/\alpha)_8$-barrel domain is labeled for improved understanding of the three-dimensional orientation of *Ach*Xyn30A. (*b*) Following a 90° forward rotation, a birds-eye view of the XBC of *Ach*Xyn30A. As also indicated in Fig. 1, this image shows the approximate location of the N-terminal and C-terminal sides of the XBC. Additionally, the XBC is bisected into the glycone and aglycone regions, which harbor the enzyme–carbohydrate covalent complex during catalysis and the hydrolytic leaving group, respectively. Two catalytic residues are highlighted in red to identify their locations in the XBC. (*c*) An additional 180° forward rotation shows the noncatalytic 'bottom side' of *Ach*Xyn30A. In this view, the hinge regions which go to the $(\alpha/\beta)_8$ barrel (blue) and which returns to the SβD following the end of the barrel sequence (orange) are indicated.

fold is the dual hinge connection between the $(\beta/\alpha)_8$ barrel and SβD portions of the CD. Fig. 3(*c*) shows an additional 180° rotation to view the bottom noncatalytic side of *Ach*Xyn30A. This image highlights the hinges coming from (blue) and going to (orange) the SβD. With the interface between these two parts of the CD being predominantly hydrophobic, this bipartite domain fold may be considered as a single catalytic entity (Valenzuela *et al.*, 2012). As discussed above, although *Ach*Xyn30A is a group 2 GH30 enzyme by structural definition, phylogenetic analysis of amino-acid sequences indicates that it frequently clusters with group 1 GH30 enzymes, indicating a discrepancy between structure and sequence-based classifications.

3.4. Distinguishing the structural characteristics of *Ach*Xyn30A relative to other GH30 enzymes

Structural alignment of *Ach*Xyn30A with representative protein structures of xylan-hydrolyzing GH30 subfamilies 7, 8 and 10 reveals that *Ach*Xyn30A adopts a more open conformation than the enzymes from these previously established subfamilies. Loop regions β1–α1, β3–α3, β4–α4 and β6–α6 are similarly sized in all of the xylan-active GH30 subfamilies (for reference, see Fig. 3*a*). Each of these four loops are recognized as having an important role in xylan-chain coordination, positioning and/or catalysis (St John *et al.*, 2011, 2018, 2022). The β1–α1 loop forms the distal glycone XBC floor. The β3–α3 loop is vital for specifically coordinating a xylose near the active site. The β4–α4 loop represents a central functional section, being involved in substrate coordination, catalysis and formation of the N-terminal side of the XBC (Fig. 3*b*; Urbániková *et al.*, 2011; St John *et al.*, 2011). The β6–α6 loop is also important for ligand positioning and catalysis. Amino acids in this loop help to form the C-terminal side of the aglycone XBC (Fig. 3*b*) and also serve to position the catalytic residues through stabilizing hydrogen bonds. The amino-acid side chains in these loop regions are highly conserved across xylan-active GH30 enzymes, indicating their important role in enabling xylan catalysis.

In contrast, the β2–α2, β5–α5, β7–α7 and β8–α8 loop regions of *Ach*Xyn30A differ significantly from those of other GH30 xylan-active subfamilies. The β2–α2 loop is comparatively small in *Ach*Xyn30A (Fig. 4*a*). The strict GH30_8 GX enzymes are similarly small in this region, while GH30 enzymes with dual GX/XBH (GH30_7) or strict XBH (GH30_10) activity contain a much larger β2–α2 loop. In

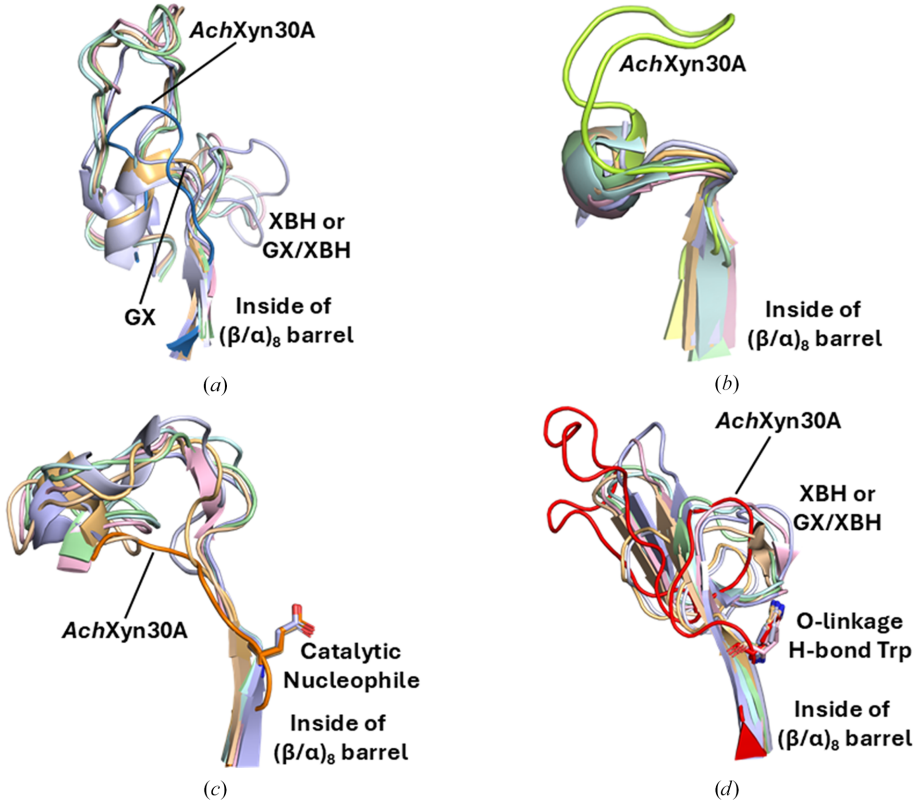


Figure 4
Differential loop regions in *Ach*Xyn30A compared with other xylan-active GH30 subfamilies. Loops are shown with the inside of the $(\beta/\alpha)_8$ barrel to the right and the extending loop to the left. (*a*) The β2–α2 loop is significantly smaller for *Ach*Xyn30A (dark blue) compared with GH30_7 and GH30_10 enzymes, which either have partial XBH activity or are strict-functioning XBH enzymes, respectively. (*b*) The β5–α5 loop in *Ach*Xyn30A (green) is significantly larger than in all of the other GH30 xylan-active enzyme structures. This may provide an extended surface area for xylan-chain interaction. (*c*) The β7–α7 loop-region comparison reveals that *Ach*Xyn30A (orange) has a much smaller loop compared with all of the other GH30 xylan-active protein structures. The location of the catalytic nucleophile is highlighted for guidance. (*d*) The β8–α8 loop shows *Ach*Xyn30A (red) to be smaller in the portion closest to the active-site cleft but larger in the distal region, where it extends much further away from the central area of the enzyme. The location of a conserved xylan-coordinating tryptophan is indicated for guidance.

enzymes having XBH activity, the β2–α2 loop extends back toward the distal glycone region, blocking the end of the channel and enabling hydrogen-bonding opportunities for the nonreducing terminal (−2 subsite) xylose. Similar to GH30_8 GXs, *Ach*Xyn30A has a small loop in this location consistent with functional characterization as a general EX, as also described for *Ac*Xyn30B (Šuchová *et al.*, 2024).

The β5–α5 loop of *Ach*Xyn30A is significantly larger, reaching higher and presenting a large region that may serve to stabilize the extending xylan chain (Fig. 4*b*). The β7–α7 loop in *Ach*Xyn30A is notably shorter than that typically observed (Fig. 4*c*). In GH30_7, GH30_8 and GH30_10 enzymes this loop contains amino acids which are thought to help coordinate the substrate (St John *et al.*, 2011, 2022; Freire *et al.*, 2016; Nakamichi *et al.*, 2023; Nakamichi, Fouquet, Ito, Watanabe *et al.*, 2019). In *Ach*Xyn30A, it appears impossible that any direct interaction could occur between the xylan main chain and amino-acid side chains in this shorter loop. The β8–α8 loop is slightly reduced in size near the substrate-binding cleft relative to the other GH30 xylan-active enzymes (Fig. 4*d*). In GH30_7 and GH30_8 xylanases the GA appendage is primarily coordinated through this loop, and similarly in GH30_7 and GH30_10 XBH enzymes this loop extends inwards to make hydrogen-bonding contacts directly with the glycone-positioned xylobiose (St John *et al.*, 2022). In *Ach*Xyn30A, as the β8–α8 loop extends to the outer region of the barrel it is comparatively much larger. It includes a β-hairpin loop in this location which engages with the same motifs from neighboring chains to form the threefold non-crystallographic symmetry described in the supporting information.

3.5. Ligand interactions in the *Ach*Xyn30A xylan-binding cleft

To envisage *Ach*Xyn30A–ligand interactions, we utilized the *PyMOL* pair fitting tool with selected atom positions (Supplementary Fig. S5 and Supplementary Table S1), rather than using the standard 3D alignments for comparison between protein structures. These positions represent amino acids which are directly involved in catalysis or XBC organization and function. Atoms chosen within each amino acid were selected based on a presumed greater positional restraint. This approach specifically targeted the alignment of the XBC and any ligands present in the structures, allowing the analysis of diverse GH30 xylan-active subfamilies. This alignment shows that the conserved XBC of these enzymes leads to near-identical positioning of xylooligosaccharide ligands from at least the −2 glycone subsite through the +1 aglycone subsite, as shown in a close-up overlay of these ligand-bound GH30 xylan-active enzymes (Fig. 5). For clarity, the carbohydrate ligands are separated from their proteins and important aspects are annotated for guidance. In this view, GH30 xylan-active enzymes are represented showing the most

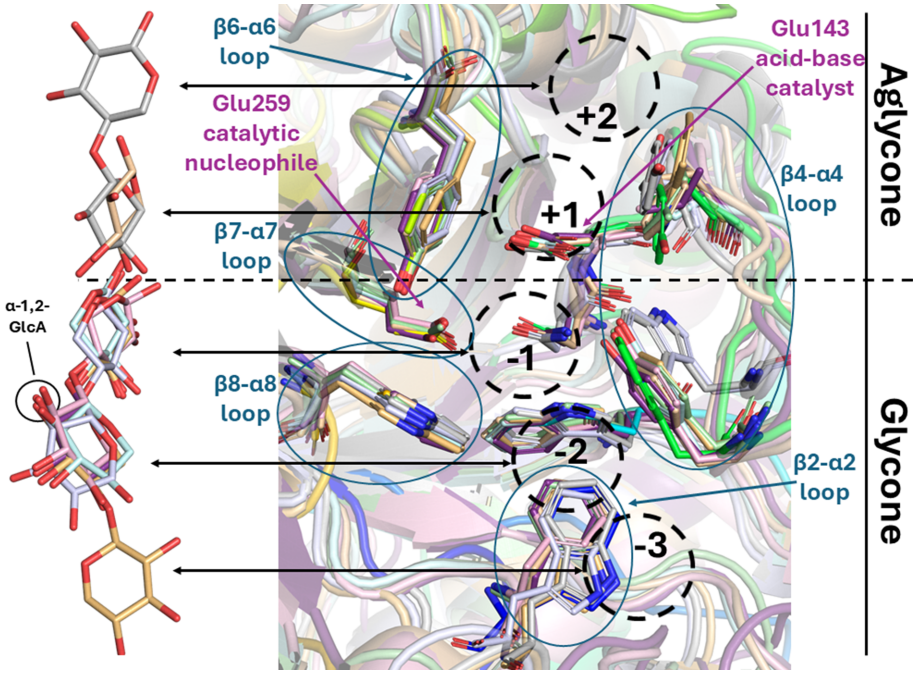


Figure 5
Amino-acid conservation in the XBC. Pair fitting alignment reveals the closeness between related GH30 subfamilies. The overlaid sugar ligands resulting from the alignment are shifted to the left side for clarity and subsites are indicated as well as the catalytic center and five loop regions. Pair fitting of *Ach*Xyn30A (in rainbow) with two GH30_7 dual-function GX/XBH enzymes [*Tc*Xyn30B (PDB entry 6krn) in pale cyan and *Tt*Xyn30A (PDB entry 7o0e) in light pink], a GH30_7 enzyme with REX activity and a GH30_7 EX with −2 subsite affinity [*Tc*Xyn30A (PDB entry 8idq) in wheat and *Tc*Xyn30C (PDB entry 6m5z) in pale green], three GH30_8 bacterial GXs [*Ec*Xyn30A (PDB entry 2y24) in light orange, *Bs*Xyn30C (PDB entry 3kl5) in blue white and *Ct*Xyn30A (PDB entry 5a6m) in gray] and lastly a GH30_10 XBH [*Ac*Xbh30A (PDB entry 7n6o) in light blue]. The glycone region ligands obtained from the listed *GX* GH30_8 enzymes each contained a GlcA substituted α-1,2 on the −2 subsite xylose. This appendage has been removed for clarity and its location is noted in the figure.

conserved XBC amino-acid side chains. The approximate locations of xylose-binding subsites and the catalytic center (dashed line), which separates the glycone and aglycone sections of the XBC, are indicated as they wrap around the N-terminal XBC side primarily formed by the large β4–α4 loop.

As indicated in Fig. 5 and shown in more detail in Fig. 6(*a*), stacking interactions between aromatic amino-acid residues that line the substrate-binding cleft are responsible for the overall position of xylan as it extends through the cleft. In the glycone region, Trp24 (*Ach*Xyn30A numbering) forms the cleft floor and provides a large stacking interface for the xyloses in the −3 and −2 subsites, while Trp151 provides the glycone N-terminal side of the XBC and stabilizes the −2 and −1 subsite xyloses. Notably, Trp151 extends upwards and creates a deeper substrate-binding cleft in *Ach*Xyn30A than that found in other GH30 xylan-active enzymes, which typically have a horizontally oriented tyrosine (Fig. 5) in this functionally conserved position. Trp88 extends the cleft floor towards the catalytic center, and on the opposite side (C-terminal side of the XBC) Trp284 acts similarly. These tryptophan residues additionally form highly conserved hydrogen bonds to the O3 hydroxyl of the −1 subsite xylose (Trp88) and to the glycosidic linkage connecting the −1 and −2 subsite xyloses (Trp284), respectively. The last conserved hydrogen bond involves Asn142 bonding to the O2 hydroxyl of the −1 subsite xylose (Fig. 6*a*). A positionally equivalent side chain of the distal Trp288 can be observed in structures of other GH30 xylan-active enzymes (PDB entries 5cxp, 8idq and 6m5z; St John *et al.*, 2018; Nakamichi, Fujii *et al.*, 2020; Nakamichi *et al.*, 2023). It forms an extended cleft wall on the C-terminal side of the glycone XBC region. Lastly, in the

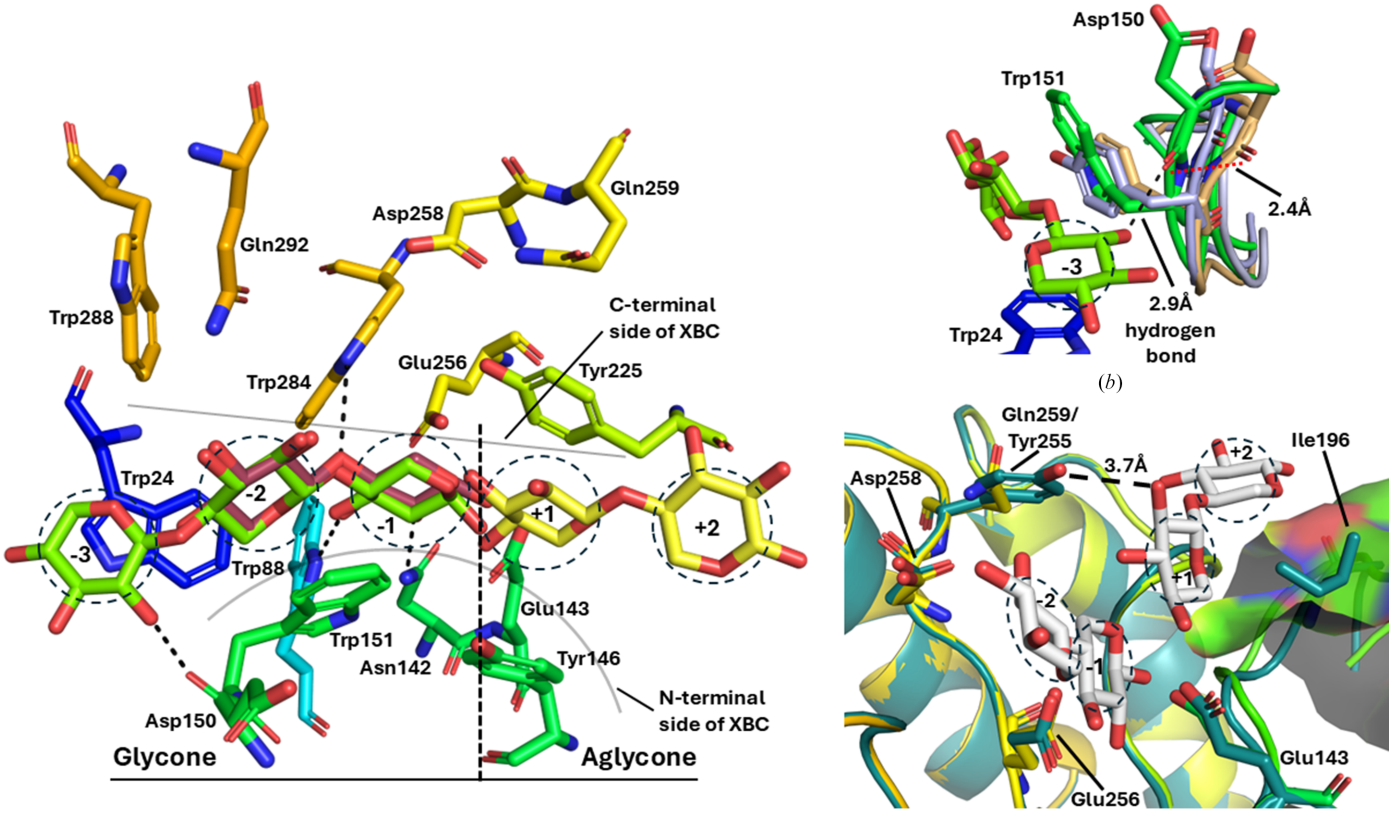


Figure 6
Active-site interactions which occur between the *Ach*Xyn30A protein structure and pair fitting alignment modeled xylooligosaccharides. (*a*) Amino-acid side chains of the *Ach*Xyn30A xylan-binding cleft with the xylan ligand positioned from −3 through +2 and subsites designated. Ligands were overlaid by the protein pair fitting alignments described in the text. The red xylobiose positioned in the −1 and −2 subsites derives from the XBH *Ac*Xbh30A (PDB entry 7n6o), the green xylotriose positioned in the −1 through −3 subsites derives from the aldotetrauronic acid (XUX) ligand from the GX *Ec*Xyn30A (PDB entry 2y24) but with the GlcA removed for clarity, and the yellow xylobiose in subsites +1 and +2 derives from the GX *Ct*Xyn30A (PDB entry 5a6m). (*b*) Close-up view of the specific β4–α4 loop in which Asp150 and Trp151 exist, showing the proposed new hydrogen bond which may account for the observed functional characteristics. The two most structurally similar GH30 enzymes (in this loop only) from the pair fitting alignment reveal that these other GH30 enzymes cannot establish a similar hydrogen bond from this loop as the main-chain carbonyls are rotated further outwards. Green, *Ach*Xyn30A, this report; light blue, *Ac*Xbh30A, PBD entry 7n6o; light orange, *Ec*Xyn30A, PDB entry 2y24. (*c*) Comparison of the XBC between *Ach*Xyn30A and *Ac*Xyn30B. In the unusually small β7–α7 loop of *Ach*Xyn30A, Gln259 (yellow) is set back too far for direct contact with the xylan occupying the +1 subsite. In *Ac*Xyn30B (cyan), this position is altered to Tyr255, which is predicted to be in proximity for a potential hydrogen bond to the O2 hydroxyl of the +1 subsite xylose. This change would also seem likely to limit +1 subsite binding of O2-substituted xylose into this position. For both *Ach*Xyn30A and *Ac*Xyn30B the β5–α5 loop is more pronounced than in other GH30 xylan-active enzymes. However, the two homologs are predicted to be different in this area as *Ac*Xyn30B has fewer amino acids in the loop. The homology-modeled structure of *Ac*Xyn30A positions an isoleucine extending upwards which would effectively disrupt any conserved +2 subsite interaction between *Ach*Xyn30A and *Ac*Xyn30B. *Ach*Xy30A, rainbow; *Ac*Xyn30B, cyan.

glycone region, Gln292 reaches towards the top side of the −2 subsite xylose but is beyond hydrogen-bonding distance from the xylan main chain. It may interact with Ara*f* substitutions on the O2 and O3 hydroxyl positions of xylose. While an O2 GlcA substitution on this xylose would clash with Gln292 if oriented as observed in GH30_8 GX enzymes (for example PDB entry 2y24), the recent biochemical characterization of *Ac*Xyn30B (Šuchová *et al.*, 2024) shows that the GlcA does not disrupt binding for catalysis, indicating then that the possible reorientation of GlcA into an alternative position and/or some structural flexibility of the region may possibly occur to accommodate substituted xylan.

In *Ac*Xyn30B, optimal activity on substituted xylooligosaccharides was observed when a xylose occupied the −3 subsite (Šuchová *et al.*, 2024). Overlaying a modeled xylotriose (from PDB entry 2y24) indicates that a hydrogen bond would be expected to occur between the O2 hydroxyl of the −3 subsite xylose and the main-chain carbonyl of Asp150 in *Ach*Xyn30A (Figs. 6*a* and 6*b*). This amino acid is conserved between *Ach*Xyn30A and *Ac*Xyn30B and would provide an additional specific stabilizing contact as the loop structure extends inwards, providing more opportunity for substrate interaction than observed in similar GH30 enzyme structures. The two visually most similar GH30 enzymes in this region of the β4–α4 loop (PDB entries 7n6o and 2y24; St John *et al.*, 2022; Urbániková *et al.*, 2011) show the equivalent main-chain carbonyl position being rotated outwards by an additional 2.4 Å (Fig. 6*b*) and therefore out of reach for hydrogen-bond formation. This is a distinguishing feature of these GH30_12 enzymes in being the first to show a structural basis for specific interactions with xylose in the −3 subsite. A previous biochemical study of the *Ca*Xyn30A noncanonical GH30_8 EX indicated a stabilizing role for this subsite, but structural evidence could not be found from the X-ray crystallographic protein structure model to account for this observation (PDB entry 5cxp; St John *et al.*, 2018).

In the aglycone region, the +1 subsite is flanked on either side by the aromatic amino acids Tyr146 and Tyr225, which position the +1 subsite xylose. No hydrogen bonds are detected for this subsite (Figs. 5 and 6*a*), and specific hydrogen-bonding interactions for the +1 subsite xylose in the related GH30 enzymes are limited, although a few have been suggested to occur (Freire *et al.*, 2016; St John *et al.*, 2022). The positioning of the +2 xylose in this alignment model is not optimal, which is not surprising given that it was obtained from a GH30_8 enzyme (PDB entry 5a6m) which is structurally different in the distal aglycone region (Freire *et al.*, 2016). The cleft widens following the +1 subsite, and the xylose in the +2 subsite is primarily bound in the model by its connection to the +1 xylose.

On the C-terminal side of the aglycone region (Figs. 5 and 6*a*), just above Tyr255, residues Asp258 and Gln259 are set back as part of the unusually short β7–α7 loop. These amino acids extend inwards towards the XBC, but not enough to establish contact with the xylan main chain (Fig. 6*a*). These residues could potentially interact with substitutions such as O2 GlcA or Ara*f*. However, the ligand modeling suggests that O3 Ara*f* substitutions on the +1 subsite xylose would clash significantly with the glycone-region xylan main chain, indicating that the +1 subsite may not be able to accommodate the common O3 Ara*f* substitution.

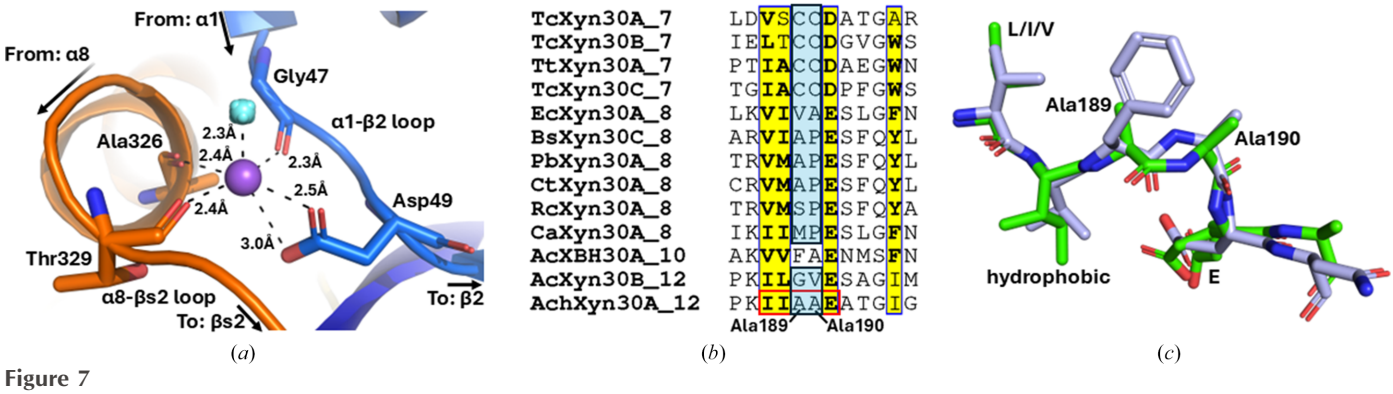
We compared the *Ach*Xyn30A crystallographic structure with a *SWISS-MODEL* (Waterhouse *et al.*, 2018) homology model of the biochemically characterized *Ac*Xyn30B (Šuchová *et al.*, 2024). Two nontrivial differences were identified among the amino-acid side chains which are thought to interface with the xylan. In the +1 subsite, in place of Gln259, *Ac*Xyn30B has Tyr255 (Fig. 6*c*). This amino acid extends an additional 3.6 Å towards the +1 subsite xylose and places the terminal side-chain hydroxyl at near-hydrogen-bonding distance (3.7 Å) to the O2 hydroxyl of the modeled +1 subsite xylose. These modeled ligand positions are not exact, and some movement would likely allow proper hydrogen-bonding contact. However, Tyr255 in *Ac*Xyn30B could possibly interfere with the O2 GlcA or Ara*f* substitutions on the +1 subsite xylose.

The second notable XBC difference between *Ach*Xyn30A and *Ac*Xyn30B is found in the β5–α5 loop. Although both enzymes have larger loops than are observed in other xylan-active GH30 subfamilies, the loop in *Ac*Xyn30B is smaller than that in *Ach*Xyn30A and is therefore predicted to have a simpler structure. In *Ac*Xyn30B, Ile196 is predicted to bulge upwards, which would alter any interactions with the +2 subsite xylose as may occur in *Ach*Xyn30A (Fig. 6*c*). Together, the possible isoleucine bulge at the +2 subsite and the Gln259→Tyr255 replacement at the +1 subsite may result in distinct xylan-hydrolysis products despite classifying together in the new GH30_12 subfamily.

3.6. Protein structure novelties in *Ach*Xyn30A

*Ach*Xyn30A has a tightly coordinated sodium ion in the hinge region (Fig. 7*a*). The same amino-acid motif seems to be fully conserved in the biochemically characterized GH30_12 *Ac*Xyn30B. This has not been observed before for any other GH30 xylan-active (or other GH30) enzyme and is thought to add rigidity to an area of the enzyme which would typically be considered to have flexibility. This sodium ion has octahedral coordination (coordination number 6), formed with five amino-acid positions and one water. As shown in Fig. 7(*a*), three of these coordination points are main-chain carbonyls and two derive from the side chain of Asp49. The water is surface-accessible. This sodium coordination system brings together the α1–β2 loop and the α8–βs2 loop (C-terminal hinge; see Fig. 1, group 2). The sodium coordination does not involve the N-terminal hinge (βs1–β1 loop) region. Gly47 and Asp49 in the α1–β2 loop each contribute one main-chain carbonyl and the two side-chain carboxyl groups of Asp49, respectively, while the very end of the α8 helix donates two main-chain carbonyls from Ala326 and Thr329. It is not clear whether the metal-binding site is opportunistic or acts as regulator in response to a metal-ion gradient.

Another notable structural feature in *Ach*Xyn30A is the presence of a rare Ala–Ala (Ala189–Ala190) *cis*-peptide right



(*a*) (*b*) (*c*)

Figure 7
Additional structural observations from *Ach*Xyn30A. (*a*) A coordinated sodium ion (purple) is observed in the hinge region of *Ach*Xyn30A. This coordination brings together the C-terminal hinge region with a section of the barrel structure, presumably increasing the rigidity of the interface. The coordination is formed by five amino-acid contacts and a surface-accessible water (cyan). *Ach*Xyn30A also has a rare Ala–Ala *cis*-peptide near its active site. (*b*) An alignment of this conserved *cis*-peptide region including many of the protein structures used for comparison throughout this study identifies conserved sequence elements (highlighted in yellow) surrounding the *cis*-peptide (highlighted in blue). The red box specifically highlights the Ala-*cis*-Ala motif in the *Ach*Xyn30A protein structure. The sequence from *Ac*Xbh30A (labeled AcXBH30A_10), as discussed in the text, is not in the *cis*-peptide conformation, although the flanking region is conserved. (*c*) Aligned structures of *Ach*Xyn30A (green) and the only GH30 xylan-active enzyme, the XBH *Ac*Xbh30A (PDB entry 7n6o), which does not have a *cis*-peptide in this region (light blue). The XBH still shows the larger conserved amino-acid sequence.

after the end of the β5 strand of the $(\beta/\alpha)_8$-barrel domain. Although an alanine residue has been observed in either the first or the second position of a *cis*-peptide, the Ala-*cis*-Ala peptide revealed in this 1.92 Å resolution structure is the first example of this amino-acid arrangement based on a reliability-filtered list of *cis*-nonPro peptides found in all known structures. An elevated presence of nonproline *cis*-peptides in carbohydrate-active enzymes has previously been reported, especially in $(\beta/\alpha)_8$ TIM barrel-containing proteins (Jabs *et al.*, 1999; Richardson *et al.*, 2025). In contrast to the typical *trans*-peptide form, *cis*-peptides are energetically less favorable. The presence of a proline in the second amino-acid position of a *cis*-peptide (*X*-*cis*-Pro) improves the energetics of *cis*-peptides; as such the *X*-*cis*-Pro arrangement represents approximately 5% of all prolines in proteins (Jabs *et al.*, 1999) and *X*-*cis*-nonPro arrangements occur in just 0.03% of peptide bonds (Richardson *et al.*, 2025). The *Ach*Xyn30A *cis*-peptide appears to be conserved in several of the GH30 xylan-active subfamilies (Fig. 7*b*). In GH30_7 enzymes this *cis*-peptide is composed of rarely observed vicinal cysteine residues (Nakamichi *et al.*, 2023; Nakamichi, Fujii *et al.*, 2020; Nakamichi, Fouquet, Ito, Watanabe *et al.*, 2019; Nikolaivits *et al.*, 2021). In the reported GH30_8 protein structures, three of the four have a proline residue for the latter amino acid, resulting in the energetically more favorable *X*-*cis*-Pro configuration (St John *et al.*, 2011, 2018; Freire *et al.*, 2016). In the new GH30_12 subfamily, as represented by *Ach*Xyn30A, the *cis* amino acids are Ala189–Ala190. In contrast, this motif in the GH30_10 subfamily of strict XBHs is not in the *cis*-peptide conformation in either the apo (PDB entry 7n6h) or the xylobiose-bound (PDB entry 7n6o) protein structure forms (Figs. 7*b* and 7*c*; St John *et al.*, 2022). Primary amino-acid sequence alignment of this region supports a level of conservation in the amino acids which flank these *cis*-peptides.

Structural informatics has identified themes involving *cis*-peptide placement and occurrence frequency (Richardson *et al.*, 2017, 2025; Pal & Chakrabarti, 1999) and much of our current understanding is built around these informatics findings. The location of the *cis*-peptide in these GH30 enzymes after the end of the β5 strand of the $(\beta/\alpha)_8$ barrel is of note as it engages directly with the catalytic motif. The neighboring β4 strand of the barrel positions the acid–base catalyst Glu143 (*Ach*Xyn30A), which is responsible for resolution of the double-displacement reaction cycle (Koshland, 1953). The Glu143-containing catalytic motif NEP (Asn-Glu143-Pro) is conserved in all of these GH30 subfamilies. Pro144 of this motif is in proximity to the second residue of the *cis*-peptide, Ala190. While no version of this *cis*-peptide (*i.e.* *cis*-vicinal Cys, *X*-*cis*-nonPro or *X*-*cis*-Pro) is directly exposed to the surface of the active-site cleft or involved in any obvious manner with catalysis, its conserved position proximal to the acid–base catalyst, its placement as the 'subfloor' of the aglycone (leaving group) side of the XBC, its rare occurrence and its unfavorable energy may all indicate it to have a role in the function of these enzymes.

4. Conclusion

The crystal structure of *Ach*Xyn30A confirms its classification as a GH30 group 2 enzyme. Phylogenetic studies support a previous determination (Šuchová *et al.*, 2024) that this enzyme classifies into a newly defined GH30_12 subfamily. The clearest structural observation concerning *Ach*Xyn30A is that the XBC appears to be less restrictive compared with other GH30 xylanase-subfamily members. The C-terminal side of both the glycone and aglycone loop regions, which are well characterized to coordinate either the xylan chain or xylan-chain appendages, are recessed from the cleft. This limits direct xylan-chain engagement, but presumably allows greater accommodation of xylan-chain appendages. These GH30_12 enzymes are the first to show an increased specificity via a hydrogen bond which occurs between the O2 hydroxyl of

the −3 subsite xylose and the β4–α4 loop residue Asp150. Although *Ach*Xyn30A shares greater than 60% amino-acid sequence identity with the recent biochemically characterized *Ac*Xyn30B, differences in the aglycone region may alter the functional specificity. GH30 enzymes are uniquely adept at hydrolysis of β-1,4-xylan polymers, as indicated by the diversity of function which has evolved around a central core protein structure.

Acknowledgements

We are grateful to members of the Structural Biology Center (SBC) at Argonne National Laboratory and Michelle Radford for facilitating the beamline time application for research access.

Funding information

We thank the Institute for Microbial and Biochemical Technology of the Forest Products Laboratory (FPL), a National Laboratory of the USDA Forest Service, for ongoing financial support. Results in this report are derived from work performed through the eBERlight and SBC programs using Sector 84 (Advanced Protein Characterization Facility), funded by the US Department of Energy, Office of Biological and Environmental Research at the Advanced Photon Source, a US DOE Office of Science User Facility, supported by the US DOE, Office of Basic Energy Sciences, under Contract No. DE-AC02-06CH11357. This research used resources of the National Synchrotron Light Source II, a US Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Brookhaven National Laboratory under Contract No. DE-SC0012704.

References

Abik, F., Palasingh, C., Bhattarai, M., Leivers, S., Ström, A., Westereng, B., Mikkonen, K. S. & Nypelö, T. (2023). *J. Agric. Food Chem.* **71**, 2667–2683.
Almagro Armenteros, J. J., Tsirigos, K. D., Sønderby, C. K., Petersen, T. N., Winther, O., Brunak, S., von Heijne, G. & Nielsen, H. (2019). *Nat. Biotechnol.* **37**, 420–423.
Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). *J. Mol. Biol.* **215**, 403–410.
Anderson, V. R. & Perry, C. M. (2006). *Drugs*, **66**, 821–835.
Artzi, L., Morag, E., Barak, Y., Lamed, R. & Bayer, E. A. (2015). *mBio*, **6**, e00411-15.
Biely, P., Puchart, V., Stringer, M. A. & Mørkeberg Krogh, K. B. R. (2014). *FEBS J.* **281**, 3894–3903.
Chakraborty, M. & Bhowal, J. (2025). *Food Rev. Int.* **41**, 643–670.
Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D**66**, 12–21.
Chow, V., Nong, G., St John, F. J., Sawhney, N., Rice, J. D. & Preston, J. F. (2022). *J. Ind. Microbiol. Biotechnol.* **49**, kuab080.
Collins, T., Gerday, C. & Feller, G. (2005). *FEMS Microbiol. Rev.* **29**, 3–23.
Crooks, C., Bechle, N. J. & St John, F. J. (2021). *Front. Mol. Biosci.* **8**, 714238.
DeLano, W. L. (2002). *The PyMOL User's Manual*. Palo Alto: DeLano Scientific.
Drula, E., Garron, M.-L., Dogan, S., Lombard, V., Henrissat, B. & Terrapon, N. (2022). *Nucleic Acids Res.* **50**, D571–D577.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Eschenfeldt, W. H., Makowska-Grzyska, M., Stols, L., Donnelly, M. I., Jedrzejczak, R. & Joachimiak, A. (2013). *J. Struct. Funct. Genomics*, **14**, 135–144.
Fan, Y., Tan, K., Chhor, G., Butler, E. K., Jedrzejczak, R. P., Missiakas, D. & Joachimiak, A. (2015). *Protein Sci.* **24**, 1389–1400.
Fengel, D. & Wegener, G. (1979). *Hydrolysis of Cellulose: Mechanisms of Enzymatic and Acid Catalysis*, edited by R. D. Brown Jr & L. Jurasek, pp. 145–158. Washington, DC: American Chemical Society.
Freire, F., Verma, A., Bule, P., Alves, V. D., Fontes, C. M. G. A., Goyal, A. & Najmudin, S. (2016). *Acta Cryst.* D**72**, 1162–1173.
Granato, D., Reshamwala, D., Korpinen, R., Azevedo, L., Vieira do Carmo, M. A., Cruz, T. M., Marques, M. B., Wen, M., Zhang, L., Marjomäki, V. & Kilpeläinen, P. (2022). *Food Chem.* **381**, 132284.
Henrissat, B. & Davies, G. (1997). *Curr. Opin. Struct. Biol.* **7**, 637–644.
Horino, H., Fujita, T. & Tonouchi, A. (2014). *Int. J. Syst. Evol. Microbiol.* **64**, 1296–1303.
Jabs, A., Weiss, M. S. & Hilgenfeld, R. (1999). *J. Mol. Biol.* **286**, 291–304.
Jacobs, A., Larsson, P. T. & Dahlman, O. (2001). *Biomacromolecules*, **2**, 979–990.
Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A. W., Kavukcuoglu, K., Kohli, P. & Hassabis, D. (2021). *Nature*, **596**, 583–589.
Kaprelyants, L., Pozhitkova, L., Velichko, T. & Bilyk, O. (2019). *Grain Prod. Mix. Fodder's*, **19**, 16–26.
Katoh, K., Misawa, K., Kuma, K. & Miyata, T. (2002). *Nucleic Acids Res.* **30**, 3059–3066.
Katsimpouras, C., Dedes, G., Thomaidis, N. S. & Topakas, E. (2019). *Biotechnol. Biofuels*, **12**, 120.
Koshland, D. E. (1953). *Biol. Rev.* **28**, 416–436.
Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.
Laemmli, U. K. (1970). *Nature*, **227**, 680–685.
Li, N., Zhang, R., Zhou, J. & Huang, Z. (2023). *J. Agric. Food Chem.* **71**, 7961–7976.
Mikkonen, K. S. & Tenkanen, M. (2012). *Trends Food Sci. Technol.* **28**, 90–102.
Minor, W., Cymborowski, M., Otwinowski, Z. & Chruszcz, M. (2006). *Acta Cryst.* D**62**, 859–866.
Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.
Nakamichi, Y., Fouquet, T., Ito, S., Matsushika, A. & Inoue, H. (2019). *Appl. Environ. Microbiol.* **85**, e00552-19.
Nakamichi, Y., Fouquet, T., Ito, S., Watanabe, M., Matsushika, A. & Inoue, H. (2019). *J. Biol. Chem.* **294**, 4065–4078.
Nakamichi, Y., Fujii, T., Fouquet, T., Matsushika, A. & Inoue, H. (2019). *Appl. Environ. Microbiol.* **85**, e01442-19.
Nakamichi, Y., Fujii, T., Watanabe, M., Matsushika, A. & Inoue, H. (2020). *Acta Cryst.* F**76**, 341–349.
Nakamichi, Y., Watanabe, M., Fujii, T., Inoue, H. & Morita, T. (2023). *Proteins*, **91**, 1341–1350.
Nakamichi, Y., Watanabe, M., Matsushika, A. & Inoue, H. (2020). *FEBS Open Bio*, **10**, 1180–1189.
Nechita, P., Mirela, R. & Ciolacu, F. (2021). *Sustainability*, **13**, 13504.
Nelson, N. (1944). *J. Biol. Chem.* **153**, 375–380.
Nikolaivits, E., Pentari, C., Kosinas, C., Feiler, C. G., Spiliopoulou, M., Weiss, M. S., Dimarogona, M. & Topakas, E. (2021). *Carbohydr. Polym.* **273**, 118553.

Ohbuchi, T., Sakaino, M., Takahashi, T., Azumi, N., Ishikawa, K., Kawazoe, S., Kobayashi, Y. & Kido, Y. (2010). *J. Nutr. Sci. Vitaminol.* **56**, 54–59.
Pal, D. & Chakrabarti, P. (1999). *J. Mol. Biol.* **294**, 271–288.
Peña, M. J., Kulkarni, A. R., Backe, J., Boyd, M., O'Neill, M. A. & York, W. S. (2016). *Planta*, **244**, 589–606.
Petzold-Welcke, K., Schwikal, K., Daus, S. & Heinze, T. (2014). *Carbohydr. Polym.* **100**, 80–88.
Puchart, V., Šuchová, K. & Biely, P. (2021). *Biotechnol. Adv.* **47**, 107704.
Richardson, J. S., Videau, L. L., Williams, C. J., Hintze, B. J., Lewis, S. M. & Richardson, D. C. (2025). *Protein Sci.* **34**, e70157.
Richardson, J. S., Videau, L. L., Williams, C. J. & Richardson, D. C. (2017). *J. Mol. Biol.* **429**, 1321–1335.
Schuler, G. D., Epstein, J. A., Ohkawa, H. & Kans, J. A. (1996). *Methods Enzymol.* **266**, 141–162.
St John, F. J., Bynum, L., Tauscheck, D. A. & Crooks, C. (2024). *BMC Res. Notes*, **17**, 175.
St John, F. J., Crooks, C., Kim, Y., Tan, K. & Joachimiak, A. (2022). *FEBS Lett.* **596**, 2449–2464.
St John, F. J., Dietrich, D., Crooks, C., Balogun, P., de Serrano, V., Pozharski, E., Smith, J. K., Bales, E. & Hurlbert, J. (2018). *Biochem. J.* **475**, 1533–1551.
St John, F. J., González, J. M. & Pozharski, E. (2010). *FEBS Lett.* **584**, 4435–4441.
St John, F. J., Hurlbert, J. C., Rice, J. D., Preston, J. F. & Pozharski, E. (2011). *J. Mol. Biol.* **407**, 92–109.
St John, F. J., Rice, J. D. & Preston, J. F. (2006). *J. Bacteriol.* **188**, 8617–8626.
Studier, F. W. (2005). *Protein Expr. Purif.* **41**, 207–234.
Šuchová, K., Fathallah, W. & Puchart, V. (2024). *Appl. Microbiol. Biotechnol.* **108**, 312.
Šuchová, K., Puchart, V. & Biely, P. (2021). *Appl. Microbiol. Biotechnol.* **105**, 185–195.
Šuchová, K., Puchart, V., Spodsberg, N., Mørkeberg Krogh, K. B. R. & Biely, P. (2020). *Enzyme Microb. Technol.* **134**, 109484.
Tan, K., Deatherage Kaiser, B. L., Wu, R., Cuff, M., Fan, Y., Bigelow, L., Jedrzejczak, R. P., Adkins, J. N., Cort, J. R., Babnigg, G. & Joachimiak, A. (2017). *Protein Sci.* **26**, 1738–1748.
Tenkanen, M., Vršanská, M., Siika-aho, M., Wong, D. W., Puchart, V., Penttilä, M., Saloheimo, M. & Biely, P. (2013). *FEBS J.* **280**, 285–301.
Timell, T. (1967). *Wood Sci. Technol.* **1**, 45–70.
UniProt Consortium (2025). *Nucleic Acids Res.* **53**, D609–D617.
Urbániková, L., Vršanská, M., Mørkeberg Krogh, K. B., Hoff, T. & Biely, P. (2011). *FEBS J.* **278**, 2105–2116.
Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* D**66**, 22–25.
Valenzuela, S. V., Diaz, P. & Pastor, F. J. (2012). *Appl. Environ. Microbiol.* **78**, 3923–3931.
Vršanská, M., Kolenová, K., Puchart, V. & Biely, P. (2007). *FEBS J.* **274**, 1666–1677.
Wang, J., Chitsaz, F., Derbyshire, M. K., Gonzales, N. R., Gwadz, M., Lu, S., Marchler, G. H., Song, J. S., Thanki, N., Yamashita, R. A., Yang, M., Zhang, D., Zheng, C., Lanczycki, C. J. & Marchler-Bauer, A. (2023). *Nucleic Acids Res.* **51**, D384–D388.
Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., de Beer, T. A. P., Rempfer, C., Bordoli, L., Lepore, R. & Schwede, T. (2018). *Nucleic Acids Res.* **46**, W296–W303.