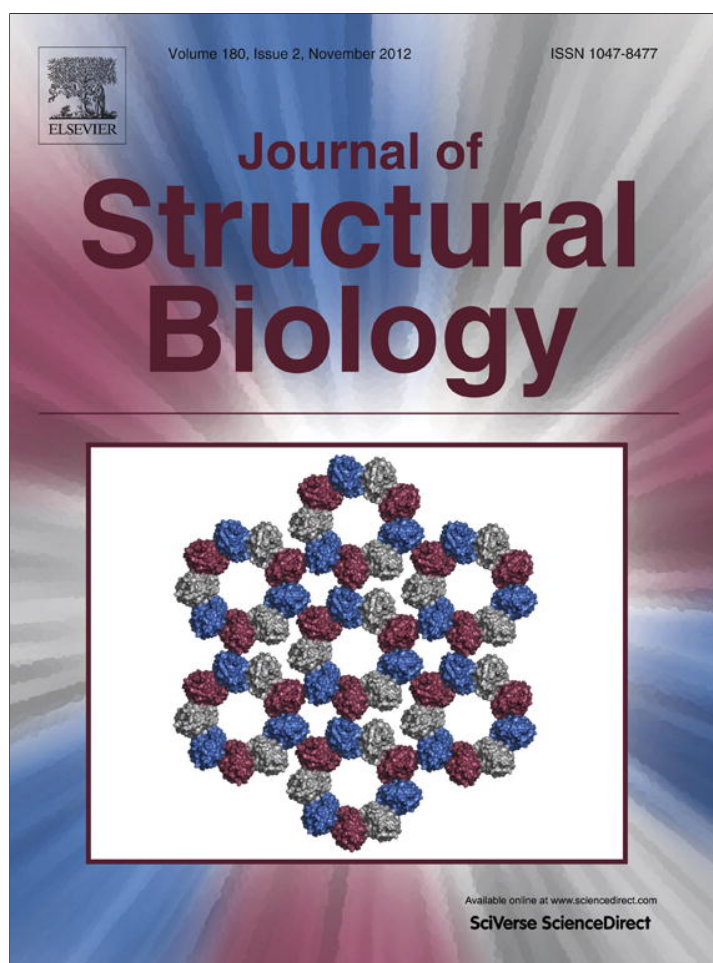




This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Novel structural features of xylanase A1 from *Paenibacillus* sp. JDR-2

Franz J. St. John^{a,b,*}, James F. Preston^c, Edwin Pozharski^a

^a Department of Pharmaceutical Sciences, University of Maryland School of Pharmacy, 20 Penn Street, Baltimore, MD 21201, USA

^b Current address: Forest Products Laboratory, United States Forest Service, United States Department of Agriculture, One Gifford Pinchot Dr. Madison, WI 53726, USA

^c Department of Microbiology and Cell Science, University of Florida, P.O. Box 110700, Gainesville, FL 32611, USA

ARTICLE INFO

Article history:

Received 13 April 2012

Received in revised form 28 August 2012

Accepted 3 September 2012

Available online 18 September 2012

Keywords:

Paenibacillus

Modular xylanase

GH10

Glycosyl hydrolase family 10

GH10 structure

ABSTRACT

The Gram-positive bacterium *Paenibacillus* sp. JDR-2 (*PbJDR2*) has been shown to have novel properties in the utilization of the abundant but chemically complex hemicellulosic sugar glucuronoxylan. Xylanase A1 of *PbJDR2* (*PbXynA1*) has been implicated in an efficient process in which extracellular depolymerization of this polysaccharide is coupled to assimilation and intracellular metabolism. *PbXynA1* is a 154 kDa cell wall anchored multimodular glycosyl hydrolase family 10 (GH10) xylanase. In this work, the 38 kDa catalytic module of *PbXynA1* has been structurally characterized revealing several new features not previously observed in structures of GH10 xylanases. These features are thought to facilitate hydrolysis of highly substituted, chemically complex xylans that may be the form found in close proximity to the cell wall of *PbJDR2*, an organism shown to have a preference for growth on polymeric glucuronoxylan.

Published by Elsevier Inc.

1. Introduction

Glycosyl hydrolase family 10 (GH10) is one of the most populous families in the CAZy database (Cantarel et al., 2009) represented primarily by a single enzymatic activity. These β -1,4-endoxylanases are found in all three domains of life including higher plants and based on database assessment seem to be the more plentiful xylanase when compared to glycosyl hydrolase family 11 xylanases which share all general catalytic features with GH10 xylanases, but consist of an alternative protein fold. Since their target of hydrolysis, xylan, is the second most abundant polysaccharide in lignocellulosic biomass it is expected that these enzymes have a prominent role in the global carbon cycle. As such, GH10 xylanases have been applied extensively in diverse industries and are major components of xylanolytic enzyme systems applied for the efficient degradation of partially disrupted biomass in the bioconversion of lignocellulosics to value-added products and

fuels (Beg et al., 2001; Collins et al., 2005; Preston et al., 2003; Saha, 2003).

GH10 xylanases have the structure of a quintessential $(\beta/\alpha)_8$ TIM barrel fold and catalyze hydrolysis of β -1,4-xylan through a double-displacement mechanism which retains the β -configuration in the newly generated reducing terminal sugar (Henrissat et al., 1995; Koshland, 1953; McCarter and Withers, 1994; Vocadlo et al., 2001). Like other carbohydrate hydrolases classified in Clan A (4/7 hydrolases) of the CAZy database classification these enzymes have two catalytic glutamate residues in a juxtaposed position on β -strands 4 and 7 which are approximately 5.5 Å apart. This spacing is necessary for the double-displacement mechanism and allows positioning of the xylosyl chain as well as a catalytic water molecule. A series of xylosyl-binding subsites are designated with a positive (+) or negative (−) integer increasing from the bond of cleavage, toward the reducing terminal and the non-reducing terminal, often referred to as the aglycone and glycone regions, respectively (Fig. 1a) (Davies et al., 1997). Subsite mapping studies have helped define the substrate binding cleft of these xylanases (Biely et al., 1981a,b, 1983; Charnock et al., 1998; Moreau et al., 1994). While the catalytic cleft of GH10 xylanases is very well conserved, affinity differences between xylose binding subsites, the overall balance of affinity between the glycone and aglycone regions and additional distal xylose binding subsites may all contribute to a significant diversity of function. Rigorous studies of several of these xylanases have detailed their substrate and product preferences and offered comprehensive insight into how GH10 xylanases are finely-tuned for various xylanolytic tasks (Andrews

Abbreviations: GH, glycosyl hydrolase; GH10, glycosyl hydrolase family 10; X₂, xylobiose; X₃, xylotriose; MeGX₃, aldohexuronic acid product of a GH10 xylanase; MeGX₂, aldopenturonic acid; MeGA, 4-O-methyl glucuronic acid; *PbJDR2*, *Paenibacillus* sp. JDR-2; *PbXynA1*, *Paenibacillus* sp. JDR-2 xylanase A1; *PbXynA1CD*, *Paenibacillus* sp. JDR-2 xylanase A1 catalytic module; CBM22, carbohydrate binding module family 22; CBM9, carbohydrate binding module family 9; SLH, surface layer homology; CD, catalytic domain.

* Corresponding author. Fax: +1 608 231 9262.

E-mail addresses: fstjohn@gmail.com (F.J. St. John), jpreston@ufl.edu (J.F. Preston), epozhars@rx.umaryland.edu (E. Pozharski).

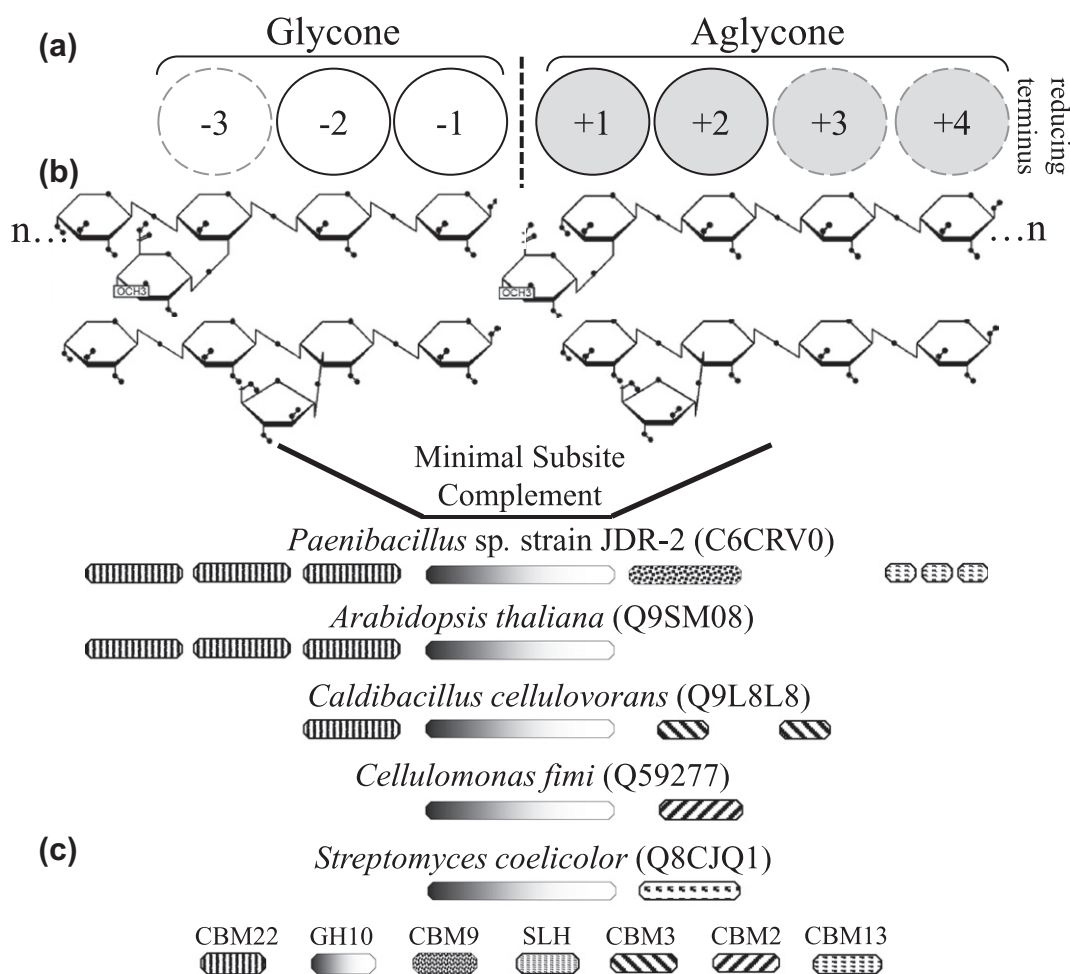


Fig. 1. GH10 xylanase active site cleft xylose binding subsite organization, subsites known to accommodate substitutions along the xylan chain and a small representation of the modular diversity found among GH10 xylanase sequences. (a) By convention, the subsites that have been characterized in GH10 xylanases and how they are organized. Subsite -2 through +2 are considered the minimal set for all endo-acting GH10 xylanases and are shown as solid-line circles. Those subsites that have been identified in only some GH10 xylanases are shown as dashed-line circles. (b) An idealized model of glucuronoxylan (top) and arabinoxylan (bottom) showing the subsite where the α -1,2-linked MeGA and arabinofuranose appendages are known to be accommodated. (c) Arrangement of modules associated with some GH10 xylanases, highlighting the diversity found in this GH family.

et al., 2000; Armand et al., 2001; Charnock et al., 1997; Notenboom et al., 1998; Suzuki et al., 2009). Two major functional categories have emerged which assign these xylanases as having a preference for polymeric xylan or small xylooligosaccharides (Pell et al., 2004a). However, highly conserved structural features in the substrate binding cleft of these enzymes do not always transfer into similar function (Charnock et al., 1997; Moreau et al., 1994), suggesting that the substrate binding cleft is best considered in its entirety rather than as individual subsites.

Studies regarding hydrolysis product profiles are consistent among all GH10 xylanases so far analyzed. Primary neutral sugar products of a limit hydrolysis include xylose, xylobiose (X_2) and xylotriose (X_3). The smallest substituted limit products of glucuronoxylan and arabinoxylan are aldohexuronic acid (MeGX₃) and arabinofuranosylobiose, respectively (Biely et al., 1997; Vardakou et al., 2003, 2005). Both the 4-O-methyl glucuronic acid (MeGA) and arabinofuranose substituted xylooligosaccharides resulting from GH10 limit hydrolysis are known to be the smallest substituted xylooligosaccharides of their type released by the two major xylanase families. Studies of GH10 xylanases have indicated that the glycone region is most important for catalysis with a requirement for xylosyl chain extension into the -2 subsite for catalysis to occur (Kolenova et al., 2006).

Paenibacillus sp. JDR-2 (*PbJDR2*) is a well characterized aggressively xylanolytic Gram-positive bacterium. The glucuronoxylan

degradation and assimilation system employed by *PbJDR2* involves several GH10 xylanases, transporters for X_2 and MeGX₃ and an intracellular α -glucuronidase, GH10 xylanase and β -xylosidase. The presumed roles of these intracellular enzymes are to process the assimilated oligomeric xylanase hydrolysis products (Chow et al., 2007; Nong et al., 2009). Growth studies have shown that this bacterium utilizes polymeric glucuronoxylan preferentially over the simple sugars glucose and xylose and also in preference to the GH10 xylanase limit hydrolysis sugars, X_2 and MeGX₃. These findings and the additional observation that there is no accumulation of GH10 xylanase hydrolysis products in the medium while *PbJDR2* grows on glucuronoxylan supports the hypothesis that *PbJDR2* preferentially utilizes polymeric glucuronoxylan through a mechanism involving vectorial or active-coupling of xylan hydrolysis with assimilation for catabolism (Nong et al., 2009; St. John et al., 2006).

Xylanase A1 of *PbJDR2* (*PbXynA1*) is a 157 kDa multimodular surface-anchored GH10 xylanase thought to be important for this glucuronoxylan utilization process. This enzyme is representative of a group of generally large, multi-modular GH10 xylanases containing a commonly observed domain architecture that consists, in a minimal description, of the domain arrangement: carbohydrate binding module (CBM) 22 (CBM22)-GH10-carbohydrate binding module 9 (CBM9). In its native form, *PbXynA1* contains consecutively arranged modules 3-CBM22/GH10/CBM9/3-surface

layer homology modules (SLH) (Fig. 1c). Modules of CBM22 have been characterized to bind soluble sugars (Najmudin et al., 2010; Xie et al., 2001) and the CBM9 module to the reducing terminus of oligosaccharides and polymers (Notenboom et al., 2001). Additionally the triplicate SLH modules are well characterized and allow surface localization through interactions with peptidoglycan secondary cell wall polysaccharides (Kern et al., 2011; St. John et al., 2006). The complete modular system represents a specialized GH10 xylanase presumably designed to allow localization of substrate near to the cell surface for rapid hydrolysis product assimilation.

This work presents a native and ligand-bound structure of the catalytic domain (CD) of *PbXynA1* (*PbXynA1CD*) from *PbJDR2*. This GH10 xylanase CD structure is unique as a representative of the SLH domain-mediated surface-anchored CBM22/GH10/CBM9 modular architectural motif commonly found in a subset of GH10 xylanases. The results of this study increase our knowledge of the diversity that exists in this large complex xylanase family and identify structural properties that may facilitate cell-surface-localized processing of complex glucuronoxylan by lignocellulose utilizing bacteria.

2. Experimental procedures

2.1. Protein preparation

Cloning *xyna1CD* and expression of the *PbXynA1CD* enzyme were previously described (St. John et al., 2006). Protein was judged by SDS–PAGE to be of sufficient purity for crystallization purposes after nickel IMAC purification. The N-terminal His-tag was removed by bovine thrombin (Calbiochem, La Jolla, CA, Cat. No. 605157) in the recommended buffer and the preparation was dialyzed to remove the cleaved His-tag and in preparation for anion exchange chromatography into 10 mM bis–tris propane, pH 6.5. His-tag free *PbXynA1CD* was resolved with a MonoQ column (GE Healthcare Bio-Sciences Corp, Piscataway, NJ). SDS–PAGE verified that thrombin digestion proceeded to completion. Purified *PbXynA1CD* was exchanged into 10 mM Tris, pH 7.5, concentrated to 12 mg/ml and filter sterilized for storage and crystal screening. The aldouronate, MeGX₃ was prepared from extracted sweetgum (*Liquidambar styraciflua*) wood methylglucuronoxylan digested with *PbXynA1CD* and purified as previously described (St. John et al., 2006).

2.2. Crystallization, data processing and analysis

Protein crystallization was performed by the sitting-drop vapor diffusion method in 24-well Cryschem plates (Charles Supper Company, Natick, MA) or 3-drop, Art Robbins (Sunnyvale, CA) 96-well Intelliplates depending on manual or robotic plate preparation, respectively. Crystals of native *PbXynA1CD* (space group *p2*₁) were obtained from a refinement of Nextal Classic II Suite (Qiagen, Valencia, CA) condition G12 (#84) which consisted of 18% PEG3350, 100 mM HEPES, pH 7.1, and 200 mM MgCl₂. Crystallization drops consisted of 2:1 protein:precipitant ratio. The crystal was cryoprotected in 15% PEG400 and flash cooled in liquid nitrogen. Diffraction data were collected on beam line 7–1 of the Stanford Synchrotron Radiation Lightsource (SSRL) at 105 K. Reflections were integrated and scaled in HKL2000 (Otwinowski and Minor, 1997) and phases determined through molecular replacement with the program Phaser (McCoy et al., 2007) using xylanase A from *Geobacillus stearothermophilus* (PDB code: 1HIZ). The model was refined using iterative rounds of isotropic restrained refinement with the program REFMAC (Murshudov et al., 2011) and real-space refinement with Coot (Emsley et al., 2010), both as part of the CCP4

suite of programs (Winn et al., 2011). The final round of refinement applied anisotropic B-factors which improved the 2Fo – Fc density map and decreased the R-factor difference. The reported maximum resolution for this native structure, 1.40 Å, is the effective resolution (Weiss, 2001) as all data were used in the refinement.

For the *c2* space group native and the MeGX₃-bound *PbXynA1CD* structures, crystallization trays were set using an Oryx Nano crystallization robot (Douglas Instruments Ltd, Berkshire, England) with drop volumes of 1 µl consisting of 50% mother liquor and *PbXynA1CD* containing ~13 mM *PbXynA1CD* generated xylooligosaccharides and aldotetrauronate mixture (an approximation based on the phenol–sulfuric acid total carbohydrate assay) (Dubois et al., 1956) or 3 mM each of X₂ and X₃. Condition 59 of the Nextal Classic Suite (Qiagen) consisting of 100 mM HEPES sodium salt, pH 7.5 and 1.5 M LiSO₄ yielded the MeGX₃-bound *PbXynA1CD* crystal used for this study, while the native crystal (space group *c2*) was obtained in the same condition as the *p2*₁ native crystal form, but for the X₂ and X₃ xylooligosaccharide addition. Crystals were cryoprotected in mother liquor containing 8% glycerol and the respective amounts of sugars used for cocrystallization. Data were collected on beam line 12–2 of the SSRL at 105 K. Reflection data for the MeGX₃-bound structure was processed, the data were phased and the model refined as described above with the *p2*₁ native *PbXynA1CD* structure. The *c2* native structure was processed similarly to the ligand-bound model, but TLS restrained refinement was applied in the final rounds of refinement. Models of *XynA1CD* were analyzed using Molprobity (Davis et al., 2007), HBPlus (McDonald and Thornton, 1994) and Ligplot (Wallace et al., 1995) and figures were prepared in PyMOL (DeLano, 2002). Hydrogen bond predictions are based upon hydrogen position assignments by the program Reduce (Word et al., 1999) as part of Molprobity.

For primary amino acid sequence comparison, the UniProt database (Jain et al., 2009) was used to collect the top 20 *PbXynA1CD* homologs and the CAZy database (Cantarel et al., 2009) was used to source the amino acid sequences of structurally characterized GH10 xylanases. Comparison of the identity levels between *PbXynA1CD* and collected GH10 sequences and modular complement and arrangement was performed using the *bl2seq* BLAST tool (Johnson et al., 2008) and Conserved Domain Database available through the NCBI (Marchler-Bauer et al., 2007), respectively. Sequence alignments were performed using Clustal (Thompson et al., 1994) as part of MEGA 4 (Kumar et al., 2008) and verified with the alignment program MAFFT (Katoh et al., 2002) as well as with 3-D alignment when available. Structures were aligned using the 'align' function in PyMOL (DeLano, 2002) or with SSM Superposition (Krissinel and Henrick, 2004).

2.3. Accession numbers

The protein sequence accession number for *PbXynA1* with all associated modules is UniProt ID: C6CRV0. Protein structure data for *PbXynA1CD* were deposited with the PDB under the accession numbers PDB ID: 3RO8, PDB ID: 4E4P and PDB ID: 3RDK for the *p2*₁ space group native structure, the *c2* space group native structure and the MeGX₃ ligand-bound structure, respectively.

3. Results

3.1. Differences between the native and ligand-bound forms of *PbXynA1CD*

The overall features of the structures presented in this work are typical of GH10 xylanases (Fig. 2a) (Collins et al., 2005). The two native *PbXynA1CD* crystals were of the *p2*₁ and *c2* space groups and the asymmetric unit of the crystals contained eight and two

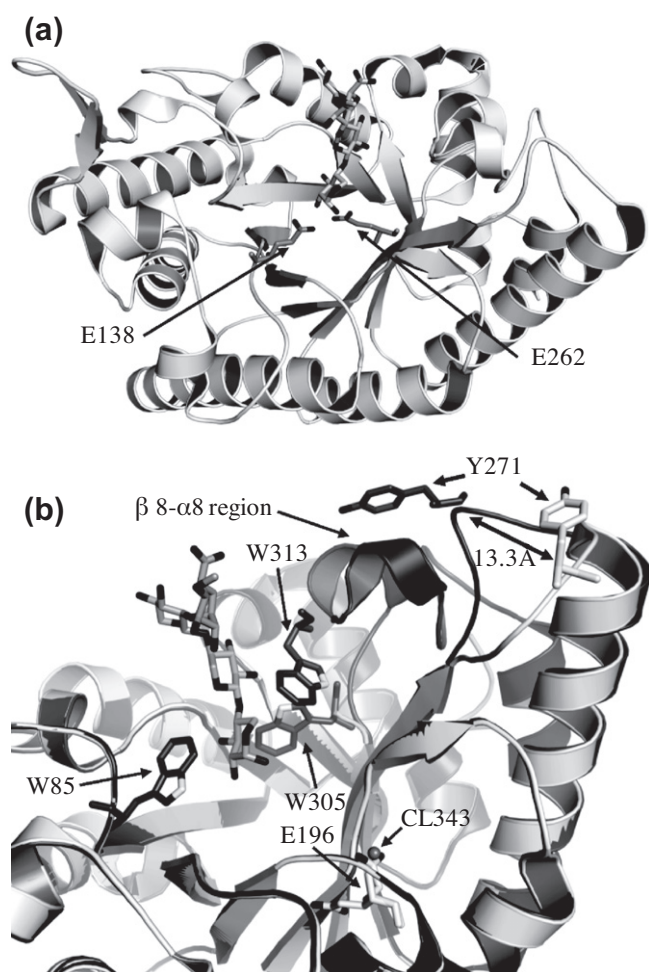


Fig. 2. Comparison of the native and ligand bound structures of *PbXynA1CD*. (a) A cartoon of *PbXynA1CD* showing MeGX₃ coordinated into the glycone region and the open catalytic cleft in the aglycone region separated by the catalytic center indicated by the two juxtaposed catalytic glutamates. (b) A superposition of the native (light gray) and MeGX₃-bound (dark gray) *PbXynA1CD* structures showing the two major differences observed. The disordered β 8-α 8 region of the native structure (missing in image) is thought to allow movement in the β 7-α 7 loop. The amino acid side chain of Tyr271 can be seen shifted 13.3 Å away. In the native structure the amino acid side chain of Glu196 can be observed in two conformations sharing the position as found in the MeGX₃ bound structure with a chloride ion.

chains, respectively, while the MeGX₃-bound *PbXynA1CD* crystals were in the space group *p*4₃2₁ and contained two chains (Table 1). MeGX₃-bound *PbXynA1CD* crystals were obtained from crystallization conditions containing approximately 13 mM *PbXynA1CD*-generated sweetgum (*L. styraciflua*) xylan limit hydrolysate, primarily consisting of X₂ and MeGX₃. Cocrystals containing MeGX₃ were not obtained from the same crystallization condition containing 6 mM purified MeGX₃. The MeGX₃ ligand is coordinated within the glycone region of the catalytic substrate binding cleft of both chains. The 2F_o – F_c map shows clear electron density for the tetrasaccharide MeGX₃ sugar ligand due in part to stabilization of the MeGA substituted xylose and the MeGA appendage each through a single hydrogen bond with the other *PbXynA1CD* chain. These contacts between the MeGX₃ ligand positioned in each of the two chains with the other chain in the unit cell establish a perfect 2-fold non-crystallographic symmetry. This finding is considered an artifact of crystallization rather than having implications for biological assembly because: (1) much of the observed interface is coordinated through the aglycone bound MeGX₃, (2) no previous GH10 xylanase has been reported being multimeric, (3) chromatographic gel filtration sizing studies support *PbXynA1CD* as a monomer in

solution and (4) the native form of XynA1 is a large multimodular enzyme of which the CD is just single modular component. The MeGX₃ ligand only makes direct contacts with *PbXynA1CD* in the –1 and –2 subsite, an observation that is generally consistent in GH10 xylanases although there is an example showing interaction in a –3 subsite (Pell et al., 2004a). Sequence studies suggest that a similar –3 subsite may occur in only a limited number of GH10 xylanases.

In each chain of both native structures, amino acids Gly306 through the C-terminal end, which accounts for the β 8-α 8 loop region and terminal α-helical secondary structure elements, are completely disordered. In contrast, the MeGX₃-bound crystal form yielded clear density in these two consecutive C-terminal motifs. All structures derive from the same *PbXynA1CD* protein preparation with crystal screening for *PbXynA1CD*-ligand cocrystals occurring several months after the original native protein crystallization (space group *p*2₁). From this later screening, the second native structure (space group *c*2) was obtained from a *PbXynA1CD*-xylooligosaccharide cocrystallization. Except for the xylooligosaccharide addition the condition was similar to the original *PbXynA1CD* native crystallization condition. The resulting data set represented a new *PbXynA1CD* crystal form in the space group *c*2 and provided a second unit cell packing arrangement which confirms the original finding concerning the disordered β 8-α 8 loop and α 8 helix motifs of the original native structure. All further discussion regarding native *PbXynA1CD* protein structure will refer to the *p*2₁ crystal form (PDB code: 3RO8) unless otherwise specified. In the region that is missing from the native structures, just following Phe304, GH10 family conserved Trp305 and Trp313 establish primary hydrophobic contact interactions for the xylose coordinated in the –1 subsite most proximal to the catalytic center (Figs. 2b and 3a). Stabilization of this region is achieved upon binding of substrate and therefore this observation offers an explanation for the disorder in the native structure. The well ordered density of the β 8-α 8 loop and α 8 helix motif in the MeGX₃-bound structure also corresponds with well ordered density for the β 7-α 7 loop that leans over and caps the β 8-α 8 loop (Fig. 2b). In the native structure the α 7 helix motif is shifted increasingly from its middle section near Tyr283 to its top where it is about 2.5 Å further inward toward the catalytic center (measured from the Cα of Glu275) (Fig. 2b) and the β 7-α 7 loop diverges from a position where Tyr271 stacks on and likely stabilizes the β 8-α 8 loop to a position 13.3 Å further toward the β 6-α 6 loop. In this position of the native structure, amino acid side chains of the β 7-α 7 loop may interact in the distal aglycone region. Lastly, in the small β 5-α 5 loop of the native *PbXynA1CD* structure, Glu196 is observed having a dual conformation, sharing a position as is observed in the ligand bound form of *PbXynA1CD* with a chloride ion. This arrangement displaces the side chain of Glu196 into a second conformation deeper into the native *PbXynA1CD* core structure (Fig. 2b). Additional differences between these two structures will be presented in more detail below.

3.2. Coordination of MeGX₃ in the glycone subsites

Coordination of MeGX₃ throughout the glycone region largely mirrors what has been described previously (Fujimoto et al., 2004; Lo Leggio et al., 2000; Pell et al., 2004b). As shown in Fig. 3a, the xylose in the –1 subsite is in a ⁴C₁ conformation with the C-1 hydroxyl positioned in its α-anomeric form with the O-1 and O-2 hydroxyls hydrogen bonding with the Oε-1 and Oε-2 of the predicted catalytic nucleophile Glu262. The O-2 oxygen hydrogen bonds with the Nδ of Asn137 and the O-3 hydroxyl of this xylose forms additional hydrogen bonds with the Nε of His81 and the Nζ of Lys48. Predicted hydrogen positions indicate that Nε of His81 is not protonated and rather acts as a hydrogen acceptor for the –1

Table 1

Data collection statistics and final model quality.

	Data sets		
<i>Identification</i>			
Crystal type	XynA1CD	XynA1CD	XynA1CD-MeGX ₃
PDB code	3RO8	4E4P	3RDK
<i>Data collection</i>			
Wavelength	0.980	0.979	0.979
Space group	<i>p</i> 2 ₁	<i>c</i> 2	<i>p</i> 4 ₃ 2 ₁ 2
<i>Unit-cell</i>			
<i>a</i> (Å)	82.07	179.52	165.11
<i>b</i> (Å)	93.17	58.53	165.11
<i>c</i> (Å)	182.95	65.02	66.36
β	99.96	108.81	
Resolution range (Å)	50.0–1.34	50.0–1.93	50.0–1.49
Redundancy	4.3 (2.4)	7.4 (7.0)	8.8 (8.4)
No. of unique reflections	523629 (25295)	48561 (4845)	148538 (14749)
<i>R</i> _{merge} (%)	7.7	10.9	4.7
Completeness (%)	87 (42) ^a	99.9 (100)	100 (100)
Mean $ I /\sigma_I$	15.1 (1.1)	15.6 (1.0)	33.7 (1.0)
Matthews coefficient	2.27	2.13	2.94
Corresponding solvent (%)	45.8	42.2	58.3
Molecules of XynC per asymmetric unit	8	2	2
<i>Refinement statistics</i>			
Resolution range	42.6–1.40 ^a	32.5–1.92	42.4–1.49
No. of reflections	496837	45941	140983
<i>R</i> _{work} / <i>R</i> _{free} (%)	14.2/18.5	25.3/29.5	14.2/18.1
No. of protein atoms	23733	4644	6364
No. of water molecules	2245	225	720
No. of ligand atoms	N/A	N/A	82
<i>Average B factors</i> (Å ²)			
Protein	21.2	45.6	29.2
Waters	37.2	23.2	42.0
Ligands	N/A	N/A	33.3
Solutes	21.9	37.2	39.0
<i>Ramachandran statistics</i>			
Preferred/allowed (%)	100	100	99.7%
Outliers (%)	0	0	0.3%
<i>RMSD from ideality</i>			
Bond lengths (Å)	0.027	0.024	0.029
Bond angles (°)	2.18	1.887	2.29

Values in parenthesis are for the highest resolution shell.

 $R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, $R_{\text{work, free}} = \sum_{hkl} |F_o(hkl) - F_c(hkl)| / \sum_{hkl} |F_o(hkl)|$.^a High resolution limit defined as the “effective resolution” (Weiss, 2001).

subsite xylose O-3 hydroxyl hydrogen. The hydroxyl oxygen in this position would then accept a hydrogen from N ζ of Lys48. Just opposite from this, on the other side of the cleft bordering the –2 subsite, the N ϵ of Trp305 hydrogen bonds with the O-2 hydroxyl oxygen of the –2 subsite xylose. Both Lys48 and Trp305 may potentially interact in an alternative hydrogen bonding configuration. In both cases, the geometry of the alternative hydrogen bond is similar to the primary hydrogen bond described above, but the hydrogen to acceptor distance is greater indicating a less likely interaction. Lastly, the N δ of Asn45 establishes a hydrogen bond with the O-3 hydroxyl oxygen of the xylose in the –2 subsite. Primary hydrophobic contacts occur between the xylose of the –1 subsite and Trp85, Trp305 and Trp313. Trp313 forms a top-side shelf perpendicular to the xylosyl ring and Trp85 is positioned similarly above the xylose, but more in-plane with the ring (Fig. 3a). No hydrophobic contacts were identified for the xylose in the –2 subsite.

Two other ligand bound GH10 structures have been reported having the α -anomer of xylose in the –1 subsite. In the case of *Cellvibrio mixtus* Xyn10B (CmXyn10B) (PDB code: 1UQZ) (Pell et al., 2004b), a catalytic nucleophile mutation to a serine prevents meaningful comparison. The glycone region of native *Thermoascus aurantiacus* xylanase A (TaXyn10A) (PDB code: 1GOR) contains X₂ with the –1 subsite xylose in an α -anomeric conformation and

similar coordination position as observed for MeGX₃-bound PbXynA1CD (Lo Leggio et al., 2001).

For the TaXyn10A structure it was concluded that if the –1 subsite xylose were in the β -configuration, a steric clash would likely induce ring strain, with implications for catalysis. The authors further suggested that the α -configuration as observed in TaXyn10A and now PbXynA1CD, is due to the enzyme selecting the anomeric configuration that most resembles the covalent intermediate step of the double-displacement mechanism (Lo Leggio et al., 2001). Our structure analysis of PbXynA1CD supports these earlier findings and their conclusions. Even in consideration of this similarity, the low level of primary amino acid sequence identity of 36% as well as other major categorical differences such as modularity (TaXyn10A consists of a CD only), thermophilicity and microbial origin strongly suggest that these two enzymes are diverse representatives of the GH10 family of xylanases.

3.3. The –2 subsite pocket

In the great majority of GH10 xylanase amino acid sequences (est. 85%) a glutamate extends under the –2 subsite and establishes a stabilizing hydrogen bond with the O-2 hydroxyl of this xylose. In the remaining minority the glutamate is substituted with a glycine which is equivalent to Gly44 in PbXynA1CD. This gluta-

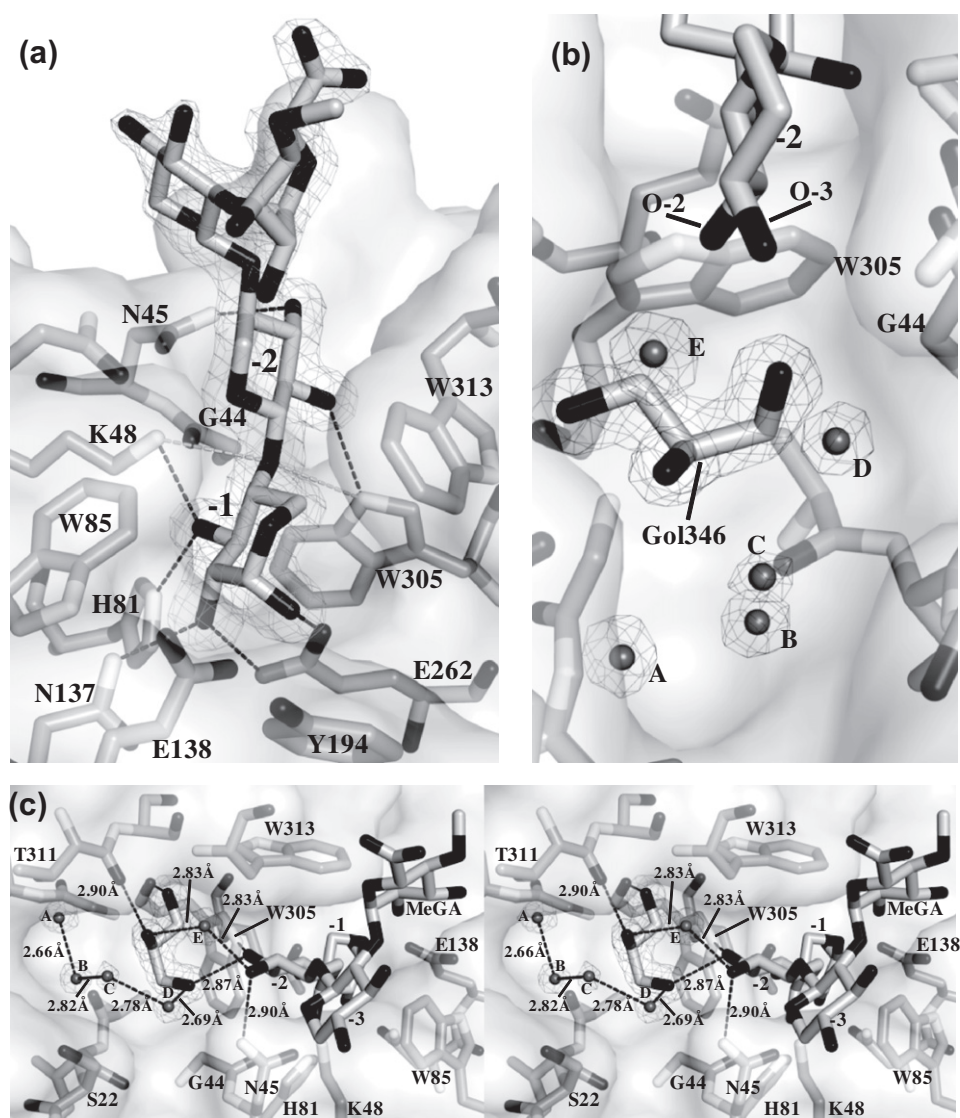


Fig. 3. Coordination of MeGX₃ in the glycone region of the substrate binding cleft of PbXynA1CD. (a) Primary (grey) and secondary (light grey) hydrogen bond interactions between the glycone region and the first two xylose residues of the MeGX₃ ligand viewed from the aglycone region. (b) An approximate 180° rotation from (a) views the MeGX₃ occupied cleft of the glycone region from the opposite direction showing the underside of the extended xylosyl chain in the -2 subsite and the -2 subsite pocket filled with the glycerol Gol346 and five waters. In most GH10 xylanases, a glutamate is in the equivalent position of Gly44 which effectively blocks the formation of the pocket. (c) In cross-eyed stereo, an approximate 90° rotation from (b) showing hydrogen bonding interactions between the -2 subsite pocket glycerol, waters and with PbXynA1CD. The -2 subsite pocket waters have been designated A–E with (A) O474; (B) O538; (C) O704; (D) O547; (E) O467. Atom and molecule numbering derive from the MeGX₃ bound model (PDB code: 3RDK).

mate to glycine (Glu–Gly) amino acid switch, first described in xylanase 10C of *Cellvibrio* (CjXyn10C, PDB code: 1US3), has been shown to result in a decreased xylose binding affinity in this subsite (Charnock et al., 1997, 1998), but was speculated to permit accommodation of an O-2 substituted acetyl moiety on this xylose and potentially allow CjXyn10C to more efficiently degrade acetylated xylans of hardwoods (Pell et al., 2004a). Substitutions on the O-3 hydroxyl are not sterically restrained since this position lies slightly above the plane of the surrounding protein surface. This pocket in CjXyn10C is considerably narrow and deep primarily due to a larger β 1– α 1 loop region and larger amino acids that extend toward the pocket from the β 2– α 2, β 3– α 3 and β 8– α 8 loops when compared to PbXynA1CD. In PbXynA1CD, this same Glu–Gly switch gives the O-2 hydroxyl of the -2 subsite xylose access to a much larger, open pocket. In contrast to that observed in CjXyn10C the β – α loop regions most proximal to the -2 subsite pocket are minimal with two small amino acids, S22 and T311

(Table S1), allowing the pocket to open lengthwise in the direction of the extending xylosyl chain (Fig. 3b and c). Throughout the GH10 family, in sequences containing the Glu–Gly switch, these two positions are not conserved and most often consist of a variety of amino acids with larger side chains or more developed, -2 subsite-proximal, β – α loops that limit the size of the -2 subsite pocket. The CjXyn10C is an example of this later type of GH10 xylanase. The specific set of amino acids that contribute to the pocket as identified in PbXynA1CD are all associated with either the genus *Paenibacillus* or *Clostridium* and are all predicted to be multimodular enzymes of the CBM22/GH10/CBM9 domain arrangement and also include SLH modules for cell surface localization (Table S1). However, this modular architecture is not exclusive to enzymes having a -2 subsite pocket as now defined for PbXynA1CD.

As depicted in Fig. 3b and c, a glycerol (Gol346) and five waters occupy the -2 subsite pocket in the MeGX₃-bound structure. The glycerol was likely derived from cryoprotection, but with the

MeGX₃ ligand coordinated in the glycone region the glycerol makes several stabilizing contacts via its three hydroxyl groups. The hydrogen from the O-1 hydroxyl of Gl346 makes a hydrogen bond with the main-chain carbonyl oxygen of Trp305. Similarly, the hydrogen of the O-2 hydroxyl hydrogen bonds with the main-chain carbonyl of Thr311. The O-2 of this hydroxyl position is predicted to accept the hydrogen from a coordinated water (O467) which, in turn, accepts the hydrogen from the O-3 hydroxyl of the xylose in the –2 subsite. More directly, the O-3 hydroxyl oxygen of Gl346 accepts a hydrogen from the O-2 hydroxyl of the xylose in the –2 subsite and additionally makes a hydrogen bond with O547. Three additional waters which fill the –2 subsite pocket interconnect through a hydrogen bond network (Fig. 3c).

Based on the hydrolysis action pattern of GH10 xylanases including *PbXynA1CD* (Biely et al., 1997; St. John et al., 2006), the smallest aldouronic acid sugar produced by these enzymes, MeGX₃, results because the O-2 hydroxyl position of the xylose in the +1 and –3 subsite positions face outward away from the enzyme so that MeGA interacts either minimally, as for the +1 subsite xylose or not at all, as for the –3 subsite xylose (Fig. 1b) (Fujimoto et al., 2004; Pell et al., 2004b). If the –2 subsite pocket of *PbXynA1* had specificity for MeGA then aldatriuronic acid (MeGX₂) would be a major hydrolytic limit product. Although biochemically, this has not been found to occur (St. John et al., 2006), Fig. S1 shows the potential of the *PbXynA1CD* –2 subsite pocket to accept a sugar the size of the MeGA which appears to fit nicely with only little manually adjusted accommodation.

3.4. The aglycone region of *PbXynA1*

During the course of *PbXynA1CD* crystallization studies no crystals were obtained containing any ligand within the aglycone region. Inspection of unit cell packing and symmetry contacts did not support ligand exclusion from this region in the MeGX₃-bound or native structures of *PbXynA1CD*. From previous structures of GH10 xylanases, MeGX₃ and various xylooligosaccharides may coordinate into this region. Based on the highly similar catalytic substrate xylan binding cleft geometry between GH10 homologs, potential contacts were analyzed by superposition of the MeGX₃ bound *CmXyn10B* xylanase (Pell et al., 2004b). These two structures aligned with a percentile-based spread (Pozharski, 2010) of just 1.17 Å. The xylotriose in the glycone region of *CmXyn10B* aligned closely with the xylotriose portion of the MeGX₃ in *PbXynA1CD* and the conserved amino acid side chains in the aglycone region primarily involved with ligand coordination also align (Fig. 4a).

Potential ligand contacts in the aglycone region, as in the glycone region, are generally well conserved among GH10 xylanases (Fig. 4b). Two contacts not observed in the +1 subsite of previous GH10 xylanase structures include a hydrogen bond between the Nε of Gln87 and the O-3 hydroxyl of xylose and with the C-6 carboxylate of the MeGA substituted on the +1 subsite xylose. All other hydrogen bonding contacts as presented in Fig. 4b have been previously described for other GH10 xylanases with aglycone coordinated sugars (Fujimoto et al., 2004; Pell et al., 2004b; Zolotnitsky et al., 2004). The aglycone region of *PbXynA1* appears minimal with xylosyl chain binding sites obvious only through the +2 subsite. This is in contrast to other GH10 xylanases which have developed distal aglycone xylose chain coordination often mediated through stacking interactions. Interestingly, in the native structure His234 was modeled in a dual conformation with the displaced conformer flipped with its Nδ extended toward the +1 subsite within hydrogen bonding distance to the xylose endocyclic oxygen and O-3 hydroxyl of the +2 subsite xylose (Fig. 4b), an observation that is verified in the c2 symmetry native structure. This finding

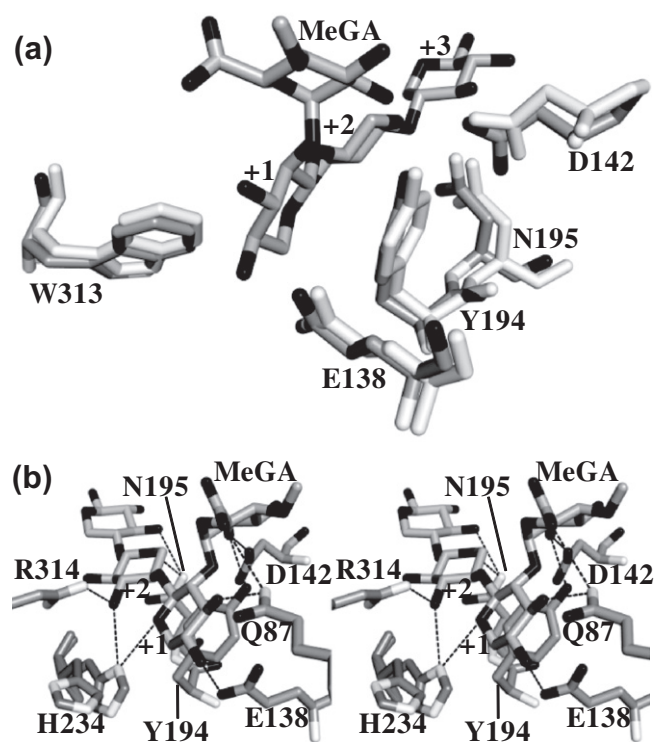


Fig. 4. MeGX₃ positioned in the aglycone region of *PbXynA1CD* (grey) through superposition with MeGX₃ bound *CmXyn10B* (white). (a) Side chains of conserved amino acids showing the generally similar positioning which orients the ligand within the aglycone clefts of both enzymes. (b) Cross-eyed stereo view of previously characterized aglycone region contacts as well as the newly predicted contacts identified through this analysis involving *PbXynA1CD* Gln87 and His234. All amino acids are conserved except for Gln87 which is only found in *PbXynA1* and similarly modular homologs (Table S1).

may have implications for substrate chain recruitment into the aglycone region.

3.5. A disordered native structure

A GH10 xylanase catalytic domain having a C-terminal region that is too disordered to be modeled has not been reported. It should be considered that this finding from the *PbXynA1CD* native structures may be an artifact of the separately crystallized CD and not a structural feature that is necessarily realized in the multi-modular *PbXynA1* enzyme. However, above average B-factors for several glycone region xylan coordinating amino acid side chains have been found in native GH10 xylanase structures (Lo Leggio et al., 2001). Two of the equivalent amino acids in native *PbXynA1CD*, Trp313 and Arg314 are completely disordered, along with the entire last two C-terminal motifs, and apparently reestablish structure upon ligand binding. The third residue in *TaXyn10A* was Glu46. This amino acid position is equivalent to Gly44 in *PbXynA1CD*, and is the residue primarily involved in creation of a –2 subsite pocket as discussed above. This Glu-Gly switch as characterized in *PbXynA1CD*, may be a central reason why the β8–α8 loop and remaining C-terminal region are disordered. This highly conserved glutamate residue, can be observed in numerous native GH10 xylanase structures (e.g., PDB code: 1I1W) stretching across to the other side of the catalytic cleft and engaging in a hydrogen bonding network with other amino acids or more commonly with coordinated waters. This coordinated water interaction may play an important role in stabilization of the β8–α8 region.

3.6. The multi-module GH10 xylanase system

The structure of PbXynA1CD represents the first GH10 xylanase CD having, in its native form, this extent of modular architecture. It is unknown how these modules might work together to implement the specific strategy of xylan utilization employed by PbJDR2. From the predicted function of the complement modules it may be considered that PbXynA1 (St. John et al., 2006) is positioned on the cell surface where the enzyme recruits soluble xylan substrate, cleaves the chain and deposits the products of hydrolysis near to the bacterial cell surface for assimilation. To some degree this presumption is correct. However, this evaluation does not offer a rationale for the other new structural aspects localized to the catalytic module of this enzyme nor does the relatively routine biochemical assessment that has been performed in previous work (Nong et al., 2009; St. John et al., 2006).

The structural observations that are new include the disordered C-terminal region of the native structure, the –2 subsite pocket that is open and potentially able to accommodate sugars appended through the O-2 hydroxyl of the –2 subsite xylose and what appears to be a greater potential for xylose and MeGA coordination in the +1 aglycone subsite. Several other findings may work in concert with these major differences to effectively hydrolyze highly substituted xylan chains.

In a recent report, xylanase 10B of *Clostridium thermocellum* was crystallized with its N-terminal CBM22 module, showing distinctive contacts and xylan binding subsite cleft alignment between the two domains. This 'bi-molecular' structure supports the theory that such domains may work in concert rather than through a simple proximity mediated benefit (Najmudin et al., 2010). It is possible that the structural observations regarding PbXynA1CD can only be explained with consideration of the two flanking CBM modules. The role of the –2 subsite pocket is a particular curiosity. While Figs. S1a and S1b suggest that accommodation of sugars appears possible the simple biochemical analysis suggests that this does not occur during limit hydrolysis of glucuronoxylan. If so, MeGA would only enter the pocket when conditions exist that force the substituted xylose destined for binding in the –2 subsite to linger longer than normal. Such a scenario may result dependent upon the extensive complementary modular architecture of this enzyme. The N-terminal CBM22 modules attract soluble xylans toward the cell-surface localized PbXynA1 enzyme, polymeric xylan interacts with the GH10 catalytic module, but fails to hydrolyze due to the complexity of the heavily substituted native substrate, the C-terminal localized CBM9 module which may bind the reducing end of soluble xylans then associates with the reducing terminus. These two CBM interactions may effectively straddle the carbohydrate chain through PbXynA1CD allowing the time necessary for the sugar to enter the pocket and hydrolysis to occur.

Acknowledgments

We would like to thank the University of Maryland X-ray Crystallography Shared Service for crystallographic resources. This project was supported in part by Biomass Research & Development Initiative Competitive Grant No. 2011-10006-30358 from the USDA National Institute of Food and Agriculture. Professional support for this work was provided by Diane Dietrich at the Forest Products Laboratory and John D. Rice at the University of Florida, Department of Microbiology and Cell Science. Portions of this research were carried out at the SSRL, a national user facility operated by Stanford University on behalf of the US Department of Energy, Office of Basic Energy Sciences. The SSRL Structural Molecular Biology Program is supported by the Department of Energy, Office of Biological and Environmental Research, and by the Na-

tional Institutes of Health, National Center for Research Resources, Biomedical Technology Program and the National Institute of General Medical Sciences.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.jsb.2012.09.007>.

References

- Andrews, S.R., Charnock, S.J., Lakey, J.H., Davies, G.J., Claeysens, M., et al., 2000. Substrate specificity in glycoside hydrolase family 10. Tyrosine 87 and leucine 314 play a pivotal role in discriminating between glucose and xylose binding in the proximal active site of *Pseudomonas cellulosa* xylanase 10A. *J. Biol. Chem.* 275, 23027–23033.
- Armand, S., Andrews, S.R., Charnock, S.J., Gilbert, H.J., 2001. Influence of the aglycone region of the substrate binding cleft of *Pseudomonas xylanase* 10A on catalysis. *Biochemistry* 40, 7404–7409.
- Beg, Q.K., Kapoor, M., Mahajan, L., Hoondal, G.S., 2001. Microbial xylanases and their industrial applications: a review. *Appl. Microbiol. Biotechnol.* 56, 326–338.
- Biely, P., Kratky, Z., Vrsanska, M., 1981a. Substrate-binding site of endo-1,4-beta-xylanase of the yeast *Cryptococcus albidus*. *Eur. J. Biochem.* 119, 559–564.
- Biely, P., Vrsanska, M., Kratky, Z., 1981b. Mechanisms of substrate digestion by endo-1,4-beta-xylanase of *Cryptococcus albidus*. Lysozyme-type pattern of action. *Eur. J. Biochem.* 119, 565–571.
- Biely, P., Vrsanska, M., Gorbacheva, I.V., 1983. The active site of an acidic endo-1,4-beta-xylanase of *Aspergillus niger*. *Biochim. Biophys. Acta* 743, 155–161.
- Biely, P., Vrsanska, M., Tenkanen, M., Kluepfel, D., 1997. Endo-beta-1,4-xylanase families: differences in catalytic properties. *J. Biotechnol.* 57, 151–166.
- Cantarel, B.L., Coutinho, P.M., Rancurel, C., Bernard, T., Lombard, V., et al., 2009. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. *Nucleic Acids Res.* 37, D233–D238.
- Charnock, S.J., Lakey, J.H., Virden, R., Hughes, N., Sinnott, M.L., et al., 1997. Key residues in subsite F play a critical role in the activity of *Pseudomonas fluorescens* subspecies *cellulosa* xylanase A against xylooligosaccharides but not against highly polymeric substrates such as xylan. *J. Biol. Chem.* 272, 2942–2951.
- Charnock, S.J., Spurway, T.D., Xie, H., Beylot, M.H., Virden, R., et al., 1998. The topology of the substrate binding clefts of glycosyl hydrolase family 10 xylanases are not conserved. *J. Biol. Chem.* 273, 32187–32199.
- Chow, V., Nong, G., Preston, J.F., 2007. Structure, function, and regulation of the alduronate utilization gene cluster from *Paenibacillus* sp. strain JDR-2. *J. Bacteriol.* 189, 8863–8870.
- Collins, T., Gerday, C., Feller, G., 2005. Xylanases, xylanase families and extremophilic xylanases. *FEMS Microbiol. Rev.* 29, 3–23.
- Davies, G.J., Wilson, K.S., Henrissat, B., 1997. Nomenclature for sugar-binding subsites in glycosyl hydrolases. *Biochem. J.* 321 (Pt 2), 557–559.
- Davis, I.W., Leaver-Fay, A., Chen, V.B., Block, J.N., Kapral, G.J., et al., 2007. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* 35, W375–W383.
- DeLano, W.L., 2002. The PyMOL User's Manual. DeLano Scientific, Palo Alto.
- Dubois, M., Gilles, K.A., Hamilton, J.K., Rebers, P.A., Smith, F., 1956. Colorimetric method for the determination of sugars and related substances. *Anal. Chem.* 28, 350–356.
- Emsley, P., Lohkamp, B., Scott, W.G., Cowtan, K., 2010. Features and development of coot. *Acta Crystallogr. D Biol. Crystallogr.* 66, 486–501.
- Fujimoto, Z., Kaneko, S., Kuno, A., Kobayashi, H., Kusakabe, I., et al., 2004. Crystal structures of decorated xylooligosaccharides bound to a family 10 xylanase from *Streptomyces olivaceoviridis* E-86. *J. Biol. Chem.* 279, 9606–9614.
- Henrissat, B., Callebaut, I., Fabrega, S., Lehn, P., Mornon, J.P., et al., 1995. Conserved catalytic machinery and the prediction of a common fold for several families of glycosyl hydrolases. *Proc. Natl. Acad. Sci. USA* 92, 7090–7094.
- Jain, E., Bairoch, A., Duvaud, S., Phan, L., Redaschi, N., et al., 2009. Infrastructure for the life sciences: design and implementation of the UniProt website. *BMC Bioinform.* 10, 136.
- Johnson, M., Zaretskaya, I., Raytselis, Y., Merezuk, Y., McGinnis, S., et al., 2008. NCBI BLAST: a better web interface. *Nucleic Acids Res.* 36, W5–W9.
- Katoh, K., Misawa, K., Kuma, K., Miyata, T., 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* 30, 3059–3066.
- Kern, J., Wilton, R., Zhang, R., Binkowski, T.A., Joachimiak, A., et al., 2011. Structure of surface layer homology (SLH) domains from *Bacillus anthracis* surface array protein. *J. Biol. Chem.* 286, 26042–26049.
- Kolenova, K., Vrsanska, M., Biely, P., 2006. Mode of action of endo-beta-1,4-xylanases of families 10 and 11 on acidic xylooligosaccharides. *J. Biotechnol.* 121, 338–345.
- Koshland, D.E., 1953. Stereochemistry and the mechanism of enzymatic reactions. *Biol. Rev.* 28, 416–436.

- Krissinel, E., Henrick, K., 2004. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr. D Biol. Crystallogr.* 60, 2256–2268.
- Kumar, S., Nei, M., Dudley, J., Tamura, K., 2008. MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. *Brief. Bioinform.* 9, 299–306.
- Lo Leggio, L., Jenkins, J., Harris, G.W., Pickersgill, R.W., 2000. X-ray crystallographic study of xylopentaose binding to *Pseudomonas fluorescens* xylanase A. *Proteins* 41, 362–373.
- Lo Leggio, L., Kalogiannis, S., Eckert, K., Teixeira, S.C., Bhat, M.K., et al., 2001. Substrate specificity and subsite mobility in *Thermoascus aurantiacus* xylanase 10A. *FEBS Lett.* 509, 303–308.
- Marchler-Bauer, A., Anderson, J.B., Derbyshire, M.K., DeWeese-Scott, C., Gonzales, N.R., et al., 2007. CDD: a conserved domain database for interactive domain family analysis. *Nucleic Acids Res.* 35, D237–D240.
- Mccarter, J.D., Withers, S.G., 1994. Mechanisms of enzymatic glycoside hydrolysis. *Curr. Opin. Struct. Biol.* 4, 885–892.
- Mccoey, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., et al., 2007. Phaser crystallographic software. *J. Appl. Cryst.* 40, 658–674.
- Mcdonald, I.K., Thornton, J.M., 1994. Satisfying hydrogen-bonding potential in proteins. *J. Mol. Biol.* 238, 777–793.
- Moreau, A., Shareck, F., Kluepfel, D., Morosoli, R., 1994. Alteration of the cleavage mode and of the transglycosylation reactions of the xylanase A of *Streptomyces lividans* 1326 by site-directed mutagenesis of the Asn173 residue. *Eur. J. Biochem.* 219, 261–266.
- Murshudov, G.N., Skubak, P., Lebedev, A.A., Pannu, N.S., Steiner, R.A., et al., 2011. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* 67, 355–367.
- Najmudin, S., Pinheiro, B.A., Prates, J.A., Gilbert, H.J., Romao, M.J., et al., 2010. Putting an N-terminal end to the *Clostridium thermocellum* xylanase Xyn10B story: crystal structure of the CBM22-1-GH10 modules complexed with xylohexaose. *J. Struct. Biol.* 172, 353–362.
- Nong, G., Rice, J.D., Chow, V., Preston, J.F., 2009. Aldouronate utilization in *Paenibacillus* sp. strain JDR-2: physiological and enzymatic evidence for coupling of extracellular depolymerization and intracellular metabolism. *Appl. Environ. Microbiol.* 75, 4410–4418.
- Notenboom, V., Boraston, A.B., Kilburn, D.G., Rose, D.R., 2001. Crystal structures of the family 9 carbohydrate-binding module from *Thermotoga maritima* xylanase 10A in native and ligand-bound forms. *Biochemistry* 40, 6248–6256.
- Notenboom, V., Birsan, C., Warren, R.A., Withers, S.G., Rose, D.R., 1998. Exploring the cellulose/xylan specificity of the beta-1,4-glycanase cex from *Cellulomonas fimi* through crystallography and mutation. *Biochemistry* 37, 4751–4758.
- Otwinowski, Z., Minor, W., 1997. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* 276, 307–326.
- Pell, G., Szabo, L., Charnock, S.J., Xie, H., Gloster, T.M., et al., 2004a. Structural and biochemical analysis of *Cellvibrio japonicus* xylanase 10C: how variation in substrate-binding cleft influences the catalytic profile of family GH-10 xylanases. *J. Biol. Chem.* 279, 11777–11788.
- Pell, G., Taylor, E.J., Gloster, T.M., Turkenburg, J.P., Fontes, C.M., et al., 2004b. The mechanisms by which family 10 glycoside hydrolases bind decorated substrates. *J. Biol. Chem.* 279, 9597–9605.
- Pozharski, E., 2010. Percentile-based spread: a more accurate way to compare crystallographic models. *Acta Crystallogr. Sect. D: Biol. Crystallogr.* 66, 970–978.
- Preston, J.F., Hurlbert, J.C., Rice, J.D., Ragunathan, A., St. John, F.J., 2003. Microbial strategies for the depolymerization of glucuronoxylan: leads to biotechnological applications of endoxylanases. *Applications of Enzymes to Lignocellulosics*. American Chemical Society, Washington, DC, p. 191–210.
- Saha, B.C., 2003. Hemicellulose bioconversion. *J. Ind. Microbiol. Biotechnol.* 30, 279–291.
- St. John, F.J., Rice, J.D., Preston, J.F., 2006. *Paenibacillus* sp. strain JDR-2 and XynA1: a novel system for methylglucuronoxylan utilization. *Appl. Environ. Microbiol.* 72, 1496–1506.
- Suzuki, R., Fujimoto, Z., Ito, S., Kawahara, S., Kaneko, S., et al., 2009. Crystallographic snapshots of an entire reaction cycle for a retaining xylanase from *Streptomyces olivaceoviridis* E-86. *J. Biochem.* 146, 61–70.
- Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22, 4673–4680.
- Vardakou, M., Katapodis, P., Samiotaki, M., Kekos, D., Panayotou, G., et al., 2003. Mode of action of family 10 and 11 endoxylanases on water-unextractable arabinoxylan. *Int. J. Biol. Macromol.* 33, 129–134.
- Vardakou, M., Flint, J., Christakopoulos, P., Lewis, R.J., Gilbert, H.J., et al., 2005. A family 10 *Thermoascus aurantiacus* xylanase utilizes arabinose decorations of xylan as significant substrate specificity determinants. *J. Mol. Biol.* 352, 1060–1067.
- Vocadlo, D.J., Davies, G.J., Laine, R., Withers, S.G., 2001. Catalysis by hen egg-white lysozyme proceeds via a covalent intermediate. *Nature* 412, 835–838.
- Wallace, A.C., Laskowski, R.A., Thornton, J.M., 1995. Ligplot – a program to generate schematic diagrams of protein ligand interactions. *Protein Eng.* 8, 127–134.
- Weiss, M.S., 2001. Global indicators of X-ray data quality. *J. Appl. Cryst.* 34, 130–135.
- Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., et al., 2011. Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* 67, 235–242.
- Word, J.M., Lovell, S.C., Richardson, J.S., Richardson, D.C., 1999. Asparagine and glutamine: using hydrogen atom contacts in the choice of side-chain amide orientation. *J. Mol. Biol.* 285, 1735–1747.
- Xie, H., Gilbert, H.J., Charnock, S.J., Davies, G.J., Williamson, M.P., et al., 2001. *Clostridium thermocellum* Xyn10B carbohydrate-binding module 22-2: the role of conserved amino acids in ligand binding. *Biochemistry* 40, 9167–9176.
- Zolotnitsky, G., Cogan, U., Adir, N., Solomon, V., Shoham, G., et al., 2004. Mapping glycoside hydrolase substrate subsites by isothermal titration calorimetry. *Proc. Natl. Acad. Sci. USA* 101, 11275–11280.