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

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

\textsuperscript{a} Department of Pharmaceutical Sciences, University of Maryland School of Pharmacy, 20 Penn Street, Baltimore, MD 21201, USA
\textsuperscript{b} Current address: Forest Products Laboratory, United States Forest Service, United Stated Department of Agriculture, One Gifford Pinchot Dr. Madison, WI 53726, USA
\textsuperscript{c} Department of Microbiology and Cell Science, University of Florida, P.O. Box 110700, Gainesville, FL 32611, USA

\textbf{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.

\section*{1. Introduction}

Glycosyl hydrolase family 10 (GH10) is one of the most populorous families in the CAZy database (Cantarel et al., 2009) represented primarily by a single enzymatic activity. These \( \beta\)-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 (\(\beta/\alpha\))\textsubscript{8} TIM barrel fold and catalyze hydrolysis of \(\beta\)-1,4-xylan through a double-displacement mechanism which retains the \(\beta\)-configuration in the newly generated reducing terminal sugar (Henrisat et al., 1995; Koshland, 1953; Mccarter and Withers, 1994; Vocaldo 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 \(\beta\)-strands 4 and 7 which are approximately 5.5 \(\AA\) 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 xylene binding substrates, the overall balance of affinity between the glycone and aglycone regions and additional distal xylose binding substrates 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 et al., 2001; Henrisat et al., 1995; Koshland, 1953; Mccarter and Withers, 1994; Vocaldo 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 \(\beta\)-strands 4 and 7 which are approximately 5.5 \(\AA\) 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 xylene binding substrates, the overall balance of affinity between the glycone and aglycone regions and additional distal xylose binding substrates 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 et al., 2001; Henrisat et al., 1995; Koshland, 1953; Mccarter and Withers, 1994; Vocaldo et al., 2001).
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 (X2) and xylotriose (X3). The smallest substituted limit products of glucuronoxylan and arabinoxylan are aldotetrauronic acid (MeGX 3) and arbinofuranoxylobiose, 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 (*Pb*JDR2) is a well characterized aggressively xylanolytic Gram-positive bacterium. The glucuronoxylan degradation and assimilation system employed by *Pb*JDR2 involves several GH10 xylanases, transporters for X2 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, X2 and MeGX3. These findings and the additional observation that there is no accumulation of GH10 xylanase hydrolysis products in the medium while *Pb*JDR2 grows on glucuronoxylan supports the hypothesis that *Pb*JDR2 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 *Pb*JDR2 (*Pb*XynA1) 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, *Pb*XynA1 contains consecutively arranged modules 3-CBM22/GH10/CBM9/3-surface

---

**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.
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 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 to 12 mg/ml and filter sterilized for storage and crystal screening. PbXynA1CD was prepared as previously described (St. John et al., 2006). Protein 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 to 12 mg/ml and filter sterilized for storage and crystal screening.

2.2. Crystallization, data processing and analysis

Protein crystallization was performed by the sitting-drop vapor diffusion method in 24-well Cryochem 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 p21) were obtained from a refinement of Nextal Classic II 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 c21) was obtained in the same condition as the p21 native crystal form, but for the X2 and X3 xyloligosaccharide 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 p21 native PbXynA1CD structure. The c2 native structure was processed similarly to the ligand-bound model, but TLS refined 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 tblseq 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-align when available. Structures were aligned using the ‘align’ function in PyMOL (DeLano, 2002) or with SSN Superposition (Krissinel and Henrick, 2004).

2.3. Accession numbers

The protein sequence accession number for PbXynA1 with all associated modules is UniProt ID: 6CGRV0. Protein structure data for PbXynA1CD were deposited with the PDB under the accession numbers PDB ID: 3R08, PDB ID: 4E4P and PDB ID: 3RDK for the p21 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 p21 and c2 space groups and the asymmetric unit of the crystals contained eight and two
chains, respectively, while the MeGX3-bound PbXynA1CD crystals were in the space group p43212 and contained two chains (Table 1). 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 p21). From this later screening, the second native structure (space group c2) 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 c2 and provided a second unit cell packing arrangement which confirms the original finding concerning the disordered β8–ω8 loop and ω8 helix motifs of the original native structure. All further discussion regarding native PbXynA1CD protein structure will refer to the p21 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 β8–ω8 loop and ω8 helix motif in the MeGX3-bound structure also corresponds with well ordered density for the β7–α7 loop that leans over and caps the β8–ω8 loop (Fig. 2b). In the native structure the α7 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 β7–α7 loop diverges from a position where Tyr271 stacks on and likely stabilizes the β8–ω8 loop to a position 13.3 Å further toward the β6–α6 loop. In this position of the native structure, amino acid side chains of the β6–α6–β7 loop may interact in the distal aglycone region. Lastly, in the small β5–ω5 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 subites

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 C1 conformation with the C-1 hydroxyl positioned in its ω-anomeric form with the O-1 and O-2 hydroxyls hydrogen bonding with the Oe-1 and Oe-2 of the predicted catalytic nucleophile Glu262. The O-2 oxygen hydrogen bonds with the Nε of Asn137 and the O-3 hydroxyl of this xylose forms additional hydrogen bonds with the Nε of His81 and the Nε of Lys48. Predicted hydrogen positions indicate that Nε of His81 is not protonated and rather acts as a hydrogen acceptor for the −1 hydrogen bond.
subsite xylose O-3 hydroxyl hydrogen. The hydroxyl oxygen in this position would then accept a hydrogen from Nf of 2 subsite pocket. 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 PbXyn1A1CD, 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 PbXyn1A1CD 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), thermostability 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 PbXyn1A1CD (Lo Leggio et al., 2001).
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 β1–α1 loop region and larger amino acids that extend toward the pocket from the β2–α2, β3–α3 and β8–α8 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, /C0 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.

As depicted in Fig. 3b and c, a glycerol (Gol346) and five waters occupy the /C0 subsite pocket in the MeGX3-bound structure. The glycerol was likely derived from cryoprotection, but with the

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).
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 G1G346 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 G1G346 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 aldouronic acid (MeGX3) 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 xyletriose 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 Gin87 and the O-3 hydroxyl of xylose and with the C-6 carbonylate 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 Nε 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 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, Thr313 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 β8–98 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 β8–98 region.
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 PbJD12. 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 though 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 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 National 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


