

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/5292589>

The Structure of the Complex between a Branched Pentasaccharide and Thermobacillus xylanilyticus GH-51 Arabinofuranosidase Reveals Xylan-Binding Determinants and Induced Fit †‡

ARTICLE *in* BIOCHEMISTRY · AUGUST 2008

Impact Factor: 3.02 · DOI: 10.1021/bi800424e · Source: PubMed

CITATIONS

23

READS

45

7 AUTHORS, INCLUDING:



Gabriel Paës

French National Institute for Agricultural R...

17 PUBLICATIONS 282 CITATIONS

[SEE PROFILE](#)



Michael J O'Donohue

French National Institute for Agricultural R...

80 PUBLICATIONS 1,275 CITATIONS

[SEE PROFILE](#)



Osman Mirza

University of Copenhagen

50 PUBLICATIONS 1,751 CITATIONS

[SEE PROFILE](#)

The Structure of the Complex between a Branched Pentasaccharide and *Thermobacillus xylanilyticus* GH-51 Arabinofuranosidase Reveals Xylan-Binding Determinants and Induced Fit^{†,‡}

Gabriel Paës,^{§,||} Lars K. Skov,[⊥] Michael J. O'Donohue,^{§,#} Caroline Rémond,[§] Jette S. Kastrup,[▽] Michael Gajhede,[▽] and Osman Mirza*,[▽]

INRA—UMR FARE 614, 8, rue Gabriel Voisin, BP 316, 51688 Reims cedex 2, France, Novozymes A/S, Krogshøjvej 36, DK-2880 Bagsværd, Denmark, and Department of Medicinal Chemistry, Faculty of Pharmaceutical Sciences, University of Copenhagen, Universitetsparken 2, 2100 Copenhagen, Denmark

Received March 12, 2008; Revised Manuscript Received May 21, 2008

ABSTRACT: The crystal structure of the family GH-51 α -L-arabinofuranosidase from *Thermobacillus xylanilyticus* has been solved as a seleno-methionyl derivative. In addition, the structure of an inactive mutant Glu176Gln is presented in complex with a branched pentasaccharide, a fragment of its natural substrate xylan. The overall structure shows the two characteristic GH-51 domains: a catalytic domain that is folded into a $(\beta/\alpha)_8$ -barrel and a C-terminal domain that displays jelly roll architecture. The pentasaccharide is bound in a groove on the surface of the enzyme, with the mono arabinosyl branch entering a tight pocket harboring the catalytic dyad. Detailed analyses of both structures and comparisons with the two previously determined structures from *Geobacillus stearothermophilus* and *Clostridium thermocellum* reveal important details unique to the *Thermobacillus xylanilyticus* enzyme. In the absence of substrate, the enzyme adopts an open conformation. In the substrate-bound form, the long loop connecting β -strand 2 to α -helix 2 closes the active site and interacts with the substrate through residues His98 and Trp99. The results of kinetic and fluorescence titration studies using mutants underline the importance of this loop, and support the notion of an interaction between Trp99 and the bound substrate. We suggest that the changes in loop conformation are an integral part of the *T. xylanilyticus* α -L-arabinofuranosidase reaction mechanism, and ensure efficient binding and release of substrate.

Harnessing of the renewable carbon reserves contained within plant biomass is regarded to be one of the major technological challenges of the 21st century. To achieve this goal several strategies are being considered. Among these, a biotechnological approach that includes the application of polysaccharide-hydrolyzing enzymes is promising, because of the enormous potential for innovation in the various aspects of the industrial biotechnology sector.

Hemicelluloses represent 20–35% of plant biomass and are the second most abundant source of renewable carbon after cellulose (1). Heteroxylans are the most common hemicelluloses and complex polysaccharides composed of

a β -D-xylose backbone that can be substituted with various neutral (arabinose, galactose) and acidic sugars (methyl-glucuronic acid), as well as with acetyl groups and phenolic acids (2). Due to the high complexity and structural variability of heteroxylans, enzymatic hydrolysis is achieved using a vast array of enzymes that include endoxylanases (EC 3.2.1.1) and debranching enzymes such as α -L-arabinofuranosidases (EC 3.2.1.55).

α -L-Arabinofuranosidases (Abf¹) are exo-acting enzymes that hydrolyze α -(1→2), α -(1→3) and/or α -(1→5) bonds and release arabinofuranosyl moieties from a wide variety of arabinose-containing substrates, including arabinan, arabinogalactans, heteroxylans and synthetic compounds such as *p*-nitrophenyl- α -L-arabinofuranoside (*p*NP- α -L-Araf) (3). According to the glycoside-hydrolase classification system, Abfs are present in six families (3, 43, 51, 54, 62 and

[†] This work was facilitated by the award of a Marie-Curie fellowship to G.P. and financial support from DANSYNC.

[‡] The structures have been deposited in the RCSB Protein Data Bank (www.pdb.org) with pdb codes 2vrk and 2vrq for the Tx-Abf-Se and Tx-Abf-Glu176Gln:X4A structure, respectively.

* To whom correspondence should be addressed: Osman Mirza, University of Copenhagen, Faculty of Pharmaceutical Sciences, Department of Medicinal Chemistry, Universitetsparken 2, 2100 Copenhagen, Denmark. E-mail: om@farma.ku.dk. Phone: +45 35336175. Fax: +45 35336100.

[§] INRA—UMR FARE 614.

^{||} Present address: METabolic EXplorer S.A., Biopôle, Clermont-Limagne, 63360 Saint-Beauzire, France.

[⊥] Novozymes A/S.

[#] Present address: INSA/INRA—UMR 792, 135, avenue de Rangueil, 31077 Toulouse cedex 04, France.

[▽] University of Copenhagen.

¹ Abbreviations: Abf, α -L-arabinofuranosidase; Ct-Abf, GH-51 α -L-arabinofuranosidase from *Clostridium thermocellum*; Gs-Abf, GH-51 arabinofuranosidase from *Geobacillus stearothermophilus*; *p*NP-Ara, *p*-nitrophenyl- α -L-arabinofuranose; *p*NP- α -L-Araf, *p*-nitrophenyl- α -L-arabinofuranoside; rms, root-mean-square; Tx-Abf, GH-51 α -L-arabinofuranosidase from *Thermobacillus xylanilyticus*; Tx-Abf-Glu176Gln, inactive Glu176Gln mutant of Tx-Abf; Tx-Abf-Glu176Gln:X4A, complex between Tx-Abf-Glu176Gln and X4A; Tx-Abf-Se, a seleno-methionyl derivative of wild type Tx-Abf; Tx-Abf-Trp248Ala, Trp248Ala mutant of Tx-Abf; vdW, van der Waals; X4A, pentasaccharide O - β -D-xylopyranosyl-(1→4)- O - α -L-arabinofuranosyl-(1→3)- O - β -D-xylopyranosyl-(1→4)- O - β -D-xylopyranosyl-(1→4)- O - β -D-xylopyranose.

Table 1: Oligonucleotide Primer Pairs Used for Site-Directed Mutagenesis^a

mutation	primer sequence (5'→3')
Trp99Ala	GTGAACACGCACGGGGCGGCGTCATCGATGACGCCCGCGTGCCTAC
Glu176Gln	GCGTCGCAACCAGAACACTGGGGCTGCGCAGCCCCAGTCTGGTTGCCGACGC
Cys180Ala	AACGAGAACTGGGGCGCAGGGCAACATGCGCGCATGTTGCCGCTGCCGCCCCAGTCTCGTT
Trp248Ala	GTTCGGGCCCCGGAGAAGAAAGGACCGGTCTTCTCCGGGGGGCCCGAAC
Δloop	AAGCGGAAGCGGGCGCGAGAACCATTCGGCACCATGATGGTTCTCGCCGCCCGCTCCGCTCTCCGTGGG

^a Δloop mutation consists of the substitution of residues Met94 to Ile103 in the β 2 α 2 loop by two Gly residues.

93) (4, 5). GH-51 and GH-54 are the two major families, containing almost exclusively Abfs.

The Abf from *T. xylanilyticus* (Tx-Abf) belongs to family GH-51 (6). This 56 kDa thermostable enzyme displays an optimum activity at 75 °C and remains active for several hours at 60 °C (6). Like all GH-51 enzymes, Tx-Abf is a retaining enzyme that catalyzes the hydrolysis of glycosidic bonds through a double displacement mechanism with Glu176 as the acid/base and Glu298 as the nucleophile (7–9). As a retaining enzyme, Tx-Abf also catalyzes a wide variety of transglycosylation reactions (10–12). With regard to specificity, combined observations from hydrolysis and transglycosylation reactions reveal that Tx-Abf specifically recognizes α -L-arabinofuranose and other configurationally related glycosides. Likewise, Tx-Abf shows hydrolytic activity on pNP- β -D-fucofuranoside and significant transglycosylation activity on pNP- β -D-galactofuranoside and pNP- β -D-xylopyranoside (10–12). However, Tx-Abf shows very low catalytic activity on compounds containing L-arabinopyranose. Among naturally occurring sugars, Tx-Abf is particularly active on arabino-xylo-oligomers, while much less activity on arabinoxylans is observed (13). In terms of bond specificity, Tx-Abf displays a preference for α -(1→2) bonds and, to a lesser extent, α -(1→3) bonds. In contrast, α -(1→5) bonds are very poorly hydrolyzed and synthesis of α -(1→5) bonds has not been clearly shown (6).

To date two structures of closely related GH-51 members have been published: Abfs from *G. stearothermophilus*, Gs-Abf (14), and from *C. thermocellum*, Ct-Abf (15). Both structures revealed that GH-51 Abfs are composed of a catalytic domain characterized by a $(\beta/\alpha)_8$ barrel and a C-terminal domain of unknown function that has the jelly roll topology. The enzyme is found to be a hexamer composed of catalytically independent monomers. Analysis of the Michaelis complex of inactivated Gs-Abf and arabinose- α -(1→3)-xylose and other complexes have allowed a detailed description of subsites -1 and +1.

Here, we describe two crystal structures of Tx-Abf, whose sequence displays only 25% identity with the previously studied GH-51 Abfs. The first structure is the apo-Abf, whereas the second is of an inactive Glu176Gln mutant (Tx-Abf-Glu176Gln) in complex with a pentasaccharide. The latter structure provides novel details of an extended substrate binding site that is composed four subsites, -1 through to +2 and +2'. Furthermore, analysis of the substrate complexed enzyme structure revealed large structural rearrangements that appear to be induced by substrate binding.

MATERIALS AND METHODS

Tx-Abf Mutagenesis, Expression and Purification. Site-directed mutagenesis experiments were performed on the

plasmid construction pET-ABF encoding Tx-Abf (6) using the QuickChange kit (Stratagene, Amsterdam, The Netherlands). The oligonucleotide primers used for preparation of Tx-Abf mutants are shown in Table 1. The correct introduction of the mutations was checked by DNA sequence analysis (MWG Biotech AG, Ebersberg, Germany).

Wild type and mutated Tx-Abf were cloned, overexpressed and purified as described previously (6). In particular, the inactivity of the mutant Glu176Gln was checked by enzymatic assay (see section below). A seleno-methionyl derivative of wild type Tx-Abf (Tx-Abf-Se) was obtained by introducing the plasmid pET-ABF into *Escherichia coli* B834 (DE3) pLysS cells (Novagen, USA). The transformed *E. coli* cells were grown on minimal media containing Se-Met according to previously established methods (16). Tx-Abf-Se was purified using the same method as for purification of wild type Tx-Abf.

Crystallization and Data Collection. Crystals of Tx-Abf-Se were obtained at room temperature using the hanging-drop vapor-diffusion method. Suitable crystals appeared within one week in drops containing 2.5 μ L of a 1.5 mg·mL⁻¹ enzyme solution and 2.5 μ L crystallization buffer consisting of 0.8 M KH₂PO₄, 0.05 M Hepes at pH 4.5. The enzyme crystallized in the hexagonal space group *P*6₅22. A data set was collected at cryogenic temperature and at the Se-absorption peak ($\lambda = 0.9789 \text{ \AA}$) on a MARMOSET225 CCD detector using synchrotron radiation of beam line BM14 at the ESRF in Grenoble, France. Crystals of Tx-Abf-Glu176Gln were grown under the same conditions as Tx-Abf-Se. To form the complex, crystals of the Glu176Gln mutant were soaked in crystallization buffer containing 10 mM of the pentasaccharide *O*- β -D-xylopyranosyl-(1→4)-*O*- α -L-arabinofuranosyl-(1→3)-*O*- β -D-xylopyranosyl-(1→4)-*O*- β -D-xylopyranosyl-(1→4)-*O*- β -D-xylopyranose (X4A) for 30 min before cryocooling in liquid nitrogen. Longer soaking times were prejudicial for crystal stability. Data were collected to 2.0 \AA at beamline ID14-4 at ESRF in Grenoble, France. The Tx-Abf-Se data set was processed using XDS (17) and SCALA (18) while the mutant data were processed with XIA-DPA (19). Data statistics are given in Table 2.

Structure Determination, Refinement and Analysis. The structure of Tx-Abf-Se was solved using the SAD data and the HKL2MAP program (20). These phases were used as input for Arp/wArp (21), which yielded a close to complete structure. The structure was refined to 2.2 \AA resolution using CNS (22) and subsequently rebuilt using COOT (23). The complex between Tx-Abf-Glu176Gln and X4A (Tx-Abf-Glu176Gln:X4A) was solved using initial phases from the Tx-Abf-Se structure and refined using CNS and finally PHENIX (24). For both structures, 1% (~1800) of the reflections were set aside for cross-validation. Residues

Table 2: Data and Refinement Statistics

	Tx-Abf-Se	Tx-Abf-Glu176Gln:X4A
Data		
space group	<i>P</i> 6 ₅ 22	<i>P</i> 6 ₅ 22
cell dimensions (Å)	158.0 158.0 380.0	156.8 156.8 378.6
wavelength (Å)	0.9789	1.0780
resolution (Å)	39.47–2.20 (2.32–2.20)	78.33–2.00 (2.11–2.00)
total reflections ^a	1934421 (221178)	2167097 (311629)
unique reflections	188045 (27020)	184568 (26538)
redundancy	10.7 (10.8)	11.7 (11.7)
completeness (%)	99.8 (99.7)	100.0 (99.9)
<i>R</i> _{sym} (%) ^b	13.1 (46.4)	11.4 (39.3)
<i>I</i> / <i>σ</i> _{<i>I</i>}	5.1 (1.6)	5.3 (1.6)
Refinement		
resolution (Å)	39.5–2.2 (2.27–2.20)	78.5–2.00 (2.05–2.00)
<i>R</i> factor (%) ^c	16.5 (17.5)	18.2 (24.8)
<i>R</i> _{free} (%) ^d	18.6 (18.8)	20.1 (27.1)
bond lengths (Å)	0.008	0.008
bond angles (deg)	1.1	1.3
dihedral angles (deg)	18.2	22.6

^a Values in parentheses are for the highest resolution shell. ^b *R*_{sym} = $\frac{1}{n} \sum |I_i(h) - \langle I(h) \rangle| / \sum I_i(h)$, where *I*_{*i*}(*h*) is the *i*th measurement. ^c *R* factor = $\frac{1}{n} \sum |F(h)_{\text{obs}} - |F(h)_{\text{calc}}|| / \sum |F(h)_{\text{obs}}|$. ^d *R*_{free} was calculated using 1% randomly selected reflections that were excluded from refinement.

Glu36 and Thr217 from both structures (in all monomers) along with Pro71 from the Tx-Abf-Glu176Gln:X4A structure (in all monomers) were found in disallowed regions of the Ramachandran plot. The refinement statistics are given in Table 2.

Structural comparisons were done using programs COOT, O (25) or PyMOL (26). All figures were prepared with PyMOL.

Bioinformatics. Sequence identities and homologies were derived from BLAST analyses using the BLAST module embedded in the UniProt server (<http://beta.uniprot.org/>). Phylogenetic analyses were performed using the Phylogeny.fr server. A CLUSTALW alignment of 28 nonredundant full length GH-51 Abf sequences was used as input data to generate an unrooted phylogenetic tree.

Enzymatic Assays. The standard assay for measuring hydrolytic activity of Tx-Abf and its mutant derivatives was performed by incubating 0.1 mL of the enzyme (concentration between 25 and 30 μM) with 0.9 mL of *p*-nitrophenyl-α-L-arabinose (pNP-Ara) (Figure 1A) in buffered conditions (5 mM pNP-Ara in 50 mM sodium acetate, pH 5.8) at 60 °C. Continuous release of pNP was measured at 401 nm. One unit of activity corresponds to the amount of enzyme releasing 1 μmol of pNP per minute. Specific activity was derived from this calculation and the concentration of the assayed enzyme. Protein concentrations (wild type and mutant enzymes) were calculated by absorbance measurement at 280 nm using a molar extinction coefficient of 85,230 M⁻¹·cm⁻¹ (7).

Kinetic parameters were determined by incubating the enzyme with pNP-Ara at concentrations ranging from 0.25 mM to 20 mM. Measurements were performed in duplicate, *k*_{cat} and *K*_M values were derived using nonlinear regression formulas implemented in the Enzyme Kinetics module of Sigma Plot 6.1 (SPSS, IL, USA).

Analysis of hydrolytic activity on a biologically relevant substrate was performed using X4A (Figure 1B). This pentasaccharide was obtained after hydrolysis of destarched wheat bran (30 g·L⁻¹) by a GH-11 xylanase (10 IU·mL⁻¹)

at 60 °C in water as previously described (27). The oligosaccharide products were separated by size-exclusion chromatography on two serially connected BioGel P-2 columns (170 × 1.5 cm) using water as the eluent (flow rate of 20 mL·h⁻¹). The X4A fraction was isolated and qualitatively analyzed using ¹H NMR and MALDI-TOF mass spectrometry. The data agreed with those previously published on X4A (28).

The kinetic parameters for X4A hydrolysis were obtained using a discontinuous method. Hydrolysis using substrate concentrations ranging from 0.625 mM to 10 mM were performed using both mutant and wild type Tx-Abf at a concentration of 0.03 μM. Reactions were stopped after various incubation times by boiling for 15 min. After centrifugation and filtering, the quantity of free L-arabinose within the samples was determined using HPAEC analysis with a PA1 column on a Dionex ED40/GP40 system (Dionex, CA, USA) with pulsed amperometric detection and fucose as the internal standard.

Protein Folding and Thermostability Investigations. To compare folding of mutant derivatives of Tx-Abf with the wild type enzyme, protein solutions (0.125 g·L⁻¹ in 50 mM K₂HPO₄, pH 6.5) were examined by circular dichroism using a J-810 spectrometer (Jasco, Hachioji, Japan). Measurements were performed at 20 °C using a quartz cell (1.0 cm path length) from 200 to 280 nm. The buffer signal was subtracted and data were averaged from two acquisitions.

To investigate thermostability, enzyme solutions (0.07 g·L⁻¹) were incubated at 80 °C for various time periods. Residual activity was quantified at 60 °C using the standard assay. Likewise, melting temperatures (*T*_m) were obtained by monitoring heat-induced changes to circular dichroism spectra at 220 nm using a J-810 spectrometer (Jasco, Hachioji, Japan) equipped with a Peltier temperature controller. The enzyme solution (0.125 g·L⁻¹ in 50 mM K₂HPO₄, pH 6.5) was placed in a quartz cell (0.2 cm path length) and the temperature was varied from 40 to 95 °C (1 °C·min⁻¹), and subsequently from 95 to 40 °C. In all experiments, unfolding was irreversible. Hence, only apparent *T*_m values, *T*_m(app), could be derived. Due to the high thermostability of Tx-Abf and its derivatives, several *T*_m(app) were measured in the presence of urea (in the range 1 to 10 M). In this case, enzyme samples were incubated in the presence of urea at room temperature for 5 min before data acquisition. Values for *T*_m(app) were obtained by extrapolating the curve *T*_m(app) vs urea concentration to a urea concentration equal to zero.

Substrate Binding Monitored by Fluorescence Spectroscopy. The ability of Tx-Abf and the Trp248Ala mutant (Tx-Abf-Trp248Ala) to bind X4A was evaluated by titration fluorescence spectroscopy. Freshly purified enzymes were thoroughly dialyzed against 2 × 100 volumes of 50 mM sodium acetate, pH 5.8. Fluorescence was measured at 5 °C using a FP-6500 spectrophotometer (Jasco, Hachioji, Japan) equipped with a Peltier temperature controller and a 1.0 × 1.0 cm quartz cell under continuous stirring. Titration of tryptophan fluorescence with the X4A ligand was performed with 1.8 mL of protein solution (0.1 μM). Tryptophan excitation was achieved at 295 nm, and emission spectra were collected in duplicate over the wavelength range 300–450 nm, using a bandwidth of 3 nm for both excitation and emission (response time of 0.2 s and scan rate of 100

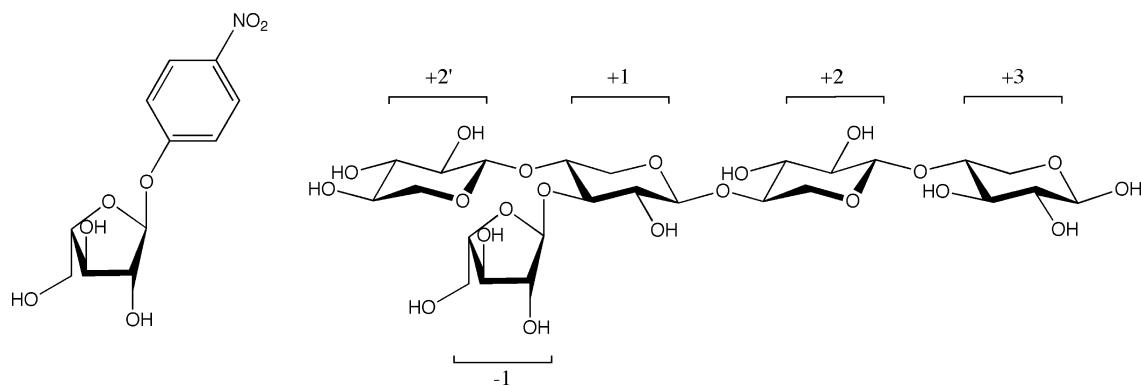
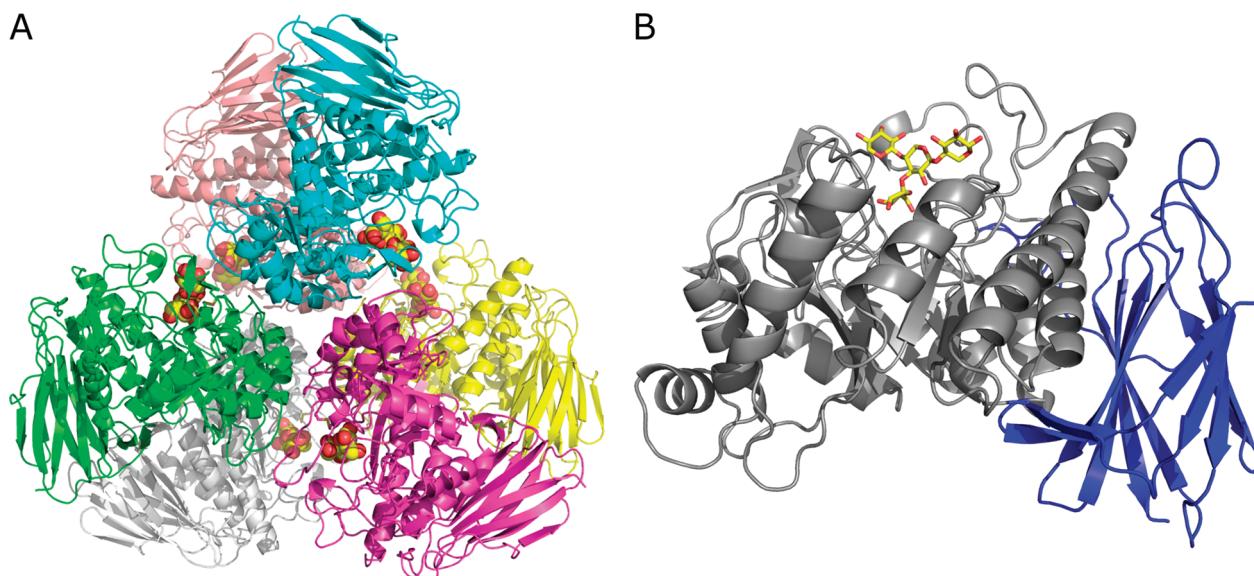


FIGURE 1: Schematic structures of the substrates pNP-Ara (left, A) and X4A (right, B) with subsites indicated.

FIGURE 2: A. The hexameric structure of Tx-Abf in complex with X4A. The hexamer is composed of a dimer of trimers. The bound X4A molecules are shown as vdW spheres. B. Overall structure of the Tx-Abf monomer in Richardson representation, with the bound X4A molecule shown in stick representation. The $(\beta/\alpha)_8$ barrel is colored gray, and the C-terminal jelly roll domain is colored blue.

$\text{nm} \cdot \text{min}^{-1}$). Fluorescence spectra were corrected for buffer fluorescence and volume changes, which did not exceed 5% (v/v). Similarly, the extent of hydrolysis of X4A by Tx-Abf at 5 °C was estimated to be less than 5% over a 15 min time period.

To calculate the dissociation constant K_d , fluorescence intensity changes were recorded at the maximum fluorescence intensity. Likewise, for each X4A concentration [L], the change in fluorescence $\Delta F = F_0 - F$ (where F_0 is the fluorescence of unbound protein) was calculated and plotted versus [L]. To obtain K_d and ΔF_{\max} , the resulting curve was fitted to the hyperbolic equation $\Delta F = \Delta F_{\max}[L]/(K_d + [L])$ using SigmaPlot 6.1 (SPSS, IL, USA).

RESULTS

Overall Structure. The Tx-Abf-Se and Tx-Abf-Glu176Gln: X4A structures contain three enzyme molecules (A, B and C) within the asymmetric unit of the crystal (Figure 2A), giving rise to a solvent content of 70%. The three molecules are related by noncrystallographic 3-fold symmetry (the trimer). Among all six monomers in the two structures, the root-mean-square (rms) deviations are less than 0.4 Å (on C α atoms excluding the $\beta 2\alpha 2$ loop residues 92 to 104 that adopt different conformations). A hexamer

(Figure 2A) is formed from two trimers by crystallographic 2-fold symmetry.

The Tx-Abf enzyme is composed of a single polypeptide chain folded into a two domain structure: a catalytic domain with the frequently encountered $(\beta/\alpha)_8$ barrel architecture and a C-terminal domain with a jelly roll topology (Figure 2B). The overall fold and multimeric organization of Tx-Abf are very similar to those of the two previously published structures of GH-51 enzymes (14, 15). A structural alignment performed using the SSM server (<http://www.ebi.ac.uk/msd-srv/ssm/cgi-bin/ssmserver>) shows an rms deviation of 1.6 Å on 433 C α atoms of Tx-Abf-Se to the structures of Gs-Abf and Ct-Abf (molecules A in all structures). In comparison, the rms deviation between Gs-Abf and Ct-Abf is 0.6 Å on 496 C α atoms. The most prominent structural differences between Tx-Abf and the other two structures arise from different lengths and conformations of the $\beta 4\alpha 4$, $\beta 5\alpha 5$, $\beta 6\alpha 6$ and $\beta 7\alpha 7$ loops. As seen in all catalytic domains that possess $(\beta/\alpha)_8$ barrel architecture (29), the active site in Tx-Abf is formed by $\beta\alpha$ loops. In Tx-Abf, the length of these loops varies from only six residues in the $\beta 3\alpha 3$ loop and up to 40 residues in the $\beta 2\alpha 2$ loop, in which three short helices are inserted.

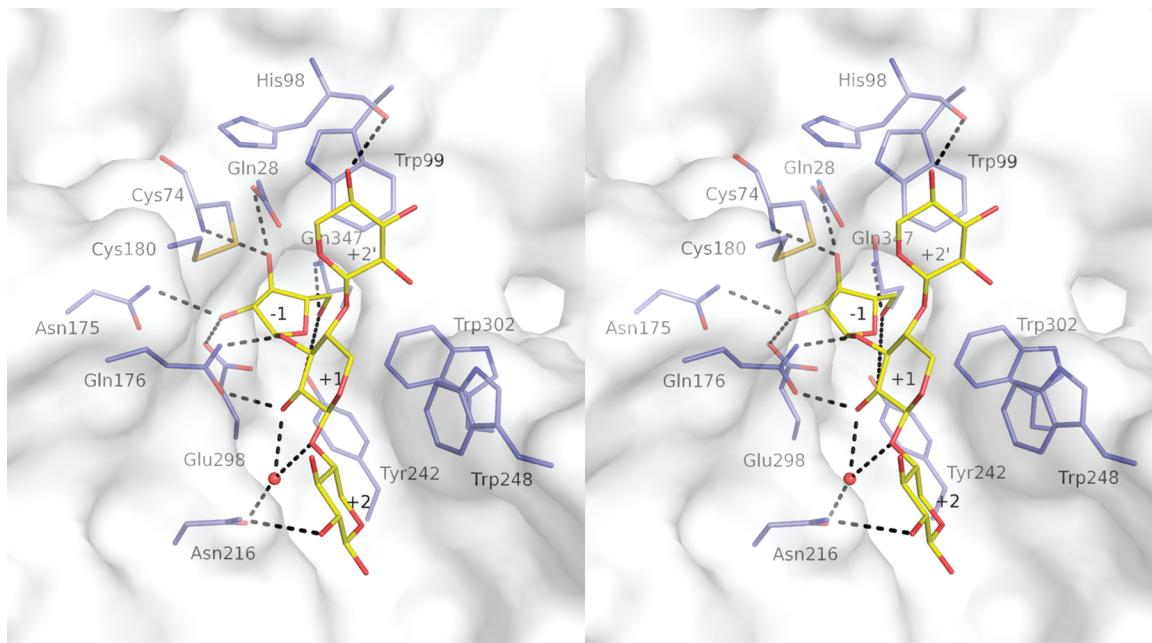


FIGURE 3: Stereoview of the X4A binding cleft of Tx-Abf (molecule A). Residues in contact are shown in stick representation. All hydrogen bonds are shown as dotted lines. Carbohydrate moieties from molecule A of the Tx-Abf-Glu179Gln:X4A are shown with yellow carbon atoms and the interacting residues are shown with light blue carbon atoms. Water molecules are shown as red spheres.

The major intermolecular contacts within the trimer in the asymmetric unit involve two distinct regions: residues from the $\alpha 3\beta 4$ and $\beta 2\alpha 2$ loops and from the $\alpha 4$ helix and the $\beta 4\alpha 4$ loop.

Structural Rearrangements of the $\beta 2\alpha 2$ Loop. The catalytic acid/base residue Glu176 is located in the $\beta 4\alpha 4$ loop, and the nucleophile Glu298 at the C-terminal of the $\beta 7$ strand. Like the catalytic residues in Gs-Abf, Glu176 and Glu298 are located on one side of a pocket (subsite -1) which is itself in the middle of a long (30 Å), nonlinear groove (Figure 3). The pocket, characterized by negatively charged electrostatic potential, is delimited by $\beta\alpha$ loops 1, 2, 4, 7 and 8, and closed by a disulfide bond formed between residues Cys74 and Cys180, which is absent in the previously determined GH51 Abf structures.

In all three molecules of the asymmetric unit of the Tx-Abf-Se, the $\beta 2\alpha 2$ loop is exclusively found in a conformation that opens access to the active site (Figure 4, open conformation in cyan). This is also the case for the native Tx-Abf structure (data not shown). The $\beta 2\alpha 2$ loop is involved in an intermolecular hydrogen bond via the backbone of Val95 (O atom) and Glu150 (N atom) from the neighboring B molecule within the trimer. In addition, Glu150 forms a salt bridge with Arg93. However, the relatively high *B*-values of atoms within this region indicate loop flexibility.

Complex formation with X4A (the Tx-Abf-Glu176Gln:X4A structure) leads to conformational rearrangements of the $\beta 2\alpha 2$ loop in molecules A and B (Figure 4, closed conformation in green). In molecule C, the $\beta 2\alpha 2$ loop adopts the open conformation that is stabilized by a salt bridge between Arg93 and Glu163 belonging to molecule A of a symmetry related hexamer. In the closed conformation of molecules A and B, residues His98 and Trp99 become part of subsite -1 by forming a much narrower pocket than in the open conformation.

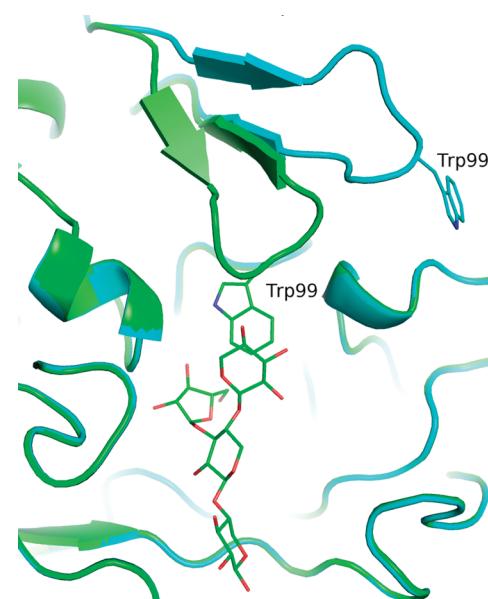


FIGURE 4: The SeMet Tx-Abf-Se (molecule A, cyan) superimposed onto the Tx-Abf-Glu179Gln:X4A complex (molecule A, green, X4A in stick representation). The view is centered on the $\beta 2\alpha 2$ loop with Trp99 shown in stick representation.

Active Site Architecture and Oligosaccharide Binding. In molecule C of the Tx-Abf-Glu176Gln:X4A complex where the $\beta 2\alpha 2$ loop adopts the open conformation, only three saccharide moieties (-1, +1 and +2) can be observed. However, in molecules A and B four saccharide moieties -1, +1, +2 and +2' could be located in the electron density (Figure 5). Therefore, this structure provides the first information on the substrate binding subsites +2, and +2' and reveals that the conformation of the $\beta 2\alpha 2$ loop influences the binding of xylose moieties in the subsite +2'. The X4A xylose moieties bind within a groove on the surface of Tx-Abf with five direct and two water-mediated hydrogen bonds between the enzyme and the xylose moieties (Figure 3). The major interactions are the following: the xylosyl-moiety

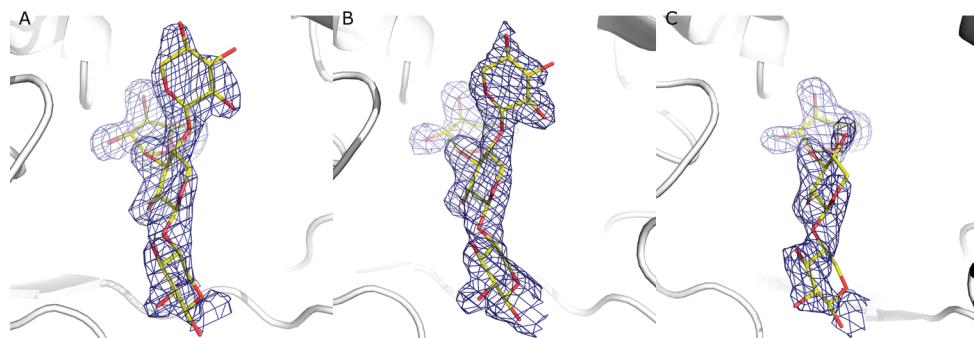


FIGURE 5: The $2F_o - F_c$ maps (blue) contoured at 1σ (A–C). Molecules A, B and C are shown in white Richardson representation with carbohydrate moieties from X4A displayed as yellow sticks.

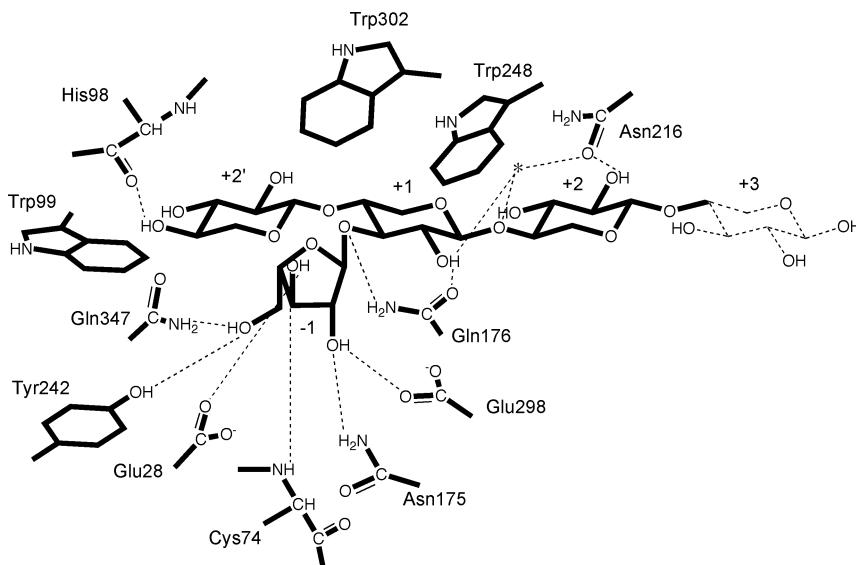


FIGURE 6: Schematic representation of the X4A interactions. Hydrogen bonds are shown as dotted lines. Water molecules are represented with an asterisk. Subsite +3 xylosyl is shown with dashed lines since it is not observed in the structures.

bound at subsite +2' in molecules A and B interacts with Trp99, possibly, through van der Waals (vdW) interactions and through a single hydrogen bond to the backbone of His98. The subsite +1 xylosyl is anchored probably through vdW interactions with Tyr242, Trp248 and Trp302, and through a direct hydrogen bond to Gln176 and an indirect hydrogen bond via a water molecule to Asn216. The subsite +2 xylosyl interacts with Trp248 and makes a hydrogen bond to Asn216. In molecule C, all interactions except with Trp99 and His98 are preserved to the moieties at subsites –1, +1 and +2.

The arabinosyl moiety extends into the subsite –1 pocket within the groove (Figure 3) and binds to the enzyme in a 4E conformation as in Gs-Abf (14). This saccharide moiety is engaged in seven direct hydrogen bonds to residues Glu28, Cys74, Asn175, Gln176, Tyr242 and Gln347. No water-mediated hydrogen bonds are observed, but vdW interactions are implied with the side chains of residues Cys74, Trp99 and Trp302. All X4A interactions are represented schematically in Figure 6. Subsite –1 is devoid of water molecules in both structures; the closest water molecule is positioned 5.3 Å away from the C1 atom of the arabinosyl moiety. Therefore, no clear candidates for the catalytic water molecule are seen in the Tx-Abf structures. In Tx-Abf-Se, Glu176 and Glu298 are positioned very close to each other, with a distance between the closest oxygen atoms (Glu176 OE2–Glu298

OE2) of 3.2 Å. In comparison, this distance is 4.4 Å in the Gs-Abf structure. This indicates that, in the Tx-Abf structure, either the catalytic nucleophile Glu298 or acid/base Glu176 or both are protonated, which can be explained by the fact that the crystals were obtained at pH 4.5. In contrast, a distance of 3.8 Å (Gln176 NE2–Glu298 OE1) is seen in the Tx-Abf-Glu176Gln:X4A structure.

Influence of the Active Site Disulfide Bridge. Examination of the Tx-Abf structure revealed the presence of one disulfide bridge Cys74–Cys180, very close to the catalytic acid/base residue (Glu176). Generally, it is considered that disulfide bridges are often important for protein folding and stability (30, 31). However, until now the presence of disulfide bridges has not been reported in GH-51 Abfs and previous data have indicated that reducing agents only slightly affect the activity of thermostable GH-51 members (32, 33). Therefore, to investigate the structural and/or functional significance of this disulfide bond, the mutation Cys180Ala was introduced in order to disrupt the bridge Cys74–Cys180. Circular dichroism analysis (data not shown) revealed that this mutant and the wild type Tx-Abf display highly similar spectra, suggesting that the mutation did not prevent protein folding. In contrast, specific activity at 60 °C shows differences, since the mutation Cys180Ala provoked a 6-fold reduction in specific activity (Table 3). Determination of the kinetic

Table 3: Activity and Thermostability of Wild Type Enzyme and Tx-Abf Mutants Tested on pNP-Ara^a

	Tx-Abf	Tx-Abf-Cys180Ala	Tx-Abf-Trp99Ala	Tx-Abf-Δloop
SA (60 °C) (IU·mg ⁻¹)	169 ± 8	27 ± 1	10.5 ± 0.1	ND
<i>k</i> _{cat} (s ⁻¹)	355 ± 15	24 ± 1	165 ± 42	35 ± 5
<i>K</i> _M (mM)	1.4 ± 0.2	3.2 ± 0.3	45 ± 16	10 ± 2
<i>k</i> _{cat} / <i>K</i> _M (s ⁻¹ ·mM ⁻¹)	254 ± 47	7.5 ± 1.0	3.6 ± 1.9	3.5 ± 1.2
<i>t</i> _{1/2} (80 °C) (h)	1.7	0.9	ND	ND

^a SA: specific activity. *t*_{1/2}: half-life time. ND: not determined.

parameters of pNP-Ara hydrolysis revealed that this reduction was mainly caused by a sharp decrease in the *k*_{cat} value (Table 3). Likewise, mutants bearing the Cys180Ala substitution displayed reduced thermostability at 80 °C (Table 3).

Probing the Importance of the β2α2 Loop. In the presence of the pentasaccharide X4A the β2α2 loop has been observed in two different conformations. In the closed conformation, Trp99 adopts a position that would enable it to form several vdW interactions with the arabinosyl moiety. To further investigate this observation, the mutation Trp99Ala was introduced into Tx-Abf and the catalytic ability of the resulting mutant Tx-Abf-Trp99Ala was evaluated using pNP-Ara as substrate. While the *k*_{cat} value of this reaction is slightly altered, the *K*_M value is drastically increased (Table 3), indicating that the arabinosyl moiety is not correctly bound in the active site when Trp99 is replaced by Ala. In a second round of experiments, a more ambitious modification of Tx-Abf was also performed. The size of the β2α2 loop was reduced by deleting 10 residues from Met94 through to Ile103. These were replaced by two Gly residues, in order to confer enough flexibility to the N- and C-terminal ends of the loop to prevent any steric constraints. In activity tests using pNP-Ara the mutant, designated Tx-Abf-Δloop, is 70-fold less active than the wild type enzyme (Table 3). This reduction in activity is due to a large decrease of *k*_{cat} (10-fold) and an increase of *K*_M (7-fold). So, unlike the point mutation Trp99Ala, the deletion of the 10 amino acids that constitute the loop has a more global effect on catalysis.

The fluorescence emission of tryptophan was used as a probe to monitor binding of X4A to Tx-Abf. In the absence of X4A, excitation at 295 nm led to strong fluorescence emission characterized by maxima at 335 and 337 nm, respectively. Upon titration with X4A the maximum emission intensity was decreased by 15%, indicating that the substrate interacts with at least one tryptophan residue (Figure 7). Examination of the structural data, in particular those of the TxAbf-Glu176Gln:X4A complex, indicates that only two tryptophan residues, Trp99 and Trp248, are likely to interact with X4A. Therefore, in order to determine whether one or both of these residues are involved in substrate binding, Trp248 was replaced by alanine. The fluorescence emission spectrum of the mutant Trp248Ala in the absence of substrate was identical to that of the wild type enzyme, and titration with X4A led to a decrease in fluorescence intensity, highly similar to that observed previously. Moreover, calculation of the binding constant *K*_d revealed that binding of X4A was characterized by similar values for both Tx-Abf (206 ± 20 μM) and Trp248Ala (183 ± 36 μM). Therefore, these results show that Trp248 does not significantly contribute to the fluorescence emission spectrum of Tx-Abf and does not play

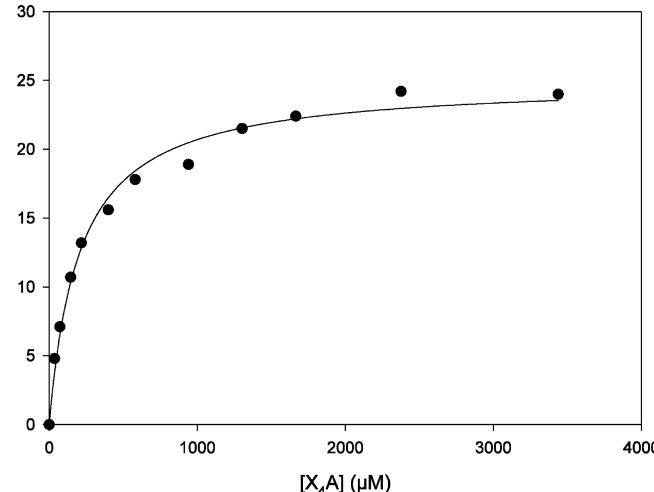


FIGURE 7: Titration of fluorescence emission from Tx-Abf using the oligosaccharide X4A. Variation in fluorescence emission at 335 nm (F) was monitored as the ligand concentration ([X4A]) was incrementally increased.

a significant role in the binding of X4A. Furthermore, it can be concluded that the Trp99 residue on the β2α2 loop provides the major contribution to the fluorescence spectrum of Tx-Abf and is a key player in the binding of the X4A substrate.

DISCUSSION

Family GH-51 in the CAZy database currently contains 128 sequences of bacterial origin. So far structural studies have been performed on two highly similar α-L-Abfs (82% sequence similarity, 68% identity) from *G. stearothermophilus* (14) and *C. thermocellum* (15), respectively, that form part of a cluster composed of at least 19 individual proteins (Figure 8). In this study, we have focused on another less similar Abf from *T. xylanilyticus* whose sequence shares only approximately 25% identity (50% similarity) with the two previously studied enzymes. Tx-Abf groups with 15 other bacterial Abfs as a UniRef50 cluster (UniProt database) whose seed sequence is the Abf from *Acidobacteria bacterium* (Ellin strain 345). The crystal structure of the *T. xylanilyticus* enzyme was determined as a seleno-methionyl derivative. In addition, the structure of an inactive mutant Glu176Gln is presented in complex with a branched pentasaccharide, X4A, which is representative of natural substrates of Tx-Abf that are generated through hydrolysis of lignocellulosic biomass. Therefore, the results presented here not only provide structural data for a previously unstudied Abf subgroup or cluster but also provide novel insight into the binding of biologically relevant substrates to GH-51 enzymes.

In the Tx-Abf-Glu176Gln:X4A complex, the pentasaccharide is bound in a groove on the surface of the enzyme,

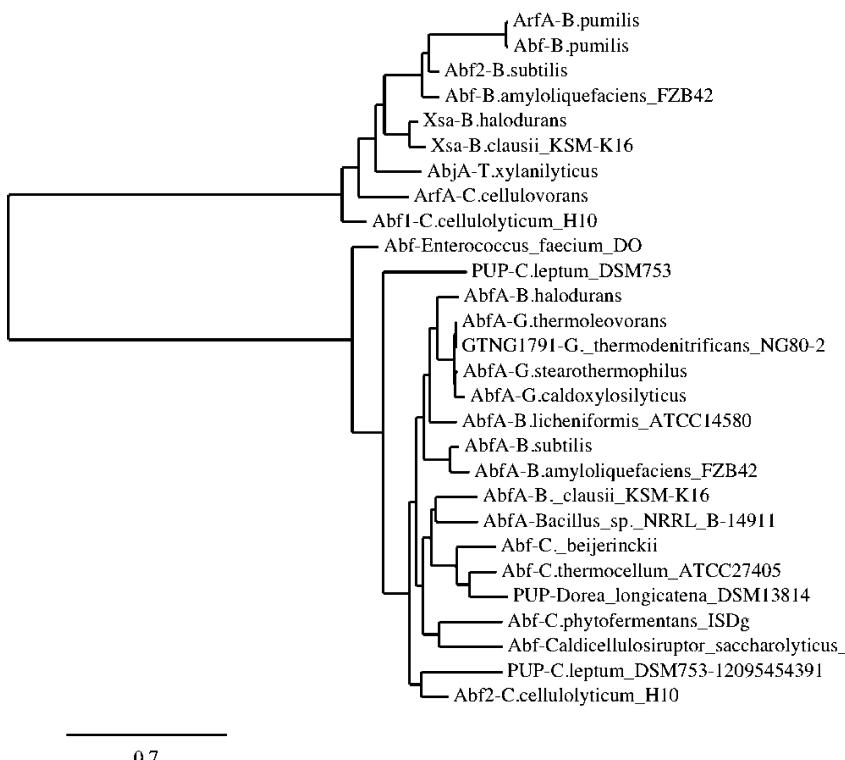


FIGURE 8: Phylogenetic representation of GH-51 Abf subgroups or clusters. Abf sequences displaying at least 60% similarity with Tx-Abf or Gs-Abf respectively were retrieved using BLAST embedded in the Uniprot server (<http://beta.uniprot.org/>). The sequences were aligned using CLUSTALW, and a tree representation was generated using the phylogeny.fr server (38). PUP is a putative uncharacterized protein, and AbjA is the protein designated Tx-ABF in this paper.

with the mono arabinosyl branch entering a pocket (subsite -1), whose topology is perfectly suited for this moiety. The arabinose is tightly bound by several hydrogen bonds, and no additional space is available for bulkier residues. This is consistent with the observation that, like most Abfs reported to date, Tx-Abf is unable to hydrolyze the bond that links a xylosyl to a 5-O-feruloyl-arabinosyl moiety. The catalytic residues, Glu176 (substituted by Gln176 in the Tx-Abf:X4A complex) and Glu298, are positioned at the entrance to the subsite -1 where they adopt conformations similar to those observed in the previously published GH-51 structures (14, 15).

When compared to Ct-Abf and Gs-Abf, the xylan binding site in Tx-Abf appears as a much narrower groove (Figure 9), suggesting that Tx-Abf would be unable to bind highly decorated arabinoxylans. This is concordant with experimental results that have shown Tx-Abf to liberate a higher proportion of arabinose from insoluble, weakly substituted wheat arabinoxylan ($A/X = 0.23$) than from water soluble, highly substituted wheat arabinoxylan ($A/X = 1.22$) (13).

A comparison of the Tx-Abf-Glu176Gln:X4A structure with those of the substrate bound complexes of Gs-Abf and Ct-Abf shows that all hydrogen-bonding partners are conserved except for Cys74. In Gs-Abf, Ct-Abf and other sequence-related Abfs, Cys74 is replaced by an asparagine and a strict equivalent of Cys180 is absent. In Tx-Abf, Cys74 forms a hydrogen bond from the backbone N atom to the O3 atom of the arabinosyl moiety. This hydrogen bond is also seen from the corresponding backbone atom of Asn74 in Gs-Abf. Furthermore, Cys74 is involved in a disulfide bridge with Cys180. Both of these side chains could make vdW interactions with the arabinose moiety. A similar situation has been reported for the GH-54 Abf from *Aspergillus*

kawachii. Indeed, it was suggested that the disulfide bond might favor recognition of the C4/C5 atoms of the arabinofuranosyl moiety via a hydrophobic interaction (34). The role of the Cys74–Cys180 disulfide bridge in Tx-Abf was investigated through mutation of Cys180 into Ala. A disruption of the disulfide bridge led to a 15-fold decrease in k_{cat} while the K_M value was only slightly increased (Table 3). Therefore, the binding interaction is not altered by the mutation Cys180Ala, contrarily to what was observed for the GH-54 Abf from *A. kawachii*. Hence, regarding the location of the disulfide bridge in the active site, the drop of the k_{cat} value in the Tx-Abf-Cys180Ala mutant is most likely due to the presence of the Cys74–Cys180 disulfide bridge that ensures correct orientation of the acid/base residue. A similar function has been attributed to a valine residue in GH-11 xylanases (35, 36). In these enzymes it is suspected that a highly conserved valine packs against the acid/base, thus blocking it into position.

The $\beta 2\alpha 2$ loop is a key feature in GH-51 arabinofuranosidases. In Tx-Abf, our results indicate that a significant movement allows Trp99 to adopt a position in the vicinity of subsite -1 where, according to the data from mutagenesis, it provides an important interaction with the bound arabinosyl moiety. Interestingly, Trp99 is highly conserved in GH-51 sequences and the structural superposition of the $\beta 2\alpha 2$ loops of Gs-Abf, Ct-Abf and the closed $\beta 2\alpha 2$ loop of Tx-Abf reveal a very similar structural organization. However, superposition of the structures of Gs-Abf and Ct-Abf onto the open conformation of Tx-Abf reveals that Gs-Abf and Ct-Abf probably cannot adopt this open conformation because of steric clashes with the $\beta 7\alpha 7$ loop, which is eight residues longer in Gs-Abf and 11 residues longer in Ct-Abf (Figure 10). Therefore, the movement of the $\beta 2\alpha 2$ loop, inferred by

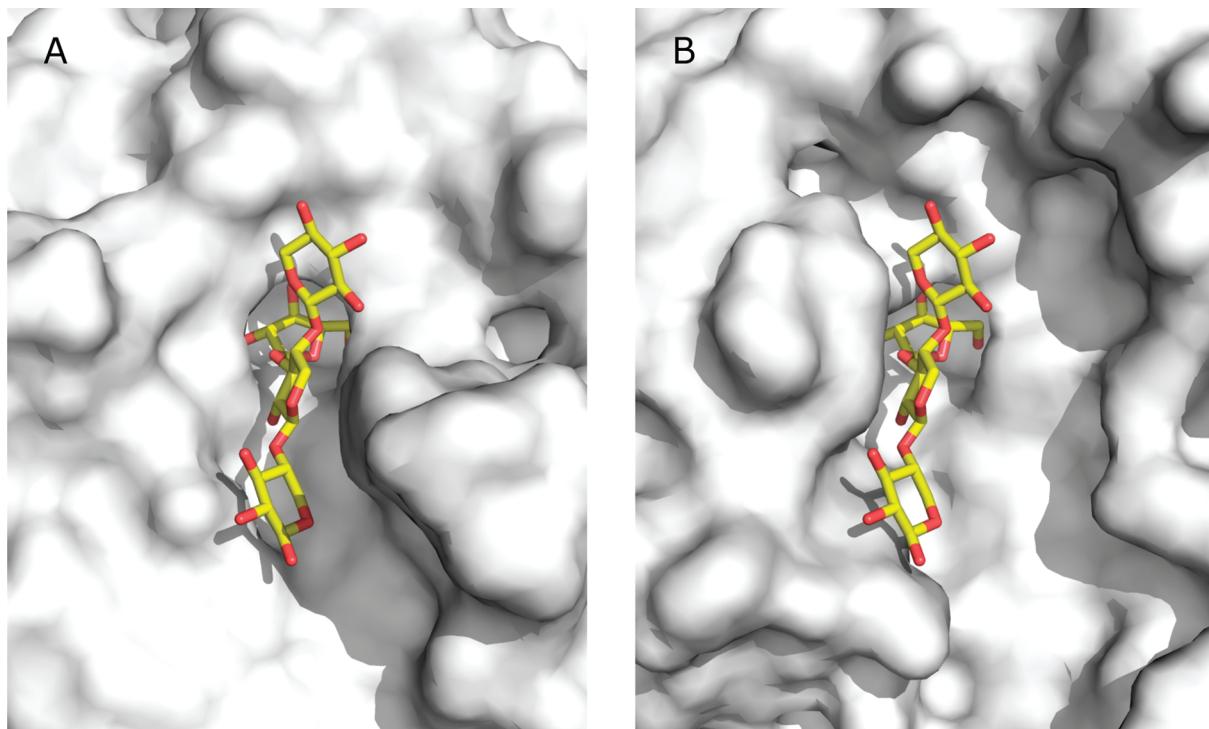


FIGURE 9: The substrate binding grooves of Tx-Abf (molecule A, left) and Ct-Abf (molecule A, right). A much narrower groove is observed in Tx-Abf compared to Ct-Abf.

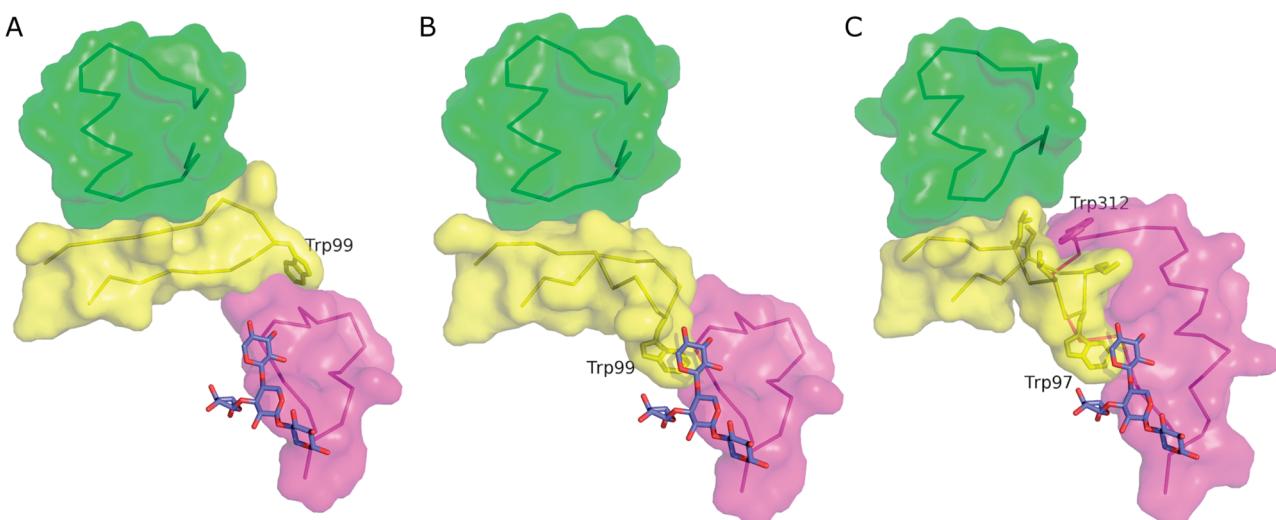


FIGURE 10: View of the $\beta 2\alpha 2$ (yellow) and $\beta 7\alpha 7$ (magenta) loops as well as of the $\beta 4\alpha 4$ loop (green) of a neighboring monomer. The loops are shown as $\text{C}\alpha$ -traces and with a transparent solvent accessible surface. In Tx-Abf, the $\beta 2\alpha 2$ and $\beta 7\alpha 7$ loops are from molecule A and the $\beta 4\alpha 4$ loop from a neighboring molecule B. In Ct-Abf, the $\beta 2\alpha 2$ and $\beta 7\alpha 7$ loops are from molecule A and the $\beta 4\alpha 4$ loop from a neighboring molecule F. A. Tx-Abf open $\beta 2\alpha 2$ loop conformation. B. Tx-Abf closed $\beta 2\alpha 2$ loop conformation with Trp99 shown in stick representation. C. Ct -Abf closed $\beta 2\alpha 2$ loop conformation. Trp97 and Trp312 as well as residues from the $\beta 2\alpha 2$ loop (Leu93, Leu95, Lys98 and Ile100, unlabeled) interacting with Trp312 are shown in stick representation.

the structural data presented here, might not be a common feature of all GH-51 members. Further structural studies and modeling will be necessary to confirm or disprove this hypothesis. Nevertheless, specifically with regard to Tx-Abf, we postulate that the conformation of the $\beta 2\alpha 2$ loop observed in molecule C of the Tx-Abf-Glu176Gln:X4A complex, trapped by crystal contacts, represents the initial binding complex. Here, Trp99 is distant from the +2' subsite and the +2' xylosyl moiety is weakly associated with the enzyme. After, the $\beta 2\alpha 2$ loop moves in and stabilizes the complex (as seen in molecules A and B), prior to the formation of the covalent intermediate. Significantly, in addition to fa-

cilitating the contribution of Trp99 to the stabilization of the complex, the loop movement also allows its neighbor, His98, to interact with the xylosyl moiety in the +2' subsite through a hydrogen bond. Unlike Trp99, His98 is not conserved throughout GH-51, although the nature of this residue appears to be conserved within closely related sequences (>60% similarity). Likewise, while Tx-Abf and its homologues display a histidine at position 98, Abfs related to Gs-Abf are characterized by an alanine at the equivalent position. Consequently, because alanine cannot provide a hydrogen bond to the xylosyl moiety occupying the +2' position, one might assume that Gs-Abf and its homologues do not strictly

speaking possess a +2' subsite and therefore hydrolyze substrates such as X4A less efficiently than Tx-Abf. Furthermore, it is tempting to correlate the presence or absence of a +2' subsite with a mobile or static β 2 α 2 loop.

In conclusion, we propose an induced fit mechanism for Tx-Abf, where flexibility of the β 2 α 2 loop facilitates substrate binding and stabilizes the formation of the covalent intermediate upon hydrolysis. More generally, our results provide an additional demonstration that certain loops in glycoside hydrolases play a vital role for substrate binding, catalysis and/or release of product. The deletion of a thumblike, mobile loop in the GH-11 xylanase from *T. xylanilyticus* has led to the near complete inactivation of the enzyme and a radical alteration in substrate selectivity (37). Therefore, engineering of such loops appears to be a relevant strategy for the alteration of the properties of glycoside hydrolases.

ACKNOWLEDGMENT

The authors would like to thank Hassan Belrhali and Flemming Hansen for help with the data collection at the ESRF (Grenoble, France), the support by MAX-laboratory (Lund, Sweden) and Nathalie Aubry, Béatrice Hermant and Kristoffer Rosenstand for their skillful technical assistance. Drs. Fabrice Fleury and Masayuki Takahashi are thanked for their assistance in fluorescence and circular dichroism measurements.

REFERENCES

- Saha, B. C. (2003) Hemicellulose bioconversion. *J. Ind. Microbiol. Biotechnol.* 30, 279–291.
- Brillouet, J. M., and Joseleau, J. P. (1987) Investigation of the structure of a heteroxylan from the outer pericarp (beeswing bran) of wheat kernel. *Carbohydr. Res.* 159, 109–126.
- Saha, B. C. (2000) α -L-arabinofuranosidases: biochemistry, molecular biology and application in biotechnology. *Biotechnol. Adv.* 18, 403–423.
- Henrissat, B. (1991) A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J.* 280, 309–316.
- Coutinho, P. M., and Henrissat, B. (1999) in *Recent Advances in Carbohydrate Bioengineering* (Gilbert, H. J., Davies, G. J., Henrissat, B., and Svensson, B., Eds.) pp 3–12, The Royal Society of Chemistry, Cambridge.
- Debeche, T., Cummings, N., Connerton, I., Debeire, P., and O'Donohue, M. J. (2000) Genetic and biochemical characterization of a highly thermostable α -L-arabinofuranosidase from *Thermobacillus xylanilyticus*. *Appl. Environ. Microbiol.* 66, 1734–1736.
- Debeche, T., Bliard, C., Debeire, P., and O'Donohue, M. J. (2002) Probing the catalytically essential residues of the α -L-arabinofuranosidase from *Thermobacillus xylanilyticus*. *Protein Eng.* 15, 21–28.
- Shallom, D., Belakhov, V., Solomon, D., S. G -G., Baasov, T., Shoham, G., and Shoham, Y. (2002) The identification of the acid-base catalyst of α -arabinofuranosidase from *Geobacillus stearothermophilus* T-6, a family 51 glycoside hydrolase. *FEBS Lett.* 514, 163–167.
- Shallom, D., Belakhov, V., Solomon, D., Shoham, G., Baasov, T., and Shoham, Y. (2002) Detailed kinetic analysis and identification of the nucleophile in α -L-arabinofuranosidase from *Geobacillus stearothermophilus* T-6, a family 51 glycoside hydrolase. *J. Biol. Chem.* 277, 43667–43673.
- Rémond, C., Plantier-Royon, R., Aubry, N., Maes, E., Bliard, C., and O'Donohue, M. J. (2004) Synthesis of pentose-containing disaccharides using a thermostable α -L-arabinofuranosidase. *Carbohydr. Res.* 339, 2019–2025.
- Rémond, C., Plantier-Royon, R., Aubry, N., and O'Donohue, M. J. (2005) An original chemoenzymatic route for the synthesis of β -D-galactofuranosides using an α -L-arabinofuranosidase. *Carbohydr. Res.* 340, 637–644.
- Euzen, R., Lopez, G., Nugier-Chauvin, C., Ferrières, V., Plusquellec, D., Rémond, C., and O'Donohue, M. J. (2005) A chemoenzymatic approach for the synthesis of unnatural disaccharides containing D-galacto- or D-fucofuranosides. *Eur. J. Org. Chem.* 22, 4860–4869.
- Rémond, C., Boukari, I., Chambat, G., and O'Donohue, M. J. (2008) Action of a GH 51 α -L-arabinofuranosidase on wheat-derived arabinoxylans and arabino-xylooligosaccharides. *Carbohydr. Polym.* 72, 424–430.
- Hövel, K., Shallom, D., Niefind, K., Belakhov, V., Shoham, G., Baasov, T., Shoham, Y., and Schomburg, D. (2003) Crystal structure and snapshots along the reaction pathway of a family 51 α -L-arabinofuranosidase. *EMBO J.* 22, 4922–4932.
- Taylor, E. J., Smith, N. L., Turkenburg, J. P., D'Souza, S., Gilbert, H. J., and Davies, G. J. (2006) Structural insight into the ligand specificity of a thermostable family 51 arabinofuranosidase, Araf 51 from *Clostridium thermocellum*. *Biochem. J.* 395 31–37.
- Hendrickson, W. A., Horton, J. R., and LeMaster, D. M. (1990) Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. *EMBO J.* 9, 1665–1672.
- Kabsch, W. (1993) Automatic processing of rotation diffraction data from crystals of internally unknown symmetry and cell constants. *J. Appl. Crystallogr.* 26, 795–800.
- Evans, P. R. (2005) Scaling and assessment of data quality. *Acta Crystallogr.* D62, 72–82.
- Bahar, M., Ballard, C., Cohen, S. X., Cowtan, K. D., Dodson, E. J., Emsley, P., Esnouf, R. M., Keegan, R., Lamzin, V., Langer, G., Levdkov, V., Long, F., Meier, C., Muller, A., Murshudov, G. N., Perrakis, A., Siebold, C., Stein, N., Turkenburg, M. G. W., Vagin, A. A., Winn, M., Winterb, G., and Wilson, K. S. (2006) SPINE workshop on automated X-ray analysis: a progress report. *Acta Crystallogr.* D62, 1170–1183.
- Britt, B. M. (2004) Understanding enzyme structure and function in terms of the shifting specificity model. *J. Biochem. Mol. Biol.* 37, 394–401.
- Perrakis, A., Harkiolaki, M., Wilson, K. S., and Lamzin, V. S. (2001) ARP/wARP and molecular replacement. *Acta Crystallogr.* D57, 1445–1450.
- Brunger, A. T., Adams, P. D., Clore, G. M., Delano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr.* D54, 905–921.
- Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr.* D60, 2126–2132.
- Adams, P. D., Grosse-Kunstleve, R. W., Hung, L.-W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K., and Terwilliger, T. C. (2002) PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr.* D58, 1948–1954.
- Jones, T. A., Cowan, S., Zou, J.-Y., and Kjeldgaard, M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr.* A47, 110–119.
- Delano, W. L. (2004) DeLano Scientific LLC, SanCarlos, CA, USA, www.pymol.org.
- Benamrouche, S., Crönier, D., Debeire, P., and Chabbert, B. (2002) A chemical and histological study on the effect of (1–4)- β -endo-xylanase treatment on wheat bran. *J. Cereal Sci.* 36, 253–260.
- Lequart, C., Nuzillard, J.-M., Kurek, B., and Debeire, P. (1999) Hydrolysis of wheat bran and straw by an endoxylanase: production and structural characterization of cinnamoyl-oligosaccharides. *Carbohydr. Res.* 319, 102–111.
- Reardon, D., and Farber, G. K. (1995) The structure and evolution of α / β barrel proteins. *FASEB J.* 9, 497–503.
- Anfinsen, C. B. (1973) Principles that govern the folding of protein chains. *Science* 181, 223–230.
- Matsumura, M., Signor, G., and Matthews, B. W. (1989) Substantial increase of protein stability by multiple disulphide bonds. *Nature* 342, 291–293.
- Schwarz, W. H., Bronnenmeir, K., Krause, B., Lottspeich, F., and Staudenbauer, W. L. (1995) Debranching of arabinoxylan: properties of the thermoactive recombinant α -L-arabinofuranosidase from *Clostridium stercorarium*. *Appl. Microbiol. Biotechnol.* 43, 856–860.
- Debeche, T. (2001) PhD Thesis, University of Reims Champagne-Ardenne, Reims, France.

34. Miyanaga, A., Koseki, T., Matsuzawa, H., Wakagi, T., Shoun, H., and Fushinobu, S. (2004) Crystal structure of a family 54 α -L-arabinofuranosidase reveals a novel carbohydrate-binding module that can bind arabinose. *J. Biol. Chem.* 279, 44907–44914.
35. Sulzenbacher, G., Shareck, F., Morosoli, R., Dupont, C., and Davies, G. J. (1997) The *Streptomyces lividans* family 12 endoglucanase: construction of the catalytic core, expression, and X-ray structure at 1.75 Å resolution. *Biochemistry* 36, 16032–16039.
36. Sabini, E., Sulzenbacher, G., Dauter, M., Dauter, Z., Jorgensen, P. L., Schulein, M., Dupont, C., Davies, G. J., and Wilson, K. S. (1999) Catalysis and specificity in enzymatic glycoside hydrolysis: a 2,5B conformation for the glycosyl-enzyme intermediate revealed by the structure of the *Bacillus agaradhaerens* family 11 xylanase. *Chem. Biol.* 6, 483–492.
37. Paës, G., Takahashi, M., Tran, V., Boukari, I., and O'Donohue, M. J. (2007) New insights into the role of the thumb-like loop in GH-11 xylanases. *Protein Eng., Des. Sel.* 20, 15–23.
38. Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., Dufayard, J. F., Guindon, S., Lefort, V., Lescot, M., Claverie, J. M., and Gascuel, O. (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res.* doi: 10.1093/nar/gkn180.

BI800424E