

The structure of a GH10 xylanase from *Fusarium oxysporum* reveals the presence of an extended loop on top of the catalytic cleft

Maria Dimarogona,^{a,b} Evangelos Topakas,^b Paul Christakopoulos^b and Evangelia D. Chrysina^{a*}

^aInstitute of Organic and Pharmaceutical Chemistry, National Hellenic Research Foundation, 48 Vassileos Constantinou Avenue, 11635 Athens, Greece, and ^bSchool of Chemical Engineering, National Technical University of Athens, 5 Iroon Polytechniou Street, Zografou Campus, 15700 Athens, Greece

Correspondence e-mail:
echrysina@eie.gr, echry@tee.gr

Received 4 November 2011
Accepted 16 February 2012

PDB Reference: FoXyn10a,
3u7b.

Xylanase enzymes have been the focus of considerable research in recent decades owing to their extensive use in a variety of biotechnological applications. Previous structural studies of a number of GH10 xylanases revealed that all GH10 family members have the $(\beta/\alpha)_8$ -barrel fold and their catalytic site is conserved. The structure of a new GH10 xylanase from *Fusarium oxysporum* (FoXyn10a) was determined at 1.94 Å resolution from crystals belonging to the tetragonal space group $P4_12_12$ with five molecules per asymmetric unit. Comparison of the structure of FoXyn10a with previously determined structures of GH10 family members indicated that most of the differences were located in the loop regions between the ordered secondary-structure elements of the barrel, as expected. However, alignment of FoXyn10a with sequence and structural homologues denoted an atypically long loop connecting strand $\beta 6b$ and helix $\alpha 6$ that was only present in one other GH10 xylanase, the structure of which is not known. This structural feature may be of functional importance, with potential implications in the catalytic efficiency of the enzyme.

1. Introduction

Endo- β -1,4-xylanases (EC 3.2.1.8) catalyze the hydrolysis of β -1,4-glycosidic bonds in heteroxylans, which are major components of hemicellulose in the biosphere. These biocatalysts are widely used in the pulp and paper, food, textile and pharmaceutical industries (Polizeli *et al.*, 2005). Addition of xylanase during the enzymatic hydrolysis of lignocellulosic materials for the production of fuels and chemicals can also improve the yield of the process by increasing biomass degradability (Alvira *et al.*, 2011). The abovementioned applications require enzymes with increased activity and stability for the economically viable production of products.

Most xylanases are classified into families 5, 8, 10 and 11 of the CAZy database (<http://www.cazy.org>; Cantarel *et al.*, 2009); xylanolytic activity has also been demonstrated for bifunctional or multimodular enzymes belonging to GH families 7, 16, 30, 43 and 62. Some of them include more than one catalytic domain as well as complementary domains such as cellulose-binding and xylan-binding domains or other distinct modules with unknown function (Collins *et al.*, 2005).

GH10 xylanases belong to the GH-A clan and have the typical $(\beta/\alpha)_8$ TIM-barrel fold, which resembles a salad bowl. The crystal structures of 20 GH10 xylanases in the native form or in complex with diverse substrates have provided valuable information on their structure–function relationship (Manikandan *et al.*, 2006; Lo Leggio *et al.*, 2001; Ihsanawati *et al.*, 2005). GH10 xylanases are retaining hydrolases that operate

Table 1Diffraction data and refinement statistics for *FoXyn10a*.

Values in parentheses are for the outermost shell.

Data-collection and processing statistics	
Station	PX beamline X13, EMBL/DESY
Wavelength (Å)	0.8123
No. of images	244
Oscillation range (°)	0.3
Space group	$P4_12_12$
Unit-cell parameters (Å, °)	$a = b = 124.2$, $c = 284.0$, $\alpha = \beta = \gamma = 90$
No. of molecules per asymmetric unit	5
No. of observations	1016139
No. of unique reflections	173419
Resolution range (Å)	17.9–1.9 (2.0–1.9)
Completeness (%)	99.3 (99.2)
$R_{\text{merge}}^{\dagger}$	0.108 (0.343)
$\langle I/\sigma(I) \rangle$	9.1 (5.0)
Multiplicity	5.9 (5.6)
B factor from Wilson plot (Å ²)	13.5
Refinement statistics and model quality	
Resolution range (Å)	17.91–1.94
No. of reflections	154871
No. of protein atoms	12700
No. of solvent molecules	1901
No. of ethylene glycol molecules	26
Chain A	7
Chain B	7
Chain C	6
Chain D	3
Chain E	3
No. of glycans	
Chain A	2 NAG, 1 BMA
Chain B	2 NAG, 1 BMA, 2 MAN
Chain C	2 NAG, 1 BMA
Chain D	2 NAG, 1 BMA, 1 MAN
Chain E	2 NAG, 1 BMA, 1 MAN
$R_{\text{free}}^{\ddagger}$	0.242 (0.360)
R^{\ddagger}	0.213 (0.316)
R.m.s. deviations	
Bond lengths (Å)	0.008
Bond angles (°)	1.069
Average B for protein residues (Å ²)	
Overall	14.6
C $^{\alpha}$, C, N, O	14.2
Side-chain atoms	15.0
Average B for heteroatoms (Å ²)	
Water molecules	25.5
Ethylene glycols	29.1
Glycans	
Chain A	NAG508, 30.7; NAG509, 34.4; BMA600, 38.6
Chain B	NAG508, 19.8; NAG509, 23.5; BMA600, 33.4; MAN601, 42.1; MAN602, 40.6
Chain C	NAG508, 23.0; NAG509, 24.7; BMA600, 26.7
Chain D	NAG508, 23.8; NAG509, 26.7; BMA600, 31.0; MAN601, 41.7
Chain E	NAG508, 19.0; NAG509, 25.2; BMA600, 35.2; MAN601, 44.5

$^{\dagger} R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $\langle I(hkl) \rangle$ and $I_i(hkl)$ are the mean and the i th measurement of intensity for reflection hkl , respectively. ‡ Crystallographic $R = \sum_{hkl} ||F_{\text{obs}}| - |F_{\text{calc}}|| / \sum_{hkl} |F_{\text{obs}}|$, where $|F_{\text{obs}}|$ and $|F_{\text{calc}}|$ are the observed and calculated structure-factor amplitudes, respectively. R_{free} is the corresponding R value for a randomly chosen 5% of the reflections that were not included in the refinement.

via a double-displacement mechanism (Davies & Henrissat, 1995). The two key catalytic glutamate residues, which function as nucleophilic and acid–base residues, respectively, lie at the carboxy-termini of β -strands 4 and 7. The biochemical

characteristics of GH10 family members are determined by the details of the three-dimensional structure rather than the overall fold; the latter is a feature also shared by other enzyme families (Davies & Henrissat, 1995). The extended substrate-binding cleft, which lies on the top face of the molecule, forms a deep groove comprising a series of subsites. The number of subsites varies from four to seven, among which the -2 , -1 and $+1$ subsites are highly conserved (Pollet *et al.*, 2010). The amino acids situated in this region determine the nature of the interactions between the enzyme and xylose units.

Here, we report the three-dimensional structure of a wild-type GH10 endo- β -1,4-xylanase (*FoXyn10a*) from *Fusarium oxysporum*, a filamentous fungus with a strong xylanolytic system from which six xylanases have been characterized to date (Moukoulis *et al.*, 2011). The recently discovered genome-sequence assembly of *F. oxysporum* (<http://www.broad.mit.edu>) revealed seven open reading frames that were recognized *in silico* and could encode putative endo- β -1,4-xylanase members of the GH11 and GH10 families. *FoXyn10a*, previously reported as xylanase III, is stable in the pH range 7–9 and is most active on oat-spelt xylan (Christakopoulos *et al.*, 1997). The crystal structure of *FoXyn10a* determined at 1.94 Å resolution confirmed that it shares the $(\beta/\alpha)_8$ -barrel fold of GH10-family members and provided insights into the role of one of the loops lying above the catalytic cleft. This loop is lined by a total of 11 amino acids, six of which are only present in one other characterized xylanase: that from *Aureobasidium pullulans* ATCC 20524 (Tanaka *et al.*, 2006).

2. Materials and methods

2.1. Xylanase production and purification

The wild-type strain of *F. oxysporum* used in this study was isolated from cumin (Christakopoulos *et al.*, 1989) and was maintained on potato dextrose agar at 277 K. For induction of xylanase activity, the fungus was grown in Braun bioreactors with a working volume of 17 l. The mineral medium was composed of corn cobs (20 g l⁻¹) and (NH₄)₂HPO₄ (10 g l⁻¹) as carbon and nitrogen sources, respectively, as described previously (Topakas & Christakopoulos, 2004). The pH of the medium was controlled automatically at pH 8.0 by the addition of 0.2 M HCl and NaOH solutions. Aeration was adjusted so that the dissolved oxygen concentration was higher than 10% and agitation was at 180 rev min⁻¹. The temperature was also controlled automatically at 303 K. *FoXyn10a* was purified to homogeneity by ion-exchange and gel-filtration chromatography according to previously established protocols (Christakopoulos *et al.*, 1997). Sample purity was assessed by SDS–PAGE, which exhibited a single band with a molecular mass of 38 kDa.

2.2. Sequence analysis

The *FoXyn10a* sequence was compared with sequences available in the general data bank (National Center for Biotechnology Information; NCBI) using *BLAST* on the NCBI server (Altschul *et al.*, 1997) and multiple sequence alignment

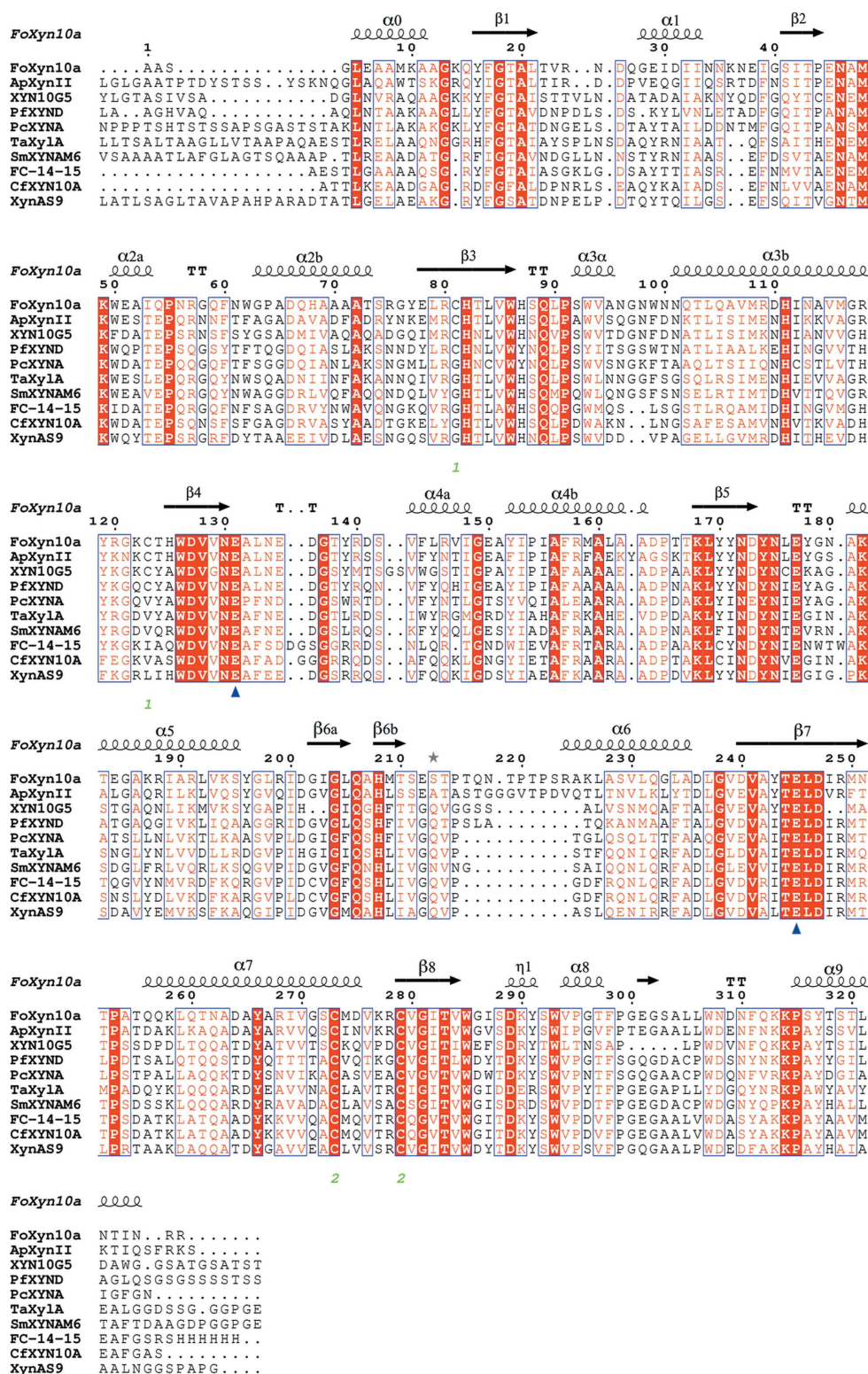


Figure 1

Multiple sequence alignment of *FoXyn10a* (at the top of the alignment) and the ten closest characterized sequence homologues: *ApXynII*, *XYN10G5* (*Phialophora* sp. CGMCC3328; GenBank accession No. ADZ99358.1), *PfxYND* (*Penicillium funiculosum*; GenBank accession No. CAG25554.1), *PcXYNA* (*Phanerochaete chrysosporium*; GenBank accession No. AAG44992.1), *TaXylA* (*Thermobifida alba*; GenBank accession No. CAB02654.1), *SmXYNAM6* (*Streptomyces megasporus*; GenBank accession No. ADE37527.1), *FC-14-15*, *CfxYn10A* and *XynAS9* (*Streptomyces* sp. S9; GenBank accession No. ABX71815.1). Identical and similar residues are shown in white on a red background and in red on a white background, respectively. The secondary-structure elements α -helices, 3_{10} -helices, β -strands and strict β -turns are denoted α , η , β and TT, respectively.

of homologous enzymes was performed with *ClustalW2* (Larkin *et al.*, 2007) on the EBI server (<http://www.ebi.ac.uk/Tools/clustalw2/>) using default values. The sequence alignment was visualized with *ESPrpt2.2* (Gouet *et al.*, 1999) and assignment of secondary structure was performed with *STRIDE* (Frishman & Argos, 1995).

2.3. Crystallization and data collection

A number of crystallization conditions were screened using the sitting-drop method in 96-well plates with the aid of an Oryx Nano crystallization robot (Douglas Instruments Ltd, England) using commercial crystallization screens. Diffracting crystals were obtained by mixing equal volumes (0.5 μ l) of protein solution (15.5 mg ml⁻¹ in 20 mM MES–NaOH pH 6.0) and reservoir solution [0.1 M ammonium acetate, 17%(w/v) PEG 10K, 0.1 M bis-Tris buffer pH 5.5; condition H6 of the JCSG-plus kit from Molecular Dimensions, UK] and equilibrating at 289 K. Prior to data collection, a single *FoXyn10a* crystal was transferred to a cryoprotectant composed of mother liquor and 18% ethylene glycol. Diffraction data were collected at 100 K using synchrotron radiation on beamline X13 ($\lambda = 0.8123$ Å) at EMBL Hamburg Outstation. The crystal diffracted to 1.9 Å resolution and preliminary characterization showed that it belonged to space group *P4₁2₁2*, with unit-cell parameters $a = b = 124.2$, $c = 284.0$ Å and five molecules per asymmetric unit. The collected data were indexed and integrated with *MOSFLM* (Leslie, 1992) and *SCALA* (Evans, 2006) as implemented in the *CCP4* program suite (Winn *et al.*, 2011). Data-collection statistics are summarized in Table 1.

2.4. Structure determination and refinement

The structure of *FoXyn10a* was solved using a model suggested by *BALBES* (Long *et al.*, 2008) based on the three-dimensional structure of *Cellulomonas fimi* xylanase/cellulase Cex (*CfXyn10A*; PDB entry 3cui; White *et al.*, 1994), which shares 43% sequence identity with *FoXyn10a* (98% coverage). Four molecules were identified using *BALBES*; however, additional density indicated the position of a fifth molecule. *Buccaneer* (Cowtan, 2006) was then used to trace the C α positions of the fifth molecule and the new model comprising five molecules was subjected to alternating rounds of refinement with *REFMAC* (Murshudov *et al.*, 2011) and manual model building with *Coot* (Emsley *et al.*, 2010). NCS restraints were used in the early stages of refinement. Additional electron density observed at the N-terminus of the enzyme clearly indicated that the amino-acid chain extended beyond the translated sequence derived from the corresponding gene

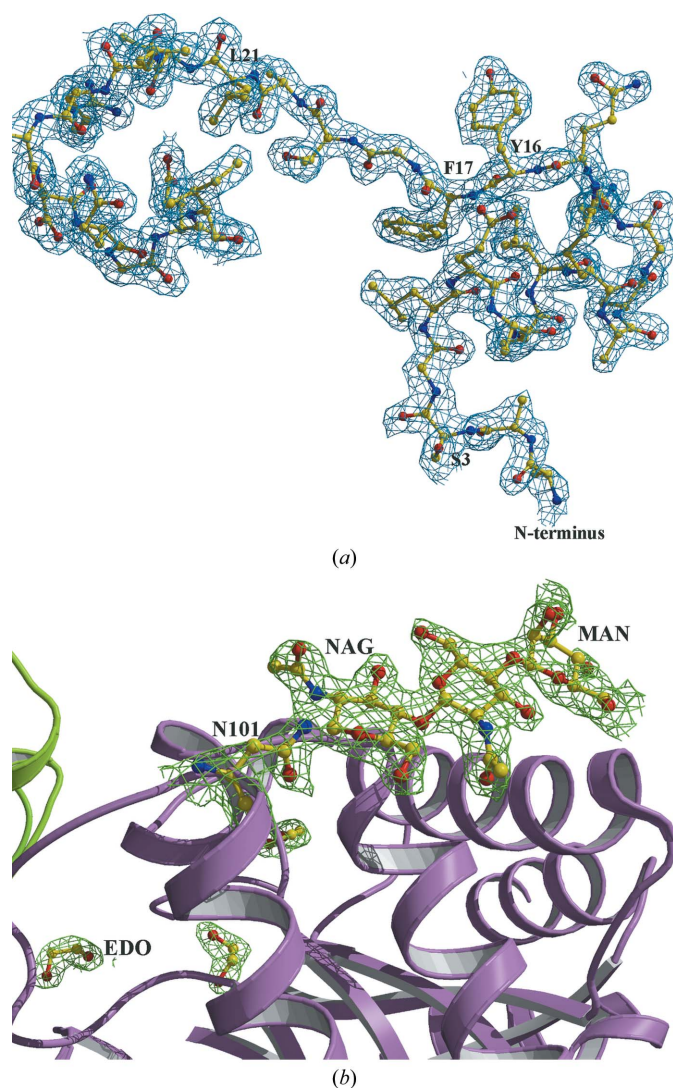


Figure 2
Schematic representation of $2F_o - F_c$ electron-density maps contoured at the 1.0σ level of the 30 N-terminal amino acids (a) and the glycosylated Asn101 (b) of monomer A. The ethylene glycol molecules (EDO) located at the interface of monomers A (mauve) and D (green) are also shown.

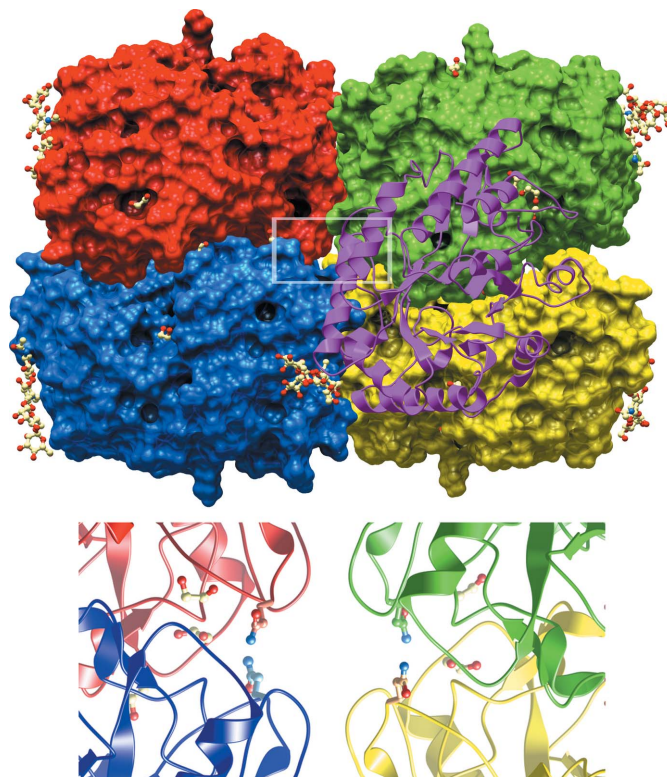


Figure 3
Packing of the five *FoXyn10a* monomers in the asymmetric unit. The molecular surfaces of monomers A (red), B (green), C (yellow) and D (blue) are shown, whereas monomer E (magenta) is drawn in cartoon representation. Inset: the 'meeting point' of subunits A, B, C and D; the positions of Asn25 and the neighbouring EDO molecules are indicated. The figure was prepared with *MolSoft*.

(foxg_17421.2) from the *Fusarium* Comparative Database (*F. oxysporum* Sequencing Project, Broad Institute of Harvard and MIT; http://www.broadinstitute.org/annotation/genome/fusarium_group/MultiHome.html). A total of 24 residues were incorporated into the crystal structure according to multiple sequence alignment of *FoXyn10a* with homologous xylanases (Fig. 1) and the $2F_o - F_c$ and $F_o - F_c$ difference electron-density maps (Fig. 2 and Supplementary Figs. S1 and S2¹). The final structure of the enzyme comprises 327 amino acids in chains A and D and 326 amino acids in chains B, C and E (Fig. 3). As the model improved, visual inspection of the difference maps indicated a large piece of continuous density extending from the side chain of Asn101 in all five monomers. The amino acids Asn101-Gln102-Thr103 formed a sequon for N-glycosylation, and a total of 15 sugars [*N*-acetylglucosamine (NAG), α -D-mannose (MAN) and β -D-mannose (BMA)] in all five monomers were built into the density according to the typical structure of N-linked glycans (Deshpande *et al.*, 2008; Fig. 2). Water molecules were also added using the automated protocol in *Coot* and were only included in the model if they formed hydrogen-bond interactions with protein residues or solvent. Additional density

¹ Supplementary material has been deposited in the IUCr electronic archive (Reference: RR5012). Services for accessing this material are described at the back of the journal.

was also observed for 26 ethylene glycol molecules (EDO), which were also incorporated in the model (Fig. 2). The final structure was refined to an R factor of 0.213 and an R_{free} of 0.242 and the refinement statistics are shown in Table 1.

2.5. Structure analysis

The overall quality of the structure was validated with *MolProbity* (Chen *et al.*, 2010), and analysis of the Ramachandran plot showed that 98.4% of all residues are in the favoured regions (100% are in allowed regions). Comparison of the *FoXyn10a* crystal structure with the structures of both *CfXyn10a* (PDB entry 3cui) as employed by *BALBES* and the β -xylanase chimera of *Streptomyces olivaceoviridis* E-86 FXYN and *C. fimi* Cex (FC-14-15; PDB entry 1v6y; Kaneko *et al.*, 2004), which is the closest sequence homologue of *FoXyn10a* with known crystal structure, was performed using *SUPERPOSE* from the *CCP4* package. The identification of structural homologues was performed with the *DALI* server (Holm & Rosenström, 2010) and multiple structural alignment was performed with *MultiProt* (Shatsky *et al.*, 2004). A schematic representation of the *FoXyn10a* crystal structure was prepared with *MolSoft* (Raush *et al.*, 2009).

2.6. PDB code

The *FoXyn10a* coordinates have been deposited in the PDB with code 3u7b.

3. Results and discussion

3.1. Sequence analysis

A sequence-homology search performed with *BLAST* (Altschul *et al.*, 1997) showed that *FoXyn10a* shares high sequence homology with a large number of both hypothetical and characterized proteins that belong to the GH10 family. The closest known sequence homologue is a *Gibberella zeae* (anamorph *F. graminearum*) hypothetical xylanase (NCBI reference sequence XP_386621.1) with 87% identity (in 98% sequence coverage), while the corresponding characterized homologue is an *A. pullulans* endo- β -1,4-xylanase (*ApXynII*; NCBI reference sequence BAE71410.1) with 62% identity in 98% coverage. A chimeric xylanase of *S. olivaceoviridis* E-86 FXYN and *C. fimi* Cex (FC-14-15; PDB entry 1v6y; Kaneko *et al.*, 2004), which shares 45% identity in 100% coverage with *FoXyn10a*, and *C. fimi* xylanase Cex (*CfXyn10a*; PDB entry 3cui; White *et al.*, 1994), with 43% identity in 98% sequence coverage, are the closest homologues with known crystal structures. The latter also appears to be the closest structural homologue as identified by the *DALI* server. Furthermore, multiple sequence alignment revealed that *FoXyn10a* has an insertion of six amino acids (219–225) that are not present in any of the xylanases characterized to date except for the *A. pullulans* endo- β -1,4-xylanase. However, this insertion has also been identified in some hypothetical xylanases annotated by genome-sequencing efforts.

3.2. Analysis of crystal structure

Diffracting crystals of *FoXyn10a* were grown in the tetragonal space group $P4_12_12$ with five molecules per asymmetric unit and the three-dimensional structure was refined to 1.94 Å resolution. A total of 24 amino acids were incorporated at the N-terminus of each monomer as indicated by the electron-density maps (Fig. 2 and Supplementary Figs. S1 and S2).

These residues are not present in the translated sequence of the foxg_17421.2 gene deposited in the *Fusarium* Comparative Database (http://www.broadinstitute.org/annotation/genome/fusarium_group/MultiHome.html), suggesting that the corresponding genome location was not sequenced properly. The crystal structure of the enzyme includes a total of ten NAG, five BMA, four MAN and 26 EDO molecules (Fig. 2), the presence of which was clearly suggested by the electron-density maps from the first rounds of refinement. N-linked glycans were covalently bonded to Asn101 followed by Gln102 and Thr103, which is a typical glycosylation motif. Binding of ethylene glycol is nonspecific; however, it seems to be conserved in most of the monomers. Although four of the five molecules seem to be in a tetrameric arrangement in the asymmetric unit, there is experimental evidence that *FoXyn10a* is in a monomeric form in solution (Christakopoulos *et al.*, 1997). This was also confirmed by *PISA* analysis for biological assemblies (Krissinel & Henrick, 2007). The fifth monomer is not related by crystallographic symmetry to any of the other monomers in the unit cell.

3.2.1. Overall structure. *FoXyn10a* folds in the classical $(\beta/\alpha)_8$ barrel (TIM barrel), retaining all of the structural features of GH10 xylanases (Fig. 3). Its secondary structure, as assigned by *STRIDE* (Frishman & Argos, 1995), comprises 13 α -helices, 13 β -strands and one 3_{10} -helix (residues 289–291). The characteristic ‘salad-bowl’ shape observed in GH10 xylanases is also present in *FoXyn10a*, with elliptical cross-sections of ~ 40 and ~ 55 Å for the maximum shorter and longer axes, respectively. The cross-sections become more rounded at the lower side of the barrel, with a diameter of ~ 35 Å. The tertiary structure of *FoXyn10a* is stabilized by disulfide bonds formed between cysteine residues Cys81–Cys123 and Cys273–Cys279.

The packing of molecules *A*, *B*, *C* and *D* follows a tetrameric arrangement (Fig. 3) and places them in similar environments, resulting in almost identical conformations as indicated by superposition of the individual monomers onto molecule *A* (average r.m.s. deviation on C^α atoms of ~ 0.13 Å for monomers *B*, *C* and *D*). Monomer *E*, which is more exposed to the solvent, differs slightly in some surface residues and surface loops (r.m.s.d. of ~ 0.25 Å; see Supplementary Material).

3.2.2. Catalytic cleft. The active site of the enzyme is the typical crevice of GH10 xylanases comprising the catalytic glutamic acid residues Glu131 (connecting loop $\beta 4$ – $\alpha 4a$) and Glu245 (strand $\beta 7$) that point towards the interior of the site and play the role of nucleophilic and acid–base catalysts, respectively. In *FoXyn10a* the two C^α and carboxylate C (C^δ) atoms are located at distances of 12.6 and 6.8 Å apart,

respectively, similar to that observed in previously determined structures. Glu131 OE1 and OE2 form hydrogen-bond interactions with Trp86 NE1 and Gln206 NE2, respectively. Similarly, the side chain of Glu245 participates in polar interactions with His208 NE2 and Asn172 ND2 through its carboxyl O atoms OE1 and OE2, respectively. Previous studies of Xyl10A from *S. lividans* (Roberge *et al.*, 1997) suggested that the interaction between the glutamic acid that belongs to strand $\beta 7$ and the corresponding histidine might be essential for the ionization state of the glutamic residue. The catalytic groove is enclosed in an aromatic cage formed by the conserved residues Trp86, Trp285 and Trp293. Both $2F_o - F_c$ and $F_o - F_c$ electron-density maps indicated that there is additional density in all monomers at the catalytic site that is sufficient to accommodate a ligand (Supplementary Fig. S3), except for molecule *E*, where only a small portion is observed. However, we could not assign this density to a known ligand present in the purification process or in the crystallization solution. However, clear evidence for three ethylene glycol molecules bound at the catalytic cleft of monomers *A*, *B*, *C* and *D* was provided, suggesting possible binding subsites for xylose subunits (Schmidt *et al.*, 1998).

3.2.3. Comparison with related xylanase structures. *FoXyn10a* shares a common fold in its overall structure with all members of the GH10 family, while differences are observed in loop regions, as in previously determined GH10 xylanase structures, and in some secondary-structure elements; however, these do not affect the overall fold. Our model was compared with both unliganded FC-14-15 (PDB entry 1v6y) and the complex of *CfXyn10a* with thio-linked xylotetraose (PDB entry 3cui) by superimposing secondary-structure elements. FC-14-15 is the closest sequence homologue with an available crystal structure (Fig. 1), while *CfXyn10a* was selected by *BALBES* to construct the molecular-replacement model.

The r.m.s.d. over all C^α atoms between the FC-14-15 and *FoXyn10a* structures, excluding loop regions, was 1.8 Å. Visual inspection of the two structures revealed considerable changes, mostly at the periphery of the molecule. More specifically, the loop lined by residues 134–138 appears to be shifted by ~ 2.5 Å in FC-14-15 owing to the presence of two additional amino acids (Fig. 1). Major alterations are also observed in the $\alpha 5$ helix formed by residues 181–196 and the neighbouring loop (shifts in the approximate range 1.5–7 Å). This new conformation could be further stabilized *via* hydrogen bonds formed to a neighbouring BMA residue; however, this additional sugar residue was not included in the final model owing to insufficient electron density. Residues 22–39 form an α -helix and a loop that also differ significantly when compared with the corresponding structural elements of FC-14-15 (C^α -atom shifts of ~ 2 –7 Å). The loop lying between helices $\alpha 3a$ and $\alpha 3b$ is more extended in the structure of *FoXyn10a* owing to the presence of two additional amino acids (Fig. 1).

With the aim of identifying changes in the amino-acid residues that are implicated in substrate binding, the functionally ‘important’ amino acids of *FoXyn10a* were assigned based on the superimposed crystal structures of homologous xylanases complexed with various substrate analogues. These residues include Glu45 and Trp285 for the -2 subsite; Lys49, His82, Trp86, Asn130, Glu131, Gln206, Glu245, Trp285 and Trp293 for the -1 subsite; Asn46 for the -3 subsite; and Tyr178 and His208 for the $+1$ subsite (Vardakou *et al.*, 2005; Schmidt *et al.*, 1999; Ducros *et al.*, 2000).

Visual inspection of *FoXyn10a* superimposed on the FC-14-15 structure did not reveal any dramatic alterations in residues located at the active site except for Trp293 and the residues in close proximity. Trp293, which is situated in the eighth $\beta\alpha$ -loop of the barrel, is a highly conserved residue that is implicated in substrate binding at subsite -1 . Its functional importance and

side-chain mobility have been reported in previous studies (Teixeira *et al.*, 2001). In the *FoXyn10a* structure the side chain of this amino acid is subjected to a considerable change in its rotameric state in comparison to the FC-14-15 structure owing to the partial binding of a ligand at the active sites of the five monomers. This change induces further rearrangement in the neighbouring residues 294–300 (the C^α atoms shift by ~ 1.5 Å on average). Rearrangement at this subsite is also supported by the tight packing interactions in the asymmetric unit (see Supplementary Material).

Comparison of the *FoXyn10a* and the *CfXyn10a* structures (r.m.s.d. over C^α atoms of 1.2 Å, excluding loop regions) showed analogous alterations to those recorded for FC-14-15 (Fig. 4). The side chain of Trp293 is subjected to less pronounced changes in its conformation

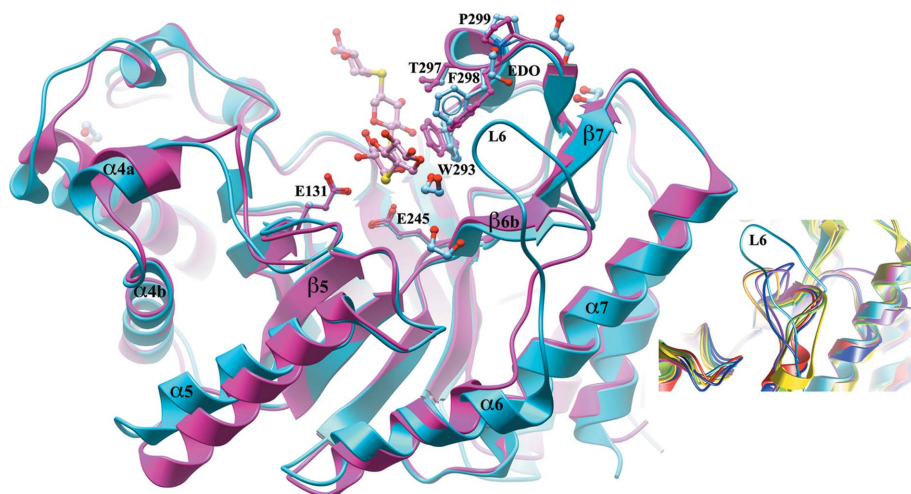


Figure 4

Superposition of the crystal structures of *FoXyn10a* (cyan) and *CfXyn10A* complexed with thio-linked xylotetraose (magenta). The two catalytic glutamic acid residues, Trp293 and amino acids in its vicinity and ethylene glycol molecules, among which EDO400 is denoted EDO, are highlighted. Inset: multiple structural alignment of *FoXyn10a* with the 11 closest structural homologues identified by the *DALI* server focused on the extended-loop (L6) region.

in the presence of thio-linked xylotetraose compared with the *FoXyn10a* crystal structure. An additional modification was identified in the backbone atoms of residues 138–150 (the C α atoms shift by ~ 4.0 Å).

Structural alignment against all GH10 xylanases with known structures revealed that although the active site seems to be almost identical in all the aligned structures, *FoXyn10a* features an extended loop (between $\beta 6b$ and $\alpha 6$) comprising residues 211–223. In the *FoXyn10a* structure this loop (L6) is involved in intersubunit interactions formed between monomers *A* and *D* as well as between *B* and *C* (residues Asn56 and Pro215). The crystal structure reveals that this loop region lies opposite Gln89, which also seems to play some role in ligand binding.

3.2.4. Loop L6 and its possible implication in enzyme function. The residues that belong to this elongated loop are located on top of the catalytic cleft, pointing towards the ‘+’ subsites, suggesting that this loop could play a functional role in forming interactions with the substrate (Fig. 4). The closed conformation is stabilized through the formation of hydrogen-bond interactions with an ethylene glycol molecule (EDO400) and Pro299. The binding of EDO might be indicative of the possibility that a xylose subunit could be accommodated here. This extended loop also induces changes in residues 190–198, which are subjected to shifts in their backbone atoms ranging from ~ 1.0 to ~ 2.5 Å.

To our knowledge, the only characterized family 10 xylanase that has a similar amino-acid insertion is XynII from *A. pullulans* (Tanaka *et al.*, 2006), which is also a very close sequence homologue of *FoXyn10a* (Fig. 1). Interestingly, according to Tanaka and coworkers XynII belongs to neither of the clusters (clusters I and II) that were previously identified as distinct phylogenetic categories of GH10 xylanases (Sato *et al.*, 1999). More specifically, XynII seems to have diverged from cluster I and features an amino-acid insertion that is absent in xylanases of both clusters I and II. Both xylanases exhibit common biochemical characteristics such as a high pI (8.9 and 9.5 for *A. pullulans* xylanase and *FoXyn10a*, respectively); the majority of fungal GH10 xylanases exhibit an acidic pI value (Wong *et al.*, 1988; Lafond *et al.*, 2011). In addition, the pH optimum of both enzymes ranges from 6 to 8 and they are also stable under alkaline conditions.

The importance of the loop in the function and biochemical properties of the xylanase has yet to be elucidated since there are no other sequence homologues with known crystal structures that bear the same loop.

4. Conclusions

The crystal structure of a new GH10 xylanase from *F. oxysporum* determined at 1.94 Å resolution adopts the characteristic (β/α) $_8$ fold of GH10 family members. The tight packing interactions of four of the five *FoXyn10a* monomers in the asymmetric unit induced minor alterations that were mainly observed in the loop regions of the barrel, in accordance with previously determined GH10 xylanases. The most striking difference between *FoXyn10a* and its sequence

homologues is an insertion comprising six amino acids in the sixth $\beta\alpha$ -loop. Superposition of *FoXyn10a* with its structural homologues showed that this extended loop region faces the aglycon side of the active-site cleft. Of the previously studied GH10 family members, only one, a xylanase from *A. pullulans*, shares the additional amino acids that compose this loop; however, its structure has not been determined. The functional implication of this loop in terms of enzyme biochemical properties requires further investigation of both *FoXyn10a* mutants and complexes with substrate analogues, and structural studies of homologous GH10 xylanases that feature this unique characteristic.

The Hellenic State Scholarships Foundation is acknowledged for funding MD's position. This work was supported by a Marie Curie Host Fellowship for the Transfer of Knowledge (ToK; contract No. MTKD-CT-2006-042776) under FP6, and the FP7 Capacities Coordination and Support Actions REGPOT-2008-1 No. 230146 ‘EUROSTRUCT’ and REGPOT-2009-1 No. 245866 ‘ARCADE’. The research leading to these results received funding from the European Community's Sixth (RII3/CT/2004/5060008, IHPP HPRI-CT-1999-00012) and Seventh Framework Programmes (FP7/2007-2013) under grant agreement No. 226716. The EMBL staff at PX beamline X13 at the DORIS storage ring, DESY, Hamburg are also acknowledged for their help during data collection.

References

- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). *Nucleic Acids Res.* **25**, 3389–3402.
- Alvira, P., Tomás-Pejó, E., Negro, M. J. & Ballesteros, A. M. (2011). *Biotechnol. Prog.* **27**, 944–950.
- Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V. & Henrissat, B. (2009). *Nucleic Acids Res.* **37**, D233–D238.
- Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* **D66**, 12–21.
- Christakopoulos, P., Macris, B. J. & Kekos, D. (1989). *Enzyme Microb. Technol.* **11**, 236–239.
- Christakopoulos, P., Nerinckx, W., Kekos, D., Macris, B. & Claeysens, M. (1997). *Carbohydr. Res.* **302**, 191–195.
- Collins, T., Gerday, C. & Feller, G. (2005). *FEMS Microbiol. Rev.* **29**, 3–23.
- Cowtan, K. (2006). *Acta Cryst.* **D62**, 1002–1011.
- Davies, G. & Henrissat, B. (1995). *Structure*, **3**, 853–859.
- Deshpande, N., Wilkins, M. R., Packer, N. & Nevalainen, H. (2008). *Glycobiology*, **18**, 626–637.
- Ducros, V., Charnock, S. J., Derewenda, U., Derewenda, Z. S., Dauter, Z., Dupont, C., Shareck, F., Morosoli, R., Kluepfel, D. & Davies, G. J. (2000). *J. Biol. Chem.* **275**, 23020–23026.
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* **D66**, 486–501.
- Evans, P. (2006). *Acta Cryst.* **D62**, 72–82.
- Frishman, D. & Argos, P. (1995). *Proteins*, **23**, 566–579.
- Gouet, P., Courcelle, E., Stuart, D. I. & Métoz, F. (1999). *Bioinformatics*, **15**, 305–308.
- Holm, L. & Rosenström, P. (2010). *Nucleic Acids Res.* **38**, W545–W549.
- Ihsanawati, Kumasaka, T., Kaneko, T., Morokuma, C., Yatsunami, R., Sato, T., Nakamura, S. & Tanaka, N. (2005). *Proteins*, **61**, 999–1009.

- Kaneko, S., Ichinose, H., Fujimoto, Z., Kuno, A., Yura, K., Go, M., Mizuno, H., Kusakabe, I. & Kobayashi, H. (2004). *J. Biol. Chem.* **279**, 26619–26626.
- Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.
- Lafond, M., Tauzin, A., Desseaux, V., Bonnin, E., Ajandouz, H. & Giardina, T. (2011). *Microb. Cell Fact.* **10**, 20.
- Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J. & Higgins, D. G. (2007). *Bioinformatics*, **23**, 2947–2948.
- Leslie, A. G. W. (1992). *Jnt CCP4/ESF-EACMB Newsl. Protein Crystallogr.* **26**.
- Lo Leggio, L., Kalogiannis, S., Eckert, K., Teixeira, S. C., Bhat, M. K., Andrei, C., Pickersgill, R. W. & Larsen, S. (2001). *FEBS Lett.* **509**, 303–308.
- Long, F., Vagin, A. A., Young, P. & Murshudov, G. N. (2008). *Acta Cryst. D* **64**, 125–132.
- Manikandan, K., Bhardwaj, A., Gupta, N., Lokanath, N. K., Ghosh, A., Reddy, V. S. & Ramakumar, S. (2006). *Protein Sci.* **15**, 1951–1960.
- Moukoulis, M., Topakas, E. & Christakopoulos, P. (2011). *N. Biotechnol.* **28**, 369–374.
- Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst. D* **67**, 355–367.
- Polizeli, M. L., Rizzatti, A. C., Monti, R., Terenzi, H. F., Jorge, J. A. & Amorim, D. S. (2005). *Appl. Microbiol. Biotechnol.* **67**, 577–591.
- Pollet, A., Delcour, J. A. & Courtin, C. M. (2010). *Crit. Rev. Biotechnol.* **30**, 176–191.
- Raush, E., Totrov, M., Marsden, B. D. & Abagyan, R. (2009). *PLoS One*, **4**, e7394.
- Roberge, M., Shareck, F., Morosoli, R., Kluepfel, D. & Dupont, C. (1997). *Biochemistry*, **36**, 7769–7775.
- Sato, Y., Niimura, Y., Yura, K. & Go, M. (1999). *Gene*, **238**, 93–101.
- Schmidt, A., Gübitz, G. M. & Kratky, C. (1999). *Biochemistry*, **38**, 2403–2412.
- Schmidt, A., Schlacher, A., Steiner, W., Schwab, H. & Kratky, C. (1998). *Protein Sci.* **7**, 2081–2088.
- Shatsky, M., Nussinov, R. & Wolfson, H. J. (2004). *Proteins*, **56**, 143–156.
- Tanaka, H., Muguruma, M. & Ohta, K. (2006). *Appl. Microbiol. Biotechnol.* **70**, 202–211.
- Teixeira, S., Lo Leggio, L., Pickersgill, R. & Cardin, C. (2001). *Acta Cryst. D* **57**, 385–392.
- Topakas, E. & Christakopoulos, P. (2004). *World J. Microbiol. Biotechnol.* **20**, 245–250.
- Vardakou, M., Flint, J., Christakopoulos, P., Lewis, R. J., Gilbert, H. J. & Murray, J. W. (2005). *J. Mol. Biol.* **352**, 1060–1067.
- White, A., Withers, S. G., Gilkes, N. R. & Rose, D. R. (1994). *Biochemistry*, **33**, 12546–12552.
- Winn, M. D. *et al.* (2011). *Acta Cryst. D* **67**, 235–242.
- Wong, K. K. Y., Tan, L. U. L. & Saddler, J. N. (1988). *Microbiol. Rev.* **52**, 305–317.