Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

The accessory domain changes the accessibility and molecular topography of the catalytic interface in monomeric GH39 β-xylosidases

Camila Ramos Santos,[a] Carla Cristina Polo,[a] Juliana Moço Corrêa,[b] Rita de Cássia Garcia Simão,[b] Flavio Augusto Vicente Seixas[c] and Mario Tyago Murakami[a]*

[a]Laboratório Nacional de Biociências (LNBio), Centro Nacional de Pesquisa em Energia e Materiais, Rua Giuseppe Maximo Scolfaro, 10000, Campinas-SP 13083-970, Brazil, [b]Centro de Ciências Médicas e Farmacêuticas, Universidade Estadual do Oeste do Parana, Rua Universitária, 2069, Cascavel-PR 85814-110, Brazil, and [c]Departamento de Bioquímica, Universidade Estadual de Maringá, Avenida Colombo, 5790, Maringá-PR 87020-900, Brazil

Correspondence e-mail: mario.murakami@lnbio.org.br

Received 12 April 2012
Accepted 23 June 2012

PDB Reference: CcXynB2, 4ekj

β-Xylosidases (EC 3.2.1.37) are among the principal glycosyl hydrolases involved in the breakdown of hemicelluloses, catalyzing the reduction of xylooligosaccharides to free xylose. All GH39 β-xylosidases structurally characterized to date display a modular multi-domain organization that assembles a tetrameric quaternary structure. In this work, the crystal structure and the SAXS molecular envelope of a new GH39 β-xylosidase from *Caulobacter crescentus* (CcXynB2) have been determined. Interestingly, CcXynB2 is a monomer in solution and comparative structural analyses suggest that the shortened C-terminus prevents the formation of a stable tetramer. Moreover, CcXynB2 has a longer loop from the auxiliary domain (the long α-helix-containing loop) which makes a number of polar and hydrophobic contacts with the parental $(\alpha/\beta)_8$-barrel domain, modifying the accessibility and the molecular topography of the catalytic interface. These interactions also maintain the accessory domain tightly linked to the catalytic core, which may be important for enzyme function and stability.

1. Introduction

Xylan, the main hemicellulosic polysaccharide in plant cell walls, is comprised of a backbone of β-1,4-linked D-xylopyranosyl residues, which can be decorated with glucuronosyl, arabinosyl and acetyl moieties. The high molecular and structural complexity of xylan represents a natural physico-chemical barrier for enzymatic degradation and its reduction to simple sugars requires the synergistic action of several hemicellulases with distinct specificities (Shallom & Shoham, 2003). In this scenario, endoxylanases (EC 3.2.1.8) and β-xylosidases (EC 3.2.1.37) play a key role by catalyzing the cleavage of internal and terminal β-1,4-glycosidic bonds in the xylan backbone, respectively.

Depending on the nature and the type of modifications that are present in xylan, other enzymes are required to complete degradation, including acetyl xylan esterases (EC 3.1.1.6), feruloyl esterases (EC 3.1.1.73), α-arabinofuranosidases (EC 3.2.1.55) and α-glucuronidases (EC 3.2.1.31) (Biely, 1985). In addition to the vital function of hemicellulases in the carbon cycle in nature, these enzymes also have tremendous potential for biotechnological applications, such as, for example, in the paper and pulp industry and in the bioconversion of lignocellulosic material to fermentable sugars (Beg *et al.*, 2001; Galbe & Zacchi, 2002).

The essential function of β-xylosidases in the final stage of xylan degradation, the reduction of xylooligosaccharides to

free xylose, has attracted much attention from both academia and industry and is reflected by the increasing number of new β-xylosidases that have been identified and characterized from bacteria and fungi (Jordan & Wagschal, 2010; Lagaert *et al.*, 2011). Recent studies have shown that the soluble xylo-oligomers released from hemicelluloses during enzymatic hydrolysis could present an additional obstacle to enzyme action by competitively inhibiting cellulase activity (Qing *et al.*, 2010), highlighting the importance of β-xylosidases in the process of the conversion of biomass into ethanol.

To date, β-xylosidases have been found to belong to ten GH families, GH1, GH3, GH30, GH39, GH43, GH51, GH52, GH54, GH116 and GH120, according to the CAZY classification (http://www.cazy.org/Glycoside-Hydrolases.html). The β-xylosidases belonging to families GH1, GH3, GH30, GH39 and GH51 display a typical $(\alpha/\beta)_8$-barrel fold, whereas GH43 and GH54 family members show fivefold β-propeller and β-trefoil architectures, respectively. The members of families GH52, GH116 and GH120 remain structurally uncharacterized.

The dimorphic bacterium *Caulobacter crescentus* predominantly inhabits freshwater ponds, lakes and rivers (Poindexter, 1964), where carbon sources from the degradation of plant materials may be relatively abundant (Hottes *et al.*, 2004). Recent studies of *Caulobacter* ssp. have led to novel findings regarding the uptake and catabolism of various carbon sources (Neugebauer *et al.*, 2005; Boutte *et al.*, 2008), unveiling new routes and molecular mechanisms that might be exploited for biotechnological purposes. Moreover, several genes related to polysaccharide degradation have been identified in these organisms; five of these genes encode β-xylosidases (Nierman *et al.*, 2001). Despite their important physiological role, biochemical and structural data are very scanty.

Here, we describe crystallographic and small-angle X-ray scattering (SAXS) analyses of a GH39 β-xylosidase from *C. crescentus* NA1000 (CcXynB2), which in contrast to the other tetrameric GH39 β-xylosidases characterized to date is a monomer in solution. Interestingly, the crystal structure revealed that the accessory domain in CcXynB2 participates in the modelling of the catalytic interface, displacing the catalytic β-hairpin motif and consequently increasing accessibility to the active-site pocket.

2. Materials and methods

2.1. Protein production and purification

The β-xylosidase encoded by the CCNA_02442 gene from *C. crescentus* NA1000 cloned into pPROEX-HTa vector (Invitrogen) was expressed in *Escherichia coli* BL21 (DE3) ΔSlyD+pRARE2 cells at 298 K. Briefly, the *E. coli* cells harbouring the expression vector were grown in selective LB medium to an $OD_{600\ nm}$ of 0.8 and 0.5 m*M* IPTG was added to induce heterologous expression for 12 h. The harvested cells were lysed and the enzyme was isolated from the soluble fraction by nickel-affinity and size-exclusion chromatography. Sample quality was assessed by polyacrylamide gel electrophoresis under denaturing conditions (Laemmli, 1970). Prior to crystallization, the sample was dialyzed against 25 m*M* Tris–HCl pH 7.5 and concentrated to 7 mg ml^{-1} using Amicon centrifugal ultrafiltration units (Millipore).

Table 1
Data-processing and refinement statistics for the CcXynB2 crystal structure.

Values in parentheses are for the highest resolution shell.

Data collection	
Space group	$P4_32_12$
Unit-cell parameters (Å)	$a = b = 71.19$, $c = 226.71$
Resolution (Å)	24.7–2.50 (2.59–2.50)
R_{merge}† (%)	12.3 (51.9)
$\langle I/\sigma(I)\rangle$	5.65 (2.05)
Completeness (%)	97.5 (97.2)
Multiplicity	4.0 (3.8)
Mosaicity (°)	0.75
Refinement	
Resolution (Å)	24.7–2.50
No. of unique reflections	20250
R_{work}/R_{free} (%)	18.8/24.8
Mean *B* factor (Å^2)	30.14
No. of residues	492
No. of ligands (sulfate ions)	5
No. of water molecules	235
R.m.s. deviations	
Bond lengths (Å)	0.009
Bond angles (°)	1.277
Ramachandran plot	
Favoured region (%)	94.3
Allowed region (%)	4.9
Disallowed region (%)	0.8

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the *i*th observation of reflection *hkl* and $\langle I(hkl)\rangle$ is the weighted average intensity for all *i* observations of reflection *hkl*.

2.2. Enzyme assay

CcXynB2 activity was tested on different xylooligosaccharides, including xylopentaose, xylotriose and xylobiose. The reaction mixture consisted of 100 µl purified enzyme solution (1 U β-xylosidase activity), 250 µl substrate (4 m*M*) and 650 µl 50 m*M* McIlvaine buffer (pH 6.0) and was incubated at 323 K for 6 h. The increase in reducing sugars released was determined using the DNS method (Miller, 1959).

2.3. X-ray crystallography

2.3.1. Crystallization. Crystallization experiments were performed by the sitting-drop vapour-diffusion method using a Cartesian HoneyBee 963 system (Genomic Solutions). 544 different formulations based on commercial crystallization kits from Hampton Research (SaltRX, Crystal Screen and Crystal Screen 2), Emerald BioSystems (Precipitant Synergy and Wizard I and II) and Qiagen/NeXtal (PACT and JCSG+) were tested. For initial screening, 0.5 µl protein solution at a concentration of 7 mg ml^{-1} in 25 m*M* Tris–HCl buffer pH 7.5 was mixed with an equal volume of the screening solution and equilibrated against a reservoir containing 80 µl of the latter solution at 291 K. Crystals appeared after 5 d in a condition consisting of 0.1 *M* MES pH 6.5, 1.8 *M* ammonium sulfate, 10 m*M* cobalt chloride. This condition was optimized and large crystals (300 × 200 × 50 µm) were obtained using 1%(*v*/*v*)

dioxane as an additive and by decreasing the precipitant concentration to 1.5 *M*.

2.3.2. X-ray diffraction data collection and processing. Crystals were transferred into a cryosolution consisting of the reservoir solution supplemented with 20%(*v*/*v*) glycerol and were flash-cooled in a nitrogen-gas stream at 100 K for data collection. X-ray diffraction data were collected on the MX2 beamline (LNLS, Campinas, Brazil) with the radiation wavelength set to 1.458 Å. A total of 360° images were collected using an oscillation angle of 0.5° and an exposure time of 60 s per image. The crystal-to-detector distance was set to 150 mm, which resulted in a maximum resolution of 2.3 Å. A MAR Mosaic 225 mm (MAR Research) charge-coupled device (CCD) was used to record the intensities. Data were indexed and scaled using *DENZO* and *SCALEPACK* from the *HKL*-2000 package (Otwinowski & Minor, 1997). Data-collection statistics are summarized in Table 1.

2.3.3. Structure determination and refinement. The crystal structure of CcXynB2 was solved by molecular-replacement calculations using the program *Phaser* (McCoy *et al.*, 2007) with the GH39 β-xylosidase Xynb1 from *Geobacillus stearothermophilus* as the template (PDB entry 1w91; Czjzek *et al.*, 2005). The best MR solution resulted in a poor-quality model with an $R_{\rm work}$ of 49% and an $R_{\rm free}$ of 55% that required extensive manual chain tracing,

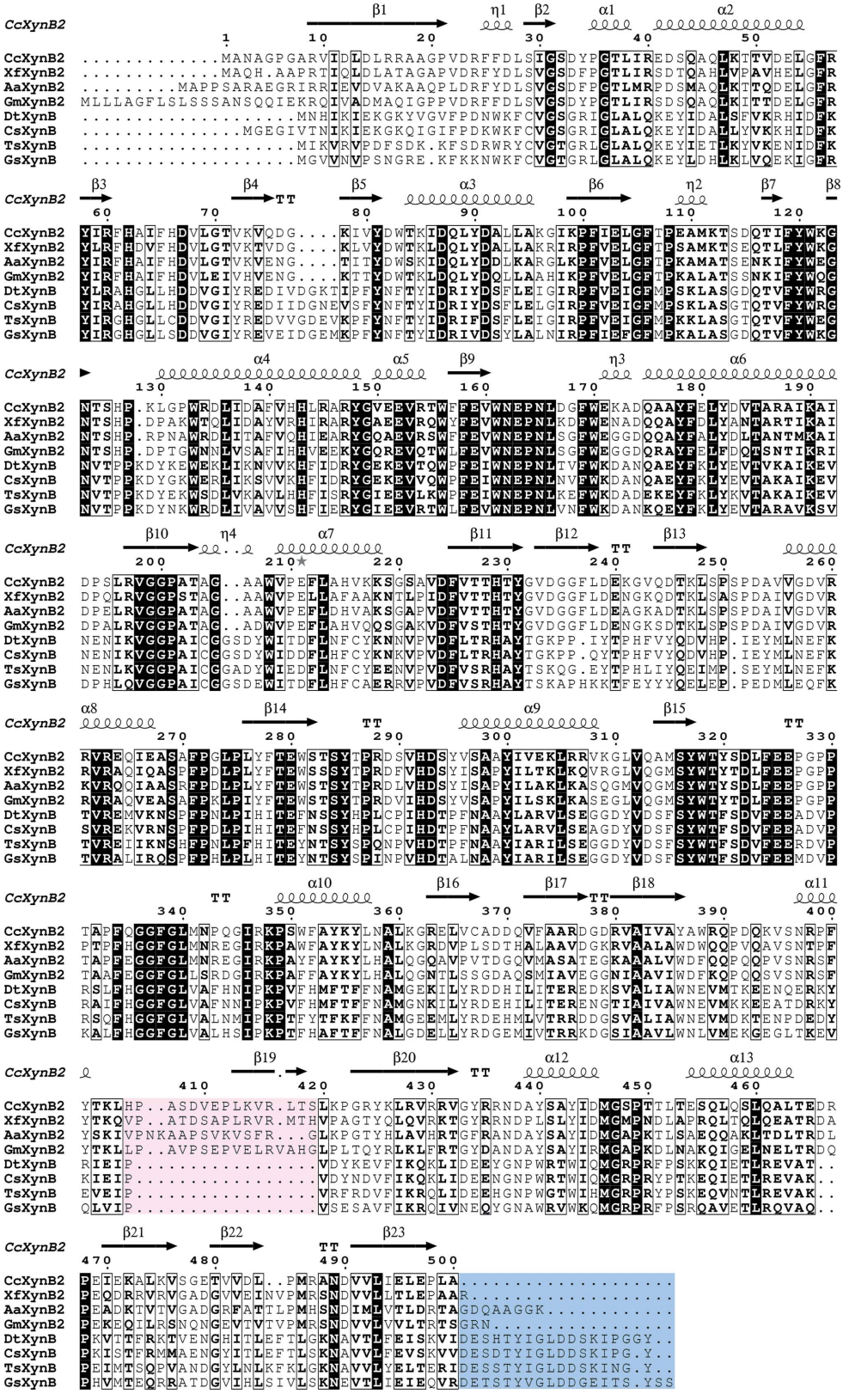

Figure 1
Sequence alignment of β-xylosidases from *Caulobacter crescentus* NA1000 (CcXynB2; GenBank ID ACL95907), *Xanthomonas fuscans* subsp. *aurantifolii* strain ICPB 10535 (XfXynB2; GenBank ID EFF48803), *Azospirillum amazonense* Y2 (AaXynB2; GenBank ID EGY01894), *Granulicella mallensis* MP5ACTX8 (GmXynB2; GenBank ID AEU37894), *Dictyoglomus thermophilum* H-6-12 (DtXynB; GenBank ID ACI18502), *Caldicellulosiruptor saccharolyticus* DSM 8903 (CsXynB; GenBank ID ABP67986), *Thermoanaerobacterium saccharolyticum* DSM 571 (TsXynB; PDB entry 1px8) and *Geobacillus stearothermophilus* (GsXynB; PDB entry 1w91). The regions shaded in magenta and light blue correspond to the long α-helix-containing loop from the accessory domain and the C-terminal extension, respectively.

mainly of the C-terminal subdomain. Refinement cycles in the resolution range 24.7–2.50 Å involved restrained and overall *B*-factor refinement using the program *REFMAC*5 (Murshudov *et al.*, 2011). After each cycle of refinement the model was inspected and manually adjusted to correspond to computed σ_A-weighted ($2F_o - F_c$) and ($F_o - F_c$) electron-density maps using the program *Coot* (Emsley & Cowtan, 2004). Water molecules were manually added at positive peaks above 2.0σ in the difference Fourier maps, taking into consideration hydrogen-bonding distances. The refined structure was evaluated using the program *MolProbity* (Chen *et al.*, 2010). Refinement statistics are summarized in Table 1.

2.3.4. PDB deposition. The structure factors and atomic coordinates of CcXynB2 have been deposited in the RCSB Protein Data Bank under entry code 4ekj.

2.4. Small-angle X-ray scattering

2.4.1. SAXS analysis. SAXS measurements for CcXynB2 at 2, 4 and 6 mg ml^{-1} in 20 m*M* Tris buffer pH 7.0 were performed on the D02A-SAXS2 beamline (LNLS, Campinas, Brazil). The radiation wavelength was set to 1.488 Å and X-ray scattering was recorded using a MAR CCD 165 mm detector (MAR Research, USA). The sample-to-detector distance was adjusted to a scattering-vector range of $0.02 < q < 0.28$ Å^{-1}, where q is the magnitude of the q-vector defined by $q = 4\pi\sin\theta/\lambda$ (2θ is the scattering angle). Two successive frames of 300 s each were recorded for each sample to monitor radiation damage. Buffer scattering was recorded and was subtracted from the corresponding protein scattering. The integration of the SAXS patterns was performed using *Fit2D* (Hammersley *et al.*, 1996).

2.4.2. Data analysis. Data were analyzed using the *GNOM* package (Svergun, 1992). The radius of gyration (R_g) was estimated by the indirect Fourier transform method and the distance distribution function $p(r)$ was calculated from the scattering curve using the maximum diameter (D_{max}) as a parameter.

2.4.3. *Ab initio* modelling. Molecular envelopes were calculated from the experimental SAXS data using the program *DAMMIN* (Svergun, 1999). Ten runs of *ab initio* shape determination yielded highly similar models [normalized spatial discrepancy (NSD) values of <1], which were then averaged using the *DAMAVER* package (Volkov & Svergun, 2003).

2.4.4. Fitting of SAXS and crystallographic structures. The theoretical scattering curve of the crystallographic structure was calculated and compared with the experimental SAXS curve using the program *CRYSOL* (Svergun *et al.*, 1995). The crystallographic structure was fitted into the SAXS molecular envelope using the program *SUPCOMB* (Kozin & Svergun, 2001).

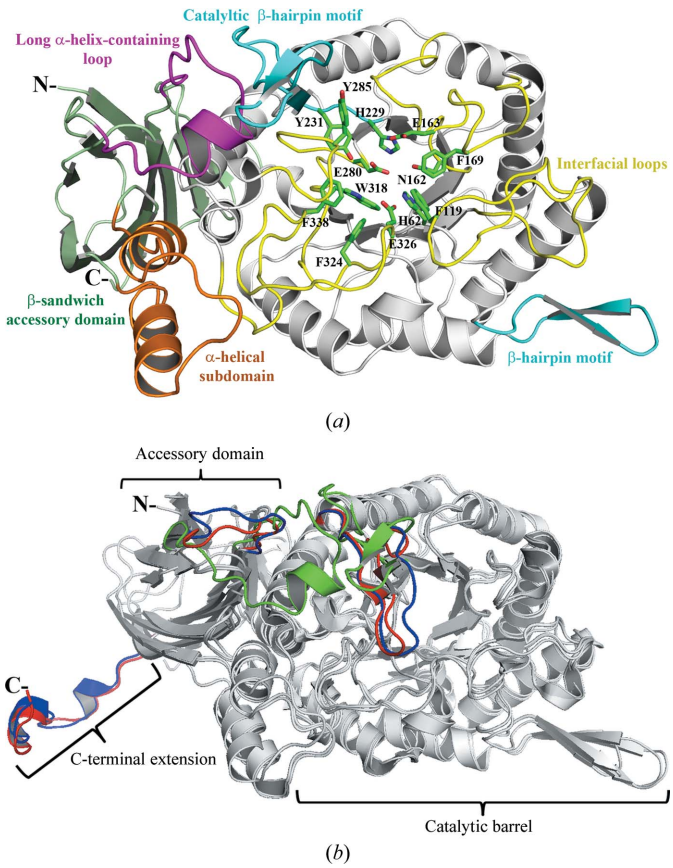

Figure 2
Crystallographic structure of CcXynB2. (*a*) Cartoon representation of the tertiary structure highlighting the catalytic domain (white), accessory domain (green), α-helical subdomain (orange), long α-helix-containing loop (magenta), catalytic β-hairpin (light blue) and interfacial loops (yellow). The residues involved in substrate recognition are represented as sticks with C atoms coloured green. (*b*) Structural superposition of CcXynB2 on GsXynB and TsXynB with significant structural differences coloured green, blue and red, respectively.

3. Results and discussion

3.1. Overall structure

Sequence analysis indicated that the protein encoded by the CC_2357 gene, also named XynB2, from *C. crescentus* NA1000 (CcXynB2) is a β-xylosidase that belongs to glycosyl hydrolase family 39 (GH39). An enzyme-activity assay confirmed that CcXynB2 was a β-xylosidase (Supplementary Table S1[1]). CcXynB2 was able to hydrolyze typical substrates of β-xylosidases such as xylobiose, xylotriose and xylopentaose with a 67, 69 and 93% increase in released reducing sugars, respectively. To date, only two structures of members of GH39 have been solved: the β-xylosidases from *Thermoanaerobacterium saccharolyticum* (TsXynB; PDB entry 1px8; Yang *et al.*, 2004) and *G. stearothermophilus* (GsXynB; PDB entry 1w91; Czjzek *et al.*, 2005), which share 36 and 37% sequence identity, respectively, with CcXynB2 (Fig. 1). GsXynB and TsXynB are very similar, sharing 65% sequence identity. The α-L-iduronidases (EC 3.2.1.76), which are exclusive to eukaryotes, also belong to the GH39 family (Stoltzfus *et al.*, 1992); however, they display a different

[1] Supplementary material has been deposited in the IUCr electronic archive (Reference: DW5019). Services for accessing this material are described at the back of the journal.

domain architecture and have very low sequence identity to β-xylosidases (<22%).

The CcXynB2 crystals belonged to space group $P4_32_12$ and diffracted to a maximum resolution of 2.5 Å. The asymmetric unit is composed of one protomer with 52.5% solvent content, which corresponds to a Matthews coefficient of 2.58 Å^3 Da^{-1}. During refinement, tetrahedron-shaped electron densities were observed at the protein surface, which were assigned to sulfate ions present in the crystallization solution. The refinement converged to a crystallographic residual of 18.8% (R_{free} = 24.8%) and the final model displayed good overall stereochemistry according to the r.m.s.d. values and Ramachandran plot (Table 1). The average temperature factor (*B* value) for all atoms is 30.1 Å^2, with a well structured catalytic core and some flexibility observed in the loops forming the accessory domain.

The CcXynB2 structure consists of an $(\alpha/\beta)_8$-barrel catalytic core decorated with two extra β-hairpin barrel motifs (segments 70–81 and 230–250) and a β-sandwich accessory domain consisting of an α-helical subdomain (434–467) and a long α-helix-containing loop (390–407) between β-strands 14 and 15 (Fig. 2*a*). The active site is located in the crevice formed by the loops 105–128, 160–174, 201–210, 229–255, 280–295 and 320–349 and is very similar to that observed for both GsXynB and TsXynB. Sequence and structural comparisons showed that all residues considered to be relevant for substrate recognition and catalysis, including His62, Phe119, Asn162, Glu163, Phe169, His229, Tyr231, Glu280, Tyr285, Trp318, Phe324, Glu326 and Phe338, are fully conserved among GH39 β-xylosidases (Figs. 1 and 2*a*). In the CcXynB2 structure Glu280 corresponds to the nucleophile, whereas Glu163 acts as the general acid/base catalytic residue (Vocadlo *et al.*, 1998; Bravman *et al.*, 2001; Yang *et al.*, 2004). Similar to other GH39 β-xylosidases, the electrostatic potential around the active site is predominantly negative owing to the presence of several acidic residues (Fig. 3).

3.2. A long α-helix-containing loop from the auxiliary domain modifies the catalytic interface

Structural comparison with TsXynB and GsXynB, the only two other GH39 members with known three-dimensional structures, resulted in r.m.s.d.s of 0.97 Å (for 326 C^α atoms) and 0.95 Å (for 333 C^α atoms), respectively. The most divergent regions are the C-terminus, the long α-helix-containing loop and the catalytic β-hairpin motif (Fig. 2*b*). In the CcXynB2 structure the C-terminal region is shortened by 18 residues when compared with both TsXynB and GsXynB, and this seems to have direct implications for the oligomerization of GH39 β-xylosidases, as discussed in the next section. Interestingly, the long α-helix-containing loop from the accessory domain is a particular feature of CcXynB2 (Fig. 2*b*). The other structurally characterized β-xylosidases present a shorter loop without α-helical structure. In CcXynB2, this motif is stabilized by polar and hydrophobic interactions with the catalytic β-hairpin from the parental catalytic domain (Supplementary Fig. S1). These interactions induce the catalytic β-hairpin to adopt an open conformation, leading to substantial changes in the molecular topography of the catalytic interface (Fig. 2*b*) and an increase of the solvation of and accessibility to the active-site pocket (Supplementary Fig. S2). Moreover, these contacts contribute to maintaining the tight attachment of the accessory domain to the catalytic barrel (Supplementary Fig. S1), which might be relevant to enzyme

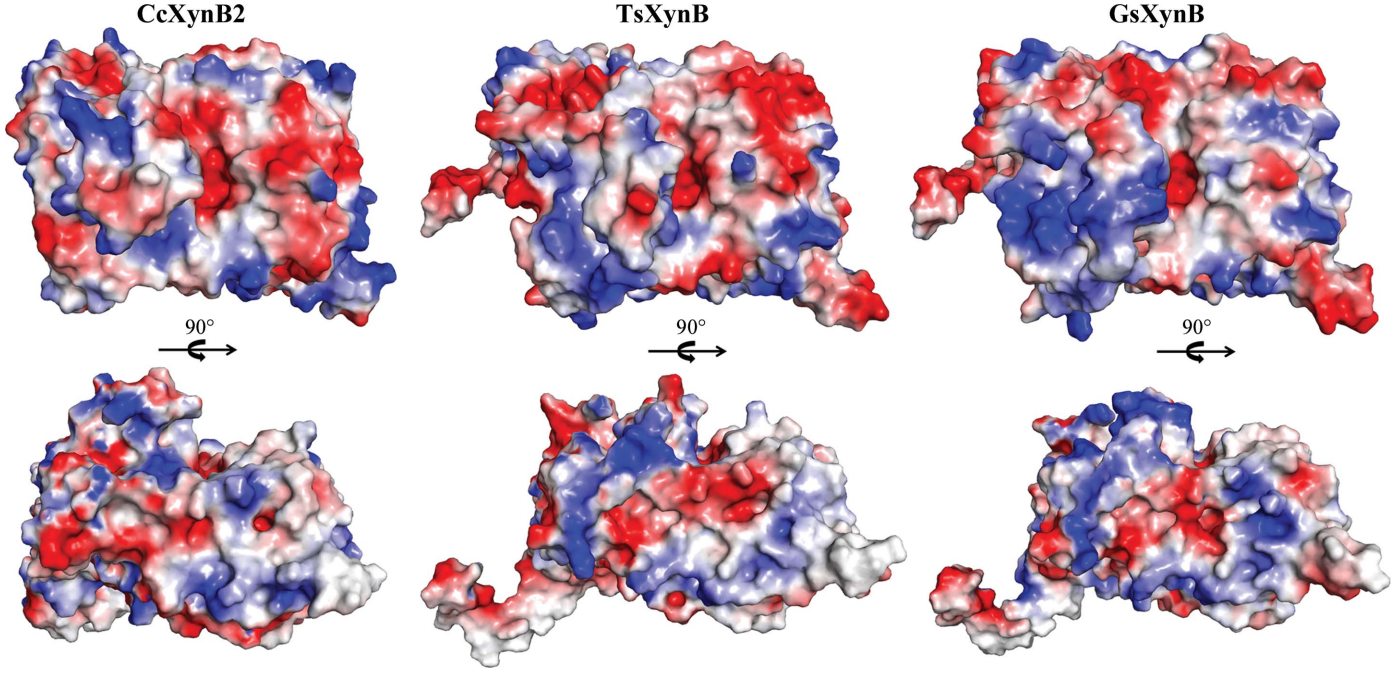

Figure 3
Surface-charge distribution of β-xylosidases from *C. crescentus*, *T. saccharolyticum* and *G. stearothermophilus*.

stability and function. In the TsXynB and GsXynB structures the C-terminal extension also exerts a stabilizing effect. A *BLASTP* search of the nonredundant NCBI database for β-xylosidases displaying high sequence identity to CcXynB2 (>60%) showed a clear correlation between the presence of the long α-helix-containing loop in the accessory domain and the shortened C-terminus, as shown in Fig. 1. The presence of the long α-helix-containing loop in β-xylosidases that have a shortened C-terminus seems to be an important structural requirement in order to keep the accessory domain intimately linked to the parental catalytic core and thus preserve the correct conformation for enzyme action and stability.

3.3. CcXynB2 is a monomer in solution and the shortened C-terminus prevents oligomerization

According to crystal structure analyses and hydrodynamic behaviour, the biological assembly of both TsXynB and GsXynB is a homotetramer (Yang *et al.*, 2004; Czjzek *et al.*, 2005). This tetrameric arrangement displays 222 symmetry and is stabilized by two dimeric interfaces, with the C-terminal extension swapped between the two monomers. The type I interface is formed by contacts between $(\beta/\alpha)_8$-barrels with 1500 Å^2 of buried area (~7%), whereas the type II interface comprises interactions between the β-sandwich domain and the α-helical subdomain with 2600 Å^2 of buried area (12%) (Yang *et al.*, 2004). The C-terminal extension makes a major contribution to the latter interface, making extensive polar and hydrophobic interactions with both the α-helical subdomain and the β-sandwich domain of the adjacent molecule. In both tetrameric structures the active site does not participate in any of the oligomeric interfaces and there are no functional data to indicate how the oligomerization could affect enzyme function.

In contrast, CcXynB contains one protomer in the asymmetric unit and protein-interface analysis of the crystalline form using the *PDBePISA* tool showed only nonsignificant interfaces with buried areas of 593, 336, 194 and 123 Å^2 corresponding to the neighbouring molecules related by the symmetry operators $(-x-1/2, y-1/2, -z-1/4)$, $(-y+1/2, x+1/2, z-1/4)$, $(y-1, x+1, -z)$ and $(y-1, x, -z)$, respectively.

In order to confirm the oligomeric state of CcXynB2 indicated by crystal structure analyses, SAXS measurements were performed. From the scattering and pair-distance distribution curves (Fig. 4*a*), it is possible to observe a maximum molecular dimension (D_{max}) of 95 Å and a radius of gyration of 30.6 Å, which correspond to a monomer. The theoretical scattering curve was calculated for the monomeric crystal structure, showing good agreement with the experimental data, which corroborates the monomeric state of CcXynB2 (Supplementary Fig. S3). Moreover, the protein envelope was determined from the SAXS data using *ab initio* molecular modelling, resulting in an elongated molecular envelope as inferred from the $p(r)$ profile. Fitting of the crystallographic monomer into the SAXS envelope indicated good shape complementarity, also supporting the monomeric state (Fig. 4*b*).

As discussed above, the C-terminal extension plays an important role in the stabilization of the tetrameric arrangement of both GsXynB and TsXynB; however, CcXynB2 does not contain the corresponding C-terminal extension involved in the type II dimeric interface. The C-terminus of CcXynB2 is shortened by 18 residues and this is probably the structural determinant that is responsible for the fact that CcXynB2 does not form a tetrameric quaternary structure. Other structural differences observed in the catalytic interface such as those in the catalytic β-hairpin and the long α-helix-containing loop are exposed to bulk solvent and do not seem to be involved in oligomerization.

Taking into account that (i) all residues considered to be relevant for substrate binding and catalysis are fully conserved in monomeric and tetrameric β-xylosidases, (ii) the active-site region is not involved in any of the oligomeric interfaces and (iii) both previously characterized tetrameric β-xylosidases are from thermophilic organisms and the monomeric β-xylosidase is from an aquatic mesophilic bacterium, we suggest

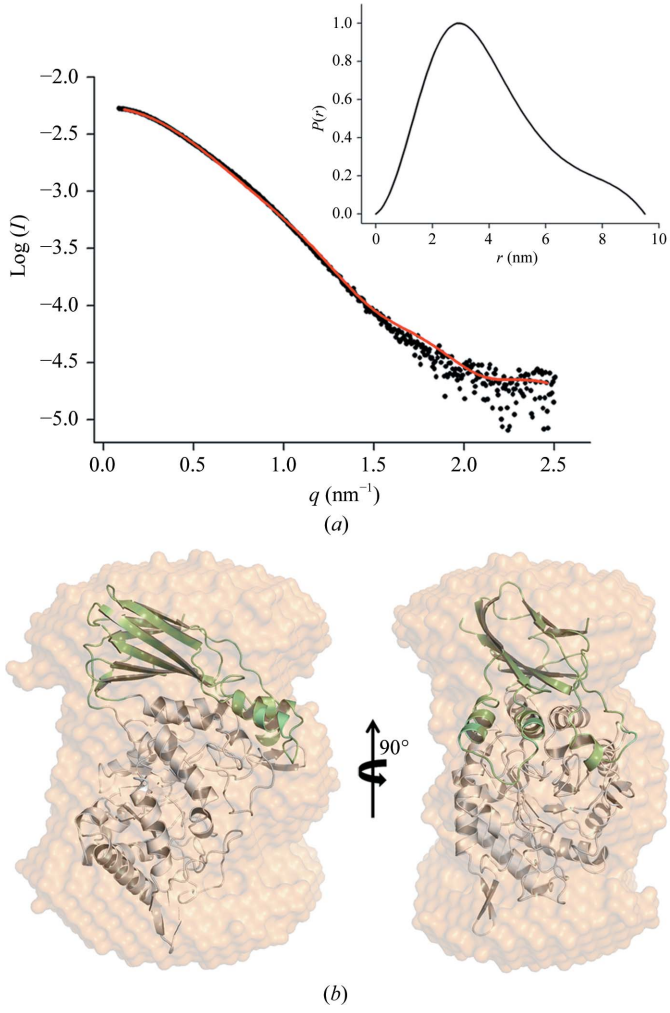

Figure 4
SAXS analysis of CcXynB2. (*a*) Experimental scattering curve (black circles) and the best intensity fitting obtained using the program *GNOM* (continuous red line). The inset shows the normalized pair-distance distribution function $p(r)$. (*b*) Different views of the crystallographic structure fitted into the *ab initio* low-resolution *DAMMIN* envelope.

that the oligomerization of β-xylosidases possibly plays an important role in enzyme stability.

4. Conclusion

In this work, a new β-xylosidase from *C. crescentus* was structurally characterized, revealing a new class of monomeric GH39 members. It was observed that the auxiliary domain modifies the catalytic interface by the presence of a long α-helix-containing loop which induces the catalytic β-hairpin motif to adopt an open conformation. In addition, this motif maintains an intimate link between the accessory domain and the parental catalytic core which may be important for enzyme function and stability. Comparative structural analyses indicated that the lack of the C-terminal extension seems to prevent the formation of a stable oligomer in CcXynB2 and that oligomerization may play a direct role in the stability of thermophilic β-xylosidases.

This work was supported by grants from the Fundação de Amparo à Pesquisa do Estado de São Paulo to MTM (10/51890-8). We also gratefully acknowledge the provision of time on the MX2 beamline (LNLS), Robolab (LNBio) and the SAXS2 beamline (LNLS) at the National Center for Research in Energy and Materials, Campinas, Brazil.

References

Beg, Q. K., Kapoor, M., Mahajan, L. & Hoondal, G. S. (2001). *Appl. Microbiol. Biotechnol.* **56**, 326–338.
Biely, P. (1985). *Trends Biotechnol.* **3**, 286–290.
Boutte, C. C., Srinivasan, B. S., Flannick, J. A., Novak, A. F., Martens, A. T., Batzoglou, S., Viollier, P. H. & Crosson, S. (2008). *PLoS Genet.* **4**, e1000310.
Bravman, T., Mechaly, A., Shulami, S., Belakhov, V., Baasov, T., Shoham, G. & Shoham, Y. (2001). *FEBS Lett.* **495**, 115–119.
Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D**66**, 12–21.
Czjzek, M., Ben David, A., Bravman, T., Shoham, G., Henrissat, B. & Shoham, Y. (2005). *J. Mol. Biol.* **353**, 838–846.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Galbe, M. & Zacchi, G. (2002). *Appl. Microbiol. Biotechnol.* **59**, 618–628.
Hammersley, A. P., Svensson, S. O., Hanfland, M., Fitch, A. N. & Häusermann, D. (1996). *High Press. Res.* **14**, 235–248.
Hottes, A. K., Meewan, M., Yang, D., Arana, N., Romero, P., McAdams, H. H. & Stephens, C. (2004). *J. Bacteriol.* **186**, 1448–1461.
Jordan, D. B. & Wagschal, K. (2010). *Appl. Microbiol. Biotechnol.* **86**, 1647–1658.
Kozin, M. B. & Svergun, D. I. (2001). *J. Appl. Cryst.* **34**, 33–41.
Laemmli, U. K. (1970). *Nature (London)*, **227**, 680–685.
Lagaert, S., Pollet, A., Delcour, J. A., Lavigne, R., Courtin, C. M. & Volckaert, G. (2011). *Appl. Microbiol. Biotechnol.* **92**, 1179–1185.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
Miller, G. L. (1959). *Anal. Chem.* **31**, 426–428.
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.
Neugebauer, H., Herrmann, C., Kammer, W., Schwarz, G., Nordheim, A. & Braun, V. (2005). *J. Bacteriol.* **187**, 8300–8311.
Nierman, W. C. *et al.* (2001). *Proc. Natl Acad. Sci. USA*, **98**, 4136–4141.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Poindexter, J. S. (1964). *Bacteriol. Rev.* **28**, 231–295.
Qing, Q., Yang, B. & Wyman, C. E. (2010). *Bioresour. Technol.* **101**, 9624–9630.
Shallom, D. & Shoham, Y. (2003). *Curr. Opin. Microbiol.* **6**, 219–228.
Stoltzfus, L. J., Sosa-Pineda, B., Moskowitz, S. M., Menon, K. P., Dlott, B., Hooper, L., Teplow, D. B., Shull, R. M. & Neufeld, E. F. (1992). *Biol. Chem.* **267**, 6570–6575.
Svergun, D. I. (1992). *J. Appl. Cryst.* **25**, 495–503.
Svergun, D. I. (1999). *Biophys. J.* **77**, 2879–2886.
Svergun, D., Barberato, C. & Koch, M. H. J. (1995). *J. Appl. Cryst.* **28**, 768–773.
Vocadlo, D. J., MacKenzie, L. F., He, S., Zeikus, G. J. & Withers, S. G. (1998). *Biochem. J.* **335**, 449–455.
Volkov, V. V. & Svergun, D. I. (2003). *J. Appl. Cryst.* **36**, 860–864.
Yang, J. K., Yoon, H.-J., Ahn, H. J., Lee, B. I., Pedelacq, J.-D., Liong, E. C., Berendzen, J., Laivenieks, M., Vieille, C., Zeikus, G. J., Vocadlo, D. J., Withers, S. G. & Suh, S. W. (2004). *J. Mol. Biol.* **335**, 155–165.