



Biophysical and biochemical studies of a major endoglucanase secreted by *Xanthomonas campestris* pv. *campestris*.



Flávio Rodolfo Rosseto^a, Livia Regina Manzone^a, Mario de Oliveira Neto^b, Igor Polikarpov^{a,*}

^a Instituto de Física de São Carlos, Universidade de São Paulo, Avenida Trabalhador São Carlense 400-Centro, São Carlos—SP 13560-970, Brazil

^b Instituto de Biociências, Universidade Estadual Paulista, Distrito de Rubião Jr. s/n, Botucatu—SP 18618-970, Brazil

ARTICLE INFO

Article history:

Received 30 December 2015

Received in revised form 14 May 2016

Accepted 20 May 2016

Available online 21 May 2016

Keywords:

Xanthomonas campestris

Endoglucanases

Hydrolysis

SAXS

ABSTRACT

Endoglucanases are the main cellulolytic enzymes secreted by the bacterium *Xanthomonas campestris* pv. *campestris* (Xcc). The major endoglucanase exported by this bacterium into an external milieu is an enzyme XccCel5A, which belongs to GH5 family subfamily 1 and is encoded by the gene engXCA. We purified XccCel5A using ammonium sulfate precipitation followed by size exclusion chromatography and identified it by zymogram analysis. Circular dichroism and fluorescence spectroscopy studies showed that XccCel5A is stable in a wide pH range and up to about 55 °C and denatures at the higher temperatures. The optimal conditions for enzyme activity were identified as T = 45 °C and pH = 7.0. Under the optimum conditions the catalytic efficiency (k_{cat}/K_M) of the enzyme was determined as $5.16 \times 10^4 \text{ s}^{-1} \text{ M}^{-1}$ using carboxymethylcellulose (CMC) as a substrate. Our SAXS studies revealed extended tadpole-shape molecular assembly, typical for cellulases, and allowed to determine an overall shape of the enzyme and a relative position of the catalytic and cellulose binding domains.

© 2016 Elsevier Inc. All rights reserved.

1. Introduction

Traditional sources of hydrocarbon energy such as coal, oil and natural gas are limited and non-renewable. In this context, the growth of the world population coupled with increasing demand for food and energy continuously intensifies interest in renewable sources of energy, such as plant biomass and agro-industrial waste and their conversion into fuels, chemicals and renewable materials [1,2]. The biomass to cellulosic ethanol conversion normally includes four major steps: biomass pre-treatment and enzymatic hydrolysis of (hemi)cellulose fractions into simple sugars; their fermentation to ethanol and, finally, ethanol separation and purification, usually by distillation [3]. The classical view of cellulose hydrolysis depicts a depolymerization process as being initiated by endoglucanases reducing the degree of polymerization of the substrate and releasing new cellulose chain ends which are targets for exoglucanases [4,5]. Due to their tunnel-shaped active site, the exoglucanases have much higher processivity than typical fungal endoglucanases, that have more open active side cleft [6,7]

and their joint enzymatic action leads to a substantial synergism in hydrolysis of crystalline cellulose substrates [8,9]. The hydrolysis is terminated by β -glucosidases that enzymatically cleave released cellobiose units into glucose. However, recent discoveries of lytic polysaccharide monoxygenases (LPMOs) reveal that the canonical view of cellulose enzymatic hydrolysis might be somewhat oversimplified and incomplete [10].

Most of the bacterial endoglucanases belong to the glycoside hydrolase (GH) families GH5, GH6, GH8, GH9 and GH12 [11,12]. Even though they hydrolyze the same cellulose substrate, they act using differing catalytic mechanisms: “inverting” for GH6, GH8 and GH9 family members and “retaining” for GH5 and GH12 endoglucanases. Bacterial endoglucanases are diverse in their molecular structure, substrate specificity and the mode of action. Some of the bacterial endoglucanases are processive and contain catalytic domain rigidly bound to cellulose binding module [13–15]. Several anaerobic cellulolytic bacteria produce multicomponent, multi-enzyme cellulosome complexes [16] whereas some other produce cellulases that are multi-domain enzymes comprised of endo- and exoglucanase catalytic domains linked to cellulose-binding modules, that “drill” into cellulose substrate instead of eroding it from the surface [17]. Finally, various endoglucanases are also enzymatically active against a number of hemicellulose polysaccharides,

* Corresponding author at: Universidade de São Paulo, Departamento de Física e Ciência Interdisciplinar, Instituto de Física de São Carlos, Av. Trabalhador São-carlense, 400, São Carlos, SP 13560-970, Brazil.

E-mail address: ipolikarpov@ifsc.usp.br (I. Polikarpov).

which could be related to their function in degradation of complex lignocellulosic biomass [18,19].

Xanthomonas campestris pv. *campestris* (Xcc) is a plant pathogenic bacterium and its genome analysis revealed presence of several genes encoding enzymes involved in the degradation of plant cell walls including pectinases, lyases, xylanases and nine cellulases [20]. This indicates that the bacterium is able to secrete various enzymes that degrade structural polysaccharides such as cellulose and hemicellulose [21]. Xcc secretes two endoglucanases and the endoglucanase XccCel5A, encoded by the gene *engXCA*, is the most abundant cellulase produced by the bacterium [22]. This endoglucanase belongs to GH5 family subfamily 1 [11] in the CAZy database [12]. BLASTP2.2.25+ program [23] analysis based on the available enzyme sequence (NCBI: NP.638867.1) shows that XccCel5A is composed by two domains: an N-terminal catalytic domain (CD) and a carbohydrate binding type-2 module (CBM), joined together by a Thr-Pro-rich linker peptide of about 30 amino acid residues long.

Given importance of Xcc endoglucanases in biotechnological applications (such as production of xanthan gum and, potentially, cellulosic ethanol) and also for plant pathogenicity of the bacteria, here we performed biophysical, biochemical and enzymatic characterization of the major endoglucanase secreted by *Xanthomonas campestris* pv. *campestris* and retrieved its low-resolution SAXS model in solution.

2. Materials and methods

2.1. Bacteria culture and production of endoglucanases

The strain of Xcc (ATCC 33913) was kindly provided by Prof. Shaker C. Farah (IQ/USP, São Paulo, Brazil). Xcc cells, maintained in 20% of glycerol, were transferred onto 100 $\mu\text{g mL}^{-1}$ ampicillin-containing LB medium plates and grown for 3–5 days at 30 °C. Then, cells were grown for 16 h in 50 mL of ampicillin-containing LB medium at 30 °C at constant agitation of 250 rpm. As a next step, the culture was transferred into a 500 mL of the same medium and kept at same conditions for further 36 h.

2.2. Identification of Xcc major endoglucanase

In order to identify the protein of interest based on its enzymatic activity, we performed zymogram analysis using 10% SDS-PAGE supplied with 0.1% of carboxymethylcellulose (CMC) [24]. Briefly, the polyacrylamide gel was separated in two parts: one containing heat-denatured protein with reducing agent and another non-heated protein without reducing agent. After electrophoresis at 10 °C, the denatured part was stained with Coomassie Brilliant Blue R-250 and the undenatured part was incubated with 2.5% (v/v) Triton X-100 solution (Sigma) for 30 min followed by 2 washes with 50 mM sodium citrate pH 5.0 and kept at 50 °C for 15 min. Finally, this part of the gel was stained with 0.1% (w/v) Congo Red and destained with 1 M NaCl until the bands were revealed. To confirm which band corresponded to XccCel5A protein, the bands were cut off from the gel and digested with trypsin (Sigma) [25]. The resulting peptides were analyzed at an Electrospray Tandem MS/MS (Q-TOF) mass spectrometer at the National Laboratory of Synchrotron Light (LNLS, Brazil) and the program MASCOT (www.matrixscience.com) was used for protein identification.

2.3. Purification of XccCel5A

Xcc culture was centrifuged and the supernatant was concentrated by tangential ultrafiltration using HollowFiber system (GE) with 10 kDa cutoff cartridge in 50 mM Tris–HCl pH 8.0. Twenty percent of ammonium sulfate was added to the extract to induce

precipitation of contaminants. After centrifugation, 40% (w/v) ammonium sulfate was added to supernatant followed by centrifugation at 13,000g for 20 min to precipitate the enzyme. The precipitate was resuspended in 50 mM Tris–HCl pH 8.0 and injected in the molecular size exclusion column Sephadex G-25 Superfine (GE) equilibrated in the same buffer. The protein concentration was determined using the theoretical extinction coefficient at 280 nm ($117225 \text{ M}^{-1} \text{ cm}^{-1}$) calculated by ProtParam tool [26].

2.4. Thermal, pH and catalytic activity assays

Optimum temperature and pH experiments were performed using carboxymethylcellulose (CMC) with degree of polymerization of 400 as a substrate. For pH-dependence studies the pH was varied from 2 to 10 and the temperature was maintained constant at 50 °C; for thermal analysis the temperature was changed from 30 to 80 °C in predetermined pH [27]. For these assays, we used final concentrations of 400 nM of enzyme, 0.25% (w/v) of CMC and 50 mM of the appropriate buffer in 100 μL volume reactions. Kinetic parameters (k_{cat} , V_{max} and K_{M}) against CMC were determined at 45 °C using 200 nM of XccCel5A in 50 mM Tris–HCl pH 7.0 by varying the concentration of low-viscosity CMC from 0.05 to 4% (w/v). The mixtures were incubated for 30 min and the released sugars were measured using 3,5-dinitrosalicylic acid (DNS) method [28]. The data represent average of triplicate experiments and their analyses were performed using SigmaPlot 10.0 program.

2.5. Circular dichroism (CD) and fluorescence spectroscopy analyzes

Temperature and pH variation assays were independently monitored by circular dichroism technique. Temperature denaturation assays were conducted in a 20–80 °C range using 2.5 μM of enzyme. Experiments in pHs 2.6, 4.0, 4.5, 6.0 (20 mM Na_2HPO_4 /citric acid), 8.0 (20 mM $\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_3$) and 10.0 (20 mM glycine) were conducted at 25 °C after incubation of the enzyme in these buffers for 16 h. CD spectra were recorded on a Jasco J-720 spectral polarimeter (Jasco, Tokyo, Japan) using 0.1 cm path length cuvette within 195–245 nm wavelength range, 100 nm min^{-1} increment step, 10 s averaging time, 1 nm bandwidth and a response time of 0.5 s. The same conditions of pH assays were used for fluorescence spectroscopy measurements using a K2 spectrofluorimeter (ISS Inc Champaign, IL–USA). The samples were excited at 295 nm and the fluorescence emission was monitored in the range from 300 to 450 nm, at 25 °C. The buffer contribution was subtracted.

2.6. SAXS experiments and data analysis

SAXS experiments were performed with XccCel5A at 3.5 mg mL^{-1} in 50 mM Tris–HCl pH 7.0 at the D02A-SAXS2 beamline of the Synchrotron Light National Laboratory (Campinas, Brazil). The measurements were conducted using a monochromatic X-ray beam with a wavelength $\lambda = 1.488 \text{ \AA}$. The sample-to-detector distance was set at 1028.4 mm, resulting in a scattering vector q range of 0.041 \AA^{-1} to 0.325 \AA^{-1} , where q is defined by $q = 4\pi \sin \theta / \lambda$ (θ is the scattering angle). All the X-ray patterns were recorded using a two-dimensional CCD detector (MarResearch, USA). Protein samples were centrifuged at 23,500g for 15 min at 4 °C to remove aggregates. For SAXS measurements, protein samples were introduced into a 1 mm path length cell with mica windows at 10 °C. Two successive frames of 300 s each were recorded for each sample to monitor radiation damage and beam stability. Buffer scattering was recorded before the sample scattering. SAXS curves were corrected for the detector response, intensity of the incident beam, sample absorption and buffer scattering. The two-dimensional patterns were integrated in 2θ resulting in an

averaged one-dimensional scattering curve. The radius of gyration, R_g was approximated using two different and independent methods: by Guinier's equation [29] and by of the inverse Fourier transform method [30] as implemented in the Gnom package. The distance distribution function $p(r)$ also was evaluated by Gnom and the maximum diameter of the protein D_{max} was obtained. Dummy atom models (DAMs) were calculated from the experimental SAXS data of XccCel5A using *ab initio* procedure implemented both in Dammin [31] and Gasbor packages [32]. For the two XccCel5A main domains (CD and CBM), tertiary structures of homologous proteins with highest amino acid identity were positioned to best fit the molecular envelope. The 3D structures of β -1,4-glycanase cellulose binding module from *Cellulomonas fimi* (PDB id: 1EXG) [33] and the catalytic domain of endoglucanase from archaeon *Pyrococcus horikoshii* (PDB id 2ZUM) [34], respectively, were used, and the relative adjustment of the structures against the experimental scattering data was calculated by SASREF [35]. The SAXS-based estimation of the molecular weight (MW) was obtained using the web tool SAXS MoW (www.ifsc.usp.br/~saxs/saxsmow.html) [36].

3. Results and discussion

3.1. Identification and purification of the XccCel5A, the major endoglucanase from *Xanthomonas campestris* pv. *campestris*

Supernatant produced by Xcc was first subjected to SDS-PAGE analysis. The theoretical molecular weight of XccCel5A computed from its amino acid sequence (NCBI: NP.638867.1) is close to 49.5 kDa, which is consistent with our experimental analyses using of SDS-PAGE and zymogram methodologies (Fig. 1A and B). To confirm the identity of the endoglucanase XccCel5A, we achieved mass spectrometry analysis of the 49.5 kDa band. The obtained mass spectra had an excellent coverage of protein amino acid sequence (approximately 55%) [25], allowing us to consistently conclude that the XccCel5A was successfully identified among the other proteins expressed in the extracellular medium.

The first purification step of the protein was performed using with 20% (w/v) of ammonium sulfate for precipitation of contaminating proteins and xanthan gum. The endoglucanase XccCel5A only precipitated after the addition of 40% ammonium sulfate (Fig. 2A). Once resuspended in 50 mM Tris-HCl pH 8.0, the enzyme was subjected to molecular size exclusion chromatography using

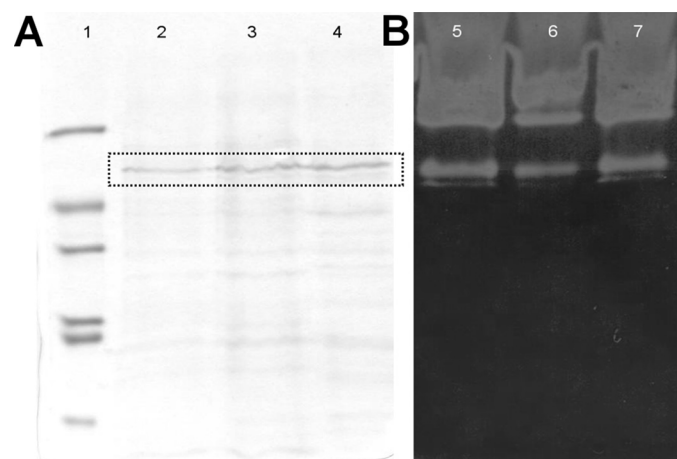


Fig. 1. Total extracellular proteins produced by Xcc. (A) Different concentrations of the samples were analyzed by SDS-PAGE (10%) stained with Coomassie Brilliant Blue. Lane 1, molecular-weight marker (66, 45, 36, 29, 24 and 14 kDa). Lane 2, 3 and 4, XccCel5A at 0.2 mg mL⁻¹. The dashed box indicates the XccCel5A protein. (B) Zymogram of extracellular proteins against CMC substrate, showing XccCel5A position. Lane 5, 6 and 7, XccCel5A at 0.2 mg mL⁻¹.

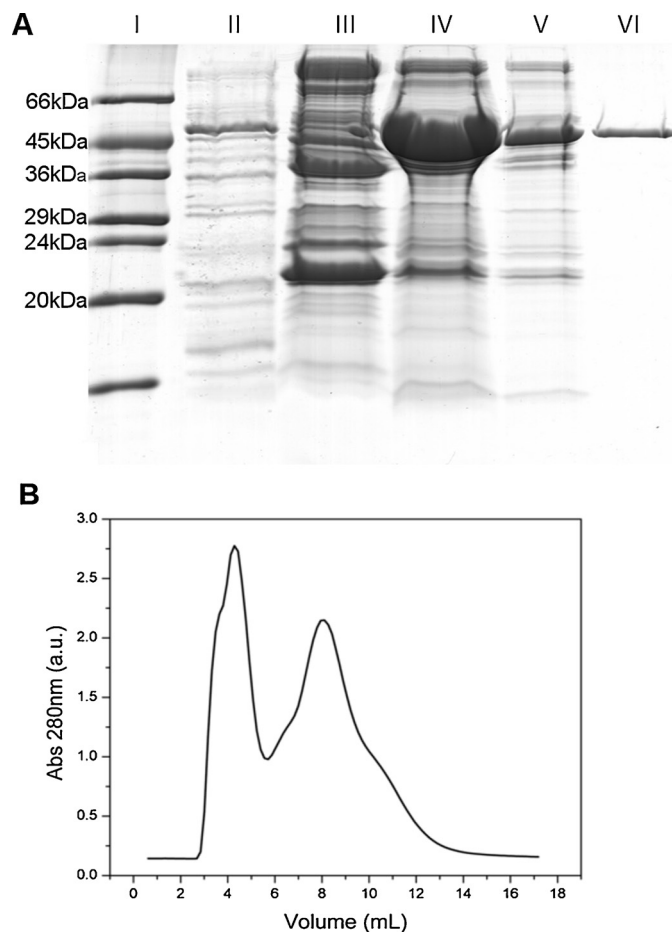


Fig. 2. XccCel5A protein purification using ammonium sulfate precipitation and size exclusion chromatography (SEC). (A) Purification steps of XccCel5A analyzed by depositing 15 μ L samples on SDS-PAGE (10%) and staining them with Coomassie Brilliant blue. Lane 1, molecular weight marker (kDa). Lane 2, total extracellular proteins of Xcc. Lane 3 and 4, protein precipitation using 20 and 40% of ammonium sulfate, respectively. Lane 5, first peak of SEC chromatography. Lane 6, second peak of SEC chromatography. (B) SEC profile of XccCel5A protein after 40% ammonium sulfate precipitation, using Sephadex G-25 superfine column (GE) as monitored at 280 nm.

Sephadex G-25 Superfine column (GE). The enzyme eluted in two peaks (as followed by absorption at 280 nm; Fig. 2B). The first elution peak contained the enzyme together with contaminants; however the second peak consisted only by pure XccCel5A. The estimated experimental yield was 1 mg of purified enzyme per liter of culture.

3.2. XccCel5A pH and temperature optima

As a next characterization step, the optimum pH and temperature conditions for XccCel5A enzymatic activity against CMC were determined. The pH dependence of XccCel5A activity (Fig. 3A) reveals that the enzyme has highest activity in neutral pHs around 7.0. This is different from pH optimum of many fungal cellulases that frequently have acidic pH optimum (typically between pHs 5.0 and 6.0) [9]. However such pH behavior is not unusual for bacterial cellulases, many of which have highest activity at neutral pHs [37,38]. Furthermore, the enzyme has significant activity (above 50% of its maximum) against CMC in a wide pH range (between 4.5–9.0). A similar assay was also used to determine the optimum temperature for catalytic activity. The temperature was varied in a range of 30–80 °C while maintaining the optimum pH 7.0. Fig. 3B shows the influence of temperature on XccCel5A activity against

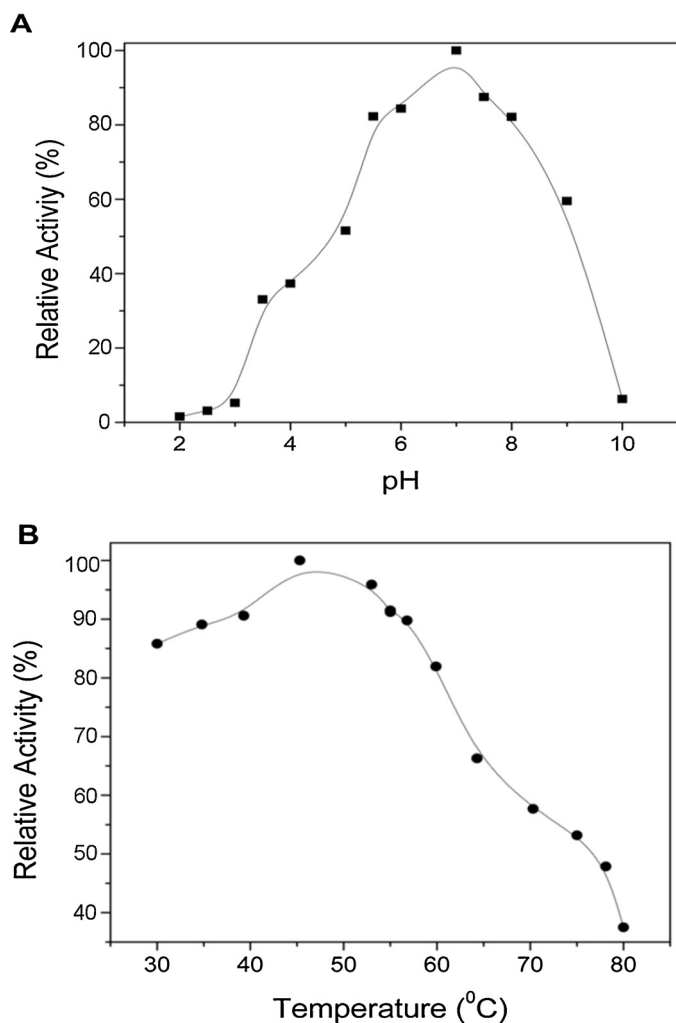


Fig. 3. Optimum pH, temperature and kinetic assays for XccCel5A activity against CMC. (A) pH dependence of XccCel5A activity. The peak at pH 7.0 indicates the best pH for the enzyme activity. (B) Thermal analysis of XccCel5A performed from 30 to 80 °C in pH 7.0 revealing the optimum temperature for XccCel5A activity (45–50 °C approximately). (C) CMC concentration effect on the reaction rate for kinetic parameters determination of XccCel5A.

CMC, revealing 45 °C as the optimum temperature for its enzymatic activity. Moreover, these results indicate that the enzyme exhibits good thermal stability since its activity decreased only by 20% in the temperature range between 30 and 60 °C. Furthermore, the enzyme was active against Avicel, indicating its activity against crystalline cellulosic substrates (results not shown).

We also studied kinetic behavior of XccCel5A against CMC under the optimum temperature and pH conditions. Kinetics of the enzyme follows Michaelis-Menten model and application of Lineweaver-Burk linearization method to the experimental data allowed the determination of kinetic parameters $K_M = 4.50 \pm 0.804 \text{ mg mL}^{-1}$ and $k_{\text{cat}} = 273.60 \text{ min}^{-1}$ (Fig. 3C). Maximum velocity (V_{max}) of the enzyme is equal to $0.05 \pm 0.00186 \mu\text{mol min}^{-1} \text{ mL}^{-1}$ and its catalytic efficiency (k_{cat}/K_M) amounts to $60.8 \text{ mL mg}^{-1} \text{ min}^{-1}$. The K_M value is comparable with $K_M = 4.8 \text{ mg mL}^{-1}$ for β -1,4-endoglucanase from a novel bacterial *Bacillus* sp. strain CTP-09 [39] and considerably lower than K_M values for *Martellella mediterranea* and *Talaromyces emersonii* endoglucanases, which were determined as 28.4 mg mL^{-1} and 31.7 mg mL^{-1} , respectively, against the same substrate [40,41]. Besides, XccCel5A catalytic efficiency is similar to (k_{cat}/K_M) of $76.9 \text{ mL mg}^{-1} \text{ min}^{-1}$ and $83.7 \text{ mL mg}^{-1} \text{ min}^{-1}$ determined for

GH5 family endoglucanases from *Martellella mediterranea* and *Thermoanaerobacter tengcongensis*, respectively [40,42].

3.3. Circular dichroism (CD) and fluorescence spectroscopy

Next, we analyzed XccCel5A stability against temperature using circular dichroism (CD) technique. By altering a temperature in the range from 20 to 90 °C and measuring CD signal, the denaturation process was found to be irreversible (data not shown), with the T_m value of 69.2 °C (Fig. 4A). Furthermore, the temperature variation monitored by CD showed that the thermal stability of the enzyme was maintained up to approximately 55 °C, and the accentuated plateau above 75 °C indicates a complete denaturation of the enzyme. This is consistent with the results of temperature dependence of XccCel5A enzymatic activity that shows decline in the activity at temperatures above 55 °C, which probably reflects beginning of the protein unfolding. Analysis of CD spectra for different pHs (Fig. 4B) shows that the enzyme retains its folding similar to the native in basic pHs values. Under extremely acidic pHs (i.e. 2.6) there was a considerable decrease in ellipticity values indicative of unfolding of the macromolecule and loss of the secondary structure.

The tertiary structure of XccCel5A was also assessed by tryptophan intrinsic fluorescence. The emission spectrum of tryptophan is very sensitive to the local environment thus providing information about protein folding at the tertiary structure level. Upon excitation at 295 nm, the XccCel5A fluorescence emission spectra have maximum at the wavelength of approximately 335 nm for the most pHs tested (Fig. 4C). However there was a shift to 328 nm in the fluorescence emission maximum when the protein was exposed to extremely acidic pHs, indicating changes of the tryptophans local environments.

3.4. Small angle X-ray scattering studies of XccCel5A

To determine oligomeric state of XccCel5A in solution and to retrieve its low-resolution molecular envelope, we performed its SAXS analysis. Experimental SAXS curve and simulated scattering curves from the rigid body (RBM) and dummy atom model (DAM) analysis are shown in Fig. 5A. The $p(r)$ curve for XccCel5A is elongated, defining a D_{max} as $115.00 \pm 0.50 \text{ \AA}$ (Fig. 5B). Using SAXS MoW package [36] we calculated the approximate molecular weight of the enzyme using experimental SAXS data. This analysis estimated a molecular weight of XccCel5A as 50.9 kDa, which is almost exactly equal to the theoretically calculated molecular weight based on the amino acid sequence of the enzyme with a difference of only 2.8%.

The three-dimensional DAM was generated *ab initio* from the experimental SAXS data. To check the uniqueness of the model, independent simulations starting from different initial parameters were performed. With these models generated and superimposed with known homologous crystallographic structures, we confirmed that the enzyme has an elongated tadpole-like molecular shape with two distinct portions accommodating, respectively, the catalytic domain and the CBM with a linker region between them (Fig. 6). Similar molecular shapes have already been reported for other cellulases [43–49]. Structural parameters for the XccCel5A obtained from SAXS analysis are summarized in Table 1.

The size of the XccCel5A ($D_{\text{max}} = 115 \text{ \AA}$) is almost the same as estimated from SAXS studies for *T. harzianum* Cel7A ($D_{\text{max}} = 110 \text{ \AA}$) [46] and its molecular model is quite similar in its shape to the latter enzyme. Earlier reports described significantly larger molecular sizes for *C. fimi* endoglucanase ($D_{\text{max}} = 210 \text{ \AA}$), *T. reesei* Cel7A ($D_{\text{max}} = 180 \text{ \AA}$) and *T. reesei* Cel6A ($D_{\text{max}} = 215 \text{ \AA}$) [43–45], although more recent electron microscopy studies defined *T. reesei* Cel7A and Cel6A D_{max} as $151 \pm 13 \text{ \AA}$ and $134 \pm 26 \text{ \AA}$, respectively [48]. Why the earlier reports show considerably larger molecular sizes for cellulases is not completely clear, although a small amount of aggregation

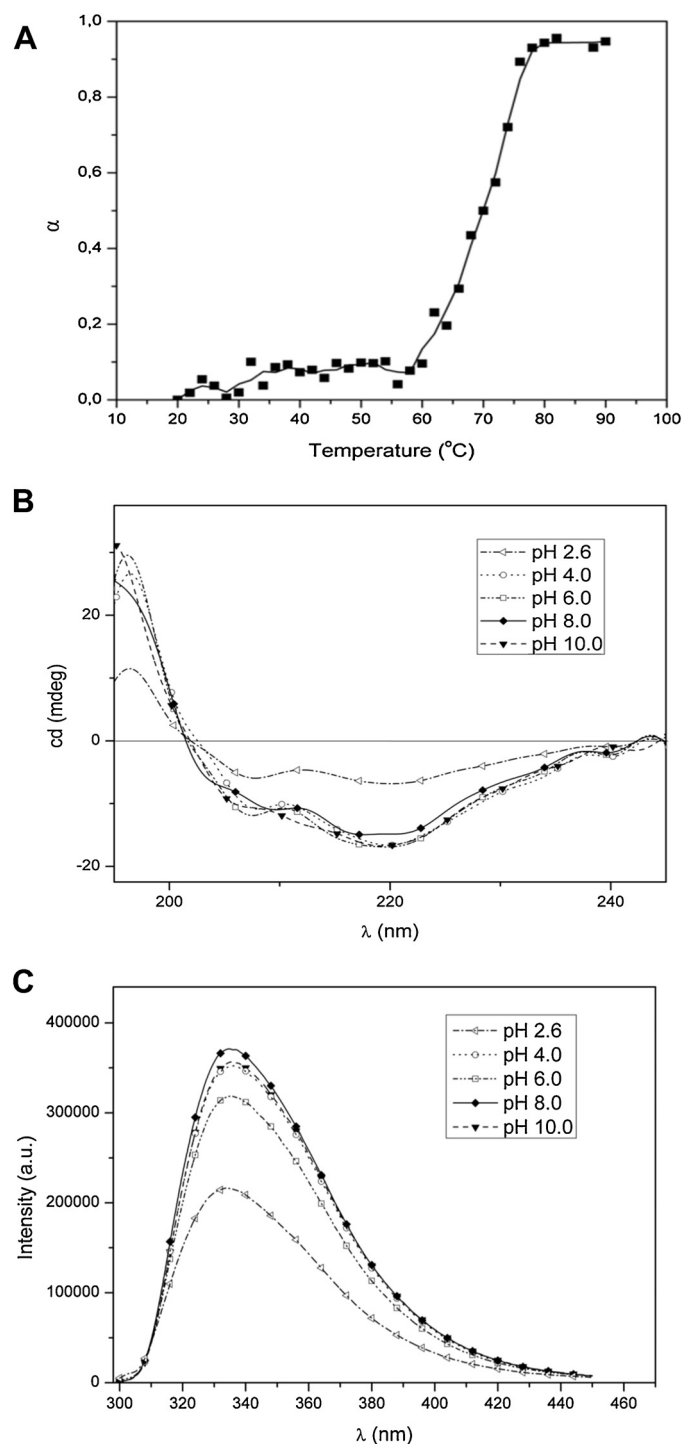


Fig. 4. Biophysical analysis of XccCel5A. (A) Transition curve of XccCel5A thermal denaturation monitoring changes in CD signal at 215 nm as a function of temperature (20–90 °C). (B) CD spectra of XccCel5A at different pHs at 25 °C. (C) Intrinsic fluorescence emission spectra of XccCel5A enzyme at different pHs. Emission spectra in the range of 300–450 nm were recorded at 25 °C using an excitation wavelength of 295 nm. Buffers contributions have been subtracted. The results are an average of three independent experiments.

present in the SAXS samples and/or lower fluxes of X-rays available at that time might have contributed to lesser precision of the measurements [43–45].

Furthermore, recent small-angle neutron scattering (SANS) studies of *T. reesei* Cel7A revealed pH-dependent conformational changes of the shape of the enzyme with the average D_{\max} vary-

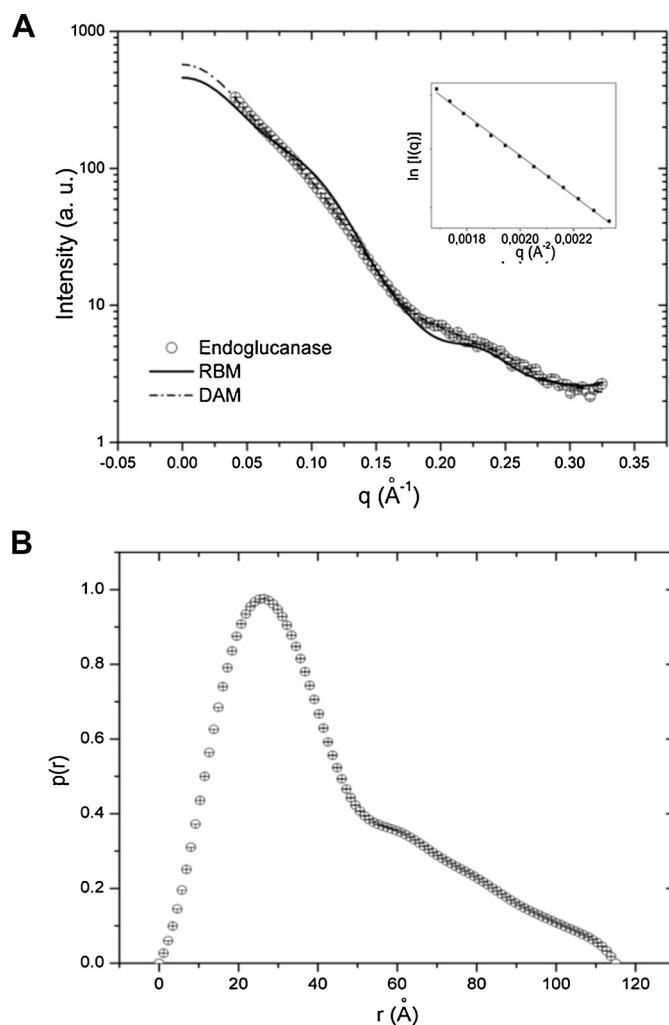


Fig. 5. SAXS analysis of XccCel5A. (A) Scattering experimental curve of Xcc major endoglucanase (○) simulated scattering intensity for the DAM (dashed line) and simulated scattering intensity for the rigid body model composed of the homologous cellulose binding module and catalytic domain (solid line). (B) Distance distribution function computed from the experimental SAXS data for XccCel5A enzyme.

Table 1
Structural parameters derived from SAXS for XccCel5A.

	Exp. ^a	DAM ^b	RBM ^c
D_{\max} (Å)	115.0 ± 0.50	116	108.8
R_g (Å)	33.98 ± 0.50	33.19	31.55
Resolution (Å)	19.31	–	–
MW (kDa)	50.9	–	–

^a Exp., calculated from experimental data.

^b DAM, Parameters of the dummy atom models from the DAMMIN package.

^c RBM, Parameters calculated from rigid body model from bunch package.

ing between 90 Å (pHs 7.0 and 6.0), 100 Å (pH 4.2) and 110 Å (pH 5.3) [49]. The results of these SANS investigations reveal that at the acidic pH optimum (pH 4.2), the enzyme adopts a range of partially disordered structures, which might be beneficial for the activity of *T. reesei* Cel7A. Although CD spectra show no considerable changes in the enzyme secondary structure content at different pHs (from pH 4.2–7.0), $p(r)$ functions and Kratky plots corresponding to pHs 5.3 and 4.2 are clearly different from the ones determined at higher pHs, indicating changes in the conformation and/or flexibility of the enzyme. In particular, changes at pH 4.2 are likely to be attributed to more flexible state of the catalytic domain. Cal-

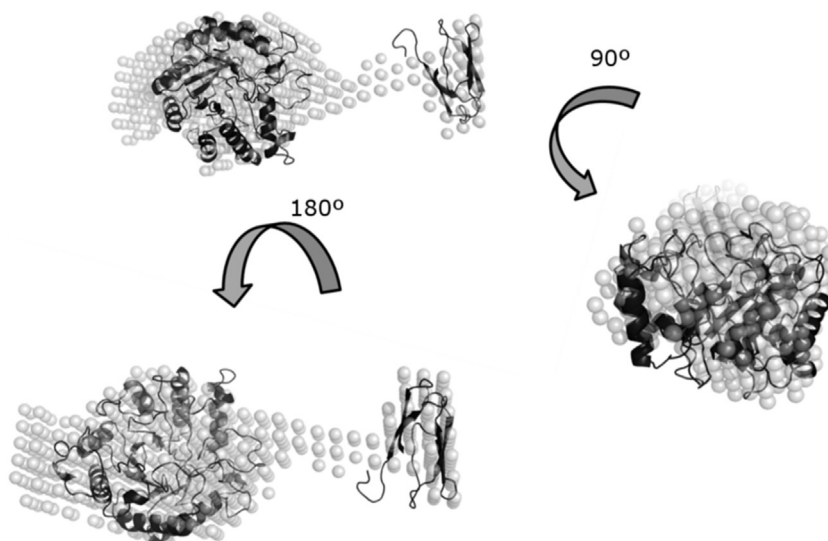


Fig. 6. Superposition of the XccCel5A dummy atoms model with the 3D structure of the cellulose binding module of *Cellulomonas fimi* β -1,4-glycanase (PDB id: 1EXG) [22] and the catalytic domain of endocellulase from archaeon *Pyrococcus horikoshii* (PDB id: 2ZUM) [23]. The individual domains positions and orientations were mutually adjusted as rigid bodies to provide the best fit against the experimental SAXS curves. The final combined models were rotated by 90° with respect to the originally chosen orientation (top left) around y and x axes (bottom left and right, respectively).

culations of the net charge of the enzyme catalytic domain show an abrupt change (from -10 to 0) at pH 4.2, close to the catalytic domain calculated pI equal to 4.3, which may be the reason for increased flexibility of the catalytic domain at this pH. Augmented flexibility of the catalytic core of *T. reesei* Cel7A might be beneficial for enhancing substrate access to the enzyme active tunnel and could indicate that conformational selection plays a role in the enzyme function [49]. Finally, not all endoglucanases have extended molecular conformation in solution. Recent report on *Neurospora crassa* Cel45A endoglucanase revealed a monkey-wrench molecular shape enzyme with the maximum dimension of 75 Å [50]. Clearly, more expensive low-resolution studies of the cellulases in solution are needed in order to understand in more detail their full-length molecular shape, conformation and flexibility.

In conclusion, our studies permitted the identification, purification, structural, biochemical, and enzymatic characterization of the major endoglucanase from *Xanthomonas campestris* pv. *campestris* and paved the way for further molecular and enzymatic studies of this enzyme.

Authors' contributions

Flávio Rodolfo Rosseto—has taken part in the planning, identification and characterization of XccCel5A;

Livia Regina Manzine—contributed to the production and purification of XccCel5A and drafted this paper;

Mario de Oliveira Neto—contributed to SAXS experiments and data analysis;

Igor Polikarpov—planned the experiments, discussed the results and participated in the writing of the manuscript.

Conflict of interest

The authors declare that no conflict of interest exists.

Acknowledgments

This work was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) via grant numbers 2008/56255-9, 2009/54035-4, 2010/08370-3, 2010/16542-9; Conselho Nacional

de Desenvolvimento Científico e Tecnológico (CNPq) via grant number 482166/2010-0 and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). We also thank Laboratório Nacional de Luz Síncrotron (LNLS) for beam time and help with SAXS data collection.

References

- [1] E. Hood, R. Love, J. Lane, J. Bray, R. Clough, K. Pappu, C. Drees, K.R. Hood, S. Yoon, A. Ahmad, J.A. Howard, Subcellular targeting is a key condition for high-level accumulation of cellulase protein in transgenic maize seed, *Plant Biotechnol. J.* 5 (2007) 709–719.
- [2] J.M. Otero, G. Panagiotou, L. Olsson, Fueling industrial biotechnology growth with bioethanol, *Adv. Biochem. Eng. Biotechnol.* 108 (2007) 1–40.
- [3] P.C. Badger, Trends in New Crops and New Uses, in: J. Janick, A. Whipkey (Eds.), *Ethanol from cellulose: A general review*, ASHS Press, Alexandria, VA, 2002.
- [4] G.J. Kleywegt, J.Y. Zou, C. Divne, G.J. Davies, I. Sinning, J. Stahlberg, T. Reinikainen, M. Srisodsuk, T.T. Teeri, T.A. Jones, The crystal structure of the catalytic core domain of endoglucanase I from *Trichoderma reesei* at 3.6 Å resolution, and a comparison with related enzymes, *J. Mol. Biol.* 272 (1997) 383–397.
- [5] V.I. Serpa, I. Polikarpov, Enzymes in bioenergy, in: M.S.S. Buckneridge, G.H.H. Goldman (Eds.), *Routes to Cellulosic Ethanol*, Springer, 2011, pp. 97–113.
- [6] K. Igarashi, T. Uchihashi, A. Koivula, M. Wada, S. Kimura, T. Okamoto, M. Penttilä, T. Ando, M. Samejima, Traffic jams reduce hydrolytic efficiency of cellulase on cellulose surface, *Science* 333 (2011) 1279–1282.
- [7] M. Kuršin, P. Våljamäe, Processivity of cellobiohydrolases is limited by the substrate, *J. Biol. Chem.* 286 (2011) 169–177.
- [8] J. Jalak, M. Kurašin, H. Teugas, P. Våljamäe, Endo-exo synergism in cellulose hydrolysis revisited, *J. Biol. Chem.* 287 (2012) 28802–28815.
- [9] C.M. Payne, B.C. Knott, H.B. Mayes, H. Hansson, M.E. Himmel, M. Sandgren, J. Stahlberg, G.T. Beckham, Fungal cellulases, *Chem. Rev.* 115 (2015) 1308–1448.
- [10] V.G.H. Eijsink, G. Vaaje-Kolstad, K.M. Varum, S.J. Horn, Towards new enzymes for biofuels: lessons from chitinase research, *Trends Biotechnol.* 26 (2008) 228–235.
- [11] B.L. Cantarel, P.M. Coutinho, C. Rancurel, T. Bernard, V. Lombard, B. Henrissat, The carbohydrate-active enzymes database (CAZy): an expert resource for glycogenomics, *Nucleic Acids Res.* 37 (2009) D233–D238.
- [12] <http://www.cazy.org/>.
- [13] J. Sakon, D. Irwin, D.B. Wilson, P.A. Karplus, Structure and mechanism of endo/exocellulase E4 from *Thermomonospora fusca*, *Nat. Struct. Biol.* 4 (1997) 810–818.
- [14] Y. Li, D.C. Irwin, D.B. Wilson, Increased crystalline cellulose activity via combinations of amino acid changes in the family 9 catalytic domain and family 3c cellulose binding module of *Thermobifida fusca* Cel9A, *Appl. Environ. Microbiol.* 76 (2010) 2582–2588.
- [15] D.B. Wilson, Microbial diversity of cellulose hydrolysis, *Curr. Opin. Microbiol.* 14 (2011) 259–263.

- [16] E.A. Bayer, J.-P. Belaich, Y. Shoham, R. Lamed, The cellulosomes: multienzyme machines for degradation of plant cell wall polysaccharides, *Ann. Rev. Microbiol.* 58 (2004) 521–554.
- [17] R. Brunecky, M. Alahuhta, Q. Xu, B.S. Donohoe, M.F. Crowley, I.A. Kataeva, S.-J. Yang, M.G. Resch, M.W.W. Adams, V.V. Lunin, M.E. Himmel, Y.J. Bomble, Revealing nature's cellulase diversity: the digestion mechanism of *Caldicellulosiruptor bescii* CelA, *Science* 342 (2013) 1513–1516.
- [18] E. Vlasenko, M. Schülein, J. Cherry, F. Xu, Substrate specificity of family 5 G, 7, 9, 12, and 45 endoglucanases, *Bioresour. Technol.* 101 (2010) 2405–2411.
- [19] V.O.A. Pellegrini, V.I. Serpa, A.S. Godoy, C.M. Camilo, A. Bernardes, C.A. Rezende, N. Pereira Jr., J.P.L. Franco Cairo, F.M. Squina, I. Polikarpov, Recombinant *Trichoderma harzianum* endoglucanase I (Cel7B) is a highly acidic and promiscuous carbohydrate-active enzyme, *Appl. Microbiol. Biotechnol.* 99 (2015) 9591–9604.
- [20] M.A. Van Sluys, C.B. Monteiro-Vitorello, L.E. Camargo, C.F. Menck, A.C. Da Silva, J.A. Ferro, M.C. Oliveira, J.C. Setubal, J.P. Kitajima, A.J. Simpson, Comparative genomic analysis of plant-associated bacteria, *Annu. Rev. Phytopathol.* 40 (2002) 169–189.
- [21] I.W. Sutherland, Xanthan, in: J.G. Swings, E.L. Civerolo (Eds.), *Xanthomonas*, London, Chapman & Hall, 1993, pp. 363–388.
- [22] C.L. Gough, J.M. Dow, J. Keen, B. Henrissat, M.J. Daniels, Nucleotide sequence of the engXCA gene encoding the major endoglucanase of *Xanthomonas campestris* pv. *campestris*, *Gene* 89 (1990) 53–59.
- [23] S.F. Altschul, T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller, D.J. Lipman, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, *Nucleic Acids Res.* 25 (1997) 3389–3402.
- [24] S.M. Holt, P.A. Hartman, A Zymogram method to detect endoglucanases from *Bacillus subtilis*, *Myrothecium verrucaria* and *Trichoderma reesei*, *J. Ind. Microbiol.* 13 (1994) 2–4.
- [25] F.R. Rosseto, A.C. Puhl, M.O. Andrade, I. Polikarpov, Crystallization and preliminary diffraction analysis of the catalytic domain of major extracellular endoglucanase from *Xanthomonas campestris* pv. *campestris*, *Acta Crystallogr. Sect. F: Struct. Biol. Cryst. Commun.* 69 (2013) 137–140.
- [26] E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, M.R. Wilkins, R.D. Appel, A. Bairoch, Protein identification and analysis tools on the ExPASy Server, in: J.M. Walker (Ed.), *The Proteomics Protocols Handbook*, Humana Press Inc., Totowa, NJ, 2005, pp. 571–607.
- [27] J. Kaur, B.S. Chadha, B.A. Kumar, H.S. Saini, Purification and characterization of two endoglucanases from *Melanocarpus* sp. MTCC 3922, *Bioresour. Technol.* 98 (2007) 74–81.
- [28] G.L. Miller, Use of dinitrosalicylic acid reagent for determination of reducing sugars, *Anal. Chem.* 31 (1959) 426–428.
- [29] A. Guinier, G. Fournet, *Small-angle of X-Rays*, Wiley, New York, 1955.
- [30] D.I. Svergun, Determination of the regularization parameter in indirect transform methods using perceptual criteria, *J. Appl. Cryst.* 25 (1992) 495–503.
- [31] D.I. Svergun, Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing, *Biophys. J.* 76 (1999) 2879–2886.
- [32] D.I. Svergun, M.V. Petoukhov, M.H. Koch, Determination of domain structure of proteins from X-ray solution scattering, *Biophys. J.* 80 (2001) 2946–2953.
- [33] G.Y. Xu, E. Ong, N.R. Gilkes, D.G. Kilburn, D.R. Muhandiram, M. Harris-Brandts, J.P. Carver, L.E. Kay, T.S. Harvey, Solution structure of a cellulose-binding domain from *Cellulomonas fimi* by nuclear magnetic resonance spectroscopy, *Biochemistry* 34 (1995) 6993–7009.
- [34] H.W. Kim, K. Ishikawa, Structure of hyperthermophilic endocellulase from *Pyrococcus horikoshii*, *Proteins* 78 (2010) 496–500.
- [35] M.V. Petoukhov, D.I. Svergun, Global rigid body modelling of macromolecular complexes against small-angle scattering data, *Biophys. J.* 89 (2005) 1237–1250.
- [36] H. Fisher, M. de Oliveira Neto, H.B. Napolitano, I. Polikarpov, A.F. Craievich, Determination of the molecular weight of proteins in solution from single small-angle X-ray scattering measurement on a relative scale, *J. Appl. Cryst.* 43 (2010) 101–109.
- [37] M. Maki, K.T. Leung, W. Qin, The prospects of cellulase-producing bacteria for the bioconversion of lignocellulosic biomass, *Int. J. Biol. Sci.* 5 (2009) 500–516.
- [38] S. Sadhu, T.K. Maiti, Cellulase production by bacteria: a review, *Br. Microbiol. Res. J.* 3 (2013) 235–258.
- [39] M. Saleem, M.S. Akhtar, R. Yasmin, M. Zahid, N.N. Malik, M. Afzal, M.I. Rajoka, Production purification and characterization of β -1,4-endoglucanase from a novel bacterial strain CTP-09 of *Bacillus* sp, *Protein Pept. Lett.* 15 (2008) 402–410.
- [40] J. Dong, Y. Hong, Z. Shao, Z. Liu, Molecular cloning purification, and characterization of a novel, acidic, pH-stable endoglucanase from *Martelella mediterranea*, *J. Microbiol.* 48 (2010) 393–398.
- [41] K. Wang, H. Luo, Y. Bai, P. Shi, H. Huang, X. Xue, B. Yao, A thermophilic endo-1,4- β -glucanase from *Talaromyces emersonii* CBS394.64 with broad substrate specificity and great application potentials, *Appl. Microbiol. Biotechnol.* 98 (2014) 7051–7060.
- [42] C. Liang, Y. Xue, M. Fioroni, F. Rodríguez-Ropero, C. Zhou, U. Schwaneberg, Y. Ma, Cloning and characterization of a thermostable and halo-tolerant endoglucanase from *Thermoanaerobacter tengcongensis* MB4, *Appl. Microbiol. Biotechnol.* 89 (2011) 315–326.
- [43] M. Schmuck, I. Pilz, M. Hayn, H. Esterbauer, Investigation of cellobiohydrolase from *Trichoderma reesei* by small angle X-ray scattering, *Biotechnol. Lett.* 8 (1986) 397–402.
- [44] P.M. Abuja, M. Schmuck, I. Pilz, P. Tomme, T. Claeysens, H. Esterbauer, Structural and functional domains of cellobiohydrolase I from *Trichoderma reesei*, *Eur. Biophys. J.* 15 (1988) 339–342.
- [45] P.M. Abuja, I. Pilz, M. Claeysens, P. Tomme, Domain structure of cellobiohydrolase II as studied by small angle X-ray scattering: close resemblance to cellobiohydrolase I, *Biochem. Biophys. Res. Commun.* 156 (1988) 180–185.
- [46] L.H.F. Lima, V.I. Serpa, F.R. Rosseto, G.R. Sartori, M.O. Neto, L. Martínez, I. Polikarpov, Small-angle X-ray scattering and structural modeling of full-length: cellobiohydrolase I from *Trichoderma harzianum*, *Cellulose* 20 (2013) 1573–1585.
- [47] I. Pilz, E. Schwarz, D.G. Kilburn, R.C. Miller Jr., R.A. Warren, N.R. Gilkes, The tertiary structure of a bacterial cellulase determined by small-angle X-ray scattering analysis, *Biochem. J.* 271 (1990) 277–280.
- [48] H.J. Lee, M. Brown Jr., A comparative structural characterization of two cellobiohydrolases from *Trichoderma reesei*: a high resolution electron microscopy study, *J. Biotechnol.* 57 (1997) 127–136.
- [49] S.V. Pingali, H.M. O'Neill, J. McGaughey, V.S. Urban, C.S. Rempe, L. Petridis, J.C. Smith, B.R. Evans, W.T. Heller, Small angle neutron scattering reveals pH-dependent conformational changes in *Trichoderma reesei* cellobiohydrolase I: implications for enzymatic activity, *J. Biol. Chem.* 286 (2011) 32801–32809.
- [50] M.A. Kadowaki, C.M. Camilo, A.B. Muniz, I. Polikarpov, Functional characterization and low-resolution structure of an endoglucanase Cel45A from the filamentous fungus *Neurospora crassa* OR74A: thermostable enzyme with high activity toward lichenan and β -glucan, *Mol. Biotechnol.* 57 (2015) 574–588.