Biotechnol Lett (2011) 33:131–137
DOI 10.1007/s10529-010-0409-3

ORIGINAL RESEARCH PAPER

Functional and biophysical characterization of a hyperthermostable GH51 α-L-arabinofuranosidase from *Thermotoga petrophila*

Camila Ramos dos Santos · Fábio Márcio Squina · Andréia Meza Navarro · Daiane Patrícia Oldiges · Adriana Franco Paes Leme · Roberto Ruller · Andrew John Mort · Rolf Prade · Mário Tyago Murakami

Received: 21 July 2010 / Accepted: 1 September 2010 / Published online: 25 September 2010
© Springer Science+Business Media B.V. 2010

Abstract A hyperthermostable glycoside hydrolase family 51 (GH51) α-L-arabinofuranosidase from *Thermotoga petrophila* RKU-1 (TpAraF) was cloned, overexpressed, purified and characterized. The recombinant enzyme had optimum activity at pH 6.0 and 70°C with linear α-1,5-linked arabinoheptaose as substrate. The substrate cleavage pattern monitored by capillary zone electrophoresis showed that TpAraF is a classical exo-acting enzyme producing arabinose as its end-product. Far-UV circular dichroism analysis displayed a typical spectrum of α/β barrel proteins analogously observed for other GH51 α-L-arabinofuranosidases. Moreover, TpAraF was crystallized in two crystalline forms, which can be used to determine its crystallographic structure.

Keywords α-L-arabinofuranosidase · Glycoside hydrolase family 51 · Thermostability · Thermotoga petrophila

Electronic supplementary material The online version of this article (doi:10.1007/s10529-010-0409-3) contains supplementary material, which is available to authorized users.

The first two authors have contributed equally to this work.

C. R. dos Santos · A. M. Navarro · D. P. Oldiges · A. F. P. Leme · M. T. Murakami (✉)
Laboratório Nacional de Biociências (LNBio), Centro Nacional de Pesquisa em Energia e Materiais, Rua giuseppe maximo scolfaro, 10000, Campinas, SP 13083-970, Brazil
e-mail: mtmurakami@lnls.br

F. M. Squina · R. Ruller
Laboratório Nacional de Ciência e Tecnologia do Bioetanol (CTBE), Centro Nacional de Pesquisa em Energia e Materiais, Campinas, SP, Brazil

A. J. Mort
Department of Biochemistry, Oklahoma State University, Stillwater, OK, USA

R. Prade
Department of Microbiology and Molecular Genetics, Oklahoma State University, Stillwater, OK, USA

Introduction

α-L-arabinofuranosidases (EC 3.2.1.55) are exo-acting hemicellulases that hydrolyze the terminal α-1,2, α-1,3 and α-1,5 bonds to release arabinofuranosyl residues from different oligosaccharides and polysaccharides including arabinan, arabinogalactans and heteroxylans (Kaji 1984). According to the CAZy classification (http://www.cazy.org), these enzymes belong to glycoside hydrolase families (GHs) 3, 43, 51, 54, 62 and 93, exhibiting a diversity of three-dimensional molecular architectures. α-L-arabinofuranosidases have been employed synergistically with other glycoside hydrolases, such as arabinanases (EC 3.2.1.99) and xylanases (EC 3.2.1.8), in several agroindustrial processes including improvement of animal feedstock digestibility, delignification of pulp, increasing the aroma during wine making, and

clarification of juices (Saha 2000). Recently, these enzymes have attracted attention due to their potential application on the reduction of lignocellulosic biomass into fermentable sugars for biofuels production (Sheehan and Himmel 1999). Thermostable α-L-arabinofuranosidases are needed for many applications as high temperatures result in increased reaction velocities, reduced risk of contamination and high substrate solubility (Beguin and Aubert 1994). Therefore, thermostable α-L-arabinofuranosidases have considerable industrial interests.

Thermotoga petrophila RKU-1 is a hyperthermophilic bacterium isolated from the Kubiki oil reservoir in Niigata, Japan, which grows optimally at 80°C (Takahata et al. 2001). It produces two α-L-arabinofuranosidases: one with 484 amino-acid residues (GenBank accession number ABQ46651) and the other with 644 residues (GenBank accession number ABQ46653), both belonging to GH51. In this study, the 484 residue GH51 α-L-arabinofuranosidase from *T. petrophila* RKU-1 (TpAraF) was overexpressed in *Escherichia coli*, purified and submitted to biophysical and functional characterization. This protein was also crystallized in two crystalline states.

Materials and methods

Cloning of the TpAraF coding sequence

The sequence coding the full-length α-L-arabinofuranosidase from *T. petrophila* RKU-1 (GenBank accession number ABQ46651) was amplified from a genomic DNA of *T. petrophila* by a standard PCR method using two oligonucleotide primers (forward, 5′-GCTAG CTACAGGATAGTGGTTGATCC-3′; reverse.

5′-GGATCCTTACTCCAATTCTACCTCAATC-3′). The amplified *TpAraF* gene was cloned into pGEM-T vector and then subcloned into the restriction-enzyme sites *NheI* and *Bam*HI (endonuclease sites are underlined at the primers sequence) of the pET28a vector with a hexahistidine-tag at the *N*-terminus.

Protein expression

E. coli BL21(DE3)ΔSlyD cells harboring pRARE2 plasmid were transformed with pET28a/TpAraF plasmid and plated in selective solid LB medium containing kanamycin (50 μg/ml) and chloramphenicol (37 μg/ml) antibiotics. A single colony was grown in liquid LB-antibiotics for 16 h at 37°C and 200 rpm. The culture was diluted at 2% (v/v) into fresh LB-antibiotics and grown at same conditions to an OD_{600nm} of 0.8. Then, the temperature was changed to 30°C and the recombinant protein expression was induced by adding 0.5 mM IPTG. After 4 h the cells were harvested at 6,000×*g* and stored at −20°C.

Purification

Cells were re-suspended in lysis buffer (20 mM sodium phosphate pH 7.5, 500 mM NaCl, 5 mM imidazole, 1 mM benzamidine and 5 mM PMSF) and disrupted by lysozyme (80 μg/ml, 1 h, on ice) and then sonication (6 pulses of 30 s at 500 W; VC750 Ultrasonic Processor, Sonics Vibracell). The extract was centrifuged at 10,000×*g* for 30 min at 4°C and TpAraF was purified from the supernatant using two chromatographic steps with an ÄKTA FPLC system at 18°C. First, the supernatant (40 ml) was loaded onto a 1 ml HiTrap Chelating HP column charged with Ni^{2+} and pre-equilibrated with buffer A (20 mM sodium phosphate pH 7.5, 500 mM NaCl, 5 mM imidazole) at 1 ml/min. The column was washed with ten column volumes (CVs) of buffer A to remove unbound fractions and bound fractions were eluted with a non-linear, imidazole gradient from 5 to 500 mM in 20 CVs. Fractions containing TpAraF were pooled and concentrated to 2 ml by filtration using a 30 kDa-pore Amicon Ultra-4 filter unit (Millipore) before subsequent passage through a Superdex 200 16/60 column which had been pre-equilibrated with 20 mM sodium phosphate buffer pH 7.5 containing 150 mM NaCl at 0.5 ml/min; elution was monitored at 280 nm. The purified TpAraF was analyzed by SDS-PAGE and protein concentration was determined at 280 nm using the molar extinction coefficient (94,115 M^{-1} cm^{-1}) (Edelhoch 1967).

Mass spectrometry

The protein band was excised, reduced, alkylated and submitted to in-gel digestion with trypsin. An aliquot (4.5 μl) of the resulting peptide mixture was separated by C18 (75 μm × 100 mm) RP-nanoUPLC

(nanoAcquity, Waters) coupled with a Q-T of ultima mass spectrometer (Waters) with nano-electrospray source at 0.6 µl/min. The gradient was 2–90% (v/v) acetonitrile in 0.1% (v/v) formic acid over 45 min. The instrument was operated in the 'top three' mode, in which one MS spectrum is acquired followed by MS/MS of the top three most-intense peaks detected. The spectra were acquired using software MassLynx v.4.1 and the raw data files were converted to a peak list format (mgf) by the software Mascot Distiller v.2.3.2.0, 2009 (Matrix Science Ltd.) and searched against non-redundant protein database (NCBI nr 2009.07.20, 9,298,190 sequences) using engine Mascot v.2.3 (Matrix Science Ltd.), with carbamidomethylation as fixed modification, oxidation of methionine as variable modification, one trypsin missed cleavage and a tolerance of 0.1 Da for both precursor and fragment ions.

Circular dichroism spectroscopy

Far-UV CD spectra were recorded on a Jasco J-810 spectropolarimeter (Jasco International Co. Ltd., Tokyo, Japan) from 190 to 260 nm in a 1 mm quartz cuvette. The purified sample of TpAraF was used at 3 µM in 10 mM of sodium phosphate buffer, pH 7.0. The data collection parameters were set to scan rate of 50 nm/min, response time of 4 s, sensitivity of 100 mdeg, scan step of 0.5 nm and accumulation of 10. Baseline subtraction, smoothing and data normalization were carried out using the graphical software ORIGIN (http://www.originlab.com/). The CD data are shown as mean residue ellipticity units (deg cm^2 $dmol^{-1}$). The secondary structure contents were evaluated by deconvolution of the CD spectrum using the DichroWeb server (Whitmore and Wallace 2004).

Fluorescence spectroscopy

Intrinsic tryptophan fluorescence emission (ITFE) experiments were carried out on a Cary varian spectrofluorimeter (Varian, USA) using a 2×10 mm^2 quartz cuvette. The excitation wavelength was 295 nm and emission spectra were recorded between 305 and 400 nm. Measurements were carried out with TpAraF at 3 µM in 10 mM sodium phosphate buffer, pH 7.0. The signal was acquired for 1 s and the wavelength increment was set to 1 nm. Blank corrections were made in all spectra.

Enzymatic assay

The enzymatic assays for TpAraF were performed following Squina et al. (2009). 50 µl of substrate solution (0.5% polysaccharide content) was incubated with 25 µl diluted enzyme for 30 min at different temperature and pHs. For thermostability evaluation, the enzyme was incubated at 70, 80, 90, and 95°C and aliquots of enzyme were taken at intervals and the residual activity was measured. The enzymatic activity was estimated by the amount of reducing sugar liberated from arabinoheptaose (from Megazyme, Ireland, and Sigma-Aldrich) as determined by the DNS method. One unit of enzyme was defined as the quantity of enzyme that liberated reducing sugar (arabinose) at rate of 1 µmol/min.

Capillary zone electrophoresis (CZE)

Arabinoheptaose (from Megazyme) was derivatized with 8-aminopyreno-1,3,6-trisulfonic acid (APTS) by reductive amination as described by Naran et al. (2007). Enzymatic reactions were performed as a forementioned, except that 1 pM of APTS-labeled arabinoheptaose was used as a substrate. CZE analysis of arabinoheptaose break down products was performed on a BioFocus 2000 (Bio-Rad Laboratories, Inc.) with laser-induced fluorescence detection. A fused-silica capillary (TSP050375, Polymicro Technologies) of internal diameter 50 µm and length 31 cm was used as the separation column for oligosaccharides. Electrophoresis conditions were 15 kV/70–100 µA using 100 mM sodium phosphate (pH 2.5) as running buffer and a controlled temperature at 20°C. The capillary was rinsed with 1 M NaOH followed by running buffer to prevent carryovers. APTS-labeled oligomers were excited at 488 nm and emission was collected through a 520 nm band pass filter. Because of variations of the experimental conditions at different electrophoretic runs, the retention times can vary slightly.

Crystallization and X-ray diffraction analysis

Crystallization was performed by the sitting-drop vapor-diffusion method using the Cartesian Honey Bee 963 system (Genomic Solutions). 544 different formulations based on commercial crystallization kits

including Hampton Research (SaltRX and Crystal Screens I and II), 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 6.7 mg/ml were mixed with an equal volume of the screening solution and equilibrated over a reservoir containing 80 µl of the mother solution.

For crystal optimization, systematic grid refinement, different drop volume ratios and hanging-drop vapor-diffusion technique were employed. Best crystals were grown using the hanging-drop vapor-diffusion method by mixing 1 µl protein with an equal volume of the mother solution consisting of 100 mM bis-Tris pH 5.5, 200 mM ammonium acetate, 40% (v/v) MPD and 1% (v/v) dioxane (form I). A second refinement step was performed using 5% (v/v) glycerol as additive and changing the ammonium acetate to magnesium chloride, which resulted in a different crystalline form (form II).

Both crystalline forms were submitted to X-ray diffraction at the MX2 beamline at the Brazilian Synchrotron Light Laboratory (Campinas, Brazil). Diffraction data were collected to a maximum resolution of 3.0 Å; however depending on the crystalline state the data resolution was cut off for adequate statistics. The data were indexed, integrated and scaled using the DENZO and SCALEPACK programs from the HKL2000 package (Otwinowski and Minor 1997).

Results and discussion

Overexpression, purification and mass spectrometry protein analysis

TpAraF was expressed in BL21(DE3)ΔSlyD cells harboring pRARE2 plasmid in the soluble fraction at 30°C. The enzyme was purified using two chromatographic steps, IMAC (Fig. 1a), and SEC (Fig. 1b), yielding approximately 0.4 mg high purity protein per liter of expression. The expected molecular mass of TpAraF is 57 kDa and agrees with the single band observed in the SDS-PAGE (Fig. 1c). This band was trypsin digested and peptide analysis by mass spectrometry confirmed to be TpAraF (Supplementary Table 1).

Biophysical characterization

The fluorescence spectrum obtained for TpAraF resulted in a maximum emission at 332 nm (Fig. 2a) indicating that most of the tryptophan residues are in a non-polar environment, mainly buried in the structure. All the ten Trp residues are located at the (β/α)-barrel domain and seven of them are conserved in *Geobacillus stearothermophilus* protein (GenBank accession number ACE73682).

Far-UV CD spectrum of TpAraF showed a positive peak at 194 nm and a minimum negative peak at 220 nm (Fig. 2b). Deconvolution of CD data using

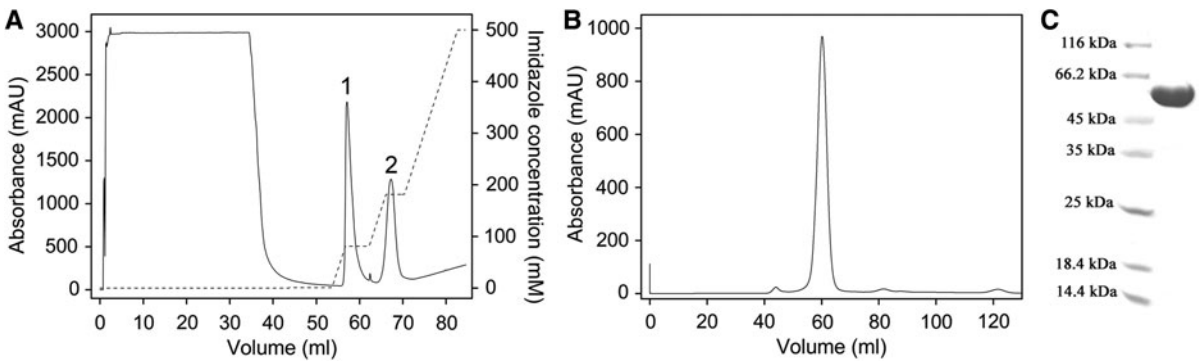

Fig. 1 **a** Elution profile of IMAC on a 1 ml HiTrap Chelating HP column. Bound fractions were eluted using non-linear imidazole gradient (*dashed line*) ranging from 5 to 500 mM in buffer A (20 mM sodium phosphate pH 7.5, 500 mM NaCl, 5 mM imidazole) at 1 ml/min. The peak 2 corresponds to TpAraF. **b** Size-exclusion chromatogram of TpAraF on a Superdex 200 16/60 column in 20 mM sodium phosphate buffer pH 7.5 containing 150 mM NaCl at a flow rate of 0.5 ml/min. The eluants were monitored at 280 nm (*solid line*). **c** SDS-PAGE of purified TpAraF. *Lane 1* molecular weight markers and *lane 2* TpAraF

Table 1 Data collection statistics

	Form I	Form II
Data collection		
Temperature (K)	100	100
Radiation source	Brazilian synchrotron light laboratory	Brazilian synchrotron light laboratory
Beamline	W01B-MX2	W01B-MX2
Wavelength used (Å)	1.458	1.458
Detector	MAR Mosaic 225 mm	MAR Mosaic 225 mm
Space group	$P2_1$	$H3$
Unit-cell parameters (Å, °)	a = 106.09, b = 187.78 and c = 181.14; β = 90.73	a = b = 110.34 and c = 251.57
Resolution range (Å)	35.0–3.10 (3.21–3.10)	30.0–3.20 (3.31–3.20)
R_{merge} (%)[a]	12.8 (40.5)	9.2 (45.2)
$<I/\sigma(I)>$	7.7 (2.6)	14.9 (2.5)
Data completeness (%)	98.2 (96.8)	96.4 (85.2)
Number of unique reflections	125, 287	18, 176
Data analysis		
Vm ($Å^3$ Da^{-1})	2.65	2.62
Solvent content (%)	53.7	53.1
Molecules per asymmetric unit	12	2

[a] $R_{merge} = \Sigma_{hkl} \Sigma_i | I_i(hkl) - < I(hkl) > |/\Sigma_{hkl} \Sigma_i I_i(hkl)$, where $I_i(hkl)$ is the ith observation of reflection *hkl* and $< I(hkl) >$ is the weighted average intensity for all observations *i* of reflection *hkl*

In parentheses are values for the last resolution shell

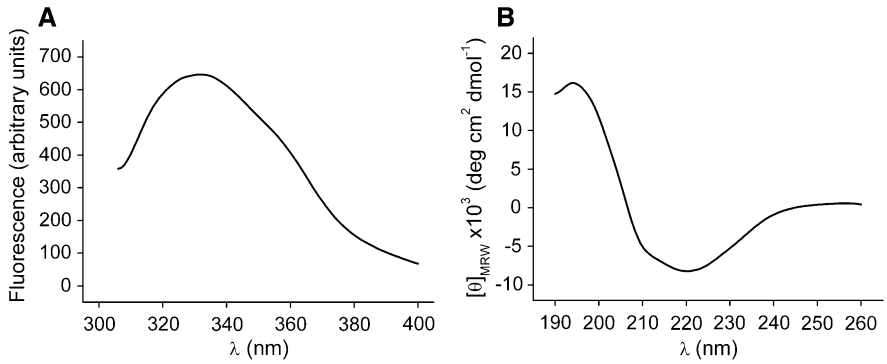

Fig. 2 ITFE **a** and Far-UV CD **b** spectra of recombinant TpAraF. The fluorescence emission spectrum was recorded after excitation of the Trp residues at 295 nm. Both experiments were carried out with 3 μM TpAraF in 10 mM sodium phosphate buffer at 20°C

Contin method available in the Dichro web server (Whitmore and Wallace 2004) resulted in secondary structure (SS) contents of 19% α-helix, 33% β-sheet and 48% random coil. These values are in full agreement with SS contents observed for other GH51 arabinofuranosidases from *G. stearothermophilus* (PDB codes: 1PZ2, 1PZ3, 1QW8, and 1QW9), *Thermobacillus xylaniloyticus* (PDB codes: 2VRQ, and 2VRK) and *Clostridium thermocellum* (PDB codes: 2C7F, and 2C8N).

Enzymatic characterization of recombinant TpAraF

The optimum pH and temperature for enzymatic activity of purified recombinant TpAraF were 6.0 and 70°C, respectively (Fig. 3a and b). The enzyme demonstrated to be very stable with virtually 100% of residual activity after 20 h at 90°C (Fig. 3c).

To confirm the exo-type activity of TpAraF over terminal arabinofuranosyl moieties, the reaction

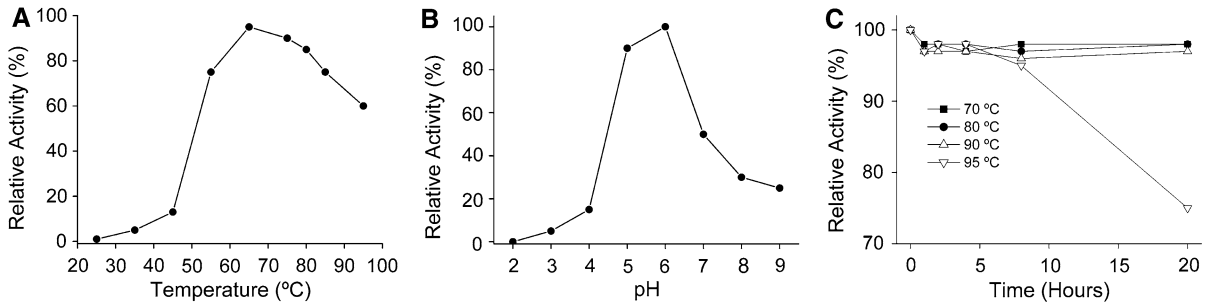

Fig. 3 Enzymatic characterization of TpAraF. Effect of temperature **a** and pH **b** on the TpAraF activity using arabinoheptaose as substrate. **c** Thermostability assay of TpAraF at 70°C (*inverted open triangle*), 80°C (*open triangle*), 90°C (*Filled square*) and 95°C (*Filled circle*)

products of enzymatic action on arabinoheptaose were analyzed by CZE (Supplementary Fig. 1). In the absence of TpAraF enzyme (Supplementary Fig. 1a), heptaose was predominant. During the enzymatic reaction, the first product formed was hexaose, with consumption of heptaose (Supplementary Fig. 1b), indicating that the hydrolysis occurred at the terminal residue. The hexaose was further hydrolyzed producing pentaose, and then consecutively until the result in arabinose was the final end-product.

X-ray diffraction analysis

Crystal form I (Fig. 4a) diffracted to 3.1 Å resolution and was indexed in the primitive monoclinic crystal system with unit cell parameters a = 106.09, b = 187.78, and c = 181.14 Å, and $\beta = 90.73^{\circ}$. An examination of the systematic absences indicated that the crystals belonged to space group $P2_1$. Crystal form II (Fig. 4b) diffracted to 3.2 Å resolution and belonged to rhombohedral lattice system with unit cell parameters a = b = 110.34, and c = 251.57 Å. In this case, space group ambiguity between $H3$ and $H3_2$ cannot be solved by scaling and merging of intensities. Taking into consideration the molecular weight of 57,000 Da for the monomer, 12 molecules are present in the asymmetric unit for the crystalline form I (solvent content of 53.7% and Matthews coefficient of 2.65 $Å^3$/Da), whereas for the crystalline form II only two molecules are present (solvent content of 53.1% and Matthews coefficient of 2.62 $Å^3$/Da) (Matthews 1968). Data-processing statistics of both data sets are presented in Table 1.

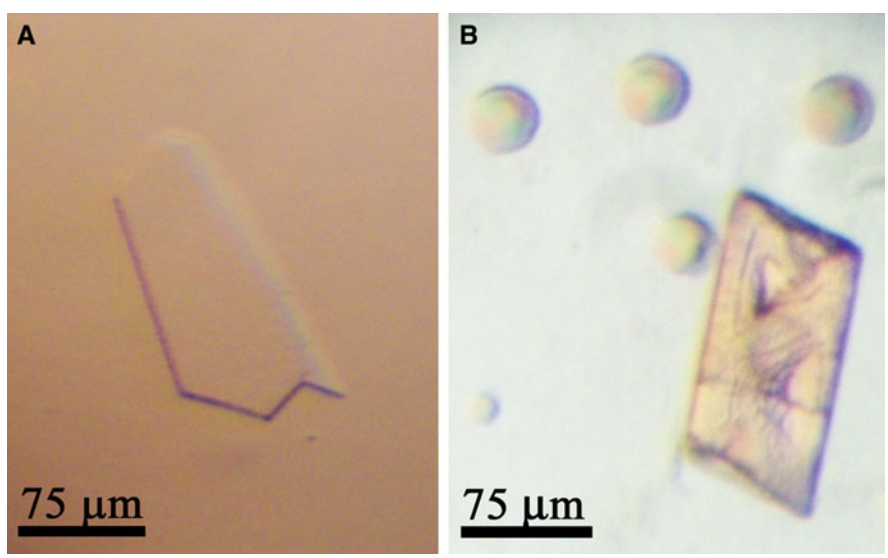

Fig. 4 Microphotographs of TpAraF crystals obtained from conditions **a** 100 mM bis–tris pH 5.5, 200 mM ammonium acetate, 40% (v/v) MPD and 1% (v/v) dioxane (*form I*), **b** 100 mM bis-Tris pH 5.5, 200 mM magnesium chloride, 55% (v/v) MPD and 5% (v/v) glycerol (*form II*). The approximate dimensions of the *form I* and *form II* were 150 × 75 × 20 µm and 150 × 75 × 75 µm, respectively

Conclusions

A hyperthermostable GH51 α-L-arabinofuranosidase from *T. petrophila* RKU-1 that is promising for biotechnological applications was overexpressed, purified and characterized by biochemical and biophysical methods. The recombinant TpAraF is a classical exo-acting enzyme with optimum activity at pH 6.0 and 70°C. Spectroscopic studies indicated a well-folded protein with a typical CD spectrum of α/β barrel proteins similar to other GH51 α-L-arabinofuranosidases. Moreover, TpAraF was crystallized in two crystalline forms, which will be employed to determine its crystallographic structure.

Acknowledgments This research was supported by grants from Fundação de Amparo a Pesquisa do Estado de São Paulo (FAPESP, 08/58037-9) to FMS and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, 478059/2009-4) to MTM. RAP has funding from the Department of Energy, awards 06103-OKL and ZDJ-7-77608-01.

References

Beguin P, Aubert JP (1994) The biological degradation of cellulose. FEMS Microbiol Lett 13:25–58

Edelhoch H (1967) Spectroscopic determination of tryptophan and tyrosine in proteins. Biochemistry 6:1948–1954

Kaji A (1984) L-Arabinosidases. Adv Carbohydr Chem Biochem 47:383–394

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680–685

Matthews BW (1968) Solvent content of protein crystals. J Mol Biol 33:491–497

Naran R, Pierce ML, Mort AJ (2007) Detection and identification of rhamnogalacturonan lyase activity in intercellular spaces of expanding cotton cotyledons. Plant J 50: 95–107

Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. Method Enzymol 276:307–326

Saha BC (2000) Alpha-L-arabinofuranosidases: biochemistry, molecular biology and application in biotechnology. Biotechnol Adv 18:403–423

Sheehan J, Himmel M (1999) Enzymes, energy, and the environment: a strategic perspective on the U.S. department of energy's research and development activities for Bioethanol. Biotechnol Progr 15:817–827

Squina FM, Mort AJ, Decker SR et al (2009) Xylan decomposition by *Aspergillus clavatus* endo-xylanase. Protein Expres Purif 68:65–71

Takahata Y, Nishijima M, Hoaki T et al (2001) *Thermotoga petrophila* sp. nov. and *Thermotoga naphthophila* sp. nov., two hyperthermophilic bacteria from the kubiki oil reservoir in Niigata, Japan. Int J Syst Evol Micr 51: 1901–1909

Whitmore L, Wallace BA (2004) DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. Nucleic Acids Res 32:W668–W673