The linker region plays a key role in the adaptation to cold of the cellulase from an Antarctic bacterium

Guillaume K Sonan, Véronique Receveur-Brechot, Colette Duez, Nushin Aghajari, Mirjam Czjzek, Richard Haser, Charles Gerday

Biochemical Journal Immediate Publication. Published on 18 Jul 2007 as manuscript BJ20070640

The linker region plays a key role in the adaptation to cold of the cellulase from an Antarctic bacterium

Guillaume K. Sonan[*], Véronique Receveur-Brechot[†], Colette Duez[*], Nushin Aghajari[‡], Mirjam Czjzek[§], Richard Haser[‡] and Charles Gerday[*1]

[*]*Laboratoire de Biochimie et Centre d'Ingénierie des Protéines, Institut de Chimie B6*
Université de Liège, B-4000 Liège Sart-Tilman, Belgium

[†]*Architecture et Fonction des Macromolécules Biologiques*
UMR 6098, CNRS et Universités d'Aix-Marseille I et II
163 avenue de Luminy, F-13488 Marseille cedex, France

[‡]*Laboratoire de Bio Cristallographie, Institut de Biologie et Chimie des Protéines*
CNRS et Université Claude Bernard Lyon 1, UMR 5086
IFR 128 "Biosciences Lyon-Gerland", 7 Passage du Vercors, F-69367 Lyon Cedex 07
France

[§]*Station Biologique de Roscoff*
Végétaux Marins et Biomolécules UMR 7139
Place George Teissier, BP 74, F-29682 Roscoff Cedex, France

1: Corresponding author

Gerday Charles
e-mail: ch.gerday@ulg.ac.be

Short title: Cold-adapted cellulase from *Pseudoalteromonas. haloplanktis*

Synopsis

The psychrophilic cellulase, Cel5G, from the Antarctic bacterium *Pseudoalteromonas haloplanktis* is composed of a catalytic module (CM) joined to a carbohydrate binding module (CBM) by an unusually long, extended and flexible linker (LR) containing three loops closed by three disulfide bridges. To evaluate the possible role of this region in cold adaptation, the linker was sequentially shortened by protein engineering successively deleting one and two loops of this module whereas the last disulfide bridge was also suppressed by replacing the last two cysteine by two alanine residues. The kinetic and thermodynamic properties of the mutants were compared to those of the full-length enzyme, also to those of the cold-adapted catalytic module alone and to those of the mesophilic homologous enzyme, Cel5A, from *Erwinia chrysanthemi*. The thermostability of the mutated enzymes as well as their relative flexibility were evaluated by differential scanning calorimetry and fluorescence quenching respectively. The topology of the structure of the shortest mutant was determined by small angle X-ray scattering (SAXS). The data indicate that the sequential shortening of the linker induces a regular decrease of the specific activity towards macromolecular substrates, reduces the relative flexibility and concomitantly increases the thermostability of the shortened enzymes. This demonstrates that the long linker of the full-length enzyme favours the catalytic efficiency at low and moderate temperatures by rendering the structure less compact but also less stable and plays a crucial role in the adaptation to cold of this cellulolytic enzyme.

Keywords: cellulase, psychrophile, *Pseudoalteromonas haloplanktis,* linker, SAXS, enzyme.

Introduction

Cellulose is a natural component composed of linear polymers of thousands of glucose residues linked by β-1,4 glycosidic bonds. Cellulose differs from other polysaccharides by its insoluble and rigid structure, leading to a natural resistance to biological degradation [1]. To exploit the energy and carbon available in cellulose, organisms such as fungi and bacteria produce mixtures of synergistically acting cellulases. Most fungal cellulases are glycoproteins and exist in multiple forms whereas bacteria produce mainly non-glycosylated endoglucanases [2]. Microbial cellulases are often made of a three-domain structure consisting of a catalytic module (CM) and a cellulose-binding module (CBM), separated by a

distinct linker region (LR) comprising a simple or repetitive sequence often rich in proline, threonine, serine or glycine [1; 3-8]. Recent studies have shown that linker domains can play an essential role in catalysis by favouring cooperative inter-domain interactions [9-13].

The Gram-negative Antarctic bacterium *Pseudoalteromonas haloplanktis* A23 collected from sea-water secretes a psychrophilic cellulase, Cel5G (EC 3.2.1.4) belonging to family 5 of glycosyl hydrolases and subfamily 2 (GH5-2). This cellulase is composed of a catalytic module long of 292 residues organized in a $(\beta/\alpha)_8$ barrel and a carbohydrate-binding module of 61 residues separated by a flexible and unusually extended 109 amino acids long linker region [14, 15]. This cold-adapted enzyme displays the usual properties of enzymes produced by psychrophilic organisms i.e. a high specific activity at low and moderate temperatures, significantly higher than that of the mesophilic counterparts and a rather high thermosensitivity induced by a decrease of the intramolecular interactions [16,17].

The crystal structure of the CM of Cel5G and a model structure of the CBM have been compared in details with similar parts of the mesophilic cellulase Cel5A from *Erwinia chrysanthemi* [17-19]. The catalytic module and the carbohydrate module of Cel5G respectively share 64% and 57% of sequence identity with their mesophilic counterparts. However the linker regions are very different. Indeed, in Cel5G, the 109 residues linker is rich in Gly, Val, Ser, Thr, Cys, Asp and Asn while the linker of the mesophilic counterpart, Cel5A, is a peptide of 34 residues rich in Ser and Thr. In addition, the Cel5G linker contains six cysteine residues engaged in disulfide bridges, resulting in the formation of three transversal loops of 13 residues [17]. The structural comparison of the catalytic modules of Cel5G, solved by X-ray crystallography, [17] and that of the mesophilic counterpart Cel5A [20, 21], associated with the analysis of the entire conformation of the psychrophilic cellulase obtained by small angle x-ray scattering (SAXS), suggests that the long linker is an extended and highly flexible structure , and that the cold adaptation of Cel5G could not only arises from the specific properties of the CM but also from the unusual properties of the linker [17].

In the present report this hypothesis was checked by studying the kinetic, thermodynamic, and stability properties of three variants of Cel5G, resulting from the gradual shortening of the linker through the sequential removal of two of the loops closed by the disulfide bridges, as well as of the last disulfide bond. The relative flexibility of the three mutants was also checked by fluorescence quenching using acrylamide as a quencher, and by SAXS.

Experimental

Strains and plasmids

Escherichia coli XL1-blue (Stratagene), RR1 (Stratagene), and JM110 (Invitrogen) were used for DNA cloning. *E. coli* BL21(DE3) (Novagen) and *E. coli* XL1-Blue MRF' (Stratagene) were used to produce the psychrophilic and mesophilic cellulase respectively.

The gene coding for the *P. haloplanktis* A23 Cel5G cellulase was cloned in the expression plasmid pET22b(+) (Novagen), between the NdeI and SacI restriction sites yielding pET22*cel5G* [14]. The pSNAB1 plasmid carrying the *cel5A* gene from the mesophilic *Erwinia chrysanthemi* 3937, has been previously described [19,20].

DNA techniques

Standard procedures for recombinant DNA technology were used as described by Sambrook *et al.* [22]. DNA fragments were isolated from 1% agarose gels using the Qiaquick Purification Kit (Qiagen, Westburg, The Netherlands). The deletions were introduced by inverse PCR using Vent DNA polymerase (New England Biolabs, Beverly, MA, USA) that possesses a proofreading activity. Cys 396 was changed for Ala by inverse PCR while the Cys382Ala mutant was obtained using a site-directed mutagenesis kit (Stratagene). All mutant constructs were verified by restriction mapping, and subsequently the accuracy of the genetic information in the modified fragments was controlled by their complete sequencing.

Construction of the truncated psychrophilic cellulase, Cel5G$_{CM}$

To reduce the size of the template used in inverse PCR, the *cel5G* gene was recovered from the pET22*cel5G* plasmid by digestion with NdeI and NotI and subcloned in the corresponding sites of the 1.8 kb smaller pET-23a (+) vector. A termination TGA codon was introduced at the 3' end of the catalytic module (CM) encoding sequence by using oligonucleotides 1 and 2 (Table 1) as primers. The expected plasmid was digested with NdeI and XhoI and the fragment coding for the truncated cellulase was isolated and replaced in the initial expression vector yielding pET22*cel5G$_{CM}$*.

Introduction of in-frame deletions in the Cel5G linker region. Construction of the Cel5GΔ1P (ΔAsp312-Val342), and Cel5GΔ2P (ΔAsp312-Ile379) truncated proteins.

The 2.6 kb NdeI-XhoI fragment containing the *cel5G* gene was isolated from the pET22*cel5G* plasmid and subcloned in the small pSP73 cloning vector (Promega, Madison, USA) to give pSP*cel5G*. The Cel5GΔ1P and Cel5GΔ2P mutants were obtained by inverse PCR using the primers 3 and 4 or 5 and 6 (Table 1) respectively. The inserts were completely sequenced with the *cel5G* specific primers: 5'-GGAACGCCTACGTGGTCGCAAG-3' and 5'-GGTAGTAGCAATGTTGATTTAGATAGCG-3'.

Construction of the Cel5GΔ2PmutCC mutant

The pET22*cel5GΔ2P* plasmid was used as template for site-directed mutagenesis. The experiment was performed in two steps. The oligonucleotides 7 and 8 (Table 1) were used to exchange Cys382 for Ala and the resulting plasmid carrying this single mutated cysteine served as template to change Cys396 for Ala using the oligonucleotides 9 and 10 (Table 1). The plasmid containing the two mutated cysteine residues was named pET22*cel5GΔ2PmutCC.*

Production of the various psychrophilic proteins

E. coli BL21(DE3) cells transformed with the different recombinant expression plasmids were grown overnight at 25°C on LB agar medium containing 100 µg.ml^{-1} ampicillin, 100 µg.ml^{-1} of trypan blue and 5 g.l^{-1} carboxymethylcellulose. A positive cellulase activity appeared as clear halos on the blue background. The production of all mutants was carried out as described for the full-length cellulase [14]. Freshly transformed positive colonies were picked from LB agar plates and used to inoculate 4 ml of LB broth containing 100 µg.ml^{-1} ampicillin. After 20 hours culture at 18°C, a 10 µl sample was used to inoculate 30 ml of the same medium and the new culture was kept at 18°C up to an A_{550nm}=5-6. One ml of this second pre-culture was then added to 300 ml of LB broth in a 1-l sterile flask and cells were grown up to A_{550nm}=3-4. The enzyme expression was then induced with 0.1 mM isopropyl-1-thio-β-D-galactoside (IPTG) and the cells were collected after 10 hours of induction at 18°C.

Production of the mesophilic cellulase, Cel5A

One colony of freshly transformed *E. coli* XL1-Blue/pSNAB1 was picked from an LB plate to inoculate 2 ml of LB broth supplemented with ampicillin. The culture was incubated at 30°C under agitation. One ml of this pre-culture was used to inoculate 100 ml of the same medium in 500 ml sterile flask and the culture was pursued at 30°C up to reaching an absorbance at 550 nm of 4-5. An aliquot of 8 ml of this pre-culture was used to inoculate 300 ml of LB broth in 1l flask supplemented with 100 µg.ml^{-1} ampicillin and 50 µg.ml^{-1} IPTG. The culture was grown at 30°C under agitation until reaching an absorbance at 550 nm of 5-6.

Extraction of proteins

Cells were harvested by centrifugation at 9000×g for 15 min at 4°C and were subjected to osmotic shock [23]. The cell pellet was suspended in the osmotic shock solution (1/10 volume of the culture broth) containing 20 % sucrose, 0.5 mM EDTA-Na_2, 100 mM Tris-HCl, pH 8.

The suspension was supplemented with 0.1 mM PMSF (Phenylmethylsulfonylfluoride) and mixed in a rotary shaker at 180 rpm. After 10 min, the mixture was centrifuged at 4°C for 15 min at 13000×g. The supernatant was eliminated and the drained pellet was rapidly mixed with a volume (v/v) of cold water containing 1 mM $MgCl_2$. The suspension was mixed for 10 min, centrifuged, and the supernatant, corresponding to the periplasmic fraction, was recovered, supplemented with 0.04 % NaN_3 and stored at -20°C.

Purification of proteins

All manipulations were performed at 4°C. The periplasmic fractions were brought to 75 % $(NH_4)_2SO_4$ saturation, stirred for 1 hour, and the suspensions were then centrifuged at 14500×g for 30 min. The supernatants were discarded and the pellets were dissolved in buffer A: 25 mM Pipes, pH 6.5 containing 0.04 % NaN_3 and 1 mM PMSF.

Purification of the catalytic domain of the psychrophilic cellulase, $Cel5G_{CM}$

The protein solution was diluted twice in buffer B: 50 mM Pipes, pH 6.5, 50 % $(NH_4)_2SO_4$ saturation and loaded onto a Phenyl Sepharose column (3x20 cm, Biorad) equilibrated in buffer C: 25 mM Pipes pH 6.5, 25 % $(NH_4)_2SO_4$ saturation. The column was washed with buffer C and the proteins were eluted at a flow rate of 2 $ml.min^{-1}$ with a linear gradient of $(NH_4)_2SO_4$ (500ml, 25% salt-500ml, 0 % salt) in buffer A. The fractions with cellulase activity were combined and dialysed against buffer 1 (10 mM Hepes, pH 7.5, 0.04% NaN_3, 1 mM PSMF) using an Amicon ultrafiltration unit (Millipore) fitted with a PTGC membrane (10 kDa cut-off, Millipore). The concentrated sample was then loaded on a Macro-prep High Q column (3 x 20 cm, Biorad) equilibrated in buffer 1 and the proteins were eluted with a KCl linear gradient (250 ml buffer 1-250 ml of a 300 mM KCl solution in the same buffer). The active fractions were collected, dialysed against buffer 1, filtered using a 0.45 µm membrane (Millipore) and loaded onto a Mono Q HR 5/5 column (0.5 x 5 cm, Pharmacia Biotech) equilibrated in buffer 1. The column was then washed with buffer 1 and the proteins were eluted at a flow rate of 1 $ml.min^{-1}$ with a KCl linear gradient (30 ml-30 ml, 300 mM KCl). The active fractions were pooled, dialysed against buffer 1, concentrated by ultrafiltration, supplemented with 20 µM Pefabloc [4-(2-amino-ethyl)-benzene-sulfonylfluoride hydrochloride] (Roche) and stored at –70°C. The homogeneity of the purified enzyme was checked by SDS-PAGE using a 12 % polyacrylamide gel.

Purification of the full-length psychrophilic cellulase, Cel5G, its mutants, and the mesophilic cellulase, Cel5A

The protein solutions obtained from the $(NH4)_2SO4$ fractionation of the periplasmic samples were diluted twice in buffer B: 50 mM Pipes, pH 6.5, 50 % $(NH_4)_2SO_4$ saturation and loaded onto a cotton-wool column equilibrated in buffer C (the support was sterilised in water and packed into a 5 x 12 cm Pharmacia column). The column was washed with buffer C and the proteins were eluted with cold Milli-Q water. The fractions containing the active cellulases were collected and dialysed against Buffer 1. Then the cellulase preparations, except the one corresponding to the mesophilic cellulase that was already fully purified from the cotton-wool column, were loaded onto a Macro-prep High Q column and eluted as described for the psychrophilic cellulase $Cel5G_{CM}$.

Determination of protein concentrations

Protein concentrations were determined for some samples with the bicinchoninic acid method [24] using the "Pierce Kit BCA Protein Assay Reagent" (Perbio Science Belgium, N.V.) and bovine serum albumin as standard and by absorbance measurement at 280 nm for the purified proteins, using the extinction coefficients.cm^{-1} of 95650 M^{-1} for Cel5G, Cel5GΔ1P, Cel5GΔ2P, Cel5GΔ2PmutCC, 69900 M^{-1} for $Cel5G_{CM}$ and 102455 M^{-1} for Cel5A.

Assays for Cellulase Activity

p-nitrophenyl-β-D-cellobioside

This was carried out as already described [15].

Macromolecular substrates: Carboxymethylcellulose and amorphous cellulose

50 µl of the cellulase solutions were mixed with 450 µl of a 0.05-8 % carboxymethylcellulose (CMC) [25] or with a 2 % amorphous cellulose solution in 25 mM Pipes, pH 6.5, 0.04 % NaN_3, 1 mM PMSF. After 10 min incubation at temperatures from 4 to 70°C the released amount of reducing sugar was measured by the dinitrosalicylic acid method [26]. One unit of activity (IU) was defined as the amount of enzyme necessary to release 1 µmole of reducing sugar per min.

Differential scanning calorimetry

The experiments were carried out using a microcalorimeter MicroCal VP-DSC (Northampton, MA, USA) and protein samples at concentrations ranging from 0.7 to 1 $mg.ml^{-1}$. The samples

were first dialysed against 30 mM MOPS buffer, pH 7.5, and eluted on a PD 10 column (Pharmacia, Uppsala, Sweden) equilibrated in the buffer supplemented with 500 mM sulphobetaine [3-(1-pyridinio)-1-propane sulphonate] in order to prevent protein aggregation upon heating [27]. The thermograms were obtained using a temperature gradient of 1°C/min and analysed according to a non-two state model. The melting points T_m and the calorimetric enthalpy ΔH_{cal}, of the individual transitions were fitted independently using the MicroCal Original software, version 4.1. The source and the magnitude of errors on the T_m and enthalpy values have been discussed elsewhere [28]. The reversibility of the thermal transitions was assessed by the rescanning of the samples after cooling them at 10°C.

Thermodependence of the activity

The activity of the various cellulases as a function of temperature (4°C-35°C) was measured at saturated substrate concentration, in 25 mM Pipes buffer, pH 6.5. In the case of the small synthetic substrate, pNPC, the samples were incubated at the respective temperatures for 5 min and for 30 min with the large size substrate carboxymethylcellulose (CMC). The activation energy E_a was calculated from the slopes of the Arrhenius plots: $\ln k_{cat}$ vs $1/T$ according to the Arrhenius equation, $k_{cat}=A\, e^{-Ea/RT}$ in which k_{cat} is the reaction rate, A the pre-exponential factor, R the gas constant (8.314 J mol^{-1} K^{-1}) and T the absolute temperature in Kelvin. The other thermodynamic parameters were calculated as previously described [15, 29].

Fluorescence quenching

The relative flexibility of the proteins was evaluated by fluorescence quenching using acrylamide as a quencher. The fluorescence was recorded with an SML-Aminco Model 8100 spectrofluorimeter (Spectronic Instruments, Rochester, NY, USA) using an excitation wavelength of 280 nm and emission wavelengths of 340 nm and 334 nm for the psychrophilic and mesophilic cellulase respectively. The samples were prepared in 25 mM Pipes buffer, pH 6.5 using a protein concentration that gave an A_{280nm} close to 0.1. The relative fluorescence intensity was recorded after 30 s incubation in the presence of acrylamide concentrations ranging from 5 to 120 mM at 10°C and 25°C. The experiments were performed in triplicate. The raw data were corrected for the dilution effect and for the screening effect caused by acrylamide (ε_{280nm}= 4.3 M^{-1} cm^{-1}). The relative fluorescence F_0/F represents the ratio of the fluorescence intensity in the absence of quencher (F_0) on the measured intensity in the presence of quencher (F). It was plotted as a function of acrylamide concentration. The data

were fitted according to the Stern-Volmer equation: $F_0/F=1+K_{SV}[Q]$ in which, K_{SV} is the Stern-Volmer constant and [Q] the concentration of the quencher [30].

Proton induced X-ray emission - PIXE

This sensitive technique [31, 32] was used to determine the amount and nature of metal ions possibly bound to the cold-adapted cellulase, especially calcium as in the case of thrombospondins [33]. The element-specific X-ray emission is generated through proton beams directed on the sample. Element traces level can be detected but no C, O and N nor elements with atomic number lower than 11. A proton beam of 2.5 and 3.0 MeV was generated with a Van de Graff accelerator belonging to the "Institut de Physique Nucléaire Expérimentale (INPE) at the University of Liège". The emitted X-ray energies (1-30 keV) were detected with a low energy Germanium detector commonly used to allow the simultaneous collection of X-ray radiations generated by collision of the proton beam with the samples in an helium filled chamber at a standard pressure of 760 mm Hg. The acquisition of the spectra was carried out using a GUPIX spectral analysis programme. A non linear least-square analysis of peak intensities was performed using a library of measured X-ray intensities for each element with a standard deviation of 5 %. For acquisition of the spectra, the native cellulase samples (1.6 mg. ml^{-1}) in the presence or not of 5 mM EGTA was prepared in MilliQ water eluted on a Chelex 100 column in order to remove calcium as much as possible. The pH was adjusted to 6.5 with NaOH. Essentially plastic utensils, previously washed with concentrated nitric acid, and thoroughly rinsed with free-calcium water, were used throughout. The samples were all filtered, before use, through a thin nucleopore filter free of any metal contamination.

Small Angle X-ray Scattering experiments

SAXS experiments were performed at the European Synchrotron Radiation Facility (Grenoble, France) on beamline ID02 as described previously [17]. The wavelength λ was 1.0 Å. The sample-to-detector distances were set at 4.0 m and 1.0 m, resulting in scattering vectors, q ranging from 0.010 $Å^{-1}$ to 0.15 $Å^{-1}$ and 0.03 $Å^{-1}$ to 0.46 $Å^{-1}$, respectively. The scattering vector is defined as $q = 4\pi/\lambda \sin\theta$, where 2θ is the scattering angle. All experiments were performed at 20°C. The data acquired at both sample-to-detector distances of 4 m and 1 m were merged for the calculations using the entire scattering spectrum.

The protein concentration of Cel5GΔ2PmutCC (in buffer 10 mM Hepes-NaOH pH 7.5, 0.04% NaN_3, 20 μM 4-(2-amino-ethyl)-benzene-sulphonyl-fluoride hydrochloride with 5% glycerol as radiation scavenger) was varied from 1.2 to 7.5 mg/ml in order to check for interparticle interactions. The protein concentration of Cel5G full length in the presence of 10mM $CaCl_2$ (same buffer) was varied from 1.5 to 7.3 mg/ml. The radius of gyration R_G was derived by the Guinier approximation $I(q) = I(0) \exp(-q^2R_G^2/3)$ for $qR_G < 1.0$. The distance distribution function *P(r)* was calculated on the merged curve by the Fourier inversion of the scattering intensity *I(q)* using GNOM [34]. This approach also features the maximum dimension of the macromolecule, D_{max} [35].

The overall shape of the variant cellulase was restored from the experimental data using the program GASBOR [36]. Several independent fits were run with no symmetry restriction and the stability of the solution was checked. The atomic structures of the catalytic module and of the cellulose binding module of Cel5G [17] were then fitted in the most typical calculated shape, using SUPCOMB [37] and TURBO-FRODO [38].

Results and Discussion

Cloning and purification of the full-length psychrophilic enzyme, its mutants and the mesophilic cellulase

The molecular architectures of the full-length psychrophilic enzyme Cel5G, the three mutants Cel5GΔ1P, Cel5GΔ2P, Cel5GΔ2PmutCC, the catalytic module $Cel5G_{CM,}$ and the mesophilic cellulase Cel5A are shown in Figure 1. The three mutants correspond to a sequential shortening of the linker region; they were generated by inverse PCR using the primers indicated in Table 1. Cel5GΔ1P is the product of the in-frame deletion of 31 residues from amino acids 312 to 342. This deletion includes the first loop, the closest to the catalytic module, containing the first disulfide bridge formed between Cys 314 and Cys 328. Cel5GΔ2P corresponds to a further shortening of 37 residues from amino acid 343 to 379; it includes the second loop and the disulfide bridge formed between Cys345 and Cys359. Cel5GΔ2PmutCC is similar in length to Cel5GΔ2P but the two cysteine residues: Cys382 and Cys396 have been replaced by two alanine residues in order to evaluate the importance of the last disulfide bridge in the adaptation to cold. The fourth construct, $Cel5G_{CM}$ extending from residue 1 to residue 292, is a truncated form of Cel5G consisting of the sole catalytic module devoid of the linker and of the carbohydrate-binding modules. The mutated fragments were inserted into pET expression vectors to transform *E. coli* BL21(DE3) competent cells.

The mutants Cel5GΔ1P, Cel5GΔ2P, and Cel5GΔ2PmutCC were purified to homogeneity from the periplasmic fractions using the procedure already described for the native Cel5G [15]. About 20 mg of each protein were isolated per culture liter. The catalytic module, $Cel5G_{CM,}$ was purified following a new protocol described in Materials and Methods. This protocol allowed the recovery of 25 mg of purified protein per liter. The purification of the mesophilic cellulase Cel5A, encoded by pSNAB1 [19, 20] was carried out in one single step by affinity chromatography on cotton-wool. A yield of 14 mg per liter was obtained. The purity of the different proteins was checked by SDS-PAGE as illustrated in Figure 2. A single band was observed in each case, showing apparent molecular masses slightly different from the theoretical ones of 49.4, 46.4, 42.6, 42.5, 32.0, and 42kDa calculated with the Protparam software for Cel5G, Cel5G1Δ1P, Cel5GΔ2P, Cel5GΔ2PmutCC, $Cel5G_{CM}$ and Cel5A, respectively. The higher apparent molecular mass of the psychrophilic enzymes is typical of proteins containing intrinsically disordered regions [39], as their peculiar amino acid composition prevents the complete binding of the anionic form of SDS to the proteins, then reducing the negative charge of the complexes.

In a previous study [17], the structural determinants at atomic level of the adaptation to cold of the catalytic module of Cel5G were determined. Despite the very high sequence similarity with its mesophilic counterpart, we showed that the psychrophilic catalytic module displays the typical adaptation to cold, which consists of a weakening of intra-molecular interactions contributing to increase the flexibility of the molecular edifice. Furthermore about 50% of the amino acids of the cold-adapted catalytic module display temperature or B factors higher than those of their mesophilic counterparts from *Erwinia chrysanthemi*. A peculiar feature is also that the catalytic module of the cold-adapted enzyme shows a large excess of negative charges: $34^{-}/18^{+}$, contrarily to the mesophilic homologue, $31^{-}/29^{+}$ [16, 40-42]. While the carbohydrate-binding modules display very high sequence similarity (57% of complete identity), the linkers are strikingly different, both in length (109 residues in the cold-adapted enzyme and only 34 in the mesophilic counterpart) and in sequence. A distinctive feature of the linker of the psychrophilic enzyme is that it contains three loops of 13 residues closed by disulfide bridges constituting three non collapsible structures that are flanked by three repeated motifs DXDXDXXXDXXD, called TSP3, also found in thrombospondins [33], in which they bind Ca^{2+}.

Kinetic and thermodynamic activation parameters

The thermodependence of the activity of Cel5G, Cel5G$_{CM}$ and Cel5A towards the synthetic and small size substrate pNPC is shown in Figure 3a. One can see that, in a temperature range of 0°C to 50°C, the activity of the catalytic module is close to that of the full-length psychrophilic enzyme, both largely outperforming the activity of the mesophilic enzyme Cel5A. At 25°C, the activity of the catalytic module alone is only about 20% lower than that of the full-length psychrophilic enzyme and an order of magnitude larger than that of the mesophilic enzyme (Table 2). The results are however quite different when large size substrates such as CMC and C are used. Indeed, unlike with pNPC, the specific activity at 25°C of the cold-adapted enzyme using CMC and C as substrates is more than twice as high as that of the catalytic module. The activity of the catalytic module is, by contrast, now approaching that of the mesophilic enzyme Cel5A. The K_m values are not very different between the various enzymes. Table 2 compares the kinetic parameters of Cel5G, Cel5G$_{CM,}$ and Cel5A, determined at 25°C, using these different substrates.

To mimic as much as possible what could occur in the real environment of the bacterium, we used CMC to measure the thermodependence of the activity of Cel5G, of its mutants and of the catalytic module Cel5G$_{CM}$. The data are shown in Figure 3b. One can see that there is a regular decrease of the activity of the various forms of the cold-adapted cellulase as a function of the linker shortening. The shortest mutant, Cel5GΔ2PmutCC, devoid of any disulfide bridge, displaying catalytic performances approaching those of the catalytic domain alone.

From these curves, the Arrhenius plots were traced and the thermodynamic activation parameters were calculated [29]. The data are shown in Table 2 along with the specific activities, at 10°C, of the various forms. As expected, the values of the activation energy E_a show a regular increase as a function of the linker shortening and culminate for the shortest form, Cel5G$_{CM}$. The values of the free energy of activation ΔG^* also increase concomitantly. This is due to an increase of the activation enthalpy ΔH^* partially reinforced by a concomitant decrease of the activation entropy ΔS^* suggesting that the linker shortening also induces a higher degree of order of the truncated forms when compared to the full-length enzyme Cel5G. One can also notice that the removal of the last disulfide bridge leads to a decrease of the activation entropy ($T\Delta S^*$ of -51.4 kJ.mol-1 for Cel5GΔ2P, and of -49.6 kJ.mol-1 for Cel5GΔ2PmutCC). This decrease in the activation entropy observed upon removal of the disulfide bridge could be due to an increase of the compactness of the disulfide-free form of the truncated mutants reducing the negative entropy change leading to the activation. This

increase in compactness is presumably induced by an increase in strength or in the number of some intra-molecular interactions prevented by the presence of the disulfide bridge. This hypothesis is supported by the increase of the activation enthalpy (more bonds have to be broken to reach the activated state) and also by the increase of the calorimetric enthalpy of the disulfide-free form when compared to that of the disulfide-containing form of the truncated enzyme (see Table 3). In the same context, the origin of the increase of the heat absorbed during the activation process of the truncated forms could be ascribed to a progressive increase of the stability and rigidity of the three dimensional structures of the cold-adapted cellulases upon shortening of the linker.

Thermal stability

To check this hypothesis, the thermal stability of the various forms of the cold-adapted enzyme was evaluated by differential scanning calorimetry. The profiles are shown in Figure 4. The values of the melting temperatures, T_{m1} and T_{m2}, corresponding to the melting of the two detectable thermodynamic domains of Cel5G and of its variants, and the associated calorimetric enthalpies ΔH_{cal} of these domains, calculated from the areas under the absorption heat peaks, are displayed in Table 3. One can see that there is a limited but significant increase of the melting temperatures of the two thermodynamic domains as a function of the linker shortening of the full-length enzyme. Indeed, $Cel5G_{CM}$ shows a melting temperature, $T_{m1,}$ 6.9°C higher than that of the first transition of the full-length enzyme. This is accompanied by a regular increase of the calorimetric enthalpies suggesting, as already stated above, that the linker shortening induces a progressive increase of the weak intra-molecular interactions. As already mentioned, even the deletion of the last disulfide bridge producing Cel5GΔ2PmutCC, gives rise to an increase of the calorimetric enthalpies suggesting that the disulfide bridges are also necessary to maintain the linker in a non collapsible and extended form, preventing interactions within, and between modules as already suggested by small angle x-ray scattering experiments carried out on the full-length enzyme.

Flexibility of Cel5G mutants

Fluorescence quenching

The specific activities reported in Table 2 show that the full-length cold-adapted enzyme outperforms the efficiency of the catalytic module alone by a factor of two whereas the activity of the catalytic module is close to that of the complete mesophilic enzyme, Cel5A

when large size substrates such as carboxymethyl cellulose and amorphous cellulose were used. It is generally considered that a high flexibility of proteins is tightly correlated to an increase in thermolability [43], and our results obtained by differential scanning calorimetry show that the thermostability of the truncated forms increases with the shortening of the linker. These datas suggest that the increase in thermal stability of the truncated cellulases should be accompanied by a decrease of the flexibility of the molecular structures. In this context the technique of fluorescence quenching using acrylamide as a quencher of the tryptophan fluorescence is very useful in evaluating the relative flexibility of psychrophilic enzymes compared to their mesophilic homologues. Indeed the quenching effect reflects the relative ability of the quencher to penetrate the structure and is therefore a measure of its permeability [44, 45]. The quenching effect of acrylamide was evaluated on the different variants of the psychrophilic cellulase, at 10°C, using increasing concentrations of acrylamide. The data were expressed in terms of relative fluorescence intensity (F0/F) and Stern-Volmer constants K_{sv}, corresponding to the slopes of the obtained lines (Figure 5). One can see that the sequential shortening of the linker also induces a dramatic and progressive decrease of the flexibility of the molecular structures. This clearly indicates that the composition and structure of the linker modulate the rigidity of the catalytic module, which contains 10 of the 13 Trp residues found in the protein. It is worth noting that the linker itself does not contain any Trp residues. The implication of the linker in the flexibility of the catalytic module is very important for catalysis but rather hard to explain. A hypothesis is that the numerous negative charges in the linker could be responsible for this effect through electrostatic repulsion since one should remember that the catalytic module also contains a large excess of negative charges ($34^{-}/18^{+}$). These electrostatic repulsions probably prevent the formation of a more cohesive structure of the catalytic module. The reduction of the number of negative charges in the linker, through its shortening, probably reduces this effect, increases the cohesion and stability of the catalytic module and progressively increases the activation enthalpy (ΔH^{*}) of the truncated forms (Table 2). Furthermore the linker shortening may render possible some interactions between the CM and CBM regions. All these effects would prevent the flexibility of the linker to be transmitted to the catalytic module when the linker is shortened.

Small angle X-ray scattering

SAXS experiments have been performed on the Cel5G shortest variant, Cel5GΔ2PmutCC, devoid of any disulfide bridge. SAXS is indeed a very appropriate tool to assess the structural

properties (dimensions, overall shape) and the flexibility in solution of multi-modular proteins containing disordered regions such as linkers.

At low angles, the scattered intensities are very well approximated by the Guinier law, and the radii of gyration R_G, calculated for different protein concentrations, revealed slight attractive interactions in solution. The radius of gyration extrapolated to zero concentration is 36.1 ± 1.5 Å at 20°C. The distance distribution function is typical of multi-modular cellulases [17, 46], and the inferred maximum diameter D_{max} is 136 ± 5Å. The overall shape of Cel5GΔ2PmutCC was calculated *ab initio* by the program GASBOR. Different runs gave similar shapes, fitting the data with a very good χ^2 of ~1.2 (Figure 6a). A typical calculated shape superimposed with the structure of the catalytic module and of the cellulose-binding module of Cel5G (Figure 6b) shows that the linker occupies a rather extended volume and that it can adopt elongated conformations between the two globular domains. As in [17], we can infer that the maximum dimension attained by the linker is 66 Å, while the linker is composed of 39 residues. This corresponds to a global compactness of 0.61 residues/Å, to be compared to 0.48 residues/Å for wild type Cel5G at 20°C when one does not take into account the residues involved in the transversal disulfide-bonded loops [17]. This result indicates that the linker preferably adopts more collapsed conformations when there is no transversal loop. The disulfide loops probably induce steric constraints that prevent the linker to collapse and, the formation of intra-molecular interactions that would stabilize and rigidify the entire protein. The shortest mutant, devoid of any disulfide-bonded loop can still adopt, as shown by the data presented here, an extended conformation but it is much more compact than the wild type linker.

Role of the TSP3 Motifs

The alignment of the amino acid sequences of the linker of the cold-adapted cellulase Cel5G and of its mesophilic counterpart from *Erwinia chrysanthemi* Cel5A is shown in Figure 7. The disulfide loops are separated by three amino acid sequences, numbered in Figure 7 as follows: 293-313, 329-344, 360-381which contain a repeated motif DXDXDGXXDXXD, at positions301-312, 332-343 and 369-380. They are identical to the TSP3 motif found in thrombospondins (TSPs), a family of extracellular glycoproteins, secreted by platelets, and involved in cellular communication and extracellular matrix interactions [33]. Thrombospondins contain several motifs of this type, which each encapsulates a Ca^{2+} with an affinity not exceeding K_D=0.1 mM. The bound cations are however very important for the

conformation and function of the protein. Therefore we postulated that the linker of the cold-adapted cellulase, containing three of these motifs, could also bind calcium ions. By contrast with thrombospondins, the chelation of calcium using EGTA does not affect the catalytic performance of the cold-adapted cellulase (data not shown). We nevertheless decided to evaluate possible metal binding using the technique of Proton Induced X-Ray Emission (PIXE), a form of elemental analysis based on the nature of the X-ray emission that follows the collision of the emitted protons with the metal ions presumably associated with the protein or present in solution. The spectra were recorded on the full-length enzyme, Cel5G, in the absence and in the presence of 5 mM EGTA added in order to concentrate the Ca^{2+}, potentially present, on a single component so as to facilitate detection. No peak, typical of Ca^{2+} or of any other metal, exception being from Na^{+} originating from the added EGTA, was detected (not illustrated).

SAXS experiments were also performed on wild type Cel5G in the presence of calcium (sea water concentration: 10 mM Ca^{2+}). The radius of gyration and the maximum diameter of Cel5G obtained in the presence of calcium are 53.8 ± 2.1 Å and 210 ± 10 Å respectively. These values are strictly identical to those obtained for Cel5G with no calcium added in the buffer solution [17]. The scattering spectra were also perfectly superimposable, indicating that Cel5G does not undergo any conformational change in the presence of calcium (data not shown).

These results show that unlike in thrombospondins, the TSP3 motifs of Cel5G do not bind calcium ions; the three loops closed by the disulfide bridges might be responsible for the fact that the three repeated motifs DXDXDXXXDXXD do not bind Ca^{2+}: indeed, these loops, through steric hindrance, might prevent the motifs from adopting the appropriate conformation required for Ca^{2+} sequestration. The negative charges of these motifs are thus not neutralized by Ca^{2+}. As a result, the linker from the psychrophilic cellulase contains altogether 23 negative charges, with no positively charged amino acids. In consequence, the negative charges of the linker also contribute, through electrostatic repulsion, to prevent the linker from collapsing, and to keep the catalytic and thecellulose binding modules well apart.

Conclusion

All the results presented here indicate that the Cel5G linker displays unique structural properties that are responsible for a key role in the adaptation to cold of the psychrophilic enzyme. Most probably, while structural constraints limit the adaptation to cold of the

catalytic domain, the linker has the possibility to explore much more conformations during evolution in such a way to adopt original structural features that confer to the full-length enzyme a higher flexibility and an efficient adaptation to cold. As demonstrated by small angle X-ray scattering [17], the long linker is an elongated and highly flexible structure. Even though transient compact conformations can always be attained, a combination of numerous repulsive negative charges, and steric hindrances due to the transversal loops prevent the stabilization of intramolecular interactions, and thus confer to the linker a higher flexibility. The role of the transversal disulfide-bonded loops of Cel5G is comparable to the glycosylation of mesophilic linkers in bimodular glycoside hydrolases [46, 47]. This higher flexibility of the linker provides a larger degree of freedom to the catalytic module and the cellulose-binding module, favouring their correct positioning with regards to the cellulose fibres even at low temperatures. It also considerably increases the area covered by the catalytic module at the surface of the cellulose and thus also favours the catalytic efficiency as previously mentioned [17, 48, 49].

Acknowledgments

The authors like to thank N. Gerardin-Otthiers and I. Thamm for their skilful technical assistance. We are also very grateful to Dr F. Barras for the generous gift of the pSNAB1 plasmid and to Dr G. Weber for his help in the PIXE experiments. We acknowledge Stéphanie Finet and the staff for the use of beamline ID02, at the ESRF synchrotron in Grenoble. We also thank Michel Desmadril for useful discussion. C. Duez is *Chercheur Qualifié* of the *Fonds National de la Recherche Scientifique* (FNRS-Belgium). This work was partly supported by the FRFC-Belgium (contract 2.4515.00) and by the EU (contract ERB-FMRX_CT 97-0131). G. Sonan is a recipient of the “Scholarship of Foreign Study from the Ministry of Education and Scientific Research of Ivory Coast.

References

1 Beguin, P. and Aubert, J. P. (1994) The biological degradation of cellulose. FEMS Microbiol Rev. **13**, 25-58

2 Wood, T. M. (1985) Properties of cellulolytic enzyme systems. Biochem. Soc. Trans. **13**, 407-410

3 Gilkes, N. R., Henrissat, B., Kilburn, D. G., Miller, R. C., Jr. and Warren, R. A. (1991) Domains in microbial beta-1, 4-glycanases: sequence conservation, function, and enzyme families. Microbiol Rev. **55**, 303-315

4 Shen, H., Schmuck, M., Pilz, I., Gilkes, N. R., Kilburn, D. G., Miller, R. C., Jr. and Warren, R. A. (1991) Deletion of the linker connecting the catalytic and cellulose-binding domains of endoglucanase A (CenA) of *Cellulomonas fimi* alters its conformation and catalytic activity. J Biol Chem. **266**, 11335-11340

5 Teeri, M. L. (1997) The roles and function of cellulose-binding domains. J. Biotechnol. **57**, 15-28

6 Tomme, P., Warren, R. A. and Gilkes, N. R. (1995) Cellulose hydrolysis by bacteria and fungi. Adv Microb Physiol. **37**, 1-81

7 Langsford, M. L., Gilkes, N. R., Singh, B., Moser, B., Miller, R. C., Jr., Warren, R. A. and Kilburn, D. G. (1987) Glycosylation of bacterial cellulases prevents proteolytic cleavage between functional domains. FEBS Lett. **225**, 163-167

8 Quentin, M., Ebbelaar, M., Derksen, J., Mariani, C. and van Der Valk, H. (2002) Description of a cellulose-binding domain and a linker sequence from Aspergillus fungi. Appl Microbiol Biotechnol. **58**, 658-662

9 Srisodsuk, M., Reinikainen, T., Penttila, M. and Teeri, T. T. (1993) Role of the interdomain linker peptide of *Trichoderma reesei* cellobiohydrolase I in its interaction with crystalline cellulose. J Biol Chem. **268**, 20756-20761

10 Argos, P. (1990) An investigation of oligopeptides linking domains in protein tertiary structures and possible candidates for general gene fusion. J Mol Biol. **211**, 943-958

11 George, R. A. and Heringa, J. (2003) An analysis of protein domain linkers: their classification and role in protein folding. Protein Eng. **15**, 871-879

12 Ferreira, L. M., Durrant, A. J., Hall, J., Hazlewood, G. P. and Gilbert, H. J. (1990) Spatial separation of protein domains is not necessary for catalytic activity or substrate binding in a xylanase. Biochem J. **269**, 261-264

13 Gokhale, R. S. and Khosla, C. (2000) Role of linkers in communication between protein modules. Curr Opin Chem Biol. **4**, 22-27

14 Violot, S., Haser, R., Sonan, G., Georlette, D., Feller, G. and Aghajari, N. (2003) Expression, purification, x-ray crystallographic studies of a psychrophilic cellulase from *Pseudoalteromonas haloplanktis*. Acta Cryst. **D59**, 1256-1258

15 Garsoux, G., Lamotte, J., Gerday, C. and Feller, G. (2004). Kinetic and structural optimization to catalysis at low temperatures in a psychrophilic cellulase from the Antarctic bacterium *Pseudoalteromonas haloplanktis*. Biochem J. **384**, 247-253

16 Feller, G. and Gerday, C. (2003) Psychrophilic enzymes: Hot topics in cold adaptation. Nature Microbiol. **1**, 200-208

17 Violot, S., Aghajari, N., Czjzek, M., Feller, G., Sonan, G. K., Gouet, P., Gerday, C., Haser, R. and Receveur-Bréchot, V. (2005) Structure of a full-length psychrophilic cellulase from *Pseudoalteromonas haloplanktis* revealed by X-ray diffraction and small angle X-ray scattering. J Mol Biol. **348**, 1211-24

18 Brun, E., Moriaud, F., Gans, P., Blackledge, M. J., Barras, F. and Marion, D. (1997) Solution structure of the cellulose-binding domain of the endoglucanase Z secreted by *Erwinia chrysanthemi*. Biochemistry **36**, 16074-86

19 Py, B., Bortoli-German, I., Haiech, J., Chippaux, M. Barras, F. (1991) Cellulase EGZ of *Erwinia chrysanthemi*: structural organization and importance of His98 and Glu133 residues for catalysis. Protein Eng **4**, 325-333

20 Py, B., Chippaux, M. & Barras, F. (1993) Mutagenesis of cellulase EGZ for studying the general protein secretory pathway in *Erwinia chrysanthemi*. Mol. Microbiol. **7**, 785-793

21 Chapon, V., Czjzek, M., El Hassouni, M., Py, B., Juy, M. and Barras, F. (2001) Type II protein secretion in gram-negative pathogenic bacteria: the study of the structure/secretion relationships of the cellulase Cel5 (formerly EGZ) from *Erwinia chrysanthemi*. J. Mol. Biol. **310**, 1055-1066

22 Sambrook, J., Fritsch, E. F.and Maniatis, T. (1989) Molecular cloning: a laboratory manual 2nd edition. New york: Cold Spring Harbor Laboratory Press.

23 Neu, H. C. and Heppel, L. A. (1965) The release of enzymes from *Escherichia coli* by osmotic shock and during the formation of spheroplasts. J Biol Chem. **240**, 3685-3692

24 Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M. D., Fujimoto, E; K., Goeke, N. M., Olson, B. J. and Klenk, D. C. (1985) Measurement of protein using bicinchoninic acid. Anal. Biochem. **150**, 76-85

25 Bhat, T. M. and Wa, K. M. (1988) Methods for measuring cellulases activities. Methods Enzymol. **160**, 87-144

26 Miller, G. L. (1959) Use of dinitrosalicylic acid reagent for determination of reducing sugars. Anal. Chem. **31**, 426-428

27 Goldberg, M. E., Expert-Bezancon, N., Vuillard, L. and Rabilloud, T. (1995) Non-detergent sulphobetaines: a new class of molecules that facilitate *in vitro* protein renaturation. Fold Des. **1**, 21-27

28 Matouschek, A., Matthews, J. M., Johnson, C. M. and Fersht, A. R. (1994) Extrapolation to water of kinetic and equilibrium data for the unfolding of barnase in urea solutions. Protein Eng. **7**, 1089-1095

29 Lonhienne, T., Gerday, C. and Feller, G. (2000) Psychrophilic enzymes: revisiting the thermodynamic parameters of activation may explain local flexibility. Biochim Biophys Acta. **1543**, 1-10

30 Lakowicz, J. (1983) Quenching of fluorescence. In Principles of Fluorescence Spectroscopy. (Lakowicz J.R.), pp 257, Plenum Press, New York.

31 Lowe, T., Chen, Q., Fernando, Q., Keith, R. and Gandolfi, A. J. (1993) Elemental analysis of renal slices by Proton-Induced X-ray Emission. Environ. Health Perspect. **101**, 302-308

32 Govil, I. M. (2001) Proton Induced X-ray Emission-A tool for non-destructive trace element analysis. Curr Sci. **80**, 1542-1548

33 Kvansakul, M., Adams, J. C. and Hohenester, E. (2004) Structure of a thrombospondin C-terminal fragment reveals a novel calcium core in the type 3 repeats. Embo J. **23**, 1223-1233

34 Svergun, D. (1992) Determination of the regularization parameter in indirect transform methods using perceptual criteria. J. Appl. Cryst. **25**, 495-503

35 von Ossowski I., Eaton, J.T., Czjzek, M., Perkins, S.J., Frandsen T.P., Schülein, M., Panine, P., Henrissat, B.and Receveur-Bréchot, V. (2005) Protein disorder: Conformational distribution of a long flexible linker in a chimeric double cellulase. Biophys. J. **88**, 2823-2832

36 Svergun, D.I., Petoukhov, M.V. and Koch, M.H. (2001) Determination of domain structure of proteins from X-ray solution scattering. Biophys. J. **80**, 2946-2953

37 Kozin, M.B. and Svergun, D.I. (2001) Automated matching of high- and low-resolution structural methods. J. Appl. Crystallogr. **34**, 33-41

38 Roussel, A. and Cambillau, C. (1989) TURBO-FRODO. In Silicon Graphics Geometry Partners (Committee, S. G., ed.), pp. 77-78. Silicon Graphics, Mountain View, California.)

39 Receveur-Bréchot V**.**, Bourhis J.M., Uversky V., Canard B. and Longhi S. (2006) Assessing protein disorder and induced folding. Proteins, **62**: 24-45

40 Russell, N. J. (2000) Toward a molecular understanding of cold activity of enzymes. Extremophiles **4**, 83-90

41 Siddiqui, K. S. and Cavicchioli, R. (2006) Cold-adapted enzymes. Ann Rev Biochem.**75**, 403-433

42 D'Amico, S., Collins, T., Marx, J. C., Feller, G. and Gerday, C. (2006) Psychrophilic microorganisms: challenges for life. EMBO Rep. **7**, 385-389

43 Tehei, M; Franzetti, B., Madern, D., Ginzburg,M., Ginzburg, B.Z., Giudici-Orticoni, M.T., Bruschi, M., and Zaccai, G. (2004) Adaptation to extreme environments: macromolecular dynamics in bacteria compared *in vivo* by neutron scattering. EMBO Rep. **5**, 66-70

44 D'Amico, S., Gerday, C. and Feller, G. (2003) Activity-stability relationships in extremophilic enzymes. J. Biol. Chem. **278**, 7891-7896

45 Collins, T., Meuwis, M. A., Gerday, C. and Feller, G. (2003) Activity, stability and flexibility in glycosidases adapted to extreme thermal environments. J Mol Biol.**328**, 419-428

46 Receveur V**.**, Czjzek M., Schülein M., Panine P.and Henrissat B. (2002) Dimension, shape and conformational flexibility of a two-domain fungal cellulase in solution probed by small angle X-ray scattering J. Biol. Chem, **277**, 40887-40892

47 Poon, D.K, Withers, S.G and McIntosh L.P. (2007) Direct demonstration of the flexibility of the glycosylated proline-threonine linker in the *Cellulomonas fimi* Xylanase Cex through NMR spectroscopic analysis. J Biol Chem. **282**, 2091-2100

48 Black, G. W., Rixon, J. E., Clarke, J. H., Hazlewood, G. P., Theodorou, M. K., Morris, P. and Gilbert, H. J. (1996) Evidence that linker sequences and cellulose -binding domains enhance the activity of hemicellulase against complex substrates. Biochem. J. **319**, 515-520

49 Black, G. W., Rixon, J. E., Clarke, J. H., Hazlewood, G. P., Ferreira, L. M., Bolam, D. N. and Gilbert, H. J. (1997). Cellulose binding domains and linker sequences potentiate the activity of hemicellulases against complex substrates. J Biotechnol. **57**, 59-69

50 Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic Ac. Res. **22**, 4673-4680

Table 1. Oligonucleotides used as primers in inverse PCR or site-directed mutagenesis experiments.

Mutant	N°	Primer	Sequence	Location	Direction
Cel5G$_{CM}$	1	Cel5G$_{CM}$_F	5'-AAGCTTGCGGCCGCACTCGAGCAC-3'	pET-23a(+)	Forward
	2	Cel5G$_{CM}$_Rev	5'-GGTCAACCACCCCAACCTTGAATGAT-3'	954-975	Reverse
Cel5GΔ1P	3	Cel5GΔ1P_F1	5'-GATCAATGTCCTAATACACCAGTAGGTGAAACTG-3'	1123-1158	Forward
(ΔAsp312-Val342)	4	Cel5G_Rev	5'-AAGGCTGTCACTTACGCCATCCCCGTC-3'	1029-1003	Reverse
Cel5GΔ2P	5	Cel5GΔ2P_F2	5'-GATCAGTGCCCAGATACACCCGCTG-3'	1234-1258	Forward
(ΔAsp312-Ile379)	6	Cel5G_Rev	5'-AAGGCTGTCACTTACGCCATCCCCGTC-3'	1029-1003	Reverse
Cel5GΔ2PmutCC					
(Cys382Ala)	7	Δ2P_C382A_F1	5'-GTAAGTGACAGCCTTGATCAGGCCCCAGATACACC-3'	1038-1072	Forward
(Cys382Ala)	8	Δ2P_C382A_R1	5'-GGTGTATCTGGGGCCTGATCAAGGCTGTCACTTAC-3'	1072-1038	Reverse
(Cys396Ala)	9	Δ2P_C396A_F2	5'-GCCAGTGTTGTAAGCTCAACAGATTGTAACGG-3'	1078-1109	Forward
(Cys396Ala)	10	Δ2P_C396A_R2	5'-TCCGTTTGTATCAACACTTGTACCAG-3'	1077-1052	Reverse

The underlined TCA triplet is complementary to a stop codon introduced at the end of the CM domain in Cel5G.

The underlined GCC triplets code for an alanine and the underlined GGC sequence is complementary to an alanine codon

Table 2. Comparison of the kinetic parameters, determined at 25°C, of the full-length psychrophilic cellulase, Cel5G, its catalytic domain, $Cel5G_{CM}$ and the mesophilic cellulase, Cel5A from *E. chrysanthemi* and comparison at 10°C of the catalytic constants and of the thermodynamic parameters, using carboxymethyl cellulose (CMC) as substrate, of the full-length psychrophilic cellulase, and of its mutants Cel5GΔ1P, Cel5GΔ2P, Cel5GΔ2PmutCC and $Cel5G_{CM}$. (pNPC, *p*-nitrophenyl-β-D-cellobioside, CMC, carboxymethylcellulose, and C, amorphous cellulose).

25°C					
	pNPC		***CMC***		***C***
Cellulases	$k_{cat} \times 10^{-2}\ (s^{-1})$	$K_m\ (mM)$	$k_{cat}\ (s^{-1})$	$K_m\ (mM)$	$k_{cat}\ (s^{-1})$
Cel5G	30.0 ± 0.01	3.5 ± 0.4	16.0 ± 0.02	0.21 ±0.02	23.5 ± 0.02
$Cel5G_{CM}$	23.0 ± 0.01	3.5 ± 0.3	8.0 ± 0.01	0.27 ± 0.03	9.70 ± 0.01
Cel5A	3.0 ± 0.01	2.4 ± 0.3	5.0 ± 0.05	0.20 ± 0.02	9.50 ± 0.05
10°C (CMC)					
	$k_{cat}(s^{-1})$	E_a (kJ.mol^{-1})	*ΔG** (kJ.mol^{-1})	*ΔH** (kJ.mol^{-1})	*TΔS** (kJ.mol^{-1})
Cel5G	12 ± 0.01	13	63.3	10.6	-52.7
Cel5GΔ1P	10 ± 0.02	14	63.8	11.6	-52.2
Cel5GΔ2P	9 ± 0.01	15	64.0	12.6	-51.4
Cel5GΔ2PmutCC	8 ± 0.02	17	64.2	14.6	-49.6
$Cel5G_{CM}$	6 ± 0.02	20	65.0	17.6	-47.4

Biochemical Journal Immediate Publication. Published on 18 Jul 2007 as manuscript BJ20070640

Table 3. Microcalorimetric parameters of thermal unfolding of the native full-length psychrophilic, Cel5G, cellulase and its mutants, Cel5GΔ1P, Cel5GΔ2P, Cel5GΔ2PmutCC, and $Cel5G_{CM}$.

Cellulase	T_{m1} (°C)	T_{m2} (°C)	ΔH_{cal1} (kcal.mol^{-1})	ΔH_{cal2} (kcal.mol^{-1})
Cel5G	43.3 ± 0.3	47.9 ± 0.1	63.2 ± 7.8	52.5 ±7.7
Cel5GΔ1P	45.3 ± 0.3	48.3 ± 0.1	63.8 ± 7.7	53.9± 7.5
Cel5GΔ2P	45.6 ± 0.4	48.8 ± 0.1	64.2 ± 8.2	54.7 ± 8.0
Cel5GΔ2PmutCC	47.7 ± 0.3	50.3 ± 0.1	64.8 ± 10.6	57.2 ± 10.4
$Cel5G_{CM}$	50.2 ± 0.1	53.8 ± 0.1	66.5 ± 8.0	70.5 ± 7.1

Legends of figures

Figure1.

Schematic representation of the structure of the native full-length psychrophilic cellulase Cel5G. (the numbering corresponds to the amino acid sequence of the mature protein), its mutants Cel5GΔ1P, Cel5GΔ2P, Cel5GΔ2PmutCC, the catalytic module, $Cel5G_{CM}$ and the full-length native mesophilic cellulase, Cel5A. The signal peptides correspond to the dotted boxes, the catalytic modules to the black boxes, the linker regions to the solid lines and the carbohydrate-binding modules (CBM) to the boxes with diagonal lines. The rectangles on the linker correspond to the disulfide bridges and loops closed by them.

Figure 2.

(**a**) SDS-PAGE (12% acrylamide) of the full-length psychrophilic cellulase, Cel5G (lane 2), its catalytic module,$Cel5G_{CM}$ (lane 3), and the mesophilic cellulase, Cel5A (lane 4). Lane 1: standard proteins (phosphorylase b, 97 kDa; bovine serum albumin, 62.2 kDa; ovalbumin, 45kDa; carbonic anhydrase, 31kDa, soybean trypsin inhibitor, 21.5 kDa, and lysoyme, 14.4 kDa.

(**b**) SDS-PAGE (12% acrylamide) of the full-length psychrophilic cellulase, Cel5G (lane 2); the mutants of deletion, Cel5GΔ1P (lane 3), Cel5GΔ2P (lane 4) and the mutant Cel5GΔ2PmutCC (lane 5). Lane 1: standard proteins as above.

Figure 3

Thermodependence of the activity of the various cellulases: (**a**) As measured using pNPC as substrate: black circles, full-length cold-adapted cellulase, Cel5G; diamonds, catalytic module of the cold-adapted cellulase, $Cel5G_{CM}$; black triangles, mesophilic cellulase from *Erwinia chrysanthemi*, Cel5A.

(**b**) Thermodependence of the cold-adapted cellulases using carboxymethylcellulose as substrate: black circles, full-length cold-adapted cellulase, Cel5G; open circles, mutant Cel5GΔ1P; black squares, mutant Cel5GΔ2P; white squares, mutant Cel5GΔ2PmutCC; diamonds, catalytic module, $Cel5G_{CM}$.

Figure 4

Thermograms obtained by Differential Scanning Calorimetry (DSC) of the full-length cold-adapted enzyme, Cel5G and of its mutants of deletion in the linker. Each peak represents the

heat absorbed during the unfolding of the 3D structure. From left to right: Cel5G (solid line), Cel5GΔ1P (long dash line), Cel5GΔ2P (dash dot dot line), Cel5GΔ2PmutCC (dotted line), and catalytic module, $Cel5G_{CM}$ (dash dot line).

Figure 5

Stern-Volmer plots corresponding to the quenching of the Trp fluorescence, at 10°C of the various cold-adapted cellulases using acrylamide as a quencher. The relative fluorescence, F0/F, is plotted as a function of acrylamide concentration. F0 is the fluorescence obtained in the absence of quencher and F the fluorescence obtained in the presence of acrylamide. The Stern-Volmer constants, K_{SV}, represent the slopes of the different lines; they have been calculated from the Stern-Volmer equation: $F0/F=1+K_{SV}$ [Q]. Black circles, Cel5G-K_{SV}= 0.89; white circles, Cel5GΔ1P-K_{SV}= 0.59; black squares, Cel5GΔ2P-K_{SV}= 0.49; white squares, Cel5GΔ2PmutCC-K_{SV}= 0.34; diamonds, Cel5GCM-K_{SV}= 0.29, and black triangles, Cel5A-K_{SV}= 0.19

Figure 6

Ab initio shape restoration of mutant Cel5GΔ2PmutCC by small angle X-ray scattering (**a**) experimental spectrum (black circles) and fit to the data given by GASBOR (grey line). (**b**) Typical calculated shape obtained by GASBOR (grey envelope), superimposed with the atomic structures of the catalytic module (left) and of the CBM (right) of Cel5G. The atomic structures are in black ribbon representation.

Figure 7

Sequence alignment, using the programme ClustalW [50], of the linker regions (LR) of the cold-adapted cellulase, Cel5G (formerly named CelG, Genbank accession no. **Y17552**) and of the mesophilic cellulase, Cel5A (formerly named CelZ, Genbank accession no. **Y00540**). The linker of the psychrophilic enzyme is long of 109 residues and that of the mesophilic cellulase is long of 34 residues. The few dots between the alignment correspond to identical or similar amino acid residues and the underlined sequences to the three TSP 3 motifs.

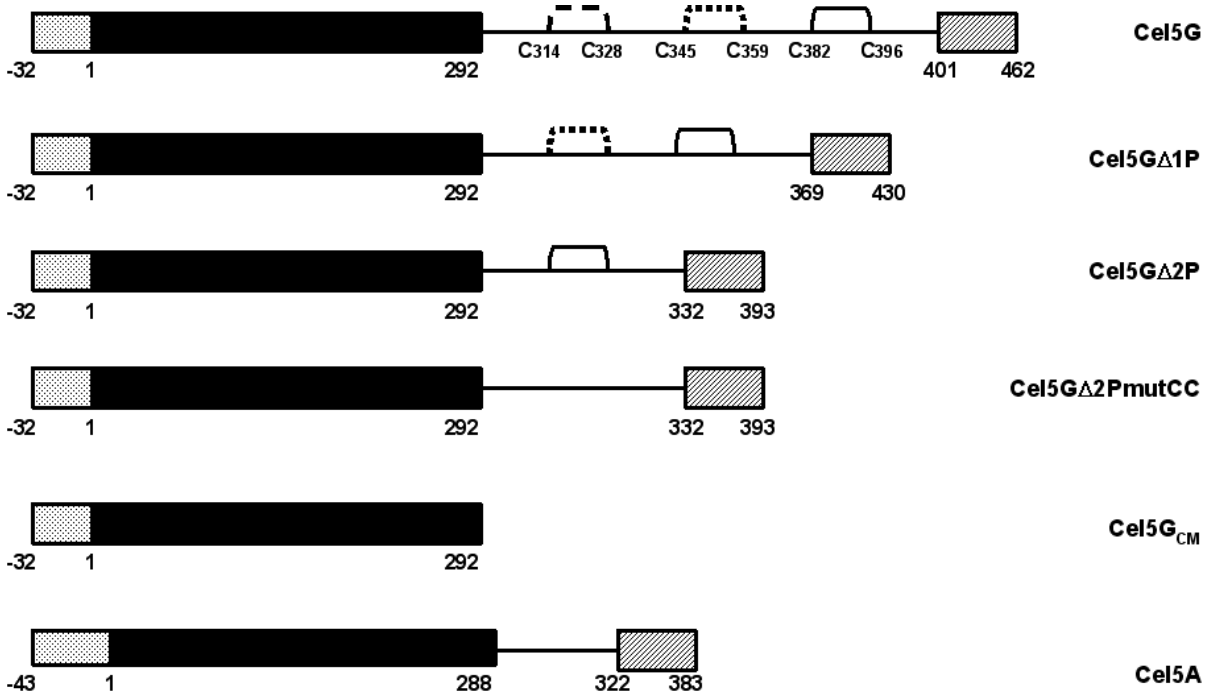

Fig.1

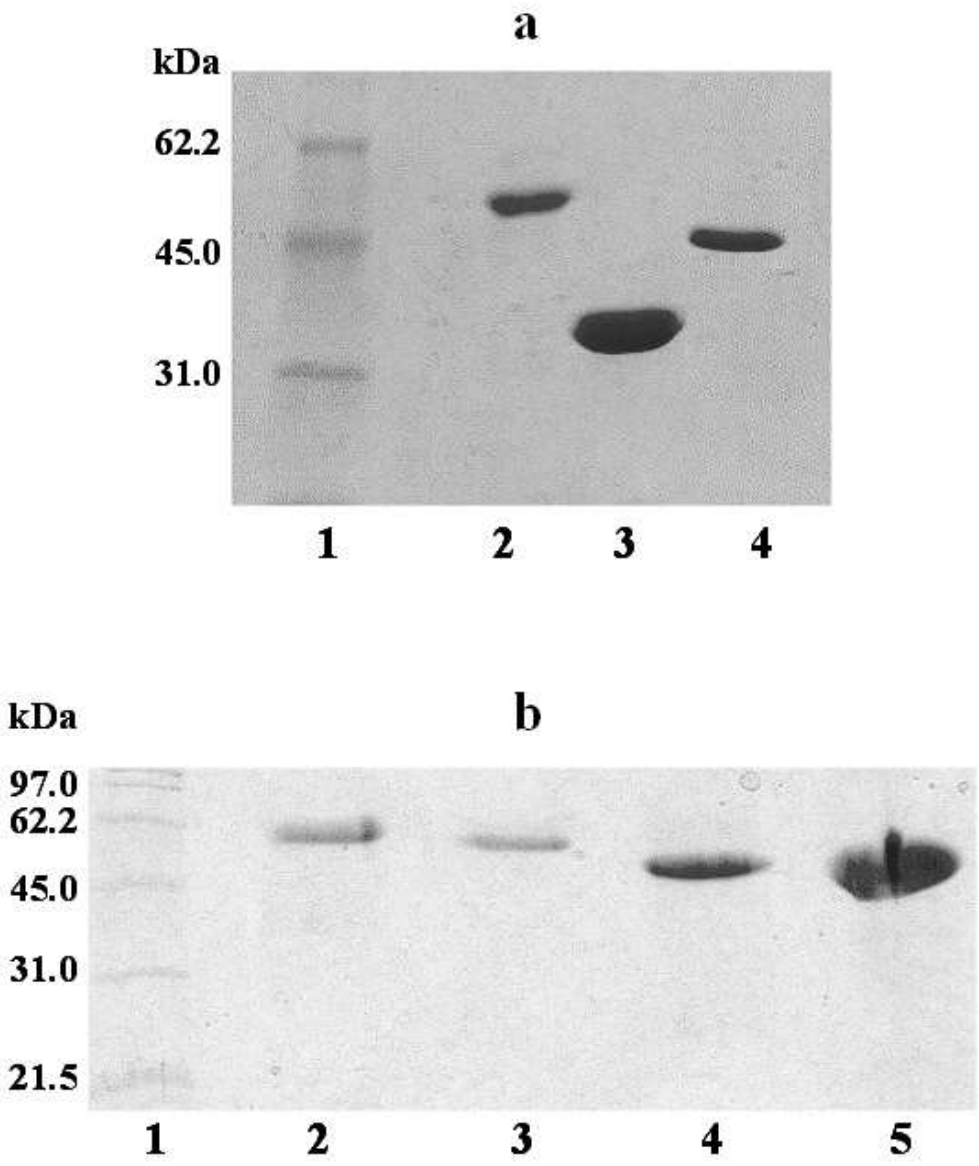

Fig. 2

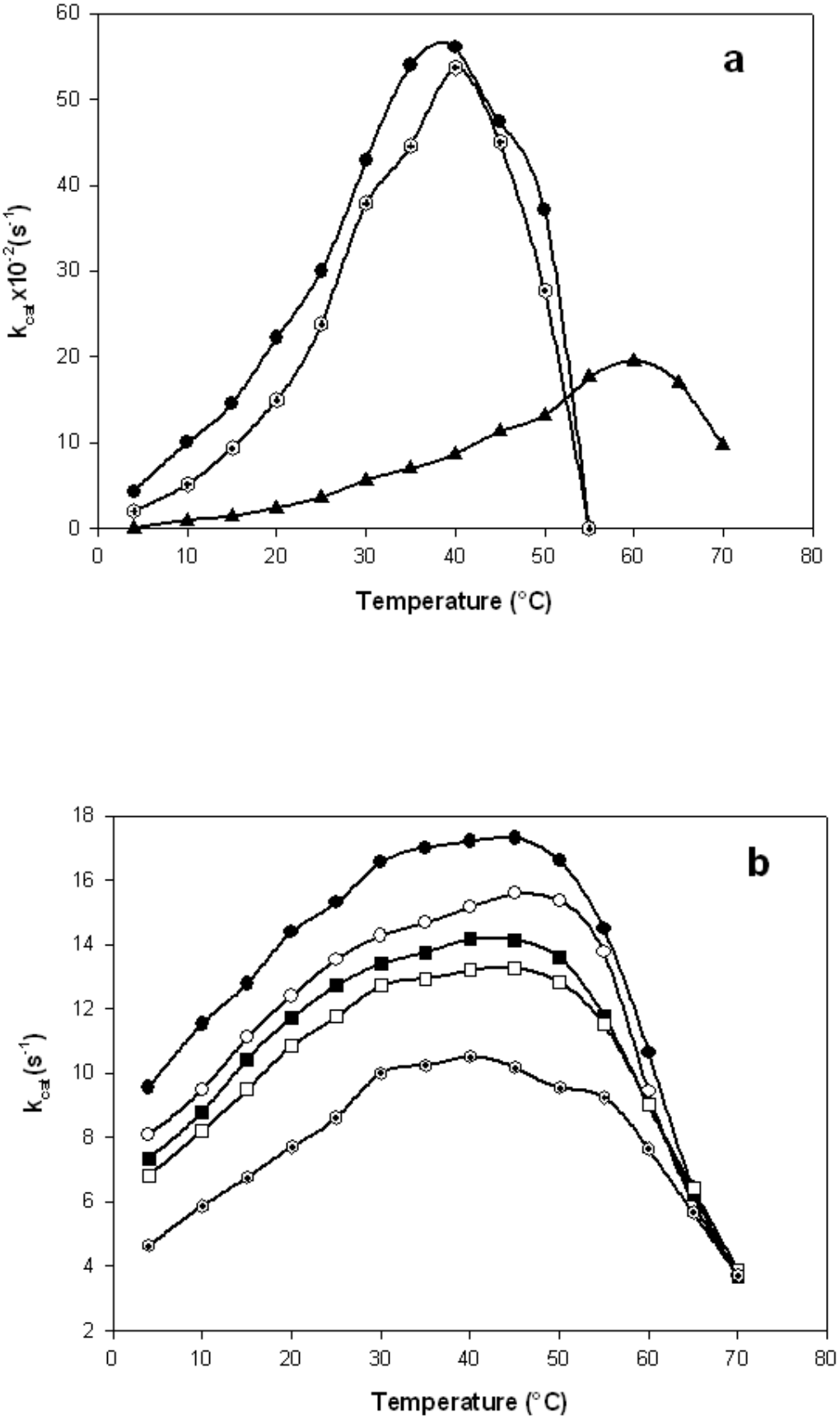

Fig. 3

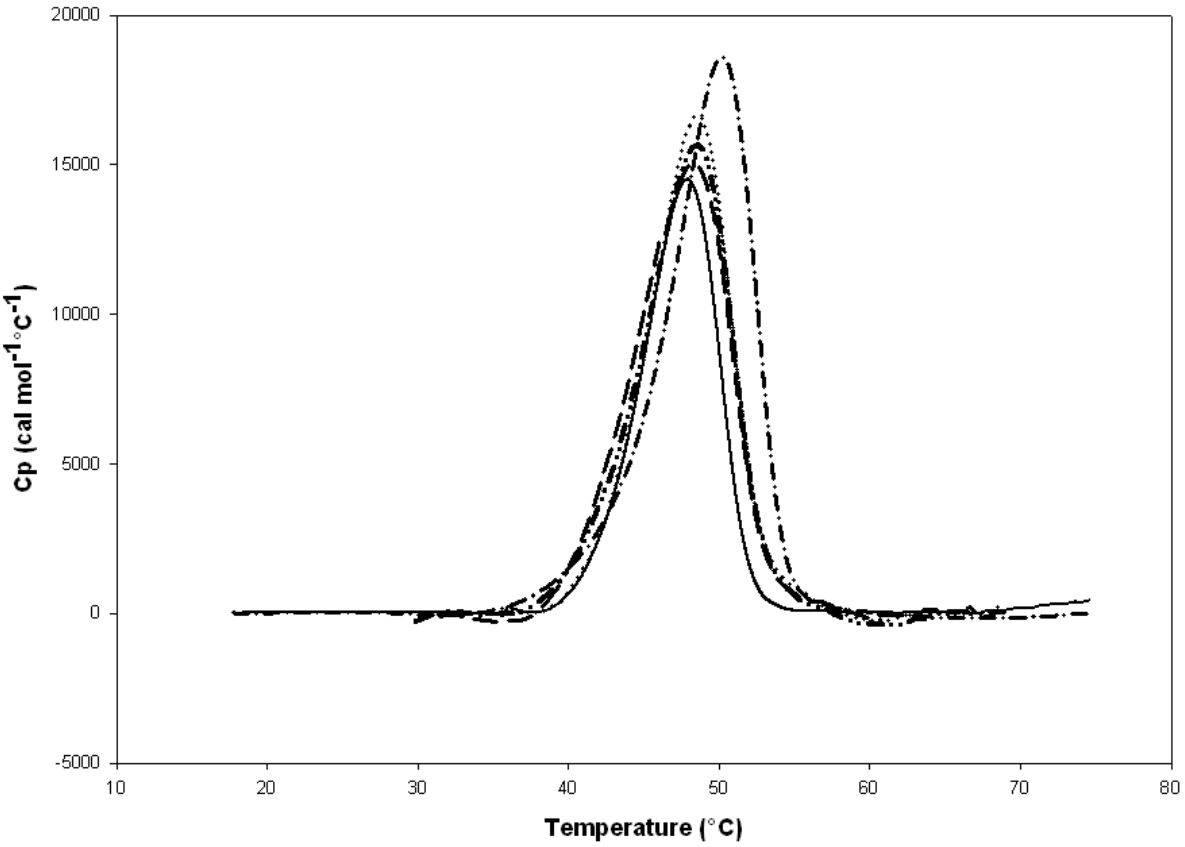

Fig. 4

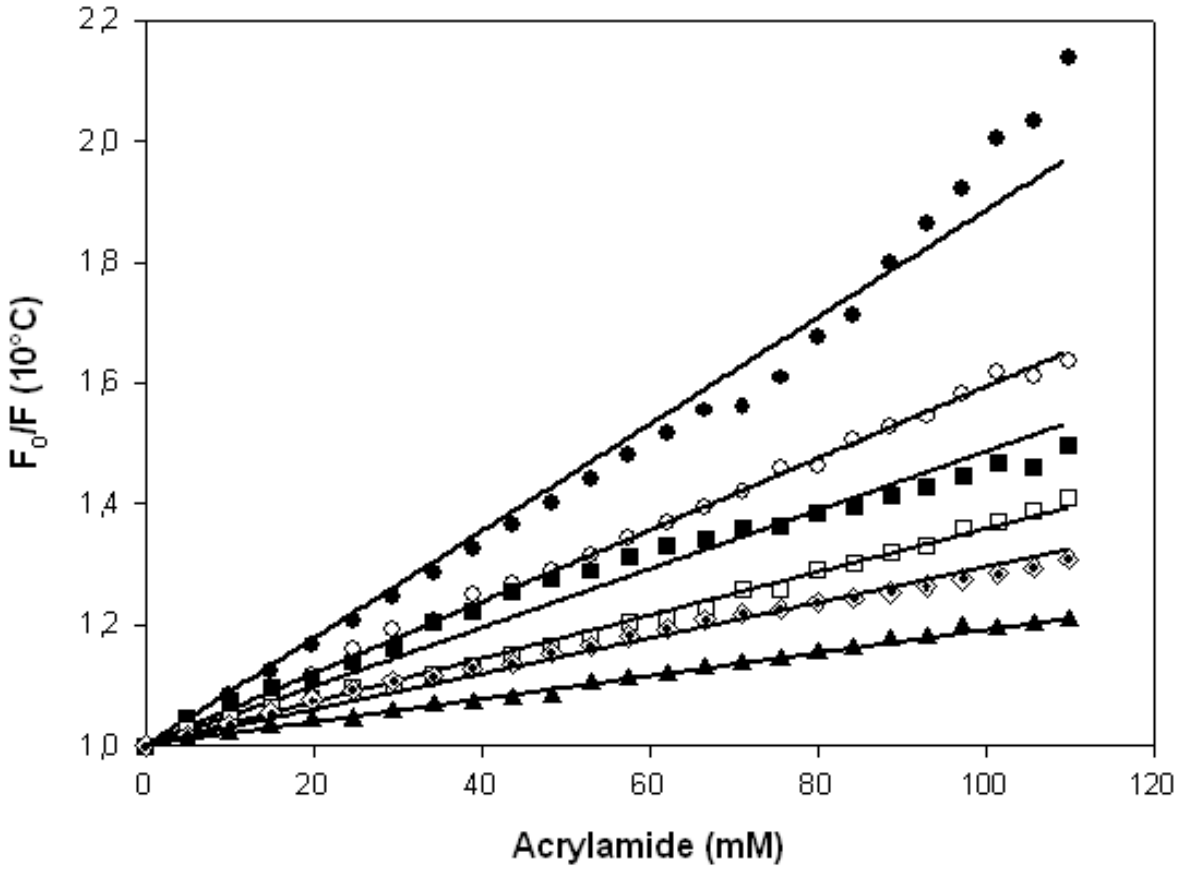

Fig. 5

Biochemical Journal Immediate Publication. Published on 18 Jul 2007 as manuscript BJ20070640

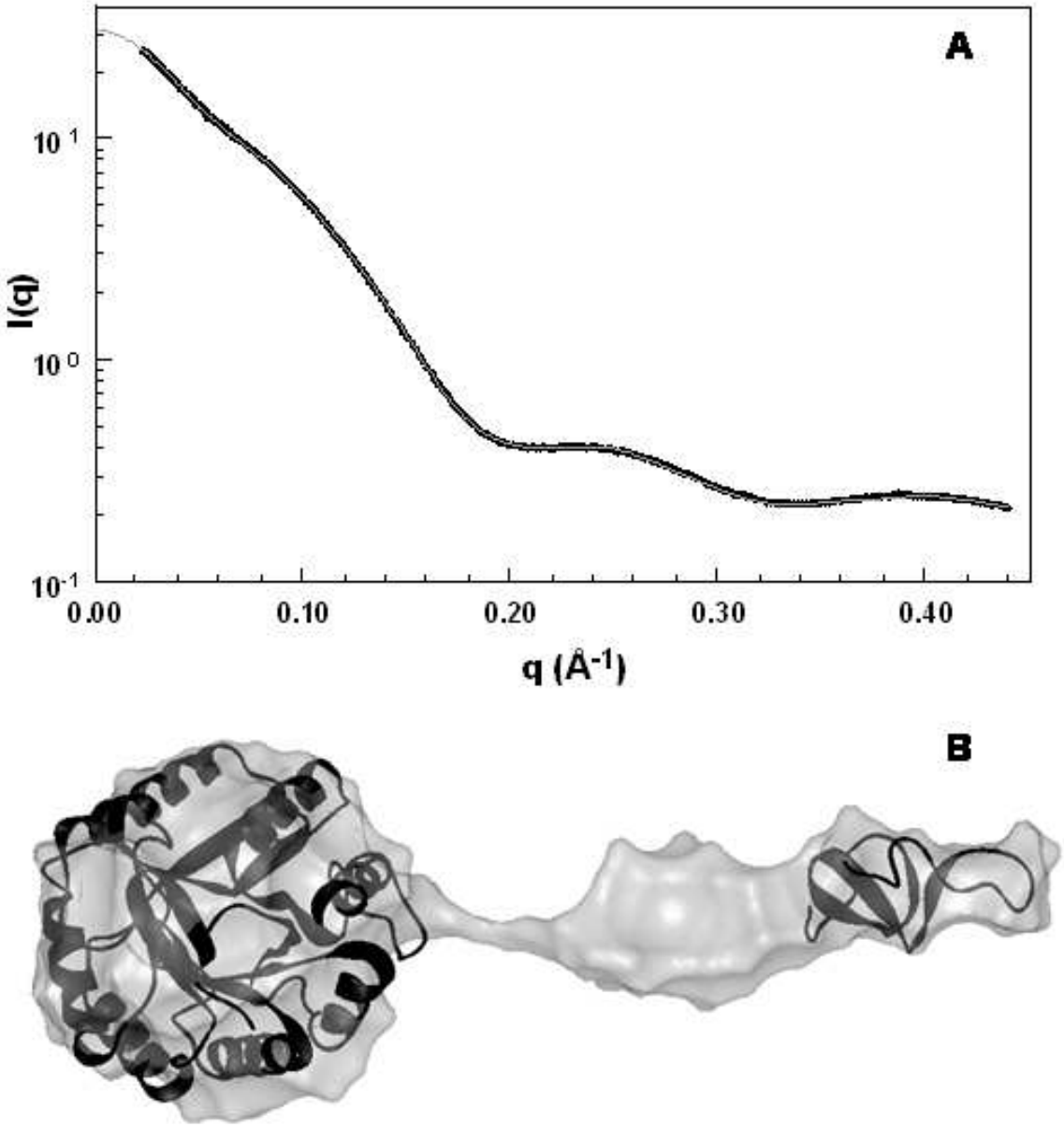

Fig. 6

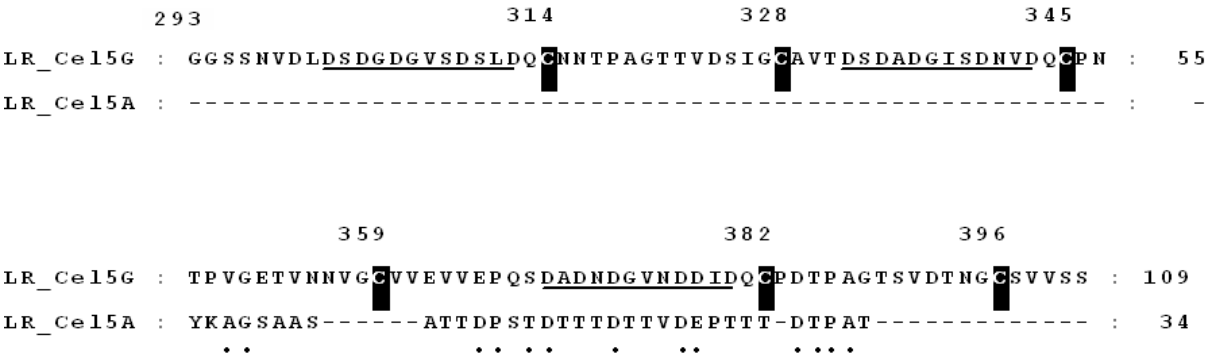

Fig. 7