BJ www.biochemj.org

BJ Structure

Biochem. J. (2007) **407**, 293–302 (Printed in Great Britain) doi:10.1042/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, and §Station Biologique de Roscoff, Végétaux Marins et Biomolécules, UMR 7139, Place George Teissier, BP 74, F-29682 Roscoff cedex, France

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 region (LR) containing three loops closed by three disulfide bridges. To evaluate the possible role of this region in cold adaptation, the LR 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 residue by two alanine residues. The kinetic and thermodynamic properties of the mutants were compared with those of the full-length enzyme, and also with those of the cold-adapted CM alone and with those of the homologous mesophilic 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 SAXS (small-angle X-ray scattering). The data indicate that the sequential shortening of the LR 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 LR of the full-length enzyme favours the catalytic efficiency at low and moderate temperatures by rendering the structure not only less compact, but also less stable, and plays a crucial role in the adaptation to cold of this cellulolytic enzyme.

Key words: cellulase, linker, *Pseudoalteromonas haloplanktis*, psychrophile, small-angle X-ray scattering (SAXS).

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 carbohydrate-binding module (CBM), separated by a distinct linker region (LR) comprising a simple or repetitive sequence often rich in proline, threonine, serine or glycine residues [1,3–8]. Previous studies have shown that LR domains can play an essential role in catalysis by favouring co-operative interdomain interactions [9–13].

The Gram-negative Antarctic bacterium *Pseudoalteromonas haloplanktis* A23 collected from seawater secretes a psychrophilic cellulase, Cel5G (EC 3.2.1.4), belonging to glycosyl hydrolase family 5, subfamily 2 (GH5-2). This cellulase is composed of a CM 292 residues in length, organized in a $(\beta/\alpha)_8$ barrel structure, and a CBM of 61 residues separated by a flexible and unusually extended 109-amino-acid-long LR [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 detail with similar parts of the mesophilic cellulase Cel5A from *Erwinia chrysanthemi* [17–19]. The CM and the CBM of Cel5G respectively share 64 and 57 % sequence identity with their mesophilic counterparts. However, the LRs are very different. Indeed, in Cel5G, the 109-residue LR is rich in glycine, valine, serine, threonine, cysteine, aspartate and asparagine residues, whereas the LR of the mesophilic counterpart, Cel5A, is a peptide of 34 residues rich in serine and threonine. In addition, the Cel5G LR contains six cysteine residues engaged in disulfide bridges, resulting in the formation of three transverse loops of 13 residues [17]. The structural comparison of the CMs 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 SAXS (small-angle X-ray scattering), suggests that the long LR is an extended and highly flexible structure, and that the cold adaptation of Cel5G could arise not only from the specific properties of the CM, but also from the unusual properties of the LR [17].

In the present study, this hypothesis was tested by studying the kinetic, thermodynamic and stability properties of three variants of Cel5G, resulting from the gradual shortening of the LR

Abbreviations used: BCA, bicinchoninic acid; CBM, carbohydrate-binding module; CM, catalytic module; CMC, carboxymethylcellulose; IPTG, isopropyl β-D-thiogalactoside; LB, Luria–Bertani; LR, linker region; PIXE, proton-induced X-ray emission; pNCP, *p*-nitrophenyl-β-D-cellobioside; SAXS, small-angle X-ray scattering; TSP3, thrombospondin 3.

[1] To whom correspondence should be addressed (email ch.gerday@ulg.ac.be).

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

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 residue, and the underlined GGC sequence is complementary to an alanine codon.

Mutant	Number	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
(ΔAsp^{312}–Val^{342})	4	Cel5G_Rev	5′-AAGGCTGTCACTTACGCCATCCCCGTC-3′	1029–1003	Reverse
Cel5GΔ2P	5	Cel5GΔ2P_F2	5′-GATCAGTGCCCAGATACACCCGCTG-3′	1234–1258	Forward
(ΔAsp^{312}–Ile^{379})	6	Cel5G_Rev	5′-AAGGCTGTCACTTACGCCATCCCCGTC-3′	1029–1003	Reverse
Cel5GΔ2PmutCC					
(C382A)	7	Δ2P_C382A_F1	5′-GTAAGTGACAGCCTTGATCAGGCCCCAGATACACC-3′	1038–1072	Forward
(C382A)	8	Δ2P_C382A_R1	5′-GGTGTATCTGGGGCCTGATCAAGGCTGTCACTTAC-3′	1072–1038	Reverse
(C396A)	9	Δ2P_C396A_F2	5′-GCCAGTGTTGTAAGCTCAACAGATTGTAACGG-3′	1078–1109	Forward
(C396A)	10	Δ2P_C396A_R2	5′-TCCGTTTGTATCAACACTTGTACCAG-3′	1077–1052	Reverse

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 determined 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 cellulases respectively.

The gene coding for the *P. haloplanktis* A23 Cel5G cellulase was cloned into 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 described previously [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). The deletions were introduced by inverse PCR using Vent DNA polymerase (New England Biolabs) that possesses a proofreading activity. Cys^{396} was changed to alanine by inverse PCR, whereas the C382A 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 into the corresponding sites of the 1.8 kb smaller pET-23a(+) vector. A termination TGA codon was introduced at the 3′ end of the 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 LR; construction of the Cel5GΔ1P (ΔAsp^{312}–Val^{342}) and Cel5GΔ2P (ΔAsp^{312}–Ile^{379}) truncated proteins

The 2.6 kb NdeI/XhoI fragment containing the *cel5G* gene was isolated from the pET22*cel5G* plasmid and subcloned into the small pSP73 cloning vector (Promega) to give pSP*cel5G*. The Cel5GΔ1P and Cel5GΔ2P mutants were obtained by inverse PCR using primers 3 and 4 or primers 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 a template for site-directed mutagenesis. The experiment was performed in two steps. Primers 7 and 8 (Table 1) were used to exchange alanine for Cys^{382}, and the resulting plasmid carrying this single mutated cysteine residue served as a template to change Cys^{396} to alanine using primers 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 (Luria–Bertani) agar medium containing 100 $\mu g \cdot ml^{-1}$ ampicillin, 100 $\mu g \cdot ml^{-1}$ Trypan Blue and 5 $mg \cdot ml^{-1}$ CMC (carboxymethylcellulose). Positive cellulase activity appeared as a clear halo 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 $\mu g \cdot ml^{-1}$ ampicillin. After 20 h of 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 a theoretical D_{550} of 5–6. A 1 ml aliquot of this second pre-culture was then added to 300 ml of LB broth in a 1 litre sterile flask, and cells were grown up to a theoretical D_{550} of 3–4. The enzyme expression was then induced with 0.1 mM IPTG (isopropyl β-D-thiogalactoside), and the cells were collected after 10 h 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. A 1 ml aliquot of this pre-culture was used to inoculate 100 ml of the same medium in a 500 ml sterile flask, and the culture was incubated at 30 °C until it reached a theoretical D_{550} of 4–5. An 8 ml aliquot of this pre-culture was used to inoculate 300 ml of LB broth in a 1 litre flask supplemented with 100 $\mu g \cdot ml^{-1}$ ampicillin and 50 $\mu g \cdot ml^{-1}$ IPTG. The culture was grown at 30 °C under agitation until reaching a theoretical D_{550} of 5–6.

Extraction of proteins

Cells were harvested by centrifugation at 9000 ***g*** for 15 min at 4 °C and were then subjected to osmotic shock [23]. The cell pellet was suspended in the osmotic shock solution (1/10 vol. of the culture broth) containing 20 % sucrose, 0.5 mM sodium EDTA and 100 mM Tris/HCl, pH 8. The suspension was supplemented with 0.1 mM PMSF and mixed in a rotary shaker at 180 rev./min. After 10 min, the mixture was centrifuged at 13 000 ***g*** for 15 min at 4 °C. The supernatant was eliminated, and the drained pellet was rapidly mixed with an equal volume (v/v) of cold water containing 1 mM $MgCl_2$. The suspension was mixed for 10 min, centrifuged at 13 000 ***g*** for 10 min at 4 °C, and the supernatant, corresponding to the periplasmic fraction, was recovered, supplemented with 0.04 % sodium azide 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 h, and the suspensions were then centrifuged at 14 500 ***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 CM of the psychrophilic cellulase, Cel5G$_{CM}$

The protein solution was diluted twice in buffer B [50 mM Pipes, pH 6.5, and 50 % saturated $(NH_4)_2SO_4$] and loaded on to a phenyl-Sepharose column (3 cm × 20 cm) (Bio-Rad) equilibrated in buffer C [25 mM Pipes, pH 6.5, and 25 % saturated $(NH_4)_2SO_4$]. The column was washed with buffer C, and the proteins were eluted at a flow rate of 2 $ml \cdot min^{-1}$ with a linear gradient of 25–0 % saturated $(NH_4)_2SO_4$ in buffer A. The fractions with cellulase activity were combined and dialysed against buffer 1 (10 mM Hepes, pH 7.5, 0.04 % sodium azide and 1 mM PMSF) using an Amicon ultrafiltration unit (Millipore) fitted with a PTGC membrane (10 kDa cut-off) (Millipore). The concentrated sample was then loaded on to a Macro-prep High Q column (3 cm × 20 cm) (Bio-Rad) equilibrated in buffer 1, and the proteins were eluted with a linear gradient of 0–300 mM KCl in buffer 1. The active fractions were collected, dialysed against buffer 1, filtered using a 0.45 μm-pore-size membrane (Millipore) and loaded on to a Mono Q HR 5/5 column (0.5 cm × 5 cm) (GE Healthcare) equilibrated in buffer 1. The column was then washed with buffer 1, and the proteins were eluted at a flow rate of 1 $ml \cdot min^{-1}$ with a linear gradient of 0–300 mM KCl in buffer 1. The active fractions were pooled, dialysed against buffer 1, concentrated by ultrafiltration, supplemented with 20 μM Pefabloc [4-(2-aminoethyl)-benzenesulfonylfluoride 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 $(NH_4)_2SO_4$ fractionation of the periplasmic samples were diluted twice in buffer B and loaded on to a cottonwool column equilibrated in buffer C (the support was sterilized in water and packed into a 5 cm × 12 cm 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. The cellulase preparations, except for the one corresponding to the mesophilic cellulase that was already fully purified from the cottonwool column, were then loaded on to a Macro-prep High Q column and was eluted as described for the psychrophilic cellulase Cel5G$_{CM}$.

Determination of protein concentrations

Protein concentrations were determined for some samples with the BCA (bicinchoninic acid) method [24] using the BCA protein assay reagent kit (Pierce) and BSA as a standard and by measuring the absorbance at 280 nm for the purified proteins, using the molar absorption coefficients of 95 650 $M^{-1} \cdot cm^{-1}$ for Cel5G, Cel5GΔ1P, Cel5GΔ2P, Cel5GΔ2PmutCC, 69 900 $M^{-1} \cdot cm^{-1}$ for Cel5G$_{CM}$ and 102 455 $M^{-1} \cdot cm^{-1}$ for Cel5A.

Assays for cellulase activity

pNPC (*p*-nitrophenyl-β-D-cellobioside)

The pNPC assay was carried out as described previously [15].

Macromolecular substrates: CMC and amorphous cellulose

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

Differential scanning calorimetry

The experiments were carried out using a MicroCal VP-DSC microcalorimeter and protein samples at concentrations ranging from 0.7 to 1 $mg \cdot ml^{-1}$. The samples were first dialysed against 30 mM Mops buffer, pH 7.5, and eluted on a PD 10 column (GE Healthcare) equilibrated in the buffer supplemented with 500 mM sulfobetaine [3-(1-pyridinio)-1-propane sulfonate] 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 point, T_m, and the calorimetric enthalpy, ΔH_{cal}, of the individual transitions were fitted independently using MicroCal Origin software, version 4.1. The source and the magnitude of errors of T_m and ΔH_{cal} have been discussed elsewhere [28]. The reversibility of the thermal transitions was assessed by the re-scanning of the samples after cooling them at 10 °C.

Thermodependence of the activity

The activity of the various cellulases as a function of temperature (4–35 °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 substrate CMC. The activation energy E_a was calculated from the slopes of the Arrhenius plots: ln k_{cat} against $1/T$ according to the Arrhenius equation, $k_{cat} = A \cdot e^{-E_a/RT}$ in which k_{cat} is the reaction rate, A is the pre-exponential factor, R is the gas constant (8.314 $J \cdot mol^{-1} \cdot K^{-1}$), and T is the absolute temperature in Kelvin. The other

thermodynamic parameters were calculated as described previously [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) using an excitation wavelength of 280 nm and emission wavelengths of 340 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_{280} close to 0.1. The relative fluorescence intensity was recorded after 30 s of incubation in the presence of acrylamide concentrations ranging from 5 to 120 mM at 10 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 ($\varepsilon_{280} = 4.3\ M^{-1} \cdot cm^{-1}$). 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 $[Q]$ is the concentration of the quencher [30].

PIXE (proton-induced X-ray emission)

PIXE is a sensitive technique [31,32] used to determine the amount and nature of metal ions possibly bound to the cold-adapted cellulase, especially Ca^{2+} as in the case of thrombospondins [33]. The element-specific X-ray emission is generated through proton beams directed on to the sample. Trace element levels can be detected, but not carbon, oxygen or nitrogen, nor elements with atomic number lower than 11. A proton beam of 2.5 and 3.0 MeV was generated with a Van de Graaff 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 a helium-filled chamber at a standard pressure of 760 mmHg (101 kPa). The acquisition of the spectra was carried out using a GUPIX spectral analysis programme. A non-linear least-squares analysis of peak intensities was performed using a library of measured X-ray intensities for each element with an S.D. of 5 %. For acquisition of the spectra, the native cellulase samples (1.6 $mg \cdot ml^{-1}$) in the presence or absence of 5 mM EGTA was prepared in Milli-Q water eluted on a Chelex 100 column in order to remove as much Ca^{2+} as possible. The pH was adjusted to 6.5 with NaOH. Essentially plastic utensils, washed previously with concentrated nitric acid, and thoroughly rinsed with Ca^{2+}-free water, were used throughout. The samples were all filtered before use through a thin nucleopore filter free from any metal contamination.

SAXS 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 (1 Å = 0.1 nm). The sample-to-detector distances were set at 4.0 and 1.0 m, resulting in scattering vectors, q, ranging from 0.010 to 0.15 $Å^{-1}$ and 0.03 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 4.0 and 1.0 m were merged for the calculations using the entire scattering spectrum.

The protein concentration of Cel5GΔ2PmutCC (in buffer consisting of 10 mM Hepes/NaOH, pH 7.5, 0.04 % sodium azide and 20 μM Pefabloc with 5 % glycerol as radiation scavenger) was varied from 1.2 to 7.5 $mg \cdot ml^{-1}$ in order to check for interparticle interactions. The protein concentration of full-length Cel5G in the presence of 10 mM $CaCl_2$ (in the same buffer) was varied from 1.5 to 7.3 $mg \cdot ml^{-1}$. The radius of gyration R_G was derived by the Guinier approximation $I(q) = I(0) \cdot \exp(-q^2R_G^2/3)$ 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 diameter 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 CM and of the CBM 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 and Cel5GΔ2PmutCC, the CM, $Cel5G_{CM}$, and the mesophilic cellulase Cel5A are shown in Figure 1. The three mutants correspond to a sequential shortening of the LR; 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 (amino acids 312–342). This deletion includes the first loop, the closest to the CM, containing the first disulfide bridge formed between Cys^{314} and Cys^{328}. Cel5GΔ2P corresponds to a further shortening of 37 residues (amino acids 343–379); it includes the second loop and the disulfide bridge formed between Cys^{345} and Cys^{359}. Cel5GΔ2PmutCC is similar in length to Cel5GΔ2P, but the two cysteine residues, Cys^{382} and Cys^{396}, have been replaced by two alanine residues in order to evaluate the importance of the last disulfide bridge in adaptation to the cold. The fourth construct, $Cel5G_{CM}$ extending from residue 1 to residue 292, is a truncated form of Cel5G consisting of the sole CM devoid of the LR and of the CBMs. 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]. Approx. 20 mg of each protein was isolated per litre of culture. $Cel5G_{CM}$ was purified following a new protocol as described in the Experimental section. This protocol allowed the recovery of 25 mg of purified protein per litre of culture. The purification of the mesophilic cellulase Cel5A, encoded by pSNAB1 [19,20], was carried out in one single step by affinity chromatography on cottonwool. A yield of 14 mg per litre of culture 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 42.0 kDa calculated with the ProtParam software for Cel5G, Cel5G1Δ1P, Cel5GΔ2P, Cel5GΔ2PmutCC, $Cel5G_{CM}$ and Cel5A respectively.

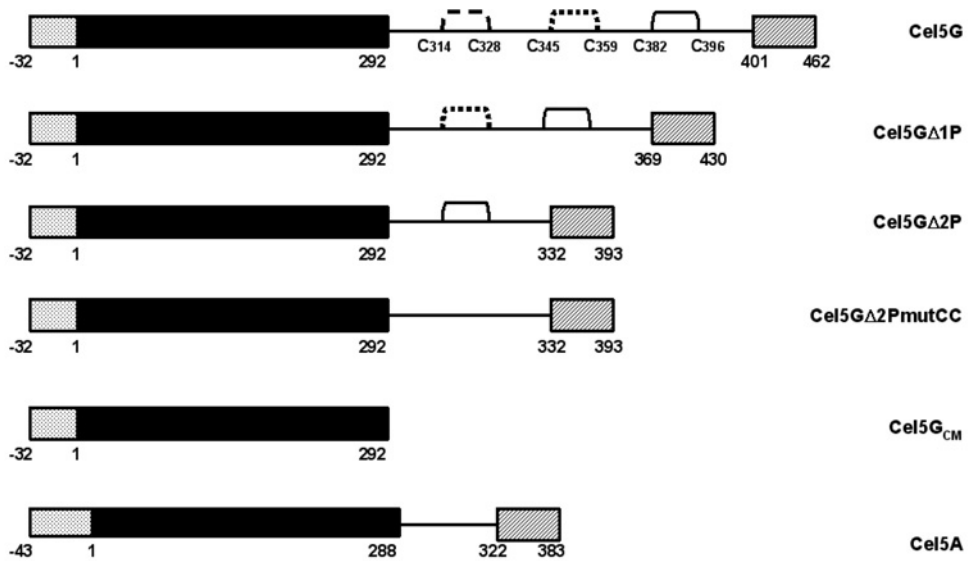

Figure 1 Schematic representation of the structures of the various cellulases

Structures of native full-length psychrophilic cellulase, Cel5G (the numbering corresponds to the amino acid sequence of the mature protein), its mutants, Cel5GΔ1P, Cel5GΔ2P, Cel5GΔ2PmutCC, the CM, Cel5G$_{CM}$, and the full-length native mesophilic cellulase, Cel5A. The signal peptides are represented by the stippled boxes, the CMs by the black boxes, the LRs by the continuous line, and the CBMs by the hatched boxes. The rectangles on the LR correspond to the disulfide bridges and loops closed by them.

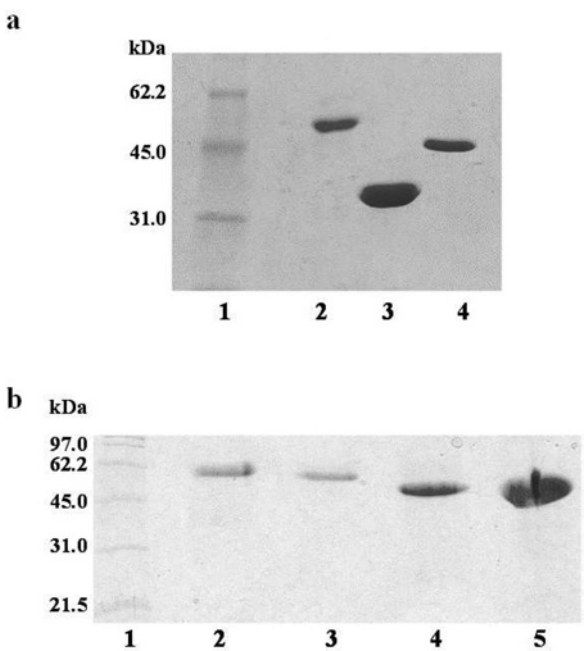

Figure 2 SDS/PAGE analysis of the psychrophilic cellulase, its deletion mutations and the mesophilic cellulase

(**a**) SDS/PAGE (12 % polyacrylamide) of the full-length psychrophilic cellulase, Cel5G (lane 2), its CM, Cel5G$_{CM}$ (lane 3) and the mesophilic cellulase, Cel5A (lane 4). Lane 1, standard proteins (phosphorylase *b*, 97 kDa; BSA, 62.2 kDa; ovalbumin, 45 kDa; carbonic anhydrase, 31 kDa; and soya bean trypsin inhibitor, 21.5 kDa). (**b**) SDS/PAGE (12 % polyacrylamide) of the full-length psychrophilic cellulase, Cel5G (lane 2); the deletion mutants, Cel5GΔ1P (lane 3), Cel5GΔ2P (lane 4) and the mutant Cel5GΔ2PmutCC (lane 5). Lane 1, standard proteins as in (**a**). Molecular masses are indicated in kDa.

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, thus 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 CM of Cel5G were determined. Despite the very high sequence similarity to its mesophilic counterpart, we showed that the psychrophilic CM displays the typical adaptation to cold, which consists of a weakening of intramolecular interactions contributing to increase the flexibility of the molecular edifice. Furthermore, approx. 50 % of the amino acids of the cold-adapted CM display temperature or *B* factors higher than those of their mesophilic counterparts from *Erwinia chrysanthemi*. A peculiar feature is also that the CM of the cold-adapted enzyme shows a large excess of negative charges over positive charges (34 compared with 18), in contrast with the mesophilic homologue (31 compared with 29) [16,40–42]. Although the CBMs display very high sequence similarity (57 % of complete identity), the LRs are strikingly different, both in length (109 residues in the cold-adapted enzyme and only 34 residues in the mesophilic counterpart) and in sequence. A distinctive feature of the LR 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 DXDXDXXXDXXD motifs, called TSP3 (thrombospondin 3) motifs, 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 substrate pNPC is shown in Figure 3(a). One can see that, in the temperature range 0–50 °C, the activity of the CM 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 CM alone is only approx. 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 substrates such as CMC and amorphous

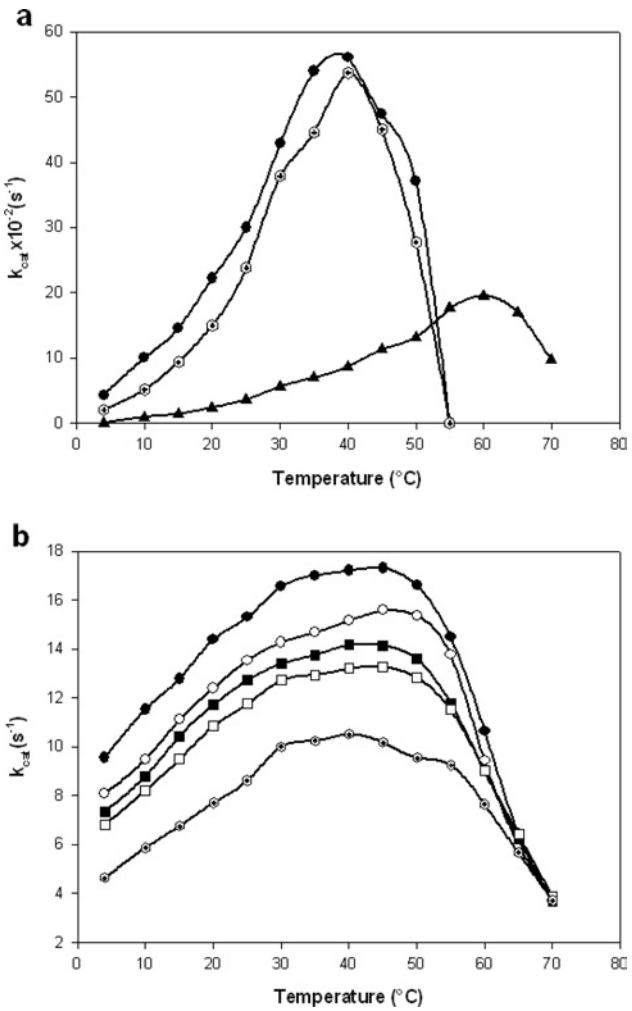

Figure 3 Thermodependence of the activity of the psychrophilic cellulase and its mutants using pNPC and CMC as substrates

Thermodependence of the activity of the various cellulases. (**a**) pNPC used as substrate: closed circles, full-length cold-adapted cellulase, Cel5G; hexagons, CM of the cold-adapted cellulase, Cel5G$_{CM}$; closed triangles, mesophilic cellulase from *Erwinia chrysanthemi*, Cel5A. (**b**) Thermodependence of the cold-adapted cellulases using CMC as substrate: closed circles, full-length cold-adapted cellulase, Cel5G; open circles, mutant Cel5GΔ1P; closed squares, mutant Cel5GΔ2P; open squares, mutant Cel5GΔ2PmutCC; hexagons, Cel5G$_{CM}$.

cellulose are used. Indeed, unlike with pNPC, the specific activity at 25 °C of the cold-adapted enzyme using CMC and amorphous cellulose as substrates is more than twice as high as that of the CM. The activity of the CM, in contrast, now approaches 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 Cel5G$_{CM}$. The data are shown in Figure 3(b). 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 LR shortening. The shortest mutant, Cel5GΔ2PmutCC, devoid of any disulfide bridge, displays catalytic performances approaching those of the CM 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 E_a show a regular increase as a function of LR 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 in the activation enthalpy, ΔH^*, partially reinforced by a concomitant decrease of the activation entropy, ΔS^*, suggesting that LR shortening also induces a higher degree of order of the truncated forms when compared with 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 -49.6 kJ · mol^{-1} for Cel5GΔ2PmutCC). This decrease in the $T\Delta S^*$ observed upon removal of the disulfide bridge could be due to an increase in 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 intramolecular interactions prevented by the presence of the disulfide bridge. This hypothesis is supported by the increase of ΔH^* (more bonds have to be broken to reach the activated state) and also by the increase in the ΔH_{cal} of the disulfide-free form when compared with that of the disulfide-containing form of the truncated enzyme (see Table 3). In the same context, the origin of the increase in the heat absorbed during the activation process of the truncated forms could be ascribed to a progressive increase in the stability and rigidity of the three-dimensional structures of the cold-adapted cellulases upon shortening of the LR.

Thermal stability

To test 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 ΔH_{cal} values 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 in the T_m values of the two thermodynamic domains as a function of LR shortening of the full-length enzyme. Indeed, Cel5G$_{CM}$ has a T_{m1} 6.9 °C higher than that of the first transition of the full-length enzyme. This is accompanied by a regular increase in the ΔH_{cal}, suggesting, as already stated above, that LR shortening induces a progressive increase in the weak intramolecular interactions. As mentioned, even the deletion of the last disulfide bridge, producing Cel5GΔ2PmutCC, gives rise to an increase in ΔH_{cal} values, suggesting that the disulfide bridges are also necessary to maintain the LR in a non-collapsible and extended form, preventing interactions within, and between, modules, as suggested by SAXS 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 CM alone by a factor of 2, whereas the activity of the CM was close to that of the complete mesophilic enzyme, Cel5A, when large substrates, such as CMC and amorphous cellulose, were used. It is generally considered that a high flexibility of proteins is closely correlated with 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 LR.

Table 2 Comparison of kinetic parameters

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 *Erwinia chrysanthemi* and at 10 °C of the catalytic constants and of the thermodynamic parameters, using CMC as the substrate, of the full-length psychrophilic cellulase, and of its mutants Cel5GΔ1P, Cel5GΔ2P, Cel5GΔ2PmutCC and Cel5G$_{CM}$.

(a) Kinetic parameters at 25 °C

	pNPC		CMC		Amorphous cellulose
Cellulase	$k_{cat} \times 10^{-2}$ (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

(b) Kinetic parameters at 10 °C (CMC)

Cellulase	k_{cat} (s^{-1})	E_a (kJ · mol^{-1})	ΔG^* (kJ · mol^{-1})	ΔH^* (kJ · mol^{-1})	$T\Delta S^*$ (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

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

Cellulase	T_{m1} (°C)	T_{m2} (°C)	ΔH_{cal1} (kJ · mol^{-1})	ΔH_{cal2} (kJ · mol^{-1})
Cel5G	45.3 ± 0.3	47.9 ± 0.1	264.4 ± 32.6	219.7 ± 32.2
Cel5GΔ1P	43.3 ± 0.3	48.3 ± 0.1	266.9 ± 32.2	225.5 ± 31.4
Cel5GΔ2P	45.6 ± 0.4	48.8 ± 0.1	268.6 ± 34.3	228.9 ± 33.5
Cel5GΔ2PmutCC	47.7 ± 0.3	50.3 ± 0.1	271.1 ± 44.4	239.3 ± 43.5
Cel5G$_{CM}$	50.2 ± 0.1	53.8 ± 0.1	278.2 ± 33.5	295.0 ± 29.7

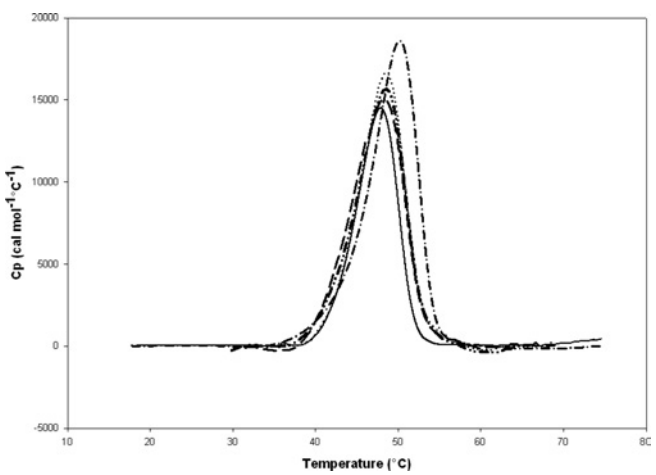

Figure 4 Differential scanning calorimetry profile of the psychrophilic cellulase, its mutants and the mesophilic cellulase

Thermograms obtained by differential scanning calorimetry of the full-length cold-adapted enzyme, Cel5G, and its LR deletion mutants. Each peak represents the heat absorbed during the unfolding of the three-dimensional structure. From left to right: Cel5G (———), Cel5GΔ1P (— — —), Cel5GΔ2P (–· ·–· ·–), Cel5GΔ2PmutCC (·····) and Cel5G$_{CM}$ (·—·—·).

These data 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,

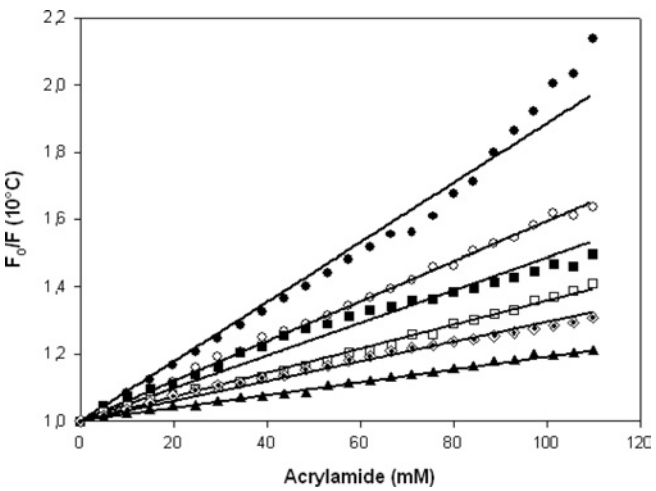

Figure 5 Fluorescence quenching at 10 °C of the psychrophilic cellulase, its mutants and the mesophilic cellulase using acrylamide as a quencher

Stern–Volmer plots corresponding to the quenching of the tryptophan fluorescence at 10 °C of the various cold-adapted cellulases using acrylamide as a quencher. The relative fluorescence, F_0/F, is plotted as a function of acrylamide concentration. F_0 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, which have been calculated from the Stern–Volmer equation, $F_0/F = 1 + K_{SV}[Q]$. Closed circles, Cel5G $K_{SV} = 0.89$; open circles, Cel5GΔ1P $K_{SV} = 0.59$; closed squares, Cel5GΔ2P $K_{SV} = 0.49$; open squares, Cel5GΔ2PmutCC $K_{SV} = 0.34$; diamonds, Cel5G$_{CM}$ $K_{SV} = 0.29$; and closed triangles, Cel5A $K_{SV} = 0.19$.

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 with 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 (F_0/F) and Stern–Volmer constants (K_{sv}), corresponding to the slopes of the lines obtained (Figure 5). One can see that the sequential shortening of the LR also induces

a dramatic and progressive decrease of the flexibility of the molecular structures. This clearly indicates that the composition and structure of the LR modulate the rigidity of the CM, which contains ten of the 13 tryptophan residues found in the protein. It is worth noting that the LR itself does not contain any tryptophan residues. The implication of the LR in the flexibility of the CM is very important for catalysis, but rather hard to explain. One hypothesis is that the numerous negative charges in the LR could be responsible for this effect through electrostatic repulsion, since one should remember that the CM also contains a large excess of negative charges (34 compared with 18 positive charges). These electrostatic repulsions probably prevent the formation of a more cohesive structure of the CM. The reduction of the number of negative charges in the LR, through its shortening, probably reduces this effect, increases the cohesion and stability of the CM and progressively increases the ΔH^* of the truncated forms (Table 2). Furthermore, LR shortening may render possible some interactions between the CM and CBM regions. All of these effects would prevent the flexibility of the LR to be transmitted to the CM when the LR is shortened.

SAXS

SAXS experiments have been performed on the shortest Cel5G variant, Cel5GΔ2PmutCC, devoid of any disulfide bridge. SAXS is indeed a very appropriate tool to assess the structural properties (dimensions and overall shape) and the flexibility in solution of multimodular proteins containing disordered regions such as LRs.

At low angles, the scattered intensities are very well approximated by the Guinier law, and the R_G values, 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 multimodular cellulases [17,46], and the inferred D_{max} is 136 ± 5 Å. 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 φ^2 of ~ 1.2 (Figure 6A). A typical calculated shape superimposed with the structure of the CM and of the CBM of Cel5G (Figure 6B) shows that the LR 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 LR is 66 Å, whereas the LR is composed of 39 residues. This corresponds to a global compactness of 0.61 residues/Å, to be compared with 0.48 residues/Å for wild-type Cel5G at 20 °C when one does not take into account the residues involved in the transverse disulfide-bonded loops [17]. This result indicates that the LR preferably adopts more collapsed conformations when there is no transverse loop. The disulfide loops probably induce steric constraints that prevent collapse of the LR and the formation of intramolecular 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 of the present study, an extended conformation, but it is much more compact than the wild-type LR.

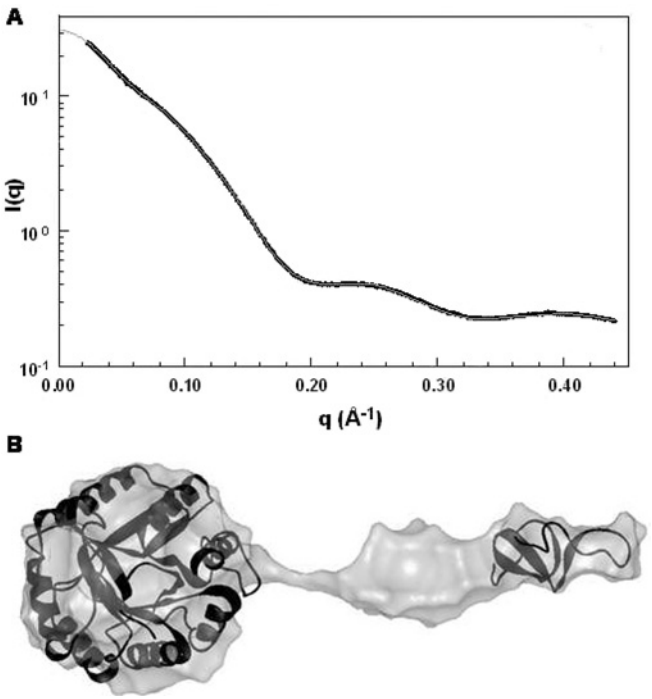

Figure 6 Structure of the shortest mutant of the psychrophilic cellulase, devoid of any disulfide bridge, as determined using SAXS

Ab initio shape restoration of mutant Cel5GΔ2PmutCC by SAXS (**A**) experimental spectrum (black circles) and fitted to the data given by GASBOR (grey line). (**B**) Typical calculated shape obtained by GASBOR (grey envelope), superimposed with the atomic structures of the CM (left) and of the CBM (right) of Cel5G. The atomic structures are in black ribbon representation.

Role of the TSP3 motifs

The alignment of the amino acid sequences of the LR 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 293–313, 329–344 and 360–381, which contain a repeated DXDXDGXXDXXD motif at positions 301–312, 332–343 and 369–380. They are identical with the TSP3 motif found in thrombospondins, 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 a K_d of 0.1 mM. The bound cations are, however, very important for the conformation and function of the protein. We therefore postulated that the LR of the cold-adapted cellulase, containing three of these motifs, could also bind Ca^{2+} ions. In contrast with thrombospondins, the chelation of Ca^{2+} using EGTA does not affect the catalytic performance of the cold-adapted cellulase (results not shown). We nevertheless decided to evaluate possible metal binding using the technique of 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 presence or absence 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 was detected that was typical of Ca^{2+} or any other metal, except from Na^+ originating from the added EGTA (results not shown).

SAXS experiments were also performed on wild-type Cel5G in the presence of Ca^{2+} (seawater concentration 10 mM). The R_G and the D_{max} of Cel5G obtained in the presence of Ca^{2+} are 53.8 ± 2.1 and 210 ± 10 Å respectively. These values are identical with those obtained for Cel5G with no Ca^{2+} 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 Ca^{2+} (results not shown).

Figure 7 Sequence alignment of the linker region of the psychrophilic and mesophilic cellulases

Sequence alignment, using the programme ClustalW [50], of the LRs of the cold-adapted cellulase, Cel5G (formerly named CelG; GenBank® accession number Y17552) and of the mesophilic cellulase, Cel5A (formerly named CelZ; GenBank® accession number Y00540). The LR of the psychrophilic enzyme is 109 residues long and that of the mesophilic cellulase is 34 residues long. The few dots between the alignment correspond to identical or similar amino acid residues, and the underlined sequences to the three TSP3 motifs.

These results show that, unlike in thrombospondins, the TSP3 motifs of Cel5G do not bind Ca^{2+} ions; the three loops closed by the disulfide bridges might be responsible for the fact that the three repeated DXDXDXXXDXXD motifs 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 LR from the psychrophilic cellulase contains 23 negative charges in total, with no positively charged amino acids. In consequence, the negative charges of the LR also contribute, through electrostatic repulsion, to prevent the LR from collapsing and to keep the CM and the CBMs well apart.

Conclusion

All of the results presented here indicate that the Cel5G LR 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 CM, the LR has the possibility to explore much more conformations during evolution in such a way to adopt original structural features that confer on the full-length enzyme a higher flexibility and an efficient adaptation to cold. As demonstrated by SAXS [17], the long LR 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 owing to the transverse loops prevent the stabilization of intramolecular interactions, and thus confer on the LR a higher flexibility. The role of the transversal disulfide-bonded loops of Cel5G is comparable with the glycosylation of mesophilic LRs in bimodular glycoside hydrolases [46,47]. This higher flexibility of the LR provides a larger degree of freedom to the CM and the CBM, favouring their correct positioning with regard to the cellulose fibres, even at low temperatures. It also considerably increases the area covered by the CM at the surface of the cellulose and thus also favours the catalytic efficiency as mentioned previously [17,48,49].

We thank N. Gerardin-Otthiers and I. Thamm for their skilful technical assistance. We are also very grateful to Dr F. Barras for the 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. D. 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 European Union (contract ERB-FMRX_CT 97-0131). G. K. S. is a recipient of the Scholarship of Foreign Study from the Ministry of Education and Scientific Research of the 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, Jr, R. C. and Warren, R. A. (1991) Domains in microbial β-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, Jr, R. C. 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, Jr, R. C., 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 Crystallogr. Sect. D Biol. Crystallogr. **59**, 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. Nat. 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**, 1211–1224
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**, 16074–16086

19 Py, B., Bortoli-German, I., Haiech, J., Chippaux, M. and Barras, F. (1991) Cellulase EGZ of *Erwinia chrysanthemi*: structural organization and importance of His^{98} and Glu^{133} residues for catalysis. Protein Eng. **4**, 325–333
20 Py, B., Chippaux, M. and 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 edn, Cold Spring Harbor Laboratory Press, Cold Spring Harbor
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., ed.), p. 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. Crystallogr. **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 Directory, pp. 77–78, Silicon Graphics, Mountain View
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. Annu. 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, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22**, 4673–4680

Received 15 May 2007/17 July 2007; accepted 18 July 2007
Published as BJ Immediate Publication 18 July 2007, doi:10.1042/BJ20070640