Kinetic and structural optimization to catalysis at low temperatures in a psychrophilic cellulase from the Antarctic bacterium *Pseudoalteromonas haloplanktis*

Geneviève GARSOUX*, Josette LAMOTTE†, Charles GERDAY* and Georges FELLER*[1]

*Laboratory of Biochemistry, Institute of Chemistry B6a, University of Liège, B-4000 Liège, Belgium, and †Center for Protein Engineering, Institute of Chemistry B6a, University of Liège, B-4000 Liège, Belgium

The cold-adapted cellulase CelG has been purified from the culture supernatant of the Antarctic bacterium *Pseudoalteromonas haloplanktis* and the gene coding for this enzyme has been cloned, sequenced and expressed in *Escherichia coli*. This cellulase is composed of three structurally and functionally distinct regions: an N-terminal catalytic domain belonging to glycosidase family 5 and a C-terminal cellulose-binding domain belonging to carbohydrate-binding module family 5. The linker of 107 residues connecting both domains is one of the longest found in cellulases, and optimizes substrate accessibility to the catalytic domain by drastically increasing the surface of cellulose available to a bound enzyme molecule. The psychrophilic enzyme is closely related to the cellulase Cel5 from *Erwinia chrysanthemi*. Both k_{cat} and k_{cat}/K_m values at 4 °C for the psychrophilic cellulase are similar to the values for Cel5 at 30–35 °C, suggesting temperature adaptation of the kinetic parameters. The thermodynamic parameters of activation of CelG suggest a heat-labile, relatively disordered active site with low substrate affinity, in agreement with the experimental data. The structure of CelG has been constructed by homology modelling with a molecule of cellotetraose docked into the active site. No structural alteration related to cold-activity can be found in the catalytic cleft, whereas several structural factors in the overall structure can explain the weak thermal stability, suggesting that the loss of stability provides the required active-site mobility at low temperatures.

Key words: carbohydrate-binding module, cellulase, extremophile, glycoside hydrolase, *Pseudoalteromonas haloplanktis*, psychrophile.

INTRODUCTION

Cellulose is the major polysaccharidic compound in plants and is the most abundant organic compound on earth. Accordingly, its turnover in the carbon cycle is of prime importance for all living organisms [1–3]. Cellulose synthesis and recycling in the marine environment account for a large proportion of the cycle because oceans cover three-quarters of the planet's surface. However, one generally forgets that these steps are performed at low temperatures, as nearly 90 % of the sea water volume is permanently below 5 °C [4,5]. This prompted us to analyse the functional and structural properties of a cold-active cellulase, involved in cellulose hydrolysis at low temperatures. This enzyme is produced by *Pseudoalteromonas haloplanktis*, a psychrophilic bacterium isolated in Antarctica, which thrives permanently at temperatures close to the freezing point of water. It is therefore expected that the cellular components of this micro-organism are adapted to life in the cold [6].

Cellulases (EC 3.2.1.4) are endoenzymes that hydrolyse randomly β-1,4-glycosidic bonds in cellulose, releasing cellobiose as the smallest product. These enzymes can be found associated with the cellulosome, an extracellular supramolecular assembly produced by some bacteria and fungi [7,8], but are generally secreted as independent enzymes by cellulolytic micro-organisms. The catalytic domain of cellulases is frequently connected to a CBM (cellulose-binding module) via a linker region [3]. These modules bind and diffuse on the substrate surface [9], while the catalytic domain hydrolyses the substrate, resulting in a proposed caterpillar-like motion of the enzyme along the cellulose chains [10]. Both the catalytic domains and the CBM are structurally diverse [3,11], and have been classified into various sequence- and structure-related families [12,13]. In the present paper, we report the cold-adaptive features of a new member of glycoside hydrolase family 5.

EXPERIMENTAL

Production and purification of the wild-type cellulase

The Gram-negative Antarctic bacterium *Ps. haloplanktis* TAB23, formerly *Alteromonas haloplanctis* A23 [14], was cultivated at 4 °C for 6 days in Erlenmeyer flasks shaken at 250 rev./min and containing 16 g/l bacto-tryptone, 16 g/l yeast extract, 20 g/l NaCl, 10 g/l sea salts, 30 g/l maltose, pH 7.6. The addition of maltose was found to induce cellulase production in the supernatant by 15–20-fold. After centrifugation at 15 000 ***g*** for 15 min, the culture supernatant was brought to 0.5 mM PMSF and 75 % saturation in $(NH_4)_2SO_4$. The precipitate collected by centrifugation at 24 000 ***g*** for 30 min was dissolved in 25 mM Pipes, 30 % $(NH_4)_2SO_4$, pH 6.5, and loaded on to a phenyl-Sepharose CL-4B column (5 cm × 6.5 cm) equilibrated and washed in the same buffer, then eluted with a decreasing gradient (1 litre) from 30 to 0 % $(NH_4)_2SO_4$. The active fractions were loaded on a DEAE-agarose column (1.5 cm × 40 cm) eluted with a linear gradient (0.5 litre) of 0–450 mM NaCl in 25 mM Pipes, pH 6.5. The active fractions were concentrated by ultrafiltration on a PTGC polysulphone membrane (Millipore) and loaded on to an Ultrogel AcA 54 column (2.5 cm × 100 cm) eluted with 25 mM Pipes, pH 6.5.

Abbreviations used: CBM, cellulose-binding module; CelG, cellulase from *Pseudoalteromonas haloplanktis*; Cel5, cellulase from *Erwinia chrysanthemi*; IPTG, isopropyl β-D-thiogalactoside; pNPC, *p*-nitrophenyl-β-D-cellobioside.

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

Gene cloning, expression and purification of the recombinant psychrophilic cellulase CelG

The genomic DNA of *Ps. haloplanktis* was partially digested by *Sau*3AI, and the resulting fragments were inserted into the single *Bam*HI site of pSP73 (Promega). The ligated DNA was used to transform *Escherichia coli* XL1-Blue MRF' Kan (Stratagene) and the cellulolytic activity of the recombinant cells was screened on L-agar plates containing 5 g/l carboxymethylcellulose and 100 mg/l Trypan Blue. The nucleotide sequence of the active clone was determined using an ALFexpress DNA Sequencer (Pharmacia LKB).

For overexpression of the psychrophilic cellulase, the nucleotide sequence containing the start codon was mutated by PCR into CATATG in order to introduce an *Nde*I restriction site. The amplified fragment was cloned in the pCRScript vector (Stratagene), and most of the coding sequence was replaced by non-amplified DNA from the original clone, using an internal *Pst*I site and the plasmid-borne *Sac*I site, before sequencing the remaining PCR-amplified region bearing the *Nde*I site. The full-length gene was then cloned in the expression vector pET22b (Novagen) at *Nde*I and *Sac*I sites and was used to transform *E. coli* BL21(DE3) cells. Cultures were performed in LB (Luria–Bertani) broth containing 100 mg/l ampicillin at 18 °C, and cellulase expression was induced by 100 μM IPTG (isopropyl β-D-thiogalactoside) at a A_{600} of approx. 3. Production was estimated to be 100 mg/l of culture. Cells were recovered by centrifugation at 15 000 ***g*** for 30 min after 10 h of induction and were submitted to osmotic shock in order to recover the periplasmic fraction. The cell pellet was dispersed in an ice-chilled buffer containing 100 mM Tris/HCl, 0.5 mM EDTA, 0.1 mM PMSF, 0.5 M saccharose, pH 8.0, and, following centrifugation at 15 000 ***g*** for 30 min, the cell pellet was dispersed further in ice-chilled Milli-Q water and 1 mM $MgCl_2$ was added. The supernatant was brought to 1 M $(NH_4)_2SO_4$ and loaded on to a cotton wool column (5 cm × 12 cm) equilibrated in 20 mM sodium phosphate and 1 M $(NH_4)_2SO_4$, pH 7.0 [15]. After extensive washing with this buffer, the cellulase was eluted with Milli-Q water. The active fractions were loaded on to a Biosquale Q2 column (0.7 cm × 5.2 mm) coupled to a FPLC system and equilibrated in 25 mM Tris/HCl, pH 8.0. The enzyme was eluted with a KCl gradient from 0 to 0.5 M. The concentrated cellulase solution was kept frozen at − 70 °C.

Expression and purification of the cellulase Cel5 from *Erwinia chrysanthemi*

The plasmid pSNAB1 [16,17], kindly provided by Dr F. Barras (CNRS, Marseille, France), was used to produce the cellulase Cel5. The protocol used was essentially similar to that described for the CelG cellulase with the following modifications: *E. coli* XL1-Blue MRF' Kan was used for Cel5 production and was cultivated at 30 °C in LB broth containing 100 mg/l ampicillin and 50 μM IPTG. Cells were harvested after 20–25 h. Production was estimated to be 60 mg/l of culture. The periplasmic extract was brought to 1 M NaCl before loading on the cotton wool column equilibrated in 20 mM sodium phosphate and 1 M NaCl, pH 7.0.

Analytical procedures

Kinetic parameters were recorded in 50 mM Hepes, pH 7.5, with increasing concentrations (0.2–30 mM) of the substrate pNPC (*p*-nitrophenyl-β-D-cellobioside) and at temperatures from 4 to 40 °C. Activities were recorded in a thermostatically controlled Uvikon 860 spectrophotometer (Kontron) and calculated on the basis of a molar absorption coefficient for *p*-nitrophenol of 12 900 $M^{-1} \cdot cm^{-1}$ at 405 nm. The kinetic parameters k_{cat} and K_m were determined by a non-linear regression fit of the saturation curves using the Michaelis–Menten equation. The standard assay was carried out at 25 °C with 5 mM pNPC. The pH-stability was recorded by incubating the enzymes in Teorell–Stenhagen buffer between pH 2 and 12, and determination of the residual activity of aliquots was achieved by the standard assay. The pH-dependence of the activity was recorded using pNPC in Teorell–Stenhagen buffer between pH 2 and 12. The reaction was stopped using 1 M Na_2CO_3, and relative activities were measured at 405 nm. Activity on 1 % carboxymethylcellulose in 50 mM Hepes, pH 7.5, was recorded on a MCS Isothermal Titration Calorimetry unit (MicroCal) as described previously [18] and calculated using a molar enthalpy of β-1,4-bond hydrolysis of 583 cal/mol of linkage.

Thermodynamic parameters of activation were calculated as described in [19] using the equations:

$$\Delta G^{\#} = RT \times [\ln(k_B T/h) - \ln k] \quad (1)$$

$$\Delta H^{\#} = E_a - RT \quad (2)$$

$$\Delta S^{\#} = (\Delta H^{\#} - \Delta G^{\#})/T \quad (3)$$

where k_B is Bolzmann's constant, h is Planck's constant, E_a is the activation energy of the reaction and R is the gas constant.

The N-terminal amino acid sequence was determined by automated Edman degradation using a pulsed-liquid-phase protein sequencer Procise 494 (Applied Biosystems). Protein concentrations were determined by the bicinchoninic acid method (Pierce) and were checked using theoretical molar absorption coefficients of 94 930 $M^{-1} \cdot cm^{-1}$ and 101 540 $M^{-1} \cdot cm^{-1}$ at 280 nm for CelG and Cel5 respectively.

Molecular modelling of cellulase CelG from *Ps. haloplanktis*

The structural model of the cellulase CelG was constructed using atomic co-ordinates of the catalytic domain (Protein Data Bank accession number 1EGZ) and of the CBM (Protein Data Bank accession number 1AIW) of the cellulase Cel5 from *E. chrysanthemi*. Amino acid sequence alignments between CelG and the templates were performed by FASTA and BESTFIT programs. Following model building using the HOMOLOGY program (MSI, San Diego, CA, U.S.A.), energy minimization was achieved by the DISCOVER program (MSI) in order to avoid bad molecular contacts. The geometric features were analysed by the INSIGHT II program (MSI) on a Silicon Graphics workstation. Cut-off distances were 3.5 Å (1 Å = 0.1 nm) for ionic interactions, 3–5.5 Å and 55–90° angles for aromatic interactions and 5 Å for hydrophobic clustering. Analysis of cavities was performed using the VOIDOO program (Uppsala Software Factory) with a 1.4 Å probe on a 0.75 Å grid. Docking the substrate cellotetraose into the catalytic cavity of CelG was performed using the atomic co-ordinates of the cellulase Cel5A from *Bacillus agaradherans* in complex with cellobiose (Protein Data Bank accession number 2A3H) and of the cellulase E1 from *Acidothermus cellulolyticus* in complex with cellotetraose (Protein Data Bank accession number 1ECE). Structurally conserved regions and the catalytic residues forming the active site of CelG and of both templates were superimposed in order to dock cellotetraose in the homology model of the *Ps. haloplanktis* cellulase.

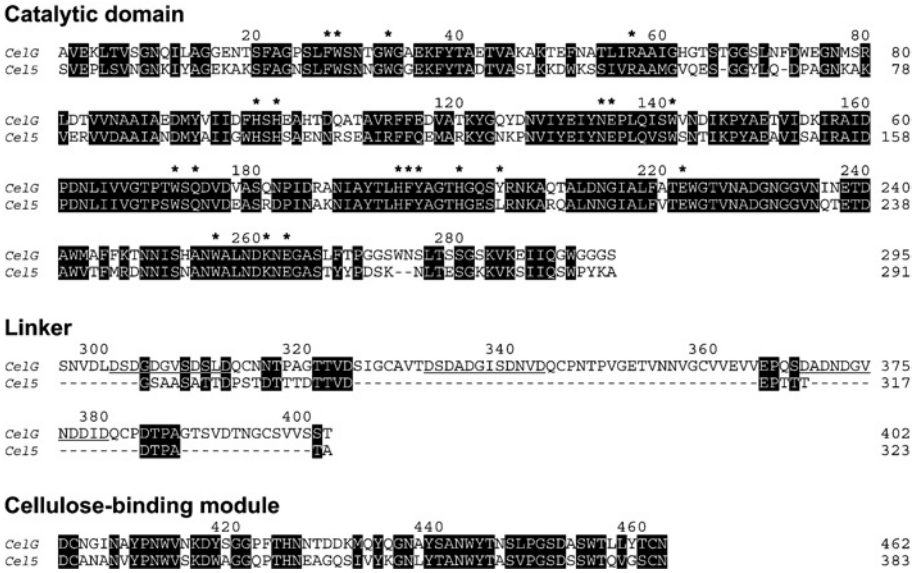

Figure 1 Amino acid sequence alignment of the psychrophilic (CelG) and mesophilic (Cel5) cellulases

The sequences are split into the three structural regions. Similarities are boxed and residues interacting with cellotetraose docked into the catalytic cleft of CelG are noted by asterisks. The three motifs of thrombospondin type 3 (T3) repeats in the linker of CelG are underlined.

RESULTS AND DISCUSSION

The wild-type cellulase from *Ps. haloplanktis*

The wild-type cellulase produced at 4 °C in the culture supernatant of *Ps. haloplanktis* has been purified to homogeneity. However, it was found that this enzyme of approx. 38 kDa is a truncated protein containing the catalytic domain, probably cleaved by the non-specific protease also secreted by the bacterium [20]. Such cleavage is frequently observed in cellulolytic micro-organisms [15,21]. This wild-type cellulase was used to locate the signal peptide cleavage site on the gene sequence (following N-terminal amino acid sequencing of the purified enzyme) and to check the catalytic properties of the recombinant, full-length cellulase produced in *E. coli* at 18 °C (see below).

Gene cloning, nucleotide sequence and domains of the psychrophilic cellulase CelG

A genomic library of *Ps. haloplanktis* has been constructed in *E. coli*, and among 10000 clones screened for cellulolytic activity on L-agar plates containing carboxymethylcellulose and Trypan Blue, one active colony has been isolated. The genomic DNA fragment from *Ps. haloplanktis* conferring a cellulolytic activity to *E. coli* has been sequenced on both strands and the nucleotide sequence is available under GenBank® accession number AHY17552. Sequence analysis reveals a coding DNA sequence of 1494 base pairs corresponding to a pre-protein of 498 amino acid residues designated CelG. The N-terminal amino acid sequence of the first 24 residues in the wild-type cellulase allowed the identification of a long signal peptide of 32 residues (formylmethionine included) in the precursor nucleotide sequence.

Further amino acid sequence analysis of the mature enzyme with data maintained at the Carbohydrate-Active Enzymes server CAZY (http://afmb.cnrs-mrs.fr/CAZY/) shows that the cellulase from *Ps. haloplanktis* is closely related to the cellulase Cel5 (formerly EGZ) secreted by *E. chrysanthemi* [16,17,22,23], a phytopathogenic Gram-negative bacterium. Figure 1 shows the amino acid sequence alignment of both enzymes. By comparison with the known structure of Cel5 [24,25], the psychrophilic cellulase CelG also contains three structurally and functionally distinct regions (Figure 1). The N-terminal domain is the catalytic domain. It belongs to glycoside hydrolase family 5 [26], subfamily 2, characterized by five strictly conserved residues in the active site [23,27], namely Arg^{57}, His^{100}, His^{194}, Glu^{135} (the acid/base residue) and Glu^{222} (the nucleophile) in CelG. These enzymes hydrolyse β-1,4-glycosidic bonds with net retention of the anomeric configuration. The catalytic domain also belongs to the GH-A clan characterized by a $(\beta/\alpha)_8$ barrel fold. The C-terminal domain is a cellulose-binding domain, belonging to family 5 of CBMs [3,13], and is responsible for binding to crystalline cellulose. Both domains are connected by a typical linker, highly flexible in solution [10]. In the case of the psychrophilic cellulase, CelG, this linker is unusually long. The amino acid sequence identity between CelG from *Ps. haloplanktis* and Cel5 from *E. chrysanthemi* was found to be 66 % for the catalytic domain and 57 % for the CBM (or 71 % and 58 % similarity respectively). Table 1 reports the amino acid composition of both cellulases, as well as that of their domains. The composition of the catalytic domains and of the CBMs is qualitatively similar; trends displayed by CelG catalytic domain will be discussed below. The most striking difference between both enzymes is the size of the linker region in CelG, which possesses 75 additional residues. As shown in Table 1, the type of these residues is not random, since the large size is mainly provided by an increase of aspartate, glycine, asparagine, serine and valine residues. This long linker is also highly negatively charged.

Kinetic properties of the psychrophilic and mesophilic cellulases

The strong primary-structural similarity between both cellulases prompted us to express and to purify the recombinant enzymes for

Table 1 Amino acid composition of psychrophilic (CelG) and mesophilic (Cel5) cellulases and of their domains

	Native		Catalytic domain		Linker		CBM	
Amino acid	CelG	Cel5	CelG	Cel5	CelG	Cel5	CelG	Cel5
Alanine	42	44	33	33	5	5	4	6
Arginine	6	10	6	10	0	0	0	0
Asparagine	45	37	25	30	10	0	10	7
Aspartic acid	43	22	18	14	20	5	5	3
Cysteine	8	2	0	0	6	0	2	2
Glutamic acid	19	19	16	17	3	1	0	1
Glutamine	15	11	9	8	4	0	2	3
Glycine	43	29	28	22	10	1	5	6
Histidine	8	5	7	4	0	0	1	1
Isoleucine	25	21	21	20	3	0	1	1
Leucine	19	13	14	12	2	0	3	1
Lysine	14	22	12	20	0	0	2	2
Methionine	4	4	3	4	0	0	1	0
Phenylalanine	14	8	13	8	0	0	1	0
Proline	16	16	7	10	6	3	3	3
Serine	37	35	20	25	12	3	5	7
Threonine	40	30	25	13	10	13	5	4
Tryptophan	13	14	10	10	0	0	3	4
Tyrosine	16	17	10	13	0	0	6	4
Valine	35	24	18	18	16	1	1	5
Aspartic acid + glutamic acid	62	41	34	31	23	6	5	4
Arginine + lysine	20	32	18	30	0	0	2	2
Total	462	383	295	291	107	32	60	60
M_r	49454	41635	32034	32101	10696	3101	6760	6469

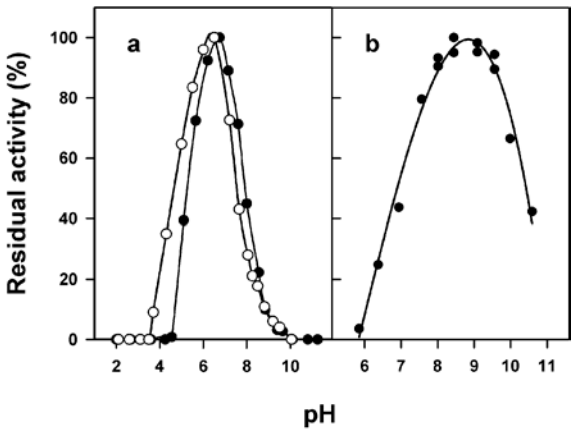

Figure 2 Effect of pH on activity and stability

(**a**) Effect of the pH of the reaction buffer on the activity of the psychrophilic CelG (●) and mesophilic Cel5 (○) cellulases. (**b**) Effect of the pH of the incubation buffer on stability of the psychrophilic cellulase.

Table 2 Thermodynamic activation parameters for pNPC hydrolysis by psychrophilic (CelG) and mesophilic (Cel5) cellulases at 4 °C

Cellulase	k_{cat} (s^{-1})	E_a ($kJ \cdot mol^{-1}$)	ΔG^* ($kJ \cdot mol^{-1}$)	ΔH^* ($kJ \cdot mol^{-1}$)	$T\Delta S^*$ ($kJ \cdot mol^{-1}$)	ΔS^* ($J \cdot mol^{-1} \cdot K^{-1}$)
CelG	0.18	48.5	71.6	46.2	−25.4	−91.8
Cel5	0.01	68.1	78.2	65.8	−12.4	−44.8

further characterization with respect to temperature adaptation. Purifications as described in the Experimental section yielded pure enzyme samples as judged by the single and homogeneous band observed in standard and overloaded SDS/PAGE (12% gels). N-terminal amino acid sequencing of the purified enzymes also indicated a single sequence and showed that the signal peptide of both cellulases was correctly cleaved in *E. coli*. As both wild-type (truncated) and recombinant (full-length) psychrophilic cellulases share similar k_{cat} and K_m values at 4 °C and 25 °C, one can conclude that expression at 18 °C in *E. coli* does not alter the kinetic properties of the cellulase CelG synthesized near 0 °C by the parent bacterium.

Both CelG and Cel5 cellulases share most of their functional properties. They readily hydrolyse the small chromogenic substrate pNPC and carboxymethylcellulose. However, the activity on the latter substrate (recorded using microcalorimetry as the heat released during β-1,4-glycosidic bond hydrolysis) is not linear and decreases rapidly with time. This has been related previously to the fast depletion of available non-substituted cellobiose units in this chemically modified cellulose [1]. The k_{cat} values have been estimated to be 0.030 s^{-1} and 0.005 s^{-1} at 25 °C for CelG and Cel5 respectively. No detectable activity was recorded on microcrystalline cellulose (Avicel). Accordingly, pNPC was used for all activity assays. As shown in Figure 2, the pH for optimal activity of CelG is found between pH 6.5 and 7.0, and this enzyme can therefore be classified as a neutral cellulase, like Cel5. The psychrophilic enzyme is optimally stable between pH 8.0 and 9.5. Its activity is not affected by 10 mM EDTA or 10 mM EGTA, indicating that CelG is not a metalloenzyme.

In contrast, the effect of temperature on both enzymes is drastically different. Like most psychrophilic enzymes [6,28], CelG is a heat-labile cellulase. For instance, its half-life during incubation at 45 °C is approx. 40 min, whereas the mesophilic Cel5 is unaffected in these conditions. The activity of CelG increases exponentially with temperature up to 40 °C (see below), but vanishes at temperatures higher than 45 °C. On the other hand, the apparent optimal temperature for activity of Cel5 is at 55 °C, whereas the psychrophilic CelG is totally inactive at this temperature. The determination of the kinetic parameters for pNPC hydrolysis reveals an interesting phenomenon of kinetic optimization at environmental temperatures, as well as the thermodynamic origin of this optimization. Indeed, the psychrophilic cellulase is at least 15 times more active at 4 °C than its mesophilic counterpart (Table 2) and, interestingly, the latter reaches the same k_{cat} value at 35 °C (Figure 3a), therefore demonstrating identical hydrolytic capacity at their respective temperatures. The increased activity at low temperature of CelG can be viewed as the main physiological adaptation to cold at the enzyme level as it compensates for the reduction of chemical reaction rates induced by low temperatures [6,28]. In contrast, the K_m value of the psychrophilic cellulase is higher at all temperatures (Figure 3b, inset) and is barely modified (5–6 mM) between 4 °C and 35 °C, whereas the K_m for Cel5 increases from 1.5 mM to 4 mM in this temperature range. As a result, the physiological efficiency or specificity constant k_{cat}/K_m (30.5 $M^{-1} \cdot s^{-1}$ and 6.5 $M^{-1} \cdot s^{-1}$ at 4 °C for CelG and Cel5 respectively) markedly increases with temperature for both enzymes (Figure 3b). Here again, the k_{cat}/K_m at 4 °C for the psychrophilic cellulase corresponds to the value reached by its mesophilic counterpart at 25–30 °C.

The thermodynamic parameters of activation for pNPC hydrolysis have been calculated from Arrhenius plots of the temperature-dependence curves of k_{cat} (Figure 3a, inset) and are given in Table 2. The lower Gibbs free energy of activation, ΔG^*, of CelG can be depicted as a lower energy barrier that has to be mastered by the ground-state enzyme–substrate complex

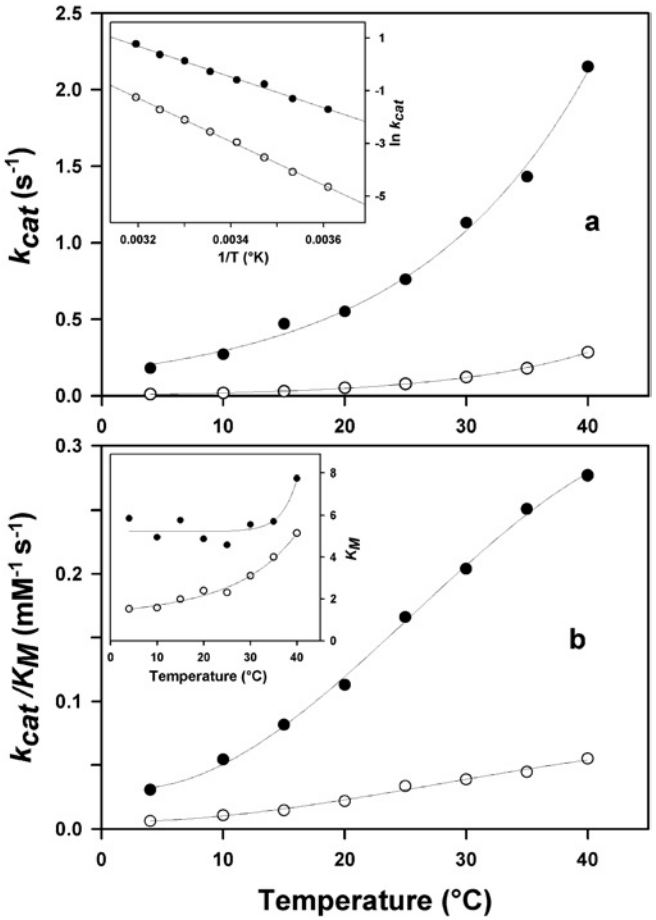

Figure 3 Temperature-dependence of the kinetic parameters

(**a**) Temperature-dependence of the activity of the psychrophilic CelG (●) and mesophilic Cel5 (○) cellulases on pNPC. Inset, Arrhenius plot of the temperature-dependence curves. (**b**) Temperature-dependence of the catalytic efficiency for pNPC, symbols as in (**a**). Inset, temperature-dependence of K_m for pNPC. Solid lines in this panel are a best-fit of the data and do not correspond to a defined function.

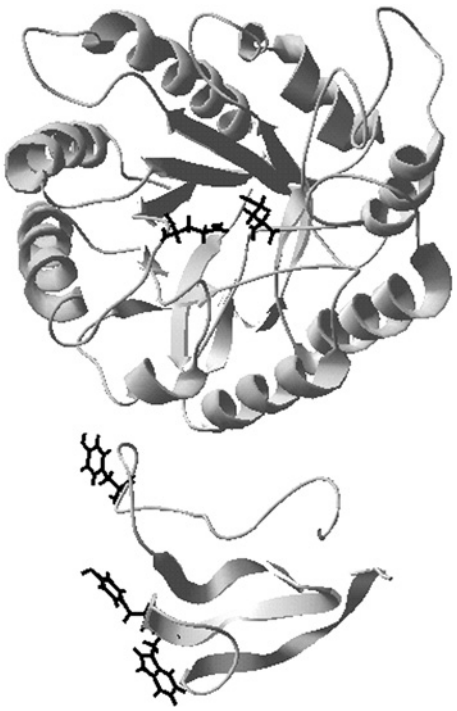

Figure 4 Structural model of the psychrophilic cellulase CelG

Upper panel, front view of the catalytic domain constructed by homology modelling using atomic co-ordinates of the Cel5 crystal structure as template. The catalytic residues Glu^{135} and Glu^{222} are in black. Lower panel, structural model of the CBM based on the atomic co-ordinates of the Cel5 module obtained by NMR. The residues Tyr^{418}, Trp^{443} and Tyr^{444} involved in cellulose binding are shown in black.

ES to reach the activated state ES* in order to react, therefore corresponding to a higher activity. However, a main determinant of the high activity at low temperatures of the psychrophilic cellulase is found in the lower enthalpic contribution, ΔH^*, to the free energy. Indeed, this low value originates from a weaker temperature-dependence of k_{cat}, as indicated by the weaker slope of the Arrhenius plot (Figure 3a, inset). Accordingly, the reduction of activity induced by low temperatures is less pronounced for the cold-active cellulase. It has been suggested that low ΔH^* values may translate a reduced number of enthalpy-driven interactions that are broken in the ES complex before reaching the activated state [6,19]. As these interactions also contribute to the active-site conformation, this should result in a heat-labile catalytic centre, as indeed noted for the psychrophilic cellulase. The higher and unfavourable entropic contribution, $T\Delta S^*$, to the free energy displayed by CelG (Table 2) can also be interpreted in this context. It has been proposed that a large entropic contribution may translate large conformational motions between a relatively disordered active site in the ES complex and the structured activated state ES* [6,19,29]. One can expect that such a heat-labile, disordered active site will bind its substrate weakly, as indeed suggested by the high K_m value of the psychrophilic cellulase. It is also worth mentioning that, in the classical scheme of the energy barrier, if the substrate binds weakly, then the ES complex falls in a less deep energy pit therefore decreasing the effective height of the energy barrier and leading to an increased activity [30,31].

Structural features of the cold-active cellulase

In an attempt to identify the structural factors potentially responsible for the specific properties of the psychrophilic cellulase, its structure was built by homology modelling. The model of the cellulase CelG was based on the three-dimensional structures of the cellulase Cel5 from *E. chrysanthemi* obtained by X-ray crystallography for the catalytic domain [25] and by NMR for the CBM [24]. No data are available for the linker in cellulases, except for a modelled shape obtained by small-angle X-ray scattering [10]. The overall fold of the catalytic domain is a $(\beta/\alpha)_8$ barrel, whereas the CBM exhibits a ski-boot-shape based mainly on a triple antiparallel β-sheet (Figure 4). More structural details can be found in the original descriptions of the templates [24,25]. In order to analyse the active site, a molecule of cellotetraose has been docked in the long substrate-binding cleft on the basis of the crystal structures of cellulases Cel5A from *B. agaradherans* and E1 from *A. cellulolyticus* in complex with cellobiose and cellotetraose respectively. In the psychrophilic cellulase, this substrate interacts with 20 residues, as indicated in Figure 1. Half of these residues are aromatic side chains that line up the cleft and perform stacking interactions with the four glucopyranose units of the substrate, a general feature in glycosidases [32]. All these residues are conserved (Figure 1) and superimpose perfectly with the structure of Cel5 from *E. chrysanthemi*, with the exception of Tyr^{204} in CelG which is a leucine residue in Cel5. No specific function has been assigned to this residue in other family

Table 3 Structural factors potentially involved in the weak thermal stability of the psychrophilic cellulase CelG

Parameters	Cel5	CelG	Expected effect
Catalytic domain			
Proline content	10	7	Loop mobility increase
Glycine content	22	28	Backbone mobility increase
Arginine content	10	6	Weak electrostatic interactions decrease
Salt bridges	13	9	Ionic interactions decrease
Arginine + lysine	30	18	Imbalance of surface charges
Solvent-accessible non-polar residues	19	14	Decreases hydrophobic effect on folding
Hydrophobic clusters	3(+2)*	5	Decreases hydrophobic effect on folding
Cavities	5	8	Folding compactness decrease
Volume of cavities (Å^3)	21	52	Folding compactness decrease
CBM			
Salt bridges	13	9	Ionic interactions decrease
Hydrophobic clusters	1	0	Decreases hydrophobic effect on folding
Hydrophobic core	(1)*	1	Decreases hydrophobic effect on folding
Cavities	0	1	Folding compactness decrease

* Weaker hydrophobicity resulting from substitution for polar side chain(s).

5 cellulases. There is therefore no obvious structural change within the active site that can be related to cold-activity in CelG.

In contrast, several structural factors that are potentially responsible for the low thermal stability of the psychrophilic cellulase are observed and are listed in Table 3. The catalytic domain of CelG lacks three proline residues in loops: this residue has only one variable dihedral angle and restricts the chain mobility between secondary structures. The domain possesses six additional glycine residues: the lack of side chain in glycine allows chain rotations and dihedral angles not available to other residues. The catalytic domain also lacks four arginine residues which are able to form up to five weak electrostatic interactions [33]. In combination with the decrease in lysine residues, the low number of basic residues has two consequences: the disappearance of four stabilizing ion pairs and an excess of negatively charged residues at the enzyme surface, also reinforced by seven additional exposed acidic residues. The latter factor is thought to destabilize the protein via improved interactions with the solvent according to the adhesive–cohesive model [34]. On the other hand, the catalytic domain of CelG exposes five additional hydrophobic residues to the solvent, an entropy-driven destabilizing factor [35]. The catalytic domains of CelG and Cel5 both contain five conserved hydrophobic core clusters. However, two of them in the psychrophilic cellulase are weakened by the replacement of a non-polar side chain for a polar one. Finally, both the number and cumulated volume of cavities in CelG appear to be higher, suggesting a less compact conformation of the catalytic domain. Most of these alterations in CelG are located near the N- and C-terminal extremities and on the $\alpha_2\beta_2$ loop of the barrel. Relaxed protein extremities have been proposed as preferential sites for unfolding [30]. The small CBM also displays some of these trends. All the above-mentioned factors have been implicated in protein stability by comparative studies of enzymes adapted to extreme temperatures [35–37].

Conclusions

The cold-active cellulase CelG displays a well-adapted ability to hydrolyse cellulose at low temperatures, as discussed above. Interestingly, the analysis of its active site with a docked substrate molecule fails to reveal structural alterations that can be related to its high k_{cat} and K_m values, because 19 residues forming the catalytic cleft are strictly conserved in its mesophilic counterpart from *E. chrysanthemi*. One can therefore propose that the properties of these conserved residues in substrate binding and hydrolysis are different in CelG. This implies that the protein structures outside the active site are responsible for cold-activity. Although the involvement of subtle, undetected mutations in CelG cannot be ruled out, several lines of evidence suggest that the loss of conformational stability provides the requested flexibility or mobility to the active site in order to perform catalysis at low temperatures [38–41]. In this respect, the structural model of CelG reveals several factors that are responsible for its low thermal stability (Table 3). It should be noted that the thermodynamic parameters of activation also point to a heat-labile, disordered active site with a low substrate affinity.

Moreover, a novel structural adaptation to catalysis at low temperatures, not previously observed in psychrophilic enzymes, can be proposed at the level of the linker region. With 107 residues, the linker of CelG is by far the largest observed in cellulases from family 5-2 and one of the largest in all cellulases, which generally have connecting regions ranging from 6 to 80 residues [21,42,43]. Considering the enzyme attached to the cellulose fibres via its CBM and a linker that can bend and extend [10], the size of the linker in CelG can considerably increase the surface of substrate available to the catalytic domain: assuming a rotation around the CBM, this surface is a square function of the linker length. Accordingly, the psychrophilic cellulase displays both a kinetic optimization of the activity and a steric optimization of substrate accessibility through its unusually long linker. Another unique feature of the linker is the presence of three DXDXDGXXDXXD motifs (Figure 1) found in calcium-binding type 3 repeats (T3) of thrombospondins. The latter are multimeric glycoproteins acting as extracellular regulators of cell–matrix interactions in animal cells [44]. If these motifs also bind calcium in the CelG linker, they can provide a structural basis for the conformational organization of the chain. Studies of the linker conformation are currently in progress.

We thank Dr Frédéric Barras and Dr Mirjam Czjzek (CNRS-Marseille, France) for providing the plasmid pSNAB1 and atomic co-ordinates of Cel5 before publication. This work was supported by the European Union, the Région Wallonne, Belgium and the Fonds National de la Recherche Scientifique, Belgium. The support of the Institut Polaire Français Paul Emile Victor is also acknowledged. We thank Nicole Gérardin, Christelle Hakin and René Marchand for expert technical assistance.

REFERENCES

1 Beguin, P. and Aubert, J. P. (1994) The biological degradation of cellulose. FEMS Microbiol. Rev. **13**, 25–58
2 Singh, A. and Hayashi, K. (1995) Microbial cellulases: protein architecture, molecular properties, and biosynthesis. Adv. Appl. Microbiol. **40**, 1–44
3 Tomme, P., Warren, R. A. and Gilkes, N. R. (1995) Cellulose hydrolysis by bacteria and fungi. Adv. Microb. Physiol. **37**, 1–81
4 Cavicchioli, R., Siddiqui, K. S., Andrews, D. and Sowers, K. R. (2002) Low-temperature extremophiles and their applications. Curr. Opin. Biotechnol. **13**, 253–261
5 Deming, J. W. (2002) Psychrophiles and polar regions. Curr. Opin. Microbiol. **5**, 301–309
6 Feller, G. and Gerday, C. (2003) Psychrophilic enzymes: hot topics in cold adaptation. Nat. Rev. Microbiol. **1**, 200–208
7 Shoham, Y., Lamed, R. and Bayer, E. A. (1999) The cellulosome concept as an efficient microbial strategy for the degradation of insoluble polysaccharides. Trends Microbiol. **7**, 275–281
8 Carrard, G., Koivula, A., Soderlund, H. and Beguin, P. (2000) Cellulose-binding domains promote hydrolysis of different sites on crystalline cellulose. Proc. Natl. Acad. Sci. U.S.A. **97**, 10342–10347
9 Jervis, E. J., Haynes, C. A. and Kilburn, D. G. (1997) Surface diffusion of cellulases and their isolated binding domains on cellulose. J. Biol. Chem. **272**, 24016–24023

10 Receveur, V., Czjzek, M., Schulein, 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

11 Bayer, E. A., Chanzy, H., Lamed, R. and Shoham, Y. (1998) Cellulose, cellulases and cellulosomes. Curr. Opin. Struct. Biol. **8**, 548–557

12 Henrissat, B. and Bairoch, A. (1996) Updating the sequence-based classification of glycosyl hydrolases. Biochem. J. **316**, 695–696

13 Tomme, P., Warren, R. A., Miller, Jr, R. C., Kilburn, D. G. and Gilkes, N. R. (1995) Cellulose-binding domains: classification and properties. In Enzymatic Degradation of Insoluble Polysaccarides (Saddler, J. N. and Penner, M., eds.), pp. 142–163, American Chemical Society, Washington

14 Feller, G., Lonhienne, T., Deroanne, C., Libioulle, C., Van Beeumen, J. and Gerday, C. (1992) Purification, characterization, and nucleotide sequence of the thermolabile α-amylase from the Antarctic psychrotroph *Alteromonas haloplanctis* A23. J. Biol. Chem. **267**, 5217–5221

15 Brun, E., Gans, P., Marion, D. and Barras, F. (1995) Overproduction, purification and characterization of the cellulose-binding domain of the *Erwinia chrysanthemi* secreted endoglucanase EGZ. Eur. J. Biochem. **231**, 142–148

16 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

17 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

18 Lonhienne, T., Baise, E., Feller, G., Bouriotis, V. and Gerday, C. (2001) Enzyme activity determination on macromolecular substrates by isothermal titration calorimetry: application to mesophilic and psychrophilic chitinases. Biochim. Biophys. Acta **1545**, 349–356

19 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

20 Feller, G., D'Amico, S., Benotmane, A. M., Joly, F., Van Beeumen, J. and Gerday, C. (1998) Characterization of the C-terminal propeptide involved in bacterial wall spanning of α-amylase from the psychrophile *Alteromonas haloplanctis*. J. Biol. Chem. **273**, 12109–12115

21 Linder, M. and Teeri, T. T. (1997) The roles and function of cellulose-binding domains. J. Biotechnol. **57**, 15–28

22 Barras, F., Bortoli-German, I., Bauzan, M., Rouvier, J., Gey, C., Heyraud, A. and Henrissat, B. (1992) Stereochemistry of the hydrolysis reaction catalyzed by endoglucanase Z from *Erwinia chrysanthemi*. FEBS Lett. **300**, 145–148

23 Bortoli-German, I., Haiech, J., Chippaux, M. and Barras, F. (1995) Informational suppression to investigate structural functional and evolutionary aspects of the *Erwinia chrysanthemi* cellulase EGZ. J. Mol. Biol. **246**, 82–94

24 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

25 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

26 Davies, G. and Henrissat, B. (1995) Structures and mechanisms of glycosyl hydrolases. Structure **3**, 853–859

27 Wang, Q., Tull, D., Meinke, A., Gilkes, N. R., Warren, R. A., Aebersold, R. and Withers, S. G. (1993) Glu^{280} is the nucleophile in the active site of *Clostridium thermocellum* CelC, a family A endo-β-1,4-glucanase. J. Biol. Chem. **268**, 14096–14102

28 Feller, G. (2003) Molecular adaptations to cold in psychrophilic enzymes. Cell. Mol. Life Sci. **60**, 648–662

29 Fields, P. A. and Somero, G. N. (1998) Hot spots in cold adaptation: localized increases in conformational flexibility in lactate dehydrogenase A4 orthologs of Antarctic notothenioid fishes. Proc. Natl. Acad. Sci. U.S.A. **95**, 11476–11481

30 Feller, G. and Gerday, C. (1997) Psychrophilic enzymes: molecular basis of cold adaptation. Cell. Mol. Life Sci. **53**, 830–841

31 Fersht, A. (1977) Enzyme Structure and Mechanism, Freeman and Co, New York

32 Vyas, N. K. (1991) Atomic features of protein–carbohydrate interactions. Curr. Opin. Struct. Biol. **1**, 732–740

33 Borders, Jr, C. L., Broadwater, J. A., Bekeny, P. A., Salmon, J. E., Lee, A. S., Eldridge, A. M. and Pett, V. B. (1994) A structural role for arginine in proteins: multiple hydrogen bonds to backbone carbonyl oxygens. Protein Sci. **3**, 541–548

34 Dadarlat, V. M. and Post, C. B. (2003) Adhesive–cohesive model for protein compressibility: an alternative perspective on stability. Proc. Natl. Acad. Sci. U.S.A. **100**, 14778–14783

35 Bell, G. S., Russell, R. J., Connaris, H., Hough, D. W., Danson, M. J. and Taylor, G. L. (2002) Stepwise adaptations of citrate synthase to survival at life's extremes: from psychrophile to hyperthermophile. Eur. J. Biochem. **269**, 6250–6260

36 Smalås, A. O., Leiros, H. K., Os, V. and Willassen, N. P. (2000) Cold adapted enzymes. Biotechnol. Annu. Rev. **6**, 1–57

37 Gianese, G., Bossa, F. and Pascarella, S. (2002) Comparative structural analysis of psychrophilic and meso- and thermophilic enzymes. Proteins **47**, 236–249

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

39 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

40 Georlette, D., Damien, B., Blaise, V., Depiereux, E., Uversky, V. N., Gerday, C. and Feller, G. (2003) Structural and functional adaptations to extreme temperatures in psychrophilic, mesophilic, and thermophilic DNA ligases. J. Biol. Chem. **278**, 37015–37023

41 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

42 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

43 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

44 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

Received 1 March 2004/23 July 2004; accepted 3 August 2004
Published as BJ Immediate Publication 3 August 2004, DOI 10.1042/BJ20040325