Calcium-induced tertiary structure modifications of endo-β-1,3-glucanase from *Pyrococcus furiosus* in 7.9 M guanidinium chloride

Roberta CHIARALUCE*[1], Giulio GIANESE*[1], Sebastiana ANGELACCIO*, Rita FLORIO*, Johan F. T. VAN LIESHOUT†, John VAN DER OOST† and Valerio CONSALVI*[2]

*Dipartimento di Scienze Biochimiche 'A Rossi Fanelli', Università 'La Sapienza', P.le A. Moro 5, 00185 Rome, Italy, and †Laboratory of Microbiology, Wageningen University, H. van Suchtelenweg 4, 6703 CT Wageningen, The Netherlands

The family 16 endo-β-1,3 glucanase from the extremophilic archaeon *Pyrococcus furiosus* is a laminarinase, which in 7.9 M GdmCl (guanidinium chloride) maintains a significant amount of tertiary structure without any change of secondary structure. The addition of calcium to the enzyme in 7.9 M GdmCl causes significant changes to the near-UV CD and fluorescence spectra, suggesting a notable increase in the tertiary structure which leads to a state comparable, but not identical, to the native state. The capability to interact with calcium in 7.9 M GdmCl with a consistent recovery of native tertiary structure is a unique property of this extremely stable endo-β-1,3 glucanase. The effect of calcium on the thermodynamic parameters relative to the GdmCl-induced equilibrium unfolding has been analysed by CD and fluorescence spectroscopy. The interaction of calcium with the native form of the enzyme is studied by Fourier-transform infrared spectroscopy in the absorption region of carboxylate groups and by titration in the presence of a chromophoric chelator. A homology-based model of the enzyme is generated and used to predict the putative binding site(s) for calcium and the structural interactions potentially responsible for the unusual stability of this protein, in comparison with other family 16 glycoside hydrolases.

Key words: calcium, endo-β-1,3-glucanase, guanidinium chloride, homology modelling, protein stability.

INTRODUCTION

Laminarinase endo-β-1,3 glucanase (EC 3.2.1.39) from the hyperthermophilic archaeon *Pyrococcus furiosus* (*pf*LamA) maintains residual tertiary structure and intact secondary structure elements in 7.9 M GdmCl (guanidinium chloride) [1]. The presence of residual tertiary interactions under such extreme denaturing conditions is a peculiar property of this protein, which makes it an attractive model to investigate the role played by non-covalent interactions as structural determinants of protein stability.

*pf*LamA belongs to family 16 glycoside hydrolases [2] (see http://afmb.cnrs-mrs.fr/~cazy/CAZY/index.html). Enzymes in this family show a number of catalytic activities and are characterized by a similar β-jelly roll fold, which is better conserved than the primary structure. A similar fold has also been demonstrated in proteins that have been classified in different families of glycoside hydrolases [3]. A considerable number of primary structures of family 16 are available, but among the few crystal structures solved, detailed structural information on the laminarinase subfamily is lacking [4].

The presence of at least one metal-binding site is one of the common features in the crystal structures of family 16 glycoside hydrolases [2,4] and calcium has been reported to protect *pf*LamA from heat-induced inactivation [5]. Indeed, previous sequence comparisons of the deduced *pf*LamA primary structure have revealed that those residues presumably involved in metal binding are conserved, in particular Asp-287 that is present in most of the family 16 enzymes, as well as residues Glu-170 and Glu-175 that are involved in catalysis (*pf*LamA amino acid numbering used throughout manuscript) [5]. However, despite the sequence similarities with all the enzymes of family 16, the unusual resistance to structural loss in 7.9 M GdmCl is a peculiar property of *pf*LamA. To gain more information about the possible structural determinants of the unusual stability of *pf*LamA, we performed multiple sequence alignments by applying Hidden Markov Model using the SAM-T02 web pages [6]. *pf*LamA was found to be closely related to the mesophilic κ-carrageenase from *Pseudoalteromonas carrageenovora* (*pc*Car) (EC 3.2.1.83), a family 16 glycoside hydrolase enzyme which cleaves the internal β $(1 \rightarrow 4)$ linkages of carrageenans, a linear polymer of galactopyranose residues linked by alternating α $(1 \rightarrow 3)$ and β $(1 \rightarrow 4)$ linkages, and whose crystal structure has been solved at 1.54 Å (1 Å = 0.1 nm) resolution [2,7]. The homology to *pc*Car allowed us to use it as a template for molecular modelling of the *pf*LamA structure. Interestingly, κ-carrageenase in crystals binds seven cadmium ions, indicating that this fold has several potential binding sites for metals [7]. In the present study, we describe that in 7.9 M GdmCl *pf*LamA interacts with calcium with a consistent regain of tertiary structure, as indicated by near-UV CD and fluorescence spectra, without any change in the far-UV CD spectrum. The interaction of calcium with the native form of *pf*LamA becomes evident by FTIR (Fourier-transform infrared) spectroscopy in the absorption region of carboxylate groups. We analysed the binding of the metal to the native enzyme in the presence of a chromophoric chelator, and present a structural model of the protein that reveals potential binding sites for metals and that may provide clues about the structural interactions potentially responsible for the extreme stability of *pf*LamA.

EXPERIMENTAL

Data collection

The three-dimensional crystal structures of a lichenase, the hybrid 1,3-1,4-β-D-glucanase from *Bacillus amyloliquefaciens* and

Abbreviations used: ANS, 8-anilinonaphthalene-1-sulphonic acid; ATR, attenuated total reflectance; BAPTA, 1,2-bis(o-aminophenoxy)ethane-*N*,*N*,*N*′,*N*′-tetraacetic acid; DTT, 1,4-dithio-DL-threitol; FTIR, Fourier-transform infrared; GdmCl, guanidinium chloride; SVD, singular value decomposition algorithm.

[1] These authors have contributed equally to this work.

[2] To whom correspondence should be addressed (email consalvi@caspur.it).

B. macerans (*b*1,3-1,4Glc; Protein Data Bank code: 2AYH), and of the catalytic domain of *pc*Car (PDB code: 1DYP) were taken from the Brookhaven PDB [8]. The two structures were superimposed with the program CE (combinatorial extension) [9] and a sequence alignment was derived from the structural consensus. The alignment was manually corrected to optimize the arrangement of insertions and deletions. The sequence of *pf*LamA was then lined up with the two sequences of the structural alignment, by the program ClustalW [10]. The alignment obtained was manually modified to take into account the structural information such as the observed and predicted secondary structural elements, the structurally and functionally conserved amino acids, and the position of insertions and deletions in the structures. The program HOMOLOGY in the package InsightII [11] was used for the processing of alignments and structures.

A search for β-glucanase structures displaying homology with *pf*LamA for a comparative analysis was conducted in the PDB using PSI-BLAST [12] with two iterations. All the hybrid and circularly permuted structures were rejected. The final selection consisted of two lichenases: the 1,3-1,4-β-glucanases from *B. licheniformis* (*bl*1,3-1,4Glc; PDB code: 1GBG) and from *B. macerans* (*bm*1,3-1,4Glc; PDB code: 1MAC).

Model building

The three-dimensional structures of *b*1,3-1,4Glc and *pc*Car were used as templates for the construction of a homology-derived model of *pf*LamA. Homology modelling was based on the multiple sequence alignment constructed as described in the above subsection. Protein models were built with the MODELLER-4 package [13]. Ten different models at the highest optimization level were built for the target protein. The model displaying the lowest 'objective function' value, which measures the violation of constraints from the template structures, was selected [14]. Construction of slightly different models of the same protein structure can be used as an indicator of the most variable and, therefore, less reliable regions in the folding. The calcium ion was taken from the *b*1,3-1,4Glc structure and included in the model as a rigid body. The quality of the final model was assessed with the programs ProsaII [15] and PROCHECK [16].

Contact surface area for apolar atoms was calculated with the program PDB_NP_CONT [17], which computes pairwise atom contact areas between apolar atoms from structural data. The program is based on classification of points located on an interaction sphere around each atom and used a method similar to that described by Connolly [18]. The total contact surface was then normalized by the number of residues of the protein considered.

Ion pairs

Ion pair interactions were determined using the program WHAT IF [19]. Two atoms of opposite charge separated by a distance less than a defined threshold are defined as an ion pair. The distance limits of 4.0 Å, the value usually accepted [20], and 6.0 Å were selected to identify 'strong' and 'weak' ion pairs respectively. Atoms with positive charge were considered the side-chain nitrogens in arginine and lysine. Atoms with negative charge were the side-chain oxygens of aspartic and glutamic residues. Results with and without histidine residues are reported for completeness; however, the assignment of protonation state of such a residue in a protein may be difficult.

Chemicals and buffers

GdmCl, ANS (8-anilinonaphthalene-1-sulphonic acid) ammonium salt, DTT (1,4-dithio-DL-threitol), EDTA and laminarin were from Fluka (Zwijndrecht, The Netherlands). 3′,5′-Dinitrosalicylic acid was purchased from Sigma. 5,5′-Br_2-BAPTA [where BAPTA stands for 1,2-bis(o-aminophenoxy)ethane-*N*,*N*,*N*′,*N*′-tetraacetic acid] was from Molecular Probes Europe BV (Leiden, The Netherlands). Buffer solutions were filtered (0.22 μm) and carefully degassed. All buffers and solutions were prepared with ultra-high quality water (ELGA UHQ, Veolia Water Systems, High Wycombe, U.K.). Buffers for calcium titrations were prepared as described in [21].

Enzyme preparation and assay

*pf*LamA was functionally produced in *Escherichia coli* BL21-(DE3) with pLUW532 and purified according to Kaper et al. [22]. The protein concentration was determined at 280 nm using a $\varepsilon_{280} = 83\,070\ M^{-1} \cdot cm^{-1}$ calculated according to Gill and von Hippel [23]. Enzyme activity was determined by measuring the amount of reducing sugars released on incubation in 0.1 M sodium phosphate buffer (pH 6.5) containing 5 mg/ml laminarin, at 60 or 80 °C for 10 min, as described by Kaper et al. [22]. Calcium-depleted protein was obtained by extensive dialysis with 100 μM EDTA and 100 μM EGTA in 10 mM Tris/HCl (pH 7.4). All the precautions required to prevent Ca^{2+} contamination were followed during the preparation and storage of protein and buffer solutions [21]. Calcium-loaded protein refers to the protein in the presence of 40 mM $CaCl_2$.

Spectroscopic techniques

Intrinsic fluorescence emission and 90° light scattering measurements were performed with a LS50B PerkinElmer spectrofluorimeter using a 1.0 cm pathlength quartz cuvette. Fluorescence emission spectra were recorded at 300–450 nm (1 nm sampling interval) at 20 °C with the excitation wavelength set at 290 nm. To check for the presence of aggregated particles, 90° light scattering was measured at 20 °C with both excitation and emission wavelength set at 480 nm.

Far-UV (180–250 nm) and near-UV (250–320 nm) CD measurements were performed at 20 °C in 0.1–0.2 cm and 1.0 cm pathlength quartz cuvettes respectively. CD spectra were recorded on a Jasco J-720 spectropolarimeter. The results are expressed as the mean residue ellipticity [Θ] assuming a mean residue weight of 110 per amino acid residue. In all the spectroscopic measurements at pH 7.4, 100–250 μM EDTA was always present unless otherwise stated.

FTIR spectra were recorded on a Nicolet Magna 760 spectrometer (Thermo Nicolet, Madison, WI, U.S.A.) equipped with a liquid nitrogen-cooled mercury–cadmium telluride solid-state detector. ATR (attenuated total reflectance) spectra in solution were measured in a CIRCLE cell (Spectra Tech, Madison, WI, U.S.A.) at 20 °C. Protein samples (130 μl) of an 8 mg/ml protein solution in 20 mM Tris/HCl at pH 7.4, in the absence or presence of 40 mM $CaCl_2$, were placed in the CIRCLE cell with a ZnSe crystal rod. For exchange of amide protons with deuterons, *pf*LamA was freeze-dried and dissolved twice in ^2H_2O buffer at 20 °C before measurements. A total of 512 interferograms at 2 cm^{-1} resolution were collected for each spectrum, with Mertz apodization and two levels of zero filling. The sample chamber of the spectrometer was continuously purged with dry air to avoid water vapour interference on the bands of interest. The background spectra were collected immediately before the sample measurements and under the same conditions with the cell filled with everything but protein. At the end of the measurements, after prolonged washing, a spectrum of the cell was recorded to check for protein adsorption on the crystal rod [24]. Water vapour spectra were collected by reduction of the dry-air purge of the clean cell.

Difference spectra were normalized to an equal protein content by normalizing spectra measured in H_2O to an amide II absorbance of 0.04 (difference in absorbance between 1545 and 1490 cm^{-1}) and spectra measured in ^2H_2O to an amide I′ absorbance of 0.06 (difference in absorbance between 1706 and 1636 cm^{-1}). The analysis of raw spectra was performed with GRAMS.

Experiments with the fluorescent dye ANS were performed at 20 °C by incubating the protein and ANS at a 1:10 molar ratio. After 5 min, fluorescence emission spectra were recorded at 400–600 nm with the excitation wavelength set at 390 nm. The maximum fluorescence emission wavelength and the intensity of the hydrophobic probe ANS depend on the environmental polarity, e.g. on the hydrophobicity of the accessible surface of the protein [25]. Fluorescence quenching was performed by adding increasing amounts of acrylamide (0–100 mM) to *pf*LamA solution (40 μg/ml) in 7.9 M GdmCl in the absence or presence of 40 mM $CaCl_2$. Emission spectra (300–450) were recorded at 20 °C, 10 min after each acrylamide addition with the excitation wavelength set at 290 nm. The effective quenching constants were obtained from modified Stern–Vollmer plots by analysing $F_0/\Delta F$ versus 1/[acrylamide] (25 data points) [26].

GdmCl-induced unfolding and refolding

For equilibrium transition studies, *pf*LamA (final concentration 40–50 μg/ml) was incubated at 20 °C at increasing concentrations of GdmCl (0–8 M) in 20 mM Tris/HCl (pH 7.4), containing 100 μM DTT and 100 μM EDTA and, when indicated, 40 mM $CaCl_2$. After 24 h equilibrium was reached and intrinsic fluorescence emission and far-UV CD spectra (0.2-cm cuvette) were recorded in parallel at 20 °C. To test the reversibility of the unfolding, *pf*LamA was unfolded at 20 °C in 7.8 M GdmCl at 0.8 mg/ml protein concentration in 25 mM Tris/HCl (pH 7.4), containing 100 μM DTT and 100 μM EDTA, in the presence and absence of 40 mM $CaCl_2$. After 24 h, refolding was started by 20-fold dilution of the unfolding mixture, at 20 °C, into solutions of the same buffer used for unfolding containing decreasing GdmCl concentrations. The final enzyme concentration was 40 μg/ml. After 24 h, a time that was established to be sufficient to reach equilibrium, intrinsic fluorescence emission and far-UV CD spectra were recorded at 20 °C.

Data analysis

Far- and near-UV CD spectra from GdmCl and Ca^{2+} titrations were analysed by the SVD (singular value decomposition algorithm) [1,27] using the software MATLAB (MathWorks, South Natick, MA, U.S.A.). SVD is useful to find the number of independent components in a set of spectra and to remove the high-frequency noise and the low-frequency random error. CD spectra in the 210–250 nm region or in the 250–310 nm region (0.2 nm sampling interval) were placed in a rectangular matrix **A** of *n* columns, one column for each spectrum collected in the titration. The **A** matrix is decomposed by SVD into the product of three matrices: $\mathbf{A} = \mathbf{USV^T}$, where **U** and **V** are orthogonal matrices and **S** is a diagonal matrix. The columns of **U** matrix contain the basis spectra and the columns of the **V** matrix contain the denaturant or the Ca^{2+} dependence of each basis spectrum. Both **U** and **V** columns are arranged in terms of decreasing order of the relative weight of information, as indicated by the magnitude of the singular values in **S**. The diagonal **S** matrix contains the singular values that quantify the relative importance of each vector in **U** and **V**. The signal-to-noise ratio is very high in the earliest columns of **U** and **V** and the random noise is mainly accumulated in the latest **U** and **V** columns. The wavelength averaged spectral changes induced by increasing denaturant or Ca^{2+} concentrations are represented by the columns of matrix **V**, hence the plot of the columns of **V** versus the denaturant or Ca^{2+} concentrations provides information about the observed transition.

GdmCl-induced equilibrium unfolding was analysed by fitting baseline and transition region data to a two-state linear extrapolation model [28] according to

$$\Delta G_{\text{unfolding}} = \Delta G^{H_2O} + m_g\,[\text{GdmCl}] = -RT\ln K_{\text{unfolding}} \quad (1)$$

where $\Delta G_{\text{unfolding}}$ is the free energy change for unfolding for a given denaturant concentration, ΔG^{H_2O} the free energy change for unfolding in the absence of denaturant and m_g a slope term which quantifies the change in $\Delta G_{\text{unfolding}}$ per unit concentration of denaturant, R the gas constant, T the temperature and $K_{\text{unfolding}}$ the equilibrium constant for unfolding. The model expresses the signal as a function of denaturant concentration:

$$y_i = \frac{y_N + m_N[X]_i + (y_D + m_D[X]_i)^* \exp[(-\Delta G^{H_2O} - m_g[X]_i)/RT]}{1 + \exp[(-\Delta G^{H_2O} - m_g[X]_i)/RT]} \quad (2)$$

where y_i is the observed signal, y_N and y_D are the native and denatured baseline intercepts, m_N and m_D are the native and denatured baseline slopes, $[X]_i$ the denaturant concentration after the ith addition, ΔG^{H_2O} the extrapolated free energy of unfolding in the absence of denaturant, m_g the slope of a G unfolding versus [X] plot. The $[\text{GdmCl}]_{0.5}$ is the denaturant concentration at the midpoint of the transition and, according to eqn (1), is calculated as:

$$[\text{GdmCl}]_{0.5} = \Delta G^{H_2O}/m_g \quad (3)$$

Calcium titrations and determination of binding constant

Calcium-depleted native *pf*LamA (9–16 μM) was titrated with $CaCl_2$ in the presence of 24 μM of the chromophoric chelator 5,5′-Br_2-BAPTA [29]. 5,5′-Br_2-BAPTA concentration was determined by measuring the absorbance A_{263} using $\varepsilon_{239.5} = 1.6 \times 10^4\ M^{-1} \cdot cm^{-1}$ [21]. Titrations were performed at 20 °C in 10 mM Tris/HCl (pH 7.5) by the addition of 1–2 μl of $CaCl_2$ solutions ranging from 0.015 to 20.0 mM to a 1-ml protein solution containing 24 μM 5,5′-Br_2-BAPTA. Absorbance spectra were monitored between 200 and 450 nm after each Ca^{2+} addition. To determine protein binding constants and number of binding sites, the variation of 5,5′-Br_2-BAPTA absorbance as a function of calcium addition was fitted by non-linear analysis using the CaLigator software [21]. The quantity χ^2 calculated by the program was used as the measure of the goodness-of-fit. All the precautions required to prevent Ca^{2+} contamination were followed during the preparation and storage of protein and buffer solutions [21].

Calcium titration of calcium-depleted *pf*LamA in 7.9 M GdmCl (25 mM Tris/HCl, pH 7.4 containing 200 μM DTT and 250 μM EDTA) was performed by the addition of increasing $CaCl_2$ concentrations (0–40 mM) under continuous stirring. Near-UV CD (240–320 nm, 22 μM protein concentration) and fluorescence (300–450 nm, 1.2 μM protein concentration) spectra were recorded 5 min after each $CaCl_2$ addition at 20 °C. The spectral changes observed after each $CaCl_2$ addition were not affected by longer incubation time. The concentration of unchelated Ca^{2+} was calculated by using the program WinMaxc version 2.40 [30] (see http://www.stanford.edu/~cpatton/maxc.html).

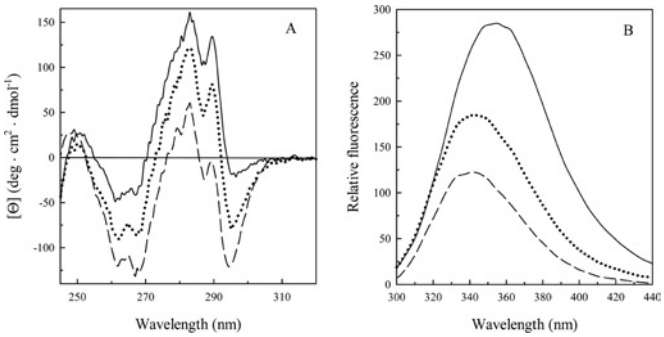

Figure 1 Effect of calcium on the spectral properties of *pf*LamA in 7.9 M GdmCl

(**A**) Near-UV CD spectra were recorded in a 1 cm quartz cuvette at 0.6 mg/ml protein concentration and (**B**) fluorescence spectra were recorded at 40 μg/ml protein concentration (290 nm excitation wavelength). All the spectra were recorded at 20 °C after 24 h incubation of the protein in 20 mM Tris/HCl at pH 7.4 (- - - - -) and in 7.9 M GdmCl at pH 7.4 in the absence (——) and presence of 40 mM $CaCl_2$ (. . . . . .).

RESULTS

Effect of calcium on *pf*LamA in 7.9 M GdmCl

Addition of $CaCl_2$ to calcium-depleted *pf*LamA in 7.9 M GdmCl, 25 mM Tris/HCl (pH 7.4) containing 200 μM DTT, 250 μM EDTA, at 20 °C, causes significant changes in the enzyme tertiary structure, as indicated by near-UV CD and intrinsic fluorescence emission spectra (Figures 1A and 1B). At 40 mM $CaCl_2$, the 295 nm band of tryptophan residue, which is significantly reduced in 7.9 M GdmCl, is restored with a concomitant regain of the negative ellipticity signal and of the fine structure in the 260–270 nm region (Figure 1A). The regain in aromatic chirality is accompanied by a 1.5-fold decrease in fluorescence intensity at 342 nm and a blue shift from 357 to 343 nm in the maximum fluorescence emission wavelength, a value close to that measured at 342 nm for the native state (Figure 1B). The far-UV CD spectrum in 7.9 M GdmCl, which is the same as that measured in the absence of denaturant [1], is not affected by the addition of $CaCl_2$ (results not shown). A titration of the enzyme in 7.9 M GdmCl with increasing amounts of $CaCl_2$, from 0.2 nM to 35 mM unchelated Ca^{2+}, and the analysis of the aromatic CD spectral changes at 295 nm indicates that, above 120 μM of unchelated Ca^{2+} concentration, no further changes are observed (Figure 2). The near-UV CD ellipticity changes at 295 nm induced by increasing $CaCl_2$ concentration (Figure 2) were analysed after removal of the high-frequency noise and the low-frequency random error by SVD. The global changes in the spectral region 250–310 nm were analysed by SVD that indicates that only two spectral components contribute to the near-UV CD spectra. The most significant singular values are 4.7×10^3, 1.0×10^3 and 0.1×10^3. All the other singular values are well below 10 % of the largest singular value and progressively decrease approaching zero. The first and the second columns of the **V** matrix ($\mathbf{V_1}$ and $\mathbf{V_2}$) show a similar dependence on increasing Ca^{2+} concentration and both confirm that saturation occurs above 120 μM (results not shown). The $[\Theta]_{295}$ data were analysed by non-linear regression analysis to define two limiting slopes, intersecting at a value which suggests that essentially 2 mol of Ca^{2+} per mol of enzyme is necessary to reach an apparent saturation effect (Figure 2). The intrinsic fluorescence emission quenching at 342 nm measured on unchelated Ca^{2+} addition from 0.2 nM to 35.0 mM remains unchanged above 120 μM Ca^{2+}, similar to that observed for the near-UV CD ellipticity changes (Figure 2). The dependence of

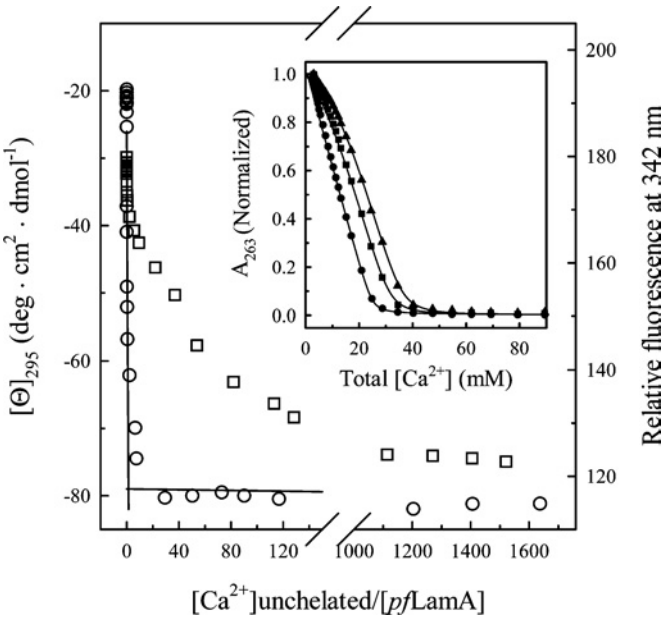

Figure 2 Interaction of calcium with *pf*LamA

$[\Theta]_{295}$ (○, 22 μM *pf*LamA in 7.9 M GdmCl) and fluorescence intensity at 342 nm (□, 1.2 μM *pf*LamA in 7.9 M GdmCl) were measured from near-UV CD and fluorescence spectra (290 nm excitation wavelength) recorded at 20 °C, 5 min after each Ca^{2+} addition. $[\Theta]_{295}$ is reported after removal of the high-frequency noise and the low-frequency random error by SVD. The two limiting slopes, calculated by non-linear regression analysis to the $[\Theta]_{295}$ data at 22 μM *pf*LamA, intersect at a point corresponding to [Ca^{2+} unchelated]/[*pf*LamA] = 2. The reported unchelated Ca^{2+} concentrations, calculated according to [30], are 0.2 nM–35.0 mM and 0.2 nM–1.82 mM for $[\Theta]_{295}$ and fluorescence changes respectively. The inset shows a plot of the decrease in 5,5′-Br_2-BAPTA A_{263} towards [Ca^{2+}] in the absence (●) and presence of 10 μM (■) and 20 μM (▲) native *pf*LamA in 10 mM Tris/HCl (pH 7.5) at 20 °C. The solid lines represent fitting of the data by non-linear analysis using the CaLigator software [21].

the spectral changes on increasing Ca^{2+} concentration is hyperbolic, suggesting that binding of two Ca^{2+} essentially produces the same signal changes either on the decrease of the intrinsic fluorescence emission at 342 nm or on the regain of negative ellipticity at 295 nm. A comparable result is obtained from the plot of the changes of the maximum fluorescence emission wavelength (results not shown).

Effect of calcium on the native *pf*LamA

The near- and far-UV CD spectra, as well as the intrinsic fluorescence emission spectrum of the calcium-depleted native *pf*LamA, in 25 mM Tris/HCl (pH 7.4) containing 200 μM DTT and 250 μM EDTA, are not affected by the addition of $CaCl_2$ up to 0.5 M (results not shown). The interaction of calcium with the native enzyme was therefore studied by titration of *pf*LamA with $CaCl_2$ in the presence of the chromophoric chelator 5,5′-Br_2-BAPTA. The effect of calcium chelation on the absorbance spectra of 5,5′-Br_2-BAPTA is characterized by a decrease in the A_{263} accompanied by an increase in absorbance at approx. 239 nm, with an isosbestic point at 247 nm [29]. A plot of the decrease in 5,5′-Br_2-BAPTA A_{263} towards [Ca^{2+}] gives a titration profile, progressively shifted to the right in the presence of increasing *pf*LamA concentration (Figure 2, inset). This indicates the affinity of the native protein for the cation that was quantitatively analysed by CaLigator software [21]. At all protein concentrations tested, best fits of the absorbance data ($\chi^2 = 9.8 \times 10^{-5}$) were obtained with two calcium-binding sites and two corresponding binding constants of 5.0×10^7 and 2.6×10^5 M^{-1} respectively were found.

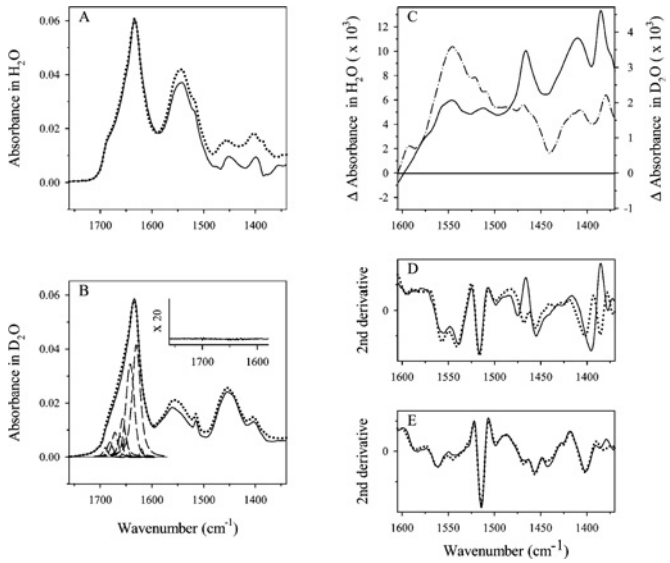

Figure 3 Effect of calcium on the IR spectra of *pf*LamA

Solution ATR-FTIR spectra of 8 mg/ml *pf*LamA were measured at 20 °C in 20 mM Tris/HCl (pH 7.4) in a CIRCLE cell with a ZnSe crystal rod. A total of 512 interferograms at 2 cm^{-1} resolution were collected for each spectrum. Absorbance spectra of amide I and II in H_2O (**A**) and in ^2H_2O (**B**) in the absence (——) and in the presence of 40 mM $CaCl_2$ (· · · · · ·). Deconvolution of the amide I′ spectrum was used to identify the individual components in the amide I′ region by curve-fitting of the raw spectrum (**B**). Inset in (**B**) shows the difference between the fitted curve and the original expanded 20 times. (**C**) Difference spectra in H_2O (——, left axis) and in ^2H_2O (- · - · - · -, right axis) were obtained by subtracting the spectrum recorded in the absence of calcium from that in the presence of 40 mM $CaCl_2$. (**D**, **E**) Second derivative of the raw spectra in the amide II region measured in H_2O (**D**) and in ^2H_2O (**E**) in the absence (——) and in the presence of 40 mM $CaCl_2$ (· · · · · ·).

Fitting the data according to only one calcium-binding site gave a single binding constant of 3.3×10^7 M^{-1} and a 3-fold higher χ^2 value.

The effect of calcium on native *pf*LamA was also investigated by FTIR spectroscopy, in H_2O and ^2H_2O, in the amide I, amide II and in the spectral region where the side chains of aspartic and glutamic residues are known to make contributions [31,32] (Figure 3). In the presence of 40 mM Ca^{2+}, FTIR absorbance spectra in H_2O show an amide I region almost identical with that observed in the absence of the cation, with the same maximum at 1634 cm^{-1}, and a shift of the amide II maximum from 1542 to 1544 cm^{-1}, which is accompanied by an increase in the amide II intensity (Figure 3A). In ^2H_2O, FTIR absorbance spectra of *pf*LamA in the presence and absence of calcium show a maximum centred at 1632 cm^{-1} and a band at 1560 cm^{-1} (Figure 3B). Similar to what was observed in H_2O, the presence of the cation does not significantly affect the amide I′ region, in agreement with the lack of any measurable change induced by the cation on the far-UV CD spectrum. In ^2H_2O, a decrease in the intensity in the amide II region and a shift of the amide II to 1455 cm^{-1} are also observed (Figure 3B), similar to what was reported for other proteins on deuteration [33,34]. However, in the presence of Ca^{2+} the decrease in intensity of the amide II′ region at 1544 cm^{-1} observed in ^2H_2O is less pronounced (Figures 3B and 3C), thus suggesting a reduced hydrogen–deuterium exchange indicative of a more compact structure [35].

The region of antisymmetric stretching vibration ν_{as} COO^- mode of the unprotonated carboxyl groups (1595–1550 cm^{-1}) and the region 1430–1390 cm^{-1} corresponding to the symmetric stretching vibration ν_s COO^- mode of the unprotonated carboxyl groups were explored [32,36]. FTIR difference absorption spectra were obtained by subtracting the absorption spectrum of the calcium-depleted from that of the calcium-loaded form of *pf*LamA. Hence positive peaks in the difference spectrum correspond to an increase in the intensities of the absorption spectrum on calcium binding, and vice versa (Figure 3C). In ^2H_2O, the binding of calcium gives rise to a main positive band centred at 1547 cm^{-1} and a minor peak at 1594 cm^{-1} (Figure 3C). In H_2O, the difference spectrum is comparable with that obtained in ^2H_2O, with a main positive peak centred at 1550 cm^{-1}, accompanied by a shoulder at approx. 1575 cm^{-1} (Figure 3C). The regions of the ν_s COO^- mode in H_2O and in ^2H_2O are also very similar, with a main positive peak at 1408 and at 1410 cm^{-1} respectively. Apart from the ν_{as} and ν_s COO^- (1595–1550 and 1430–1390 cm^{-1}), in H_2O two positive peaks are evident at 1466 and at 1385 cm^{-1} similar to the large shoulder at approx. 1468 cm^{-1} and to the positive peak at 1380 cm^{-1} observed in ^2H_2O (Figure 3C). Above 1610 cm^{-1}, in the amide I and amide I′ regions, both the difference spectra are similar, without any significant change in the presence of calcium, thus confirming the lack of changes in the secondary structure elements induced by the cation (results not shown). The 1800–1700 cm^{-1} region, corresponding to the C=O mode of protonated carboxyl groups, is not affected by the presence of calcium (results not shown).

Resolution enhancement by second derivative analysis of the FTIR spectra confirms the changes observed in the difference spectra (Figures 3D and 3E). Below the Tyr ring mode, which in ^2H_2O is downshifted from 1516 to 1514 cm^{-1}, the effect of calcium is evident, particularly in H_2O, as an upshift from 1396 to 1403 cm^{-1}, which is comparable with the upshift from 1400 to 1403 cm^{-1} observed in ^2H_2O. In the ν_{as} (1595–1550 cm^{-1}), the frequencies of the bands are not shifted by the presence of calcium, neither in H_2O nor in ^2H_2O, and an increased resolution of the band in the 1575–1550 cm^{-1} interval is observed in H_2O. In the presence of calcium, changes are also observed in frequency regions apparently not directly related to the ν_{as} and ν_s COO^-, such as a downshift from 1475 to 1468 cm^{-1} and an upshift from 1377 to 1386 cm^{-1} in H_2O, and a downshift from 1385 to 1382 cm^{-1} in ^2H_2O.

Effect of calcium on the equilibrium transition in GdmCl

Incubation of *pf*LamA at increasing GdmCl concentrations (0–8 M) in 20 mM Tris/HCl (pH 7.4) containing 100 μM DTT, 100 μM EDTA and 40 mM $CaCl_2$ for 20 h at 20 °C results in a progressive increase in the intrinsic fluorescence emission intensity at 342 nm (Figure 4), with a fluorescence maximum emission wavelength still centred at approx. 342 nm (Figure 4, inset). The process follows a two-state mechanism without any detectable intermediates, as indicated by the sigmoidal transition profiles shown in Figure 4, similar to that observed in the absence of Ca^{2+}. A plot of the relative fluorescence intensity at 342 nm as a function of GdmCl concentration shows a transition midpoint at 6.0 M GdmCl in the presence of Ca^{2+} and at 6.7 M GdmCl in the absence of Ca^{2+}. The enzyme, which is inactive in 7.9 M GdmCl in the presence and absence of $CaCl_2$, was fully reactivated on dilution of the denaturant. The fluorescence changes were reversible and the ΔG^{H_2O} and m_g values were calculated by non-linear regression fitting of the data reported in Figure 4, according to eqns (2) and (3). The ΔG^{H_2O} and m_g values were 41.8 kJ/mol and 6.99 kJ · mol^{-1} · M^{-1} and 61.5 kJ/mol and 9.20 kJ · mol^{-1} · M^{-1} in the presence and absence of Ca^{2+} respectively, suggesting an apparent decrease in protein stability in the presence of the cation. In the presence of Ca^{2+}, the analysis of *pf*LamA far-UV CD spectra in the 210–250 nm region revealed no changes on increasing GdmCl

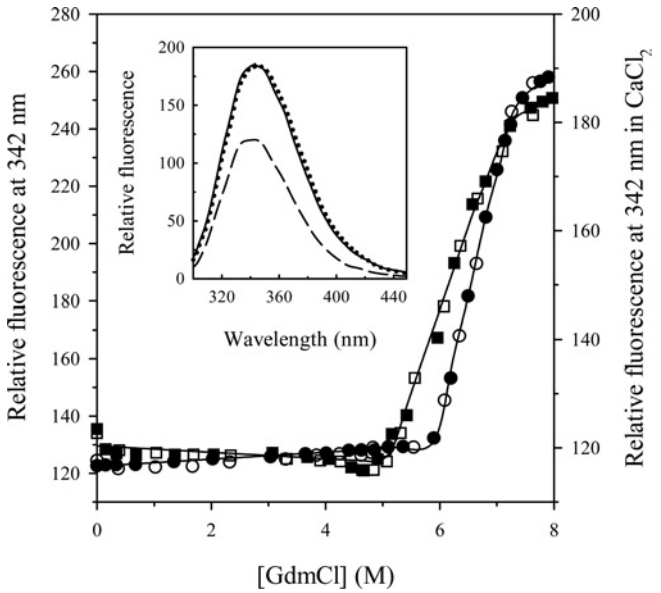

Figure 4 Effect of calcium on the GdmCl-induced fluorescence changes of *pf*LamA

Relative fluorescence emission intensity at 342 nm (290 nm excitation wavelength) was recorded at 20 °C after 24 h incubation at varying denaturant concentrations, in the absence (●, left axis) and presence of 40 mM $CaCl_2$ (■, right axis), as described in the text (see the Experimental section). Continuous lines are the non-linear regressions to eqn (2) of the fluorescence data. The reversibility points (empty symbols) were not included in the non-linear regression analysis. Inset: the intrinsic fluorescence emission spectra of *pf*LamA in 7.9 M GdmCl and 40 mM $CaCl_2$: —, the spectrum measured after 24 h incubation in 7.9 M GdmCl in the presence of 40 mM $CaCl_2$; . . . . ., the spectrum resulting from the progressive addition of $CaCl_2$ to the protein after 24 h incubation in 7.9 M GdmCl; – – –, the spectrum of the native *pf*LamA in 20 mM Tris/HCl (pH 7.4). All the spectra were recorded at 20 °C.

concentration (results not shown), similar to what was reported for the enzyme in the absence of Ca^{2+} [1]. This result was confirmed by the reconstitution of the spectra after SVD data analysis and by the random variation in magnitude and sign of the two most significant columns of the **V** matrix as a function of denaturant concentration.

The fluorescence emission spectrum measured on incubation in 7.9 M GdmCl and 40 mM $CaCl_2$ is comparable with that resulting from the progressive addition of $CaCl_2$ to the protein in 7.9 M GdmCl (Figure 4, inset).

ANS fluorescence and acrylamide quenching

The accessibility of hydrophobic residues on incubation of *pf*LamA in 7.9 M GdmCl in the presence or absence of Ca^{2+} was compared by the analysis with the fluorescent probe ANS. The fluorescence emission spectrum of ANS shows a modest, 2-fold increase in intensity in the presence of the protein in 7.9 M GdmCl, either in the presence or absence of 40 mM $CaCl_2$, without any change in the maximum fluorescence emission wavelength (results not shown). This suggests that in 7.9 M GdmCl the hydrophobic surface area of the protein is not significantly exposed.

The uncharged fluorescence quencher acrylamide was used to probe the accessibility of the hydrophobic core and the dynamic properties of *pf*LamA in 7.9 M GdmCl in the presence and absence of Ca^{2+} in comparison with that of the native enzyme. Effective acrylamide quenching constants from the modified Stern–Vollmer plots for the protein in the absence and presence of Ca^{2+} were 13.0 and 11.4 M^{-1} in 7.9 M GdmCl and 7.9 and 6.9 M^{-1} for the native enzyme respectively. These results suggest that Ca^{2+} causes a decreased accessibility of protein fluorophores to the quencher both in the native and in the 7.9 M GdmCl state.

Structural analysis

Estimation of secondary structure

A structural model for *pf*LamA was generated on the basis of an alignment of its amino acid sequence with that of two family 16 enzymes that have established crystal structures and that share the highest sequence similarity with *pf*LamA: *b*1,3-1,4Glc and *pc*Car (Figure 5). The resulting *pf*LamA model has been analysed in an attempt to unravel details of its extreme stability and of its calcium-binding site(s). The relative contribution of secondary structure elements in the *pf*LamA model was calculated by the program DSSP [37] and corresponds to 42.4 % of β-sheets, 2.6 % of α-helices, 12.4 % turns (including residues in isolated β-bridges) and 42.6 % loops. This result was compared with the deconvoluted FTIR amide I′ region in ^2H_2O (Figure 3B), which reveals nine individual peaks centred at 1620, 1630, 1642, 1651, 1658, 1662, 1672, 1682 and 1690 cm^{-1}. The main peak at 1630 cm^{-1} and the minor peak at 1620 cm^{-1} can be assigned to β-sheets and correspond to 40.0 and 1.0 % of the total amide I′ area respectively, which is in good agreement with the model and with the secondary structure prediction given by the server SSpro [38]. The peak at 1642 cm^{-1}, 29.8 %, can be attributed to unordered structures [39] as well as to flexible loops [40]. The assignment to flexible loops seems to be more satisfactory since *pf*LamA folds like the concanavalin A-like lectins [41] and a comparable band has been reported on deconvolution of the amide I′ of lectins [42]. Furthermore, the comparative analysis of the ten generated *pf*LamA models reveals that the most variable regions, potentially reflecting the most flexible elements of the protein structure, correspond to loop regions for 29.7 % of the total sequence. The band at 1651 cm^{-1}, 3.0 %, may be assigned to α-helices. The peak at 1658 cm^{-1} is 10.2 % of the amide I′ area and can be assigned to loops with dihedral angles close to α-helix [42]. Noteworthy, the Ramachandran plot of the *pf*LamA model calculated by PROCHECK [16] shows 12.4 % of residues within the region corresponding to right-handed α-helix; hence, the exclusion of residues actually involved in α-helix structures (2.6 %) results in a remaining 9.8 % of residues, which may correspond to loop regions with α-helix dihedral angles. The peaks at 1662, 1672 and 1682 cm^{-1} are assigned to β-turns [31,34,42] and correspond to 4.5, 6.1 and 3.1 % of the total amide I′ area. The peak centred at 1690 cm^{-1} can be referred to antiparallel β-sheets [42] and corresponds to 2.3 % of total amide I′ (Figure 3B). The relative content of secondary structure elements in amide I′ is in agreement with that determined by FTIR in H_2O and far-UV CD spectroscopy [1] and corresponds to that determined from the homology model of *pf*LamA.

Ion pairs

The number of ion pairs responsible for strong interactions (distance threshold of 4 Å) in the *pf*LamA model is 11, slightly higher than that observed in the lichenases *bl*1,3-1,4Glc and *bm*1,3-1,4Glc, where the ion pairs are 8 and 9 respectively, but lower than that observed in *pc*Car, which displays 15 ion pairs. The inclusion of weak ion pairs (distance threshold of 6 Å) leads to a significantly higher number of potential ion pairs in the hyperthermophilic *pf*LamA, some of which are involved in the formation of salt bridge networks. With the 6 Å threshold, the number of ion pairs in *pf*LamA increases to 45, whereas that of

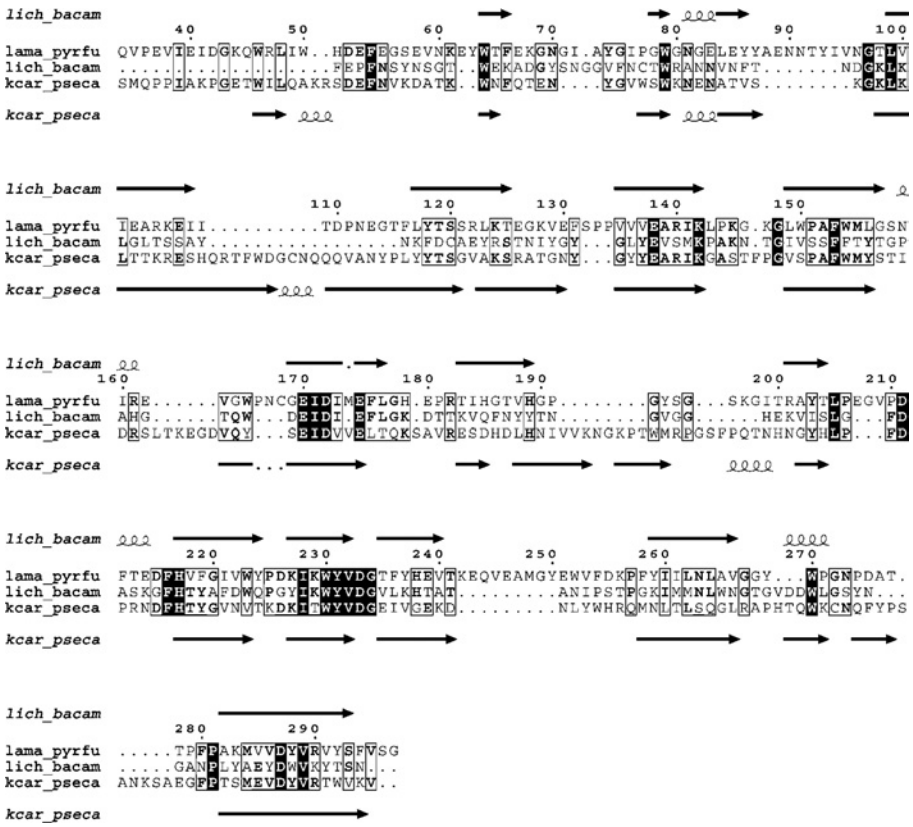

Figure 5 Multiple sequence alignment of *pf*LamA (lama_pyrfu) with structural templates of *b*1,3-1,4Glc (lich_bacam) and *pc*Car (kcar_pseca)

Single-letter code has been used for the amino acids. Dots represent deletions. Numbers above the sequences represent sequence numbering of *pf*LamA. Invariant positions are boxed in black; aligned columns displaying an amino acid identity in two of the three sequences are boxed and the most conserved residue is shown in boldface. Secondary structures of lich_bacam and kcar_pseca are reported in the first and the last line of each block respectively; α-helices and β-strands are shown as squiggles and arrows respectively. The figure was generated by using the program ESPript 2.2 [43].

*bl*1,3-1,4Glc and *bm*1,3-1,4Glc reaches 21 and 30 respectively. The number of ion pairs in *pc*Car increases more than in the two lichenases and reaches 44. The inclusion of histidine in the calculation of ion pair interactions increases the difference between the *pf*LamA and the two lichenases. The strong ion pairs increase from 8 to 9 and from 9 to 10 in *bl*1,3-1,4Glc and *bm*1,3-1,4Glc respectively, whereas they increase from 11 to 18 in *pf*LamA and from 15 to 22 in *pc*Car. If weak ion pairs are also considered, the number of total interactions reaches 63 in *pf*LamA and 33, 39 and 72 in *bl*1,3-1,4Glc, *bm*1,3-1,4Glc and *pc*Car respectively.

Calcium ions

Calcium ion in the model is bound in the convex face of the molecule, to the backbone carbonyl oxygens of Glu-53, Gly-97 and Asp-287 and to side-chain carboxylate oxygens of Glu-53 and Asp-287 (*pf*LamA amino acid numbering) (Figure 6). Residues Glu-53, Gly-97 and Asp-287 correspond respectively to Pro-9, Gly-45 and Asp-207 in *b*1,3-1,4Glc. Gly-97 (45) and Asp-287 (207) are conserved among most of the family 16 β-glucanases, whereas Glu-53 is apparently conserved only in β-glucanases from thermophilic organisms.

The cadmium ions of *pc*Car were used to investigate further potential binding sites for the calcium ions in *pf*LamA. The cadmium ion bound in the active site most probably reflects a

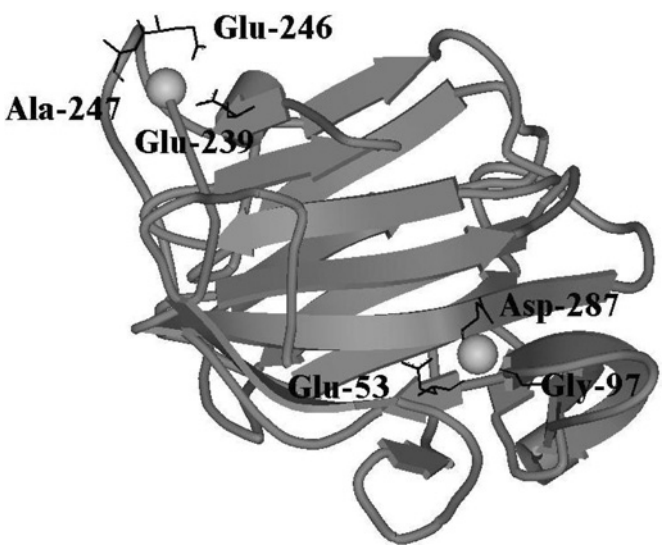

Figure 6 Schematic representation of *pf*LamA model

Calcium ions are represented as grey spheres. Amino acid residues putatively interacting with calcium are labelled and represented as sticks. The figure was prepared with DS ViewerPro 5.0 (Accelrys, San Diego, CA, U.S.A.).

crystallization artifact [7]. Another cadmium ion is bound on the convex face of the enzyme in a position corresponding to the binding site of the calcium ion, mentioned above, in *pf*LamA. The remaining five cadmium ions are bound to carboxylate oxygens of surface-exposed acidic amino acid residues. The residues of *pf*LamA at the corresponding positions of the alignment that was used to build the model do not show acidic residues except for Glu-239 (corresponding to Glu-241 in *pc*Car). This glutamic residue does not seem to be conserved in the β-glucanases of family 16, not even in those from other thermophiles. However, Glu-246 and Ala-247 in *pf*LamA are in a suitable position to interact through respectively one carboxylate and the carbonyl oxygens with the same calcium ion potentially bound to Glu-239 (Figure 6). Moreover, Glu-246 and Ala-247 are located within a loop region that is the most variable in the model (see the Experimental section). It is possible that such a loop has lower conformational restrictions so as to arrange Glu-246 and Ala-247 in the right position for the interaction with a calcium ion. Noteworthy is the fact that the loop conformation with the effective orientation of these two residues is that belonging to the model with the lowest 'objective function' value, i.e. the most reliable model.

Apolar contact surface

The extension of the apolar contact surface reveals some differences. *bl*1,3-1,4Glc, *bm*1,3-1,4Glc and *pc*Car display a total hydrophobic surface of approx. 11 959, 12 301 and 14 980 Å^2 respectively, corresponding to a surface per residue of 55.9, 58.0 and 55.3 Å^2. The apolar contact surface of the LamA model reaches an area of approx. 15 907 Å^2, corresponding to a surface per residue of 60.2 Å^2.

DISCUSSION

In 7.9 M GdmCl, the *pf*LamA tertiary structure can be consistently regained on calcium addition. The spectral properties of calcium-depleted *pf*LamA in 7.9 M GdmCl indicate the integrity of secondary structure elements [1] with the persistence of residual tertiary structure, which shows significant alterations, such as the solvent exposure of some tryptophan residues, indicated by the red-shift of the fluorescence emission maximum wavelength, and the loss of the asymmetric environment for the tryptophan residue(s) contributing to the 295 nm band of the near-UV CD spectrum. The changed spectral properties of the protein in 7.9 M GdmCl induced by calcium addition, i.e. the blue-shift of the intrinsic fluorescence emission wavelength to 343 nm and the recovery of the Trp-295 nm band in the near-UV CD spectrum, suggest a notable increase in the tertiary structure, which leads to a state very close to the native one. In the presence of Ca^{2+} in 7.9 M GdmCl, the structure of *pf*LamA is almost completely recovered, as judged from the spectral properties, and the accessibility of hydrophobic core is comparable with that of the native state, as suggested by ANS fluorescence and acrylamide quenching. However, the lack of catalytic activity confirms that this state is comparable, but not identical, with the native state.

Amongst the 380 carbohydrate active enzymes of family 16 glycoside hydrolases [2] (see http://afmb.cnrs-mrs.fr/~cazy/CAZY/index.html), 6 out of the 16 three-dimensional structures available in PDB present calcium ion(s) bound to the crystal structures. The interaction with Ca^{2+} is also common amongst the lectin-like proteins [44,45], which share a similar fold with family 16 glycoside hydrolases. The stabilizing effect of calcium on proteins is well known; however, the capability of *pf*LamA in 7.9 M GdmCl to interact with calcium with a consistent recovery of native tertiary structure is, to the best of our knowledge, unprecedented. In 7.9 M GdmCl, metal binding should be, at least in principle, highly unfavoured since it requires the presence of several lateral chains locked in the proper orientation and within the correct relative distance. In the present study, we suggest that *pf*LamA in 7.9 M GdmCl may still preserve the correct position of the residues involved in metal binding, despite the perturbation of its tertiary structure induced by the denaturant. Indeed, in 7.9 M GdmCl the interaction between *pf*LamA and Ca^{2+} displays the specificity typical for a protein–ligand interaction, e.g. the saturability. Interestingly, this interaction shows the same hyperbolic dependence on increasing calcium concentration, independent of the spectral probe used to follow the tertiary structure changes, suggesting that the process of structural regain of *pf*LamA does not proceed through any detectable structural intermediate.

The two-state GdmCl equilibrium transition of calcium-loaded *pf*LamA leads to a state indistinguishable from that obtained on the progressive addition of calcium to the calcium-depleted protein in 7.9 M GdmCl; this suggests that in the presence of calcium the state of the enzyme in 7.9 M GdmCl is the same, independent of the way it was formed. The spectral properties of this state are very similar to those of the *pf*LamA native state. Interestingly, the ΔG^{H_2O} value relative to the GdmCl-induced fluorescence changes of the calcium-loaded protein is 41.8 kJ/mol, a value relatively high if one considers the close structural similarity between the native state and the calcium-loaded state in 7.9 M GdmCl. This observation points out the remarkably high thermodynamic stability of *pf*LamA, as also supported by the ΔG^{H_2O} value of 61.5 kJ/mol relative to the partial unfolding of the calcium-depleted *pf*LamA, and suggests that much higher ΔG^{H_2O} values may reflect the complete unfolding of this 31 kDa monomeric protein. High values of free energy of stabilization are not unprecedented, and have been reported for some oligomeric [46] and for one large monomeric protein [47]; however, the thermodynamic parameters determined for *pf*LamA are associated with a state which is only partially denatured. Differences in thermodynamic parameters for the GdmCl-induced fluorescence changes in the presence and absence of calcium probably suggest that the calcium-loaded protein is less stable compared with the calcium-depleted protein. The decrease in thermodynamic stability is only apparent, since in 7.9 M GdmCl the spectral properties of the calcium-loaded state are more similar to those of the native state compared with those of the calcium-depleted state.

The relevant stability of *pf*LamA, as well as the residual tertiary structure in 7.9 M GdmCl capable of interacting with calcium, can be related to some peculiar features deduced from the comparison of its homology-based model with the structures of the mesophilic counterparts. The *pf*LamA model indicates that the number of ion pairs is almost doubled in comparison with the closely related mesophilic lichenases [7] and that the extension of the hydrophobic contact surface is significantly increased. Interestingly, these structural features are reported to be important for increasing the stability of thermophilic proteins [48,49] and may also contribute in maintaining *pf*LamA's residual structure in 7.9 M GdmCl [50].

The native state of *pf*LamA does not show any fluorescence or dichroic signal changes on Ca^{2+} addition; however, the decreased accessibility of the native protein fluorophores to acrylamide in the presence of calcium suggests a more compact structure as supported by the reduced hydrogen/deuterium exchange observed in the FTIR amide II region in the presence of the cation [35]. The result of Ca^{2+} binding to the native state can be directly monitored from the changes in intensity of the FTIR signal in the region of

the antisymmetric COO^- stretching vibration of the carboxylate moiety of the amino acid side chains of glutamic and aspartic residues, accompanied by an upshift of the corresponding symmetric bands. This evidence is in line with the possible involvement of aspartic and glutamic residues predicted by the homology modelling of *pf*LamA. The quantitative analysis of calcium binding by 5,5′-Br_2-BAPTA titration indicates that two Ca^{2+} interact with native *pf*LamA with high affinity. A correlation between protein stabilization and *in vivo* calcium concentration in *P. furiosus* is difficult to establish at the moment, also in consideration of the variability of Ca^{2+} concentration in the marine hydrothermal environment [51]. Work on the effect of different metals on the residual tertiary structure of *pf*LamA is in progress to make a comparative analysis with the structural changes induced by calcium and to gain information about the interactions of metals with non-native states of proteins.

We thank Dr R. Contestabile for a critical reading of the manuscript. This work was partially supported by a grant 'Progetti Strategici MIUR Legge 499/97' Project GENEFUN.

REFERENCES

1 Chiaraluce, R., van der Oost, J., Lebbink, J. H., Kaper, T. and Consalvi, V. (2002) Persistence of tertiary structure in 7.9 M guanidinium chloride: the case of endo-beta-1,3-glucanase from *Pyrococcus furiosus*. Biochemistry **41**, 14624–14632

2 Coutinho, P. M. and Henrissat, B. (1999) Carbohydrate-active enzymes: an integrated database approach, In Recent Advances in Carbohydrate Bioengineering (Gilbert, H. J., Davies, G., Henrissat, B. and Svensson, B., eds.), pp. 3–12, Royal Society of Chemistry, Cambridge

3 Bateman, A., Coin, L., Durbin, R., Finn, R. D., Hollich, V., Griffiths-Jones, S., Khanna, A., Marshall, M., Moxon, S., Sonnhammer, E. L. et al. (2004) The Pfam protein families database. Nucleic Acids Res. **32**, (**Database issue**) D138–D141

4 Allouch, J., Jam, M., Helbert, W., Barbeyron, T., Kloareg, B., Henrissat, B. and Czjzek, M. (2003) The three-dimensional structures of two beta-agarases. J. Biol. Chem. **278**, 47171–47180

5 Gueguen, Y., Voorhorst, W. G., van der Oost, J. and de Vos, W. M. (1997) Molecular and biochemical characterization of an endo-β-1,3-glucanase of the hyperthermophilic archaeon *Pyrococcus furiosus*. J. Biol. Chem. **272**, 31258–31264

6 Karplus, K., Karchin, R., Draper, J., Casper, J., Mandel-Gutfreund, Y., Diekhans, M. and Hughey, R. (2003) Combining local-structure, fold-recognition, and new fold methods for protein structure prediction. Proteins **53**, (Suppl. 6), 491–496

7 Michel, G., Chantalat, L., Duee, E., Barbeyron, T., Henrissat, B., Kloareg, B. and Dideberg, O. (2001) The kappa-carrageenase of *P. carrageenovora* features a tunnel-shaped active site: a novel insight in the evolution of Clan-B glycoside hydrolases. Structure **9**, 513–525

8 Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. and Bourne, P. E. (2000) The protein data bank. Nucleic Acids Res. **28**, 235–242

9 Shindyalov, I. N. and Bourne, P. E. (1998) Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. Protein Eng. **11**, 739–747

10 Thompson, J. D., Higgins, D. G. and Gibson, T. J. (1994) CLUSTALW: improving the sensitivity of progressive multiple alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22**, 4673–4680

11 InsightII, Version 2000 (2000) Molecular Modeling System, User Guide, Accelrys 2000, San Diego, CA

12 Friedberg, I., Kaplan, T. and Margalit, H. (2000) Evaluation of PSI-BLAST alignment accuracy in comparison to structural alignments. Protein Sci. **9**, 2278–2284

13 Šali, A., Potterton, L., Yuan, F., Van Vlijmen, H. and Karplus, M. (1995) Evaluation of comparative protein modeling by MODELLER. Proteins **23**, 318–326

14 Burke, D. F., Deane, C. M., Nagarajaram, H. A., Campillo, N., Martin-Martinez, M., Mendez, J. and Molina, F. (1999) An iterative structure-assisted approach to sequence alignment and comparative modeling. Proteins **3**, 55–60

15 Sippl, M. J. (1993) Recognition of errors in three-dimensional structures of proteins. Proteins **17**, 355–362

16 Laskowski, R. A., MacArthur, M. W., Moss, D. S. and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. **26**, 283–291

17 Drabløs, F. (1999) Clustering of non-polar contacts in proteins. Bioinformatics **15**, 501–509

18 Connolly, M. L. (1983) Solvent-accessible surfaces of proteins and nucleic acids. Science **221**, 709–713

19 Vriend, G. (1990) WHAT IF: a molecular modeling and drug design program. J. Mol. Graph. **8**, 52–56

20 Barlow, D. J. and Thornton, J. M. (1983) Ion-pairs in proteins. J. Mol. Biol. **168**, 867–885

21 Andre, I. and Linse, S. (2002) Measurement of Ca^{2+}-binding constants of proteins and presentation of the CaLigator software. Anal. Biochem. **305**, 195–205

22 Kaper, T., Verhees, C. H., Lebbink, J. H., van Lieshout, J. F., Kluskens, L. D., Ward, D. E., Kengen, S. W., Beerthuyzen, M. M., de Vos, W. M. and van der Oost, J. (2001) Characterization of beta-glycosylhydrolases from *Pyrococcus furiosus*. Methods Enzymol. **330**, 329–346

23 Gill, S. C. and von Hippel, P. H. (1989) Calculation of protein extinction coefficients from amino acid sequence data. Anal. Biochem. **182**, 319–326

24 Oberg, K. A. and Fink, A. L. (1998) A new attenuated total reflectance Fourier transform infrared spectroscopy method for the study of proteins in solution. Anal. Biochem. **256**, 92–106

25 Semisotnov, G. V., Rodionova, N. A., Razgulyaev, O. I., Uversky, V. N., Gripas', A. F. and Gilmanshin, R. I. (1991) Study of the 'molten globule' intermediate state in protein folding by a hydrophobic fluorescent probe. Biopolymers **31**, 119–128

26 Lehrer, S. S. (1971) Solute perturbation of protein fluorescence. The quenching of the tryptophyl fluorescence of model compounds and of lysozyme by iodide ion. Biochemistry **10**, 3254–3263

27 Ionescu, R. M., Smith, V. F., O'Neill, Jr, J. C. and Matthews, C. R. (2000) Multistate equilibrium unfolding of *Escherichia coli* dihydrofolate reductase: thermodynamic and spectroscopic description of the native, intermediate, and unfolded ensembles. Biochemistry **39**, 9540–9550

28 Santoro, M. M. and Bolen, D. W. (1988) Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl alpha-chymotrypsin using different denaturants. Biochemistry **27**, 8063–8068

29 Tsien, R. Y. (1980) New calcium indicators and buffers with high selectivity against magnesium and protons: design, synthesis, and properties of prototype structures. Biochemistry **19**, 2396–2404

30 Patton, C., Thompson, S. and Epel, D. (2004) Some precautions in using chelators to buffer metals in biological solutions. Cell Calcium **35**, 427–431

31 Jackson, M. and Mantsch, H. H. (1995) The use and misuse of FTIR spectroscopy in the determination of protein structure. Crit. Rev. Biochem. Mol. Biol. **30**, 95–120

32 Barth, A. (2000) The infrared absorption of amino acid side chains. Prog. Biophys. Mol. Biol. **74**, 141–173

33 Susi, H. (1972) Infrared spectroscopy conformation. Methods Enzymol. **26**, 455–472

34 Barth, A. and Zscherp, C. (2002) What vibrations tell us about proteins. Q. Rev. Biophys. **35**, 369–430

35 Jackson, M., Haris, P. I. and Chapman, D. (1991) Fourier transform infrared spectroscopic studies of Ca^{2+}-binding proteins. Biochemistry **30**, 9681–9686

36 Fabian, H. and Vogel, H. J. (2002) Fourier transform infrared spectroscopy of calcium-binding proteins. Methods Mol. Biol. **173**, 57–74

37 Kabsch, W. and Sander, C. (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers **22**, 2577–2637

38 Pollastri, G., Przybylski, D., Rost, B. and Baldi, P. (2002) Improving the prediction of protein secondary structure in three and eight classes using recurrent neural networks and profiles. Proteins **47**, 228–235

39 Heredia, P. and De Las Rivas, J. (2003) Calcium-dependent conformational change and thermal stability of the isolated PsbO protein detected by FTIR spectroscopy. Biochemistry **42**, 11831–11838

40 Fabian, H., Naumann, D., Misselwitz, R., Ristau, O., Gerlach, D. and Welfle, H. (1992) Secondary structure of streptokinase in aqueous solution: a Fourier transform infrared spectroscopic study. Biochemistry **31**, 6532–6538

41 Murzin, A. G., Brenner, S. E., Hubbard, T. and Chothia, C. (1995) SCOP: a structural classification of proteins database for the investigation of sequences and structures. J. Mol. Biol. **247**, 536–540

42 Chehin, R., Iloro, I., Marcos, M. J., Villar, E., Shnyrov, V. L. and Arrondo, J. L. (1999) Thermal and pH-induced conformational changes of a beta-sheet protein monitored by infrared spectroscopy. Biochemistry **38**, 1525–1530

43 Gouet, P., Courcelle, E., Stuart, D. I. and Metoz, F. (1999) ESPript: analysis of multiple sequence alignments using PostScript. Bioinformatics **15**, 305–308

44 Cioci, G., Mitchell, E. P., Gautier, C., Wimmerova, M., Sudakevitz, D., Perez, S., Gilboa-Garber, N. and Imberty, A. (2003) Structural basis of calcium and galactose recognition by the lectin PA-IL of *Pseudomonas aeruginosa*. FEBS Lett. **555**, 297–301

45 Loris, R. (2002) Principles of structures of animal and plant lectins. Biochim. Biophys. Acta **1572**, 198–208
46 Mitra, N., Srinivas, V. R., Ramya, T. N., Ahmad, N., Reddy, G. B. and Surolia, A. (2002) Conformational stability of legume lectins reflect their different modes of quaternary association: solvent denaturation studies on concanavalin A and winged bean acidic agglutinin. Biochemistry **41**, 9256–9263
47 Schoeffler, A. J., Joubert, A. M., Peng, F., Khan, F., Liu, C. C. and LiCata, V. J. (2004) Extreme free energy of stabilization of Taq DNA polymerase. Proteins **54**, 616–621
48 Yano, J. K. and Poulos, T. L. (2003) New understandings of thermostable and peizostable enzymes. Curr. Opin. Biotechnol. **14**, 360–365
49 Kumar, S. and Nussinov, R. (2001) How do thermophilic proteins deal with heat? Cell. Mol. Life Sci. **58**, 1216–1233
50 Robic, S., Guzman-Casado, M., Sanchez-Ruiz, J. M. and Marqusee, S. (2003) Role of residual structure in the unfolded state of a thermophilic protein. Proc. Natl. Acad. Sci. U.S.A. **100**, 11345–11349
51 Holden, J. F. and Adams, M. W. (2003) Microbe-metal interactions in marine hydrothermal environments. Curr. Opin. Chem. Biol. **7**, 160–165

Received 2 July 2004/2 October 2004; accepted 14 October 2004
Published as BJ Immediate Publication 14 October 2004, DOI 10.1042/BJ20041137