Calcium Binding and Thermostability of Carbohydrate Binding Module CBM4-2 of Xyn10A from *Rhodothermus marinus*†

Maher Abou-Hachem,*,‡ Eva Nordberg Karlsson,‡ Peter J. Simpson,§ Sara Linse,|| Peter Sellers,|| Michael P. Williamson,§ Stuart J. Jamieson,§ Harry J. Gilbert,⊥ David N. Bolam,⊥ and Olle Holst‡

Department of Biotechnology, Center for Chemistry and Chemical Engineering, Lund University, P.O. Box 124, SE-221 00 Lund, Sweden, Krebs Institute, Department of Molecular Biology and Biotechnology, University of Sheffield, Sheffield S10 2TN, U.K., Department of Biophysical Chemistry, Center for Chemistry and Chemical Engineering, Lund University, P.O. Box 124, SE-221 00 Lund, Sweden, and Department of Biological and Nutritional Sciences, University of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU, U.K.

Received November 26, 2001; Revised Manuscript Received February 28, 2002

ABSTRACT: Calcium binding to carbohydrate binding module CBM4-2 of xylanase 10A (Xyn10A) from *Rhodothermus marinus* was explored using calorimetry, NMR, fluorescence, and absorbance spectroscopy. CBM4-2 binds two calcium ions, one with moderate affinity and one with extremely high affinity. The moderate-affinity site has an association constant of $(1.3 \pm 0.3) \times 10^5$ M^{-1} and a binding enthalpy ΔH_a of -9.3 ± 0.4 kJ mol^{-1}, while the high-affinity site has an association constant of approximately 10^{10} M^{-1} and a binding enthalpy ΔH_a of -40.5 ± 0.5 kJ mol^{-1}. The locations of the binding sites have been identified by NMR and structural homology, and were verified by site-directed mutagenesis. The high-affinity site consists of the side chains of E11 and D160 and backbone carbonyls of E52 and K55, while the moderate-affinity site comprises the side chain of D29 and backbone carbonyls of L21, A22, V25, and W28. The high-affinity site is in a position analogous to the calcium site in CBM4 structures and in a recent CBM22 structure. Binding of calcium increases the unfolding temperature of the protein (T_m) by approximately 23 °C at pH 7.5. No correlation between binding affinity and T_m change was noted, as each of the two calcium ions contributes almost equally to the increase in unfolding temperature.

Calcium ions play a key role in a diversity of biological processes. Proteins exhibit a wide range of affinities for calcium, with affinity constants (K_a) typically in the range of 10^5-10^8 M^{-1}. Ca^{2+} plays a major role in signaling processes (*1*) as well as in stabilizing extracellular proteins such as glycoside hydrolases. For example, the vast majority of α-amylases which have an $(\alpha/\beta)_8$ fold are known to bind at least one calcium ion (*2*, *3*). A recent study showed that calcium ions stabilize this group of enzymes by bridging the two domains constituting the catalytic module, and thus maintaining the conformation of the active site (*4*). Another system where calcium ions are intimately associated with the function of glycoside hydrolases is the cellulosome, which is an efficient multienzyme complex responsible for cellulose degradation in anaerobic clostridia. The integrity of this consortium and its ability to degrade crystalline cellulose is dependent on the presence of calcium (*5*, *6*). Calcium binding is essential for the folding of dockerin domains that bind tightly to the cohesin domains of the scaffoldin protein anchoring cellulosomal proteins to it (*7*, *8*). Binding of Ca^{2+} not only is confined to dockerin domains but also extends to other components such as the endoglucanase CelD, where calcium is believed to stabilize the active site conformation of the enzyme and enhance its activity (*9*, *10*). In addition to these roles, Ca^{2+} ions are reported to alter the bond cleavage preference of *Pseudomonas fluorescens* xylanase10A (formerly XYLA), as well as protecting the enzyme from proteolytic attack (*11*, *12*). Calcium ions have also been found in glycoside hydrolases that have folds distinct from the $(\alpha/\beta)_8$ barrel. Thus, endoglucanases with an α/α barrel fold (*13*−*15*) and others with a β jellyroll structure (*16*, *17*) have been shown to contain calcium binding sites.

Glycoside hydrolases typically adopt a modular structure containing separate catalytic modules and noncatalytic modules (NCMs).[1] These are commonly but not always connected by linkers which are generally flexible (*18*−*20*). Carbohydrate binding modules (CBMs) are the most common noncatalytic modules, and they probably enhance the

† We thank the K & A Wallenberg Foundation for support of E.N.K., the Helge Axson Johnson foundation for support of M.A.-H., and the Swedish Research Council for support of S.L. We also thank the BBSRC for support of P.J.S. and S.J.J. and BBSRC and the Wellcome Trust for equipment grants.

* To whom correspondence should be addressed: Department of Biotechnology, Center for Chemistry and Chemical Engineering, Lund University, P.O. Box 124, SE-221 00 Lund, Sweden. E-mail: maher.abou_hachem@biotek.lu.se. Telephone: +46 46 222 46 26. Fax: +46 46 222 47 13.

‡ Department of Biotechnology, Center for Chemistry and Chemical Engineering, Lund University.

§ University of Sheffield.

|| Department of Biophysical Chemistry, Center for Chemistry and Chemical Engineering, Lund University.

⊥ University of Newcastle upon Tyne.

[1] Abbreviations: CBM, carbohydrate binding module; XBM, xylan binding module; TD, thermostabilizing domain; quin2, 2-({2-[bis-(carboxymethyl)amino]-5-methylphenoxy}methyl)-6-methoxy-8-[bis-(carboxymethyl)amino]quinoline; EDTA, ethylenediaminetetraacetic acid.

10.1021/bi012094a CCC: $22.00
Published on Web 04/16/2002

Table 1: Primers Used To Introduce Mutations[a]

E11A	
	F5′-AGTCATACATATGCTTGTCGCCAACATCAACGGTGGATTCG**C**ATCGACG-3′
	R5′-TACGATAAGCTAGCAATGGCCAGGCCATCAATGTAGATGG-3′
D160A	
	F5′-ACGTCATACATATGCTTGTCGCCAACATCAACGGTGG-3′
	R5′-CAGTACGCTAGCAATGGCCAGGCCA**G**CAATGTAGATGG-3′
E11A/D160A	
	F5′-AGTCATACATATGCTTGTCGCCAACATCAACGGTGGATTCG**C**ATCGACG-3′
	R5′-CAGTACGCTAGCAATGGCCAGGCCA**G**CAATGTAGATGG-3′
E23A	
	F5′-ACGGATCTGGCCG**CG**GGTGTGGAGGGC-3′
	R5′-GCCCTCCACACC**CG**CGGCCAGATCCGT-3′
E26A	
	F5′-GCCGAAGGTGTGG**CA**GGCTGGGATCTG-3′
	R5′-CAGATCCCAGCC**TG**CCACACCTTCGGC-3′
D29A	
	F5′-GTGGAGGGCTGGG**CG**CTGAACGTGGGC-3′
	R5′-GCCCACGTTCAG**CG**CCCAGCCCTCCAC-3′

[a] Restriction sites are underlined, and the altered bases are in bold. Prefixes F and R denote the forward and reverse primers, respectively.

activity of the catalytic modules by mediating prolonged enzyme−substrate proximity (*21*). CBMs have been classified into 28 families based primarily on primary sequence similarities, according to the latest update of the CAZy database (http://afmb.cnrs-mrs.fr/~pedro/CAZY/db.html) (*22*). Three-dimensional structures of CBM representatives from different families have revealed that calcium binding is a widespread property within some families (*23*, *24*), and there is some evidence that Ca^{2+} confers thermostability on a cellulose binding CBM belonging to family 4 (*25*).

The thermophilic bacterium *Rhodothermus marinus* produces a thermostable modular family 10 xylanase (Xyn10A) (*26*). In a previous study, the properties of the two N-terminal family 4 CBMs of Xyn10A from *R. marinus* have been investigated in our laboratory (*27*), and calcium ions seemed to enhance the binding of these modules to insoluble substrates, although the role of Ca^{2+} in maintaining the structural integrity of the protein remains to be established. In the study presented here, we report that CBM4-2 indeed binds two calcium ions, a novel property among family 4 CBMs. The binding constants of both sites are estimated and the locations of the sites identified by NMR and structural homology, and verified by site-directed mutagenesis. Finally, the effects of calcium binding on thermostability are evaluated.

MATERIALS AND METHODS

Cloning of CBM4-2 Mutants and Sample Preparation. Escherichia coli strains BL21(DE3) and XL1-Blue, expression vector pET25b(+) encoding CBM4-2, media, growth conditions, protein production, and purification were as previously described (*27*). To probe the locations of the two calcium binding sites, five different single-site mutations and a double mutation were introduced either by conventional PCR or using the Stratagene Quik-Change Kit (Stratagene, La Jolla, CA). The amino acid substitutions introduced into CBM4-2 were as follows: E11A, D160A, E11A/D160A, E23A, E26A, and D29A. Plasmids encoding CBM4-2 were used as templates in the mutation procedure. A set of primers was used to introduce the mutations (Table 1), and standard molecular biology procedures were used throughout. Recombinant protein production was verified by SDS−PAGE, and the samples were treated with carboxypeptidase A (Sigma, St. Louis, MO) to remove the histidine tag as previously outlined (*27*). Protein concentrations were determined by acid hydrolysis in a commercial laboratory (BMC, Uppsala, Sweden), and the values that were obtained were used to estimate the extinction coefficient of CBM4-2 at 280 nm (29 340 M^{-1} cm^{-1}) which was in good agreement with the theoretically estimated value (29 160 M^{-1} cm^{-1}). The same analysis verified that the five C-terminal histidine residues were lacking after the protease treatment. To remove metal ions from buffers, the solutions, all prepared using ultrapure Millipore water, were passed through a column packed with Chelex-100 beads (Bio-Rad, Hercules, CA). Plastic containers and vessels were used throughout for preparation and storage of buffers and protein samples, except for quartz cuvettes used in fluorescence and quin2 chelator titrations. To remove calcium from the proteins, samples were dialyzed extensively against equilibrated Chelex-100 beads. Determination of the amount of residual Ca^{2+} in protein samples, buffers, and $CaCl_2$ stocks was carried out with plasma mass spectrometry. The 10 mM Tris-HCl buffer (pH 7.5) which was used for the dilution of protein and calcium stocks and for titrations contained <0.5 μM Ca^{2+}.

Isothermal Titration Calorimetry (ITC). Titrations were performed with an OMEGA microcalorimeter (MicroCal, Inc., Northampton, MA). The stirred cell contained 600−950 μM CBM4-2 in 10 mM Tris-HCl buffer (pH 7.5), and the syringe contained 10 mM $CaCl_2$ in the same buffer. A titration experiment consisted of 25−30 injections of 5 or 10 μL with a 10 s duration and 600 s intervals each. The titration data were corrected for the small heat changes observed in control titrations of the $CaCl_2$ solution into buffer. Data were fitted to a two-site model using Origin (version 2.9, MicroCal, Inc.).

NMR Spectroscopy. NMR measurements were taken using Bruker DRX 500 and 600 spectrometers. Chemical shifts were measured using both unlabeled and ^{15}N-labeled protein [produced from growth in minimal medium containing $(^{15}NH_4)_2SO_4$ as the sole nitrogen source]. Most measurements on unlabeled protein were taken using one-dimensional spectra, while labeled protein was measured using HSQC (heteronuclear single-quantum coherence) spectroscopy. ^1H chemical shifts were referenced to internal trimethylsilyl-

propionate (0.00 ppm), and ^{15}N shifts were referenced indirectly to ^{1}H using the ratio of gyromagnetic ratios (*28*). All spectra were processed and analyzed using FELIX 2000 (Accelrys Inc., San Diego, CA). The moderate-affinity binding constant was determined by following chemical shift changes in the NMR spectrum on titration with $CaCl_2$. Chemical shift changes were fitted to a binding constant using a standard equation for a saturation isotherm in EXCEL version 5.0 (Microsoft Corp., Seattle, WA).

Fluorescence Spectroscopy. Fluorescence spectra were recorded at 25 °C using a Perkin-Elmer LS 50B fluorescence spectrometer with an excitation wavelength of 295 nm. Excitation and emission slits were 3 and 3.5 nm, respectively. Emission spectra (320−360 nm) were recorded, and each spectrum was an average of five scans. CBM4-2 solutions (2.5 mL of a 3 μM solution) were titrated with 1 μL aliquots of a $CaCl_2$ solution ranging in concentration from 0.1 μM to 100 mM in 10 mM Tris-HCl (pH 7.5). A nonlinear least-squares fitting procedure was employed to fit data (*29*), using the equation

$$F = F_0 + F_{Ca_1}[Ca^{2+}]/([Ca^{2+}] + K_{D_1}) + F_{Ca_2}[Ca^{2+}]/([Ca^{2+}] + K_{D_2})$$

where F_0, F_{Ca_1}, and F_{Ca_2} are the fluorescence of the apo species and singly and doubly loaded species, respectively. K_{D_1} and K_{D_2} are the dissociation constants of the two sites, and the model assumes two sites with vastly different affinities. The free Ca^{2+} concentration was determined using the Newton−Raphson method (*30*) from the equation

$$C_{Ca} = [Ca^{2+}] + C_P[Ca^{2+}]/([Ca^{2+}] + K_{D_1}) + C_P[Ca^{2+}]/([Ca^{2+}] + K_{D_2})$$

where C_{Ca} and C_P are the total calcium and total protein concentrations, respectively.

Chelator Method. To estimate the binding constant of the high-affinity site, Ca^{2+} titrations of the proteins were performed in the presence of the fluorescent Ca^{2+} chelator quin2 (*31*, *32*). The Ca^{2+} affinity of this chelator [log K = 8.022 (*32*)] is substantially higher than that of the moderate-affinity site, and consequently, competition for Ca^{2+} by that site is negligible. The quin2 concentration was determined by measuring the absorbance at 240 nm ($\epsilon_{240} = 4.2 \times 10^4$ M^{-1} cm^{-1}), and the initial calcium concentration in the stock was determined as described previously (*33*). The titrations were performed using a quartz cuvette containing 1 mL of 20 μM quin2 (Fluka, Buchs, Switzerland) and 10 μM CBM4-2 in 10 mM Tris-HCl (pH 7.5). Calcium aliquots (1 mM $CaCl_2$ in the same buffer) of 1 μL were added; the solution was mixed thoroughly, and the absorbance at 263 was measured after equilibration for 30 s. The binding constant was obtained from a least-squares fit of the absorbance as a function of the total calcium concentration as described previously (*34*). Corrections were made for the initial Ca^{2+} concentration and for dilutions due to Ca^{2+} additions. The protein concentration was multiplied by a correction factor, which was varied during the minimization procedure, and was in the range of 0.9−0.95. The reported results are the average of three independent titrations.

Differential Scanning Calorimetry (DSC). DSC measurements were performed on a differential scanning calorimeter (MicroCal, Inc.) with a cell volume of 0.5072 mL, at a scanning rate of 1 °C/min. Prior to the measurements, the samples were degassed for 15 min at room temperature. Background scans, collected with buffer in the reference and sample cells, were subtracted from sample scans. The reversibility of the thermal transitions was assessed by checking the reproducibility of the scan upon immediate cooling and rescanning and by varying the scan rate. The effect of pH on the stability of wild-type CBM4-2 in the presence of different calcium loads was explored. The buffers that were used were acetate buffer (pH 4.3 and 5), 2,2-bis-(hydroxymethyl)-2,2′,2″-nitrilotriethanol (pH 6 and 6.5), Tris-HCl buffer (pH 7.5), glycine-NaOH (pH 10), and CAPS [3-(cyclohexylamino)-1-propanesulfonic acid] (pH 11). Mutant proteins were only analyzed at pH 7.5 in 10 mM Tris-HCl. Protein samples with initial calcium loads of 0.5−1 molar equiv were analyzed (i) in buffer only, (ii) in the presence of 0.5 mM $CaCl_2$, or (iii) in the presence of 1 mM EDTA. All buffers used for the DSC experiments were 10 mM with respect to the buffering agent, and protein concentrations were approximately 10 μM.

RESULTS

Determination of Ca^{2+} Binding Constants

Samples of wild-type CBM4-2, which were prepared following brief Chelex treatment and/or extensive dialysis against Tris buffer, were shown by plasma mass spectrometry to contain between 0.9 and 0.95 mol of calcium/mol of protein, indicating a very high-affinity site. Various efforts to obtain the protein in the completely apo form, including urea denaturation and refolding, were unsuccessful. Exhaustive dialysis against an excess of Chelex still resulted in a protein that contained approximately 0.5 mol of Ca^{2+}/mol of protein. EDTA was avoided in these preparations, as preliminary fluorescence titrations with this chelator showed emission spectra indicating unspecific interaction with the CBM (data not shown).

Calcium titrations of samples (containing approximately 0.9−0.95 mol of Ca^{2+}/mol of protein) using fluorescence emission spectroscopy were consistent with the presence of at least two calcium binding sites (Figure 1B). Tryptophan fluorescence change (λ_{max} = 338 nm) was measured, which is consistent with the presence of tryptophan side chains reasonably close to both calcium binding sites. A rapid decrease in the emission maximum (with essentially no change in the wavelength) was observed during the first few additions of calcium (up to 1 molar equiv of Ca^{2+}), suggesting the saturation of a high-affinity site, which was followed by a smooth decrease in emission arising from titrating calcium into the lower-affinity site (Figure 1). Calcium binding analyzed by isothermal titration calorimetry (ITC) established unambiguously the presence of two calcium binding sites (Figure 2). Using CBM4-2 with 0.5 mol of Ca^{2+}/mol of protein, the association constant of the high-affinity site was estimated to be $\geq 10^9$ M^{-1}, with a large enthalpy of binding of -40.5 ± 0.5 kJ mol^{-1}. NMR titrations of the high-affinity site using a sample containing 0.5 mol of Ca^{2+}/mol of protein showed that the unloaded and singly loaded forms of the protein are in slow exchange on the

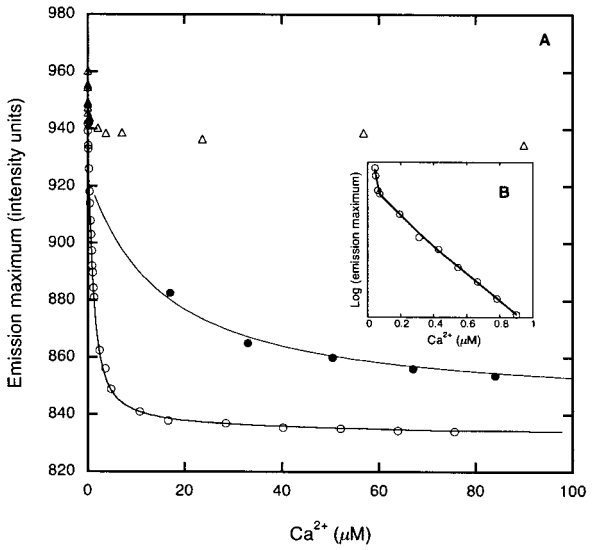

FIGURE 1: Fluorescence Ca^{2+} titrations of wild-type CBM4-2 and the E26A and D29A mutants at 25 °C, low salt buffer, and pH 7.5. (A) The empty circles (○) represent wild-type CBM4-2, and the fit of the data is represented by a solid line ($K_a = 3.3 \times 10^5$ M^{-1}). The filled circles (●) represent the E26A mutant, with the solid line representing a fit of the data to a binding affinity 5-fold lower than that of the wild-type protein. The triangles (△) represent titration of the D29A mutant. The binding affinity of the latter is too low to permit data fitting, and the initial rapid drop in intensity can be attributed to the saturation of the high-affinity site. The markedly diminished affinity of D29A for Ca^{2+} supports the postulated location of the binding site and the ligand status of the D29 side chain. (B) Lower region of the titration curve of wild-type CBM4-2 shown in panel A. To illustrate the presence of the high-affinity site, the log(emission maximum) is plotted vs calcium concentration for the first 10 additions of calcium.

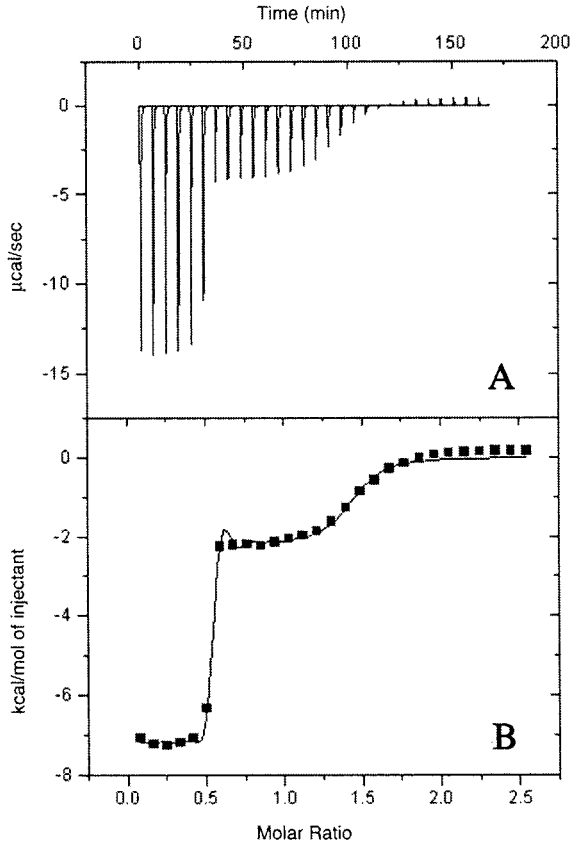

FIGURE 2: Results from ITC experiments (A and B). Calorimetric titration of 948 μM CBM4-2 with 10 mM $CaCl_2$ at 28 °C. Panel A shows the measured heat change vs time and panel B the titration curve normalized to molar concentrations. The best fit of the data to a model with two sets of binding sites obtained by nonlinear regression is shown as a solid line in panel B.

chemical shift time scale, implying a sub-micromolar affinity, and confirming the results given above. A more accurate measurement of the association constant for the high-affinity site (again using the sample with 0.5 mol of Ca^{2+}/mol of protein) was made using the chelator method, which makes use of a competition for calcium binding between the protein and quin2 (Figure 3). The binding constant for binding of quin2 to Ca^{2+} at low salt concentration was previously determined (*32*), and was used to estimate the binding constant of CBM4-2 (log $K = 10.1 \pm 0.5$).

Binding at the moderate-affinity site was measured by ITC, using CBM4-2 batches containing either 0.5 or 1 molar equiv of calcium, and yielded an association constant K_a of $(1.3 \pm 0.3) \times 10^5$ M^{-1} and a binding enthalpy ΔH_a of -9.3 ± 0.4 kJ mol^{-1} for the same site. Measurements by NMR under the same conditions (10 mM Tris-HCl, pH 7.5, and 25 °C) gave essentially the same result ($K_a = 1.8 \times 10^5$ M^{-1}). At pH 6.0, 50 mM acetate, and 37 °C, the affinity was approximately 10 times weaker. The fluorescence emission spectroscopy data (from the titrations mentioned above, using 0.9−0.95 mol of Ca^{2+}/mol of CBM4-2 in 10 mM Tris-HCl at pH 7.5 and 25 °C) were fitted to obtain a binding constant for the moderate-affinity site in the same range ($K_a = 3.3 \times 10^5$ M^{-1}).

Identification of the Calcium Binding Sites

(1) The Moderate-Affinity Site. NMR titrations of the moderate-affinity site, using CBM4-2 with 1 molar equiv of calcium, revealed large chemical shift perturbations to a number of amide groups, of which the largest shifts were (in order) for residues E23, D29, E26, and A22, with a number of other residues in the loop containing amino acids 25−30 also showing significant shifts. Calcium ions are normally coordinated by a mixture of side chain carboxyls and backbone carbonyls. Binding of calcium to a carbonyl influences the chemical shift of the amide group in the same peptide bond, which is the residue following the carbonyl. The NMR data therefore suggest that the ligands could be the backbone carbonyls of residues A22, possibly residues W28, V25, and L21, and possibly the D29, E26, and E23 side chains. Comparison with the solution structure of the protein in ref *35* shows that all of these amino acids, except E23 and E26, are well oriented to act as ligands. The D29 side chain could well coordinate the calcium via both oxygens, but the E26 and the E23 side chains appear to be too far from the calcium to act as ligands. The structure of the proposed binding site is shown in Figure 4C. NMR provides no evidence for water ligands, but it is likely that water forms at least one additional ligand.

To confirm the proposed location of the moderate-affinity calcium binding site, the mutants D29A, E23A, and E26A were produced, purified, and subjected to 5% (w/v) Chelex treatment, which resulted in proteins with 0.9−1 mol of Ca^{2+}/mol of protein. After the high-affinity site had been filled (displayed as a very slight decrease in the emission maximum in conjunction with the first additions of Ca^{2+}), the D29A mutant displayed a very low affinity that was not possible

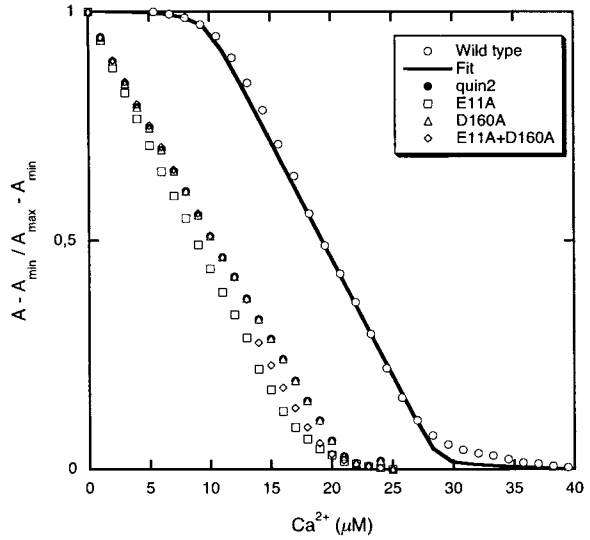

FIGURE 3: Titrations of quin2 with Ca^{2+} in the absence (control experiment) or presence of CBM4-2, E11A, D160A, or E11A/D160A. The normalized absorbance at 263 nm is plotted vs the total Ca^{2+} concentration. Protein and quin2 concentrations for this data set are 10.0 and 20.0 μM, respectively, and the initial residual Ca^{2+} concentration is 5 μM for wild-type CBM4-2. The best fits of three independent sets of data were averaged to estimate the binding constant of the wild-type CBM (log $K = 10.1 \pm 0.5$). All three mutants fail to exhibit high-affinity binding as titration curves of the mutants are similar to that of the quin2 titration in the absence of protein.

to quantify either with fluorescence titration (Figure 1) or by ITC. By contrast, the E26A mutant had an affinity for the second calcium ion only 5−10 times lower than that of the wild type (Figure 1). Thermostability results (see Table 3 below) suggest that both the E23A and E26A mutants are still able to bind two calcium ions. These results are consistent with the only side chain ligand being D29; the limited affinity change of the E26A mutant is presumably due to electrostatic effects on the protein rather than direct ligand interactions.

(2) The High-Affinity Site. As it proved to be impossible to completely remove calcium from the high-affinity site, and because partial occupancy of the high-affinity site only causes intensity changes rather than gradual frequency changes in NMR spectra, NMR failed to provide experimental evidence for the location of the site. Comparison of the structure of CBM4-2 with other CBMs did, however, lead to some interesting observations. There is a calcium binding loop on the side opposite the binding cleft in a family 22 CBM, which is in the same superfamily as CBM4-2 (*36*), and whose structure has recently been determined by X-ray crystallography (*18*). Comparison of the structures (Figure 5) shows that CBM4-2 in this region is very similar to CBM22 (*35*), with liganding groups in corresponding positions. From the comparison, the calcium ligands are expected to be the main chain carbonyls of E52 and K55 (K39 and E42, respectively, in CBM22) and the side chain carboxyls of E11 and D160 (E16 and D149, respectively, in CBM22). Hence, these latter residues were chosen as targets for site-directed mutagenesis. The two single mutants (E11A and D160A) and the double mutant (E11A/D160A) were produced, purified, and subjected to Chelex treatment as previously described, resulting in three mutant proteins, all essentially in the apo form. NMR spectra of wild-type and

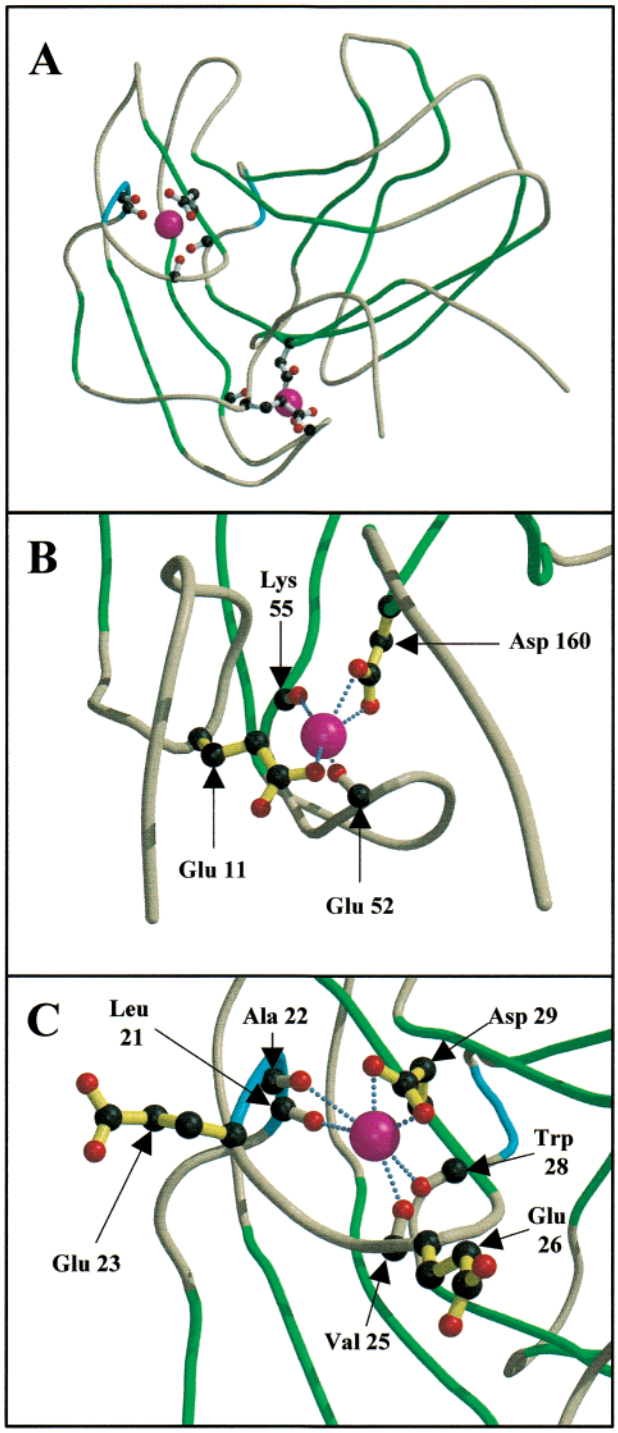

FIGURE 4: Structures of the high- and moderate-affinity sites. The coordinates are those described in ref *35*. (A) Overview of the protein. The binding site cleft is at the top of the figure; β sheets are in green, and α helical turns are in blue. The two calcium binding sites are indicated; the one nearer the binding site cleft is the moderate-affinity site. (B) Closeup of the high-affinity site. Residues identified as ligands are shown in ball-and-stick format, with likely hydrogen bonds shown as dotted lines. Residues mutated in this study are shown in yellow. (C) Closeup of the moderate-affinity site, with labeling as described above.

mutant proteins have closely similar patterns of upfield shifts, indicating that the mutations have caused only local structural changes (data not shown). Wild-type CBM4-2 (0.5 molar equiv of calcium) and the three mutants (apo form) were titrated with quin2 under the same conditions. No calcium binding could be detected for any of the three mutants in

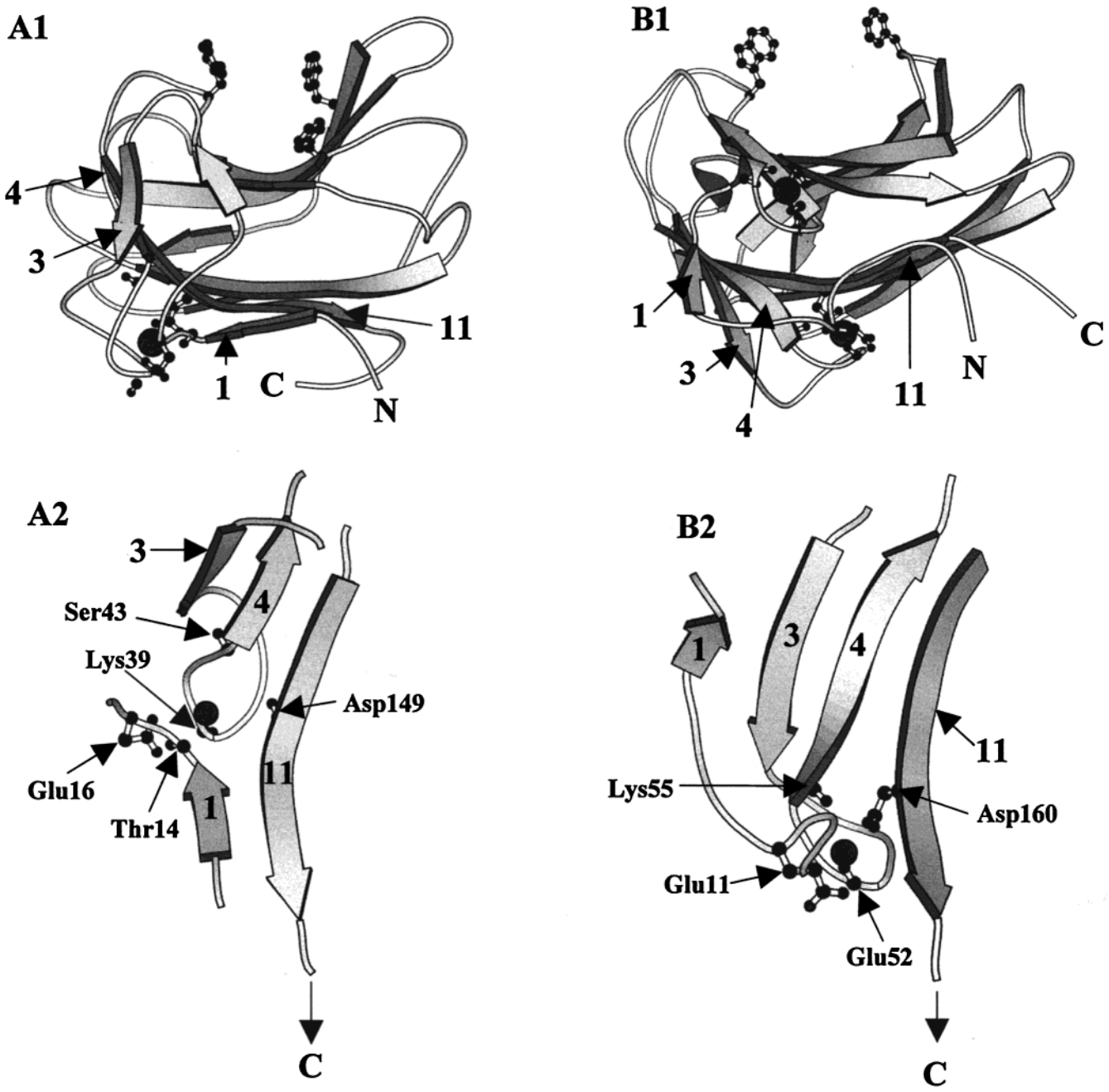

FIGURE 5: Structures of CBM4-2 (B1) and CBM22 (A1), shown oriented for the best fit of secondary structure elements. Calcium ions in the high-affinity site (bottom of B1) and moderate-affinity site (top of B1) of CBM4 and the CBM22 calcium site are indicated by balls, and the ligand binding residues are shown in ball-and-stick representation. Sheets and termini are labeled. In the bottom half of the figures, the analogous binding sites [high-affinity site in CBM4-2 (B2) and CBM22 (A2)] are displayed in expanded form, to show the calcium ligands.

the quin2 affinity range (Figure 3). ITC titrations with calcium carried out on E11A and D160A verified that the high-affinity binding was eliminated. The total calcium affinity of D160A was approximately 2-fold lower than the affinity of the moderate affinity site in wild-type CBM4-2 (data not shown). The total affinity of the E11A mutant was too low to allow a reasonable fitting of the data, suggesting a binding affinity in the low millimolar range. In light of these findings, we conclude that the E11 and D160 side chains are indeed calcium ligands, with the backbone carbonyls of E52 and K55 as other likely ligands. From the structure, it is likely that D160 binds in a bidentate manner (Figure 4B).

The high-affinity site identified here is in a position similar to that of the calcium site observed in *Cellulomonas fimi* N-terminal CBM4-1 (formerly CBD_{N1}) (*25*). The latter site was identified by homology to *Bacillus* 1,3:1,4-β-glucanases, and contained calcium ligands from the N- and C-terminal strands, in addition to a backbone carbonyl from a loop between strands β2 and β3. These are in positions analogous to those of the ligands seen here. In CBM4-1 from *C. fimi*, both the interstrand loop and the N-terminal segment are considerably shorter than those seen here and there is no direct structural homology, but it is clear that the sites are topologically equivalent.

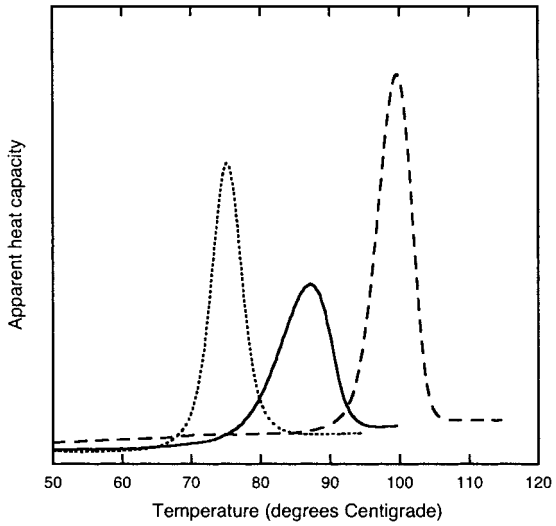

FIGURE 6: DSC thermograms of CBM4-2 in 10 mM Tris-HCl (pH 7.5) with protein concentrations of 10−13 μM and a scanning rate of 1 K/min. The dotted line shows unfolding in the presence of 1 mM EDTA (T_m = 75.1 °C); the solid line represents unfolding with only the high-affinity calcium binding site saturated (1 molar equiv of calcium, T_m = 87.4 °C), and the dashed line represents the protein with both sites saturated (0.5 mM $CaCl_2$, T_m = 97.5 °C).

Thermal Stability

(1) Dependence of the Unfolding Temperature (T_m) on pH and Calcium Content. The stability of CBM4-2 at different pH values and calcium contents was analyzed by differential scanning calorimetry (DSC). The calorimetric traces were completely irreversible at pH <6 and 11, and clearly visible aggregates were observed following the scanning analysis in the former case. Thermal unfolding was partially reversible between pH 6.5 and 10 with the best reversibility at pH 7.5. This buffering condition [10 mM Tris-HCl (pH 7.5)] was subsequently used for further analysis of the protein and its mutants. Moreover, an evident correlation between the reversibility of thermal transitions and calcium load was noted, as higher calcium loads led to a substantial decrease in reversibility at all pH values. The irreversibility of unfolding (especially when considering the differing reversibilities at different pH values and calcium loads) implies that a full thermodynamic treatment is not possible, and hence, we report here only a comparison of unfolding temperatures (T_m).

In line with the calcium binding results, three different thermal transitions related to the three different forms of the protein (apo, one site saturated, and both sites saturated) could be detected. The thermogram of CBM4-2 containing 0.5 molar equiv of calcium [in 10 mM Tris-HCl (pH 7.5)] exhibited two thermal transitions with T_m values of 74.2 and 86.7 °C for the first and second peaks, respectively (data not shown), indicating the presence of two forms of the protein (apo form and the form containing Ca^{2+} in the high-affinity site). In the thermogram of CBM4-2 with 1 molar equiv of calcium, only one transition was observed (T_m = 87.5 °C) (Figure 6), while in the presence of excess calcium (0.5 mM), a single thermal transition at a higher temperature was observed (T_m = 97.5 °C), indicative of a form with both calcium binding sites occupied. Interestingly, when the sample of 10 μM CBM4-2 containing 5 μM Ca^{2+} was analyzed in the presence of 1 mM EDTA, a single thermal transition was observed (T_m = 75.1 °C), corresponding to the T_m of the suggested apo form. Under these conditions, the high-affinity site will be approximately 17% saturated. We believe that the observation of a single peak at this temperature is explained by the fact that the scanning calorimeter changes the temperature at a slow rate (1 °C/min) compared to the off rate of Ca^{2+} from the protein (on the order of inverse seconds). Thus, when the temperature approaches T_m, the apo form is denatured first and re-equilibration between the calcium-loaded and apo form takes place sufficiently fast that the apo form is replenished, resulting in only the melting of the apo form being visible.

Table 2: Summary of DSC Data on CBM4-2 Thermal Unfolding at Different pH Values and Calcium Contents[a]

pH	T_{m1} (°C)	T_{m2} (°C)	T_{m3} (°C)	pH	T_{m1} (°C)	T_{m2} (°C)	T_{m3} (°C)
4.3	ND	65.2	75.1	7.5	75.1	87.4	97.5
5.0	ND	74.8,[b] 79[b]	87.9	10	72.1	87.5	100.6
6.0	77.4	82.1	88.4	11	70.9	86.5	99.5
6.5	76.6	85.6	94.0				

[a] DSC analysis carried out in low salt buffer at a wild-type CBM4-2 concentration of approximately 10 μM. ND, not determined; T_{m1}, T_m in the presence of 1 mM EDTA; T_{m2}, T_m in the presence of approximately 1 mol of Ca^{2+}/mol of protein; T_{m3}, T_m in the presence of excess Ca^{2+} (0.5 mM). [b] Two thermal transitions were observed at pH 5.

Calcium ion addition to CBM4-2 solutions (starting with 0.5 molar equiv) caused a significant increase in T_m under all the tested conditions (Table 2). In the presence of 1 mM EDTA, there was no major difference in unfolding temperature (T_m) between pH 6 and 8, whereas higher pH values led to a slight decrease in T_m. In contrast, the calcium forms of the protein (one site and two sites saturated) exhibited a clear trend of decreased T_m values with decreased pH values, at pH <6. At higher pH values, the T_m of CBM4-2 with one or two bound calcium ions was not greatly affected by pH variations, and the module exhibited good stability up to pH 11. Calcium binding was specific as Mg^{2+} and Cd^{2+} had little effect on the T_m of CBM4-2.

(2) Stability of CBM4-2 Mutants with Different Ca^{2+} Contents. In contrast to the wild-type protein, the mutants E11A, D160A, and E11A/D160A could be obtained in the completely Ca^{2+} free form, and no difference in T_m could be observed in the presence or absence of EDTA (data not shown) after Chelex treatment. The E11A was the least thermostable of the mutants that were produced (Table 3). These results are consistent with the proposed role of the E11 and D160 side chains as high-affinity site calcium ligands. The apo form of the double mutant (E11A/D160A) had a T_m similar to that of the apo form of D160A but higher than those of wild-type CBM and the E11A mutant (Table 3). It seems that the D160A mutation stabilizes the apo form of the CBM while resulting in a reduction of the calcium affinity of the moderate-affinity binding site as observed when the calcium affinity of the mutant was estimated by ITC.

The increase in T_m due to calcium addition to the D29A mutant (ΔT_m = 6.5 °C) was much lower than the corresponding increase noted due to filling the moderate-affinity site in the wild-type module or the two control mutants, E23A and E26A (in all cases, $\Delta T_m \geq$ 10.1 °C, Tables 2 and

Table 3: Thermal Unfolding As Determined by DSC of CBM4-2 Mutants with Different Calcium Contents[a]

	T_m (°C)		
sample	0 equiv of Ca^{2+}	1 equiv of Ca^{2+}	saturated with Ca^{2+}
wild type	74.2	87.4	97.5
E11A	75.3		82.5
D160A	80.4		82.5
E11A/D160A	80.5		86.1
E23A		85.9	96.1
E26A		85.1	97.8
D29A		87.7	94.2

[a] The calorimetric traces were determined under conditions similar to those used for the wild-type CBM. The first T_m listed (0 equiv) was measured in the presence of 1 mM EDTA (for the wild-type module, T_m was determined from thermograms with partial calcium occupancy), while the second (saturated) T_m was measured in the presence of 0.5 mM calcium chloride. For the high-affinity site mutants (E11A, D160A, and E11A/D160A), this corresponds to a singly loaded form, while for the wild type and for the mutants of the low-affinity site, it corresponds approximately to a doubly loaded form.

3). This is in line with the proposed role of D29 as part of the moderate-affinity site, but suggests that mutation of D29 may not completely abolish the low-affinity site. Mutations at E23 and E26 were intended as controls in the context of identifying the ligands of the calcium ions since they are located within 10 Å of the calcium binding sites. The DSC data show only a small reduction in the T_m of the singly occupied and saturated forms of these mutants compared to that of the wild type (Table 3), confirming that the stability and probably the calcium affinity differences between these mutants and the wild-type CBM at different calcium loads are marginal.

DISCUSSION

In this study, calcium binding to the thermostable CBM4-2 of Xyn10A from *R. marinus* was investigated. The data established the presence of two specific binding sites for Ca^{2+}, one with moderate affinity and one with extremely high affinity. The affinity of the moderate-affinity site was obtained independently by three methods with good agreement, while the affinity of the stronger site was beyond the range of most techniques. To estimate the binding constant of the high-affinity site, it proved to be necessary to use a method based on competition for calcium between CBM4-2 and a chromophoric chelator with high affinity for Ca^{2+}. This has been successfully employed for calcium binding studies on proteins with multiple calcium binding sites with K_a values in the range of 10^5-10^9 M^{-1} (*32*, *34*, *37*). The value of the high-affinity binding constant is one of the highest calcium affinities reported, and comparable to those in amylases [log $K = 11.3$ (*38*)] and thermitase (log $K = 10$) and other subtilases (*39*, *40*). Interestingly, CBM4-2, subtilase AK1, and thermitase possess some common features in the architecture of the sites. The calcium ion is totally buried, and its ligands include two acidic side chains and no water molecules. The ligands, which come from three different locations on the polypeptide, are stabilized by a network of hydrogen bonds where Ca^{2+} is the sole positively charged group in the network. It appears that by ligating residues that are distant in sequence, the tightly bound Ca^{2+} bridges the different parts of the protein together, including both termini, and stabilizes the fold. The high-affinity calcium site is in a position similar to that seen in the well-characterized *C. fimi* CBM4-1 module (*25*), as recently predicted (*41*). Deletion of the site in mutants E11A and D160A has no obvious effect on stability against digestion by proteases (data not shown). In contrast, Spurway et al. (*12*) found that the calcium binding site in *Pseudomonas fluorescens* xylanase 10A stabilized the protein against proteases.

The locations of both sites have been confirmed using mutants. Calcium titrations conducted on mutants E11A and D160A demonstrated the loss of high-affinity binding. Interestingly, the double mutant E11A/D160A seems to be more stable than either of the single mutants at similar Ca^{2+} concentrations, possibly due to the elimination of a buried charge in the double mutant with subsequent structural rearrangement. A similar conclusion was reached in the case of Xyn10A from *P. fluorescens* where replacing three Ca^{2+} ligating residues with alanine yielded a protein that matches the Ca^{2+}-saturated wild-type enzyme in its thermostability (*12*).

The only side chain ligand of the moderate-affinity site is that of D29 (Figure 4). The mutation causes changes in stability which are qualitatively in line with its ligating role. Substituting the D29 side chain still leaves four main chain calcium ligands, and it is reasonable to expect the retention of some calcium affinity. Indeed, the elevation of T_m in the DSC analysis in the presence of excess (0.5 mM) Ca^{2+} suggests that binding is still occurring.

It is of interest that we observed a 3-fold lower affinity in ITC measurements when quantifying the moderate-affinity binding site (using protein concentrations of approximately 950 μM) than in fluorescent emission measurements (using protein concentrations of 3 μM). This can be rationalized in that proteins are polyampholytes with multiple surface charges, and thus, protein concentration has an effect on the calcium binding affinity. The screening of electrostatic interactions at either a high salt or high protein concentration may reduce the calcium affinity severalfold (*39*).

The two sites of CBM4-2 differ appreciably not only in their affinity but also in their architecture, type of ligands, and their position in the sequence. Despite these differences, saturation of each site seems to make a similar contribution to the thermostability of the protein as judged by the elevation of its unfolding temperature. A thermodynamic treatment was not possible, because of irreversible unfolding and aggregation. Aggregation is typically accelerated at higher temperatures, and similar effects were reported from studies on hybrid glucanases with the capacity to bind calcium (*42*). Nonetheless, comparison of T_m values under different conditions and between different mutants of the protein gave a good indication of the stability of the module. The difference of ≈23 °C in unfolding temperature between the apo and saturated forms of CBM4-2 at pH 7.5 is remarkable, and can be compared to an increase of only 8 °C in the T_m of CBM4-1 of Cel9B from *C. fimi*.

The higher T_m values of CBM4-2 at higher pH values, with either one or both calcium binding sites saturated, are in line with the activity profile of the catalytic module which exhibited an optimum at pH 7.5 and higher activity on the alkaline side of the pH range (*43*). Both *R. marinus* CBM4 modules have low pI values between 4.3 and 4.5, and in the case of CBM4-2, many of the acidic side chains are located at the surface of the module. Low pH values cause the

protonation of carboxylates, and therefore, the affinity for calcium is expected to be lower at these conditions. The two *R. marinus* Xyn10A CBMs are homologous, sharing 88% sequence identity, with all the identified calcium ligands conserved in CBM4-1, except for D29 which was replaced by a valine in CBM4-1. Despite this substitution, we have evidence from DSC and fluorescence titrations that CBM4-1 binds two Ca^{2+} ions, although preliminary data suggest that the binding affinity and effects on thermostability differ between both modules (data not shown). The high level of identity between both modules implies conservation of fold and overall structure. It seems that subtle steric differences rather than a gross change in structure modulate these differences in calcium affinity and stability.

A survey of available CBM X-ray structures (*18*, *44*, *45*) in addition to other studies (*25*) reveals that representatives of families 3, 4, and 22, and the recently determined family 9, bind calcium ions. Family 3 CBMs contain a conserved Ca^{2+} binding site, which is solvent inaccessible and in which the ion is suggested to stabilize the protein fold (*24*), although no studies have been performed to explore this hypothesis. In contrast, the Ca^{2+} bound to CBM4-1 (previously termed CBD_{N1}) of Cel9B from *C. fimi* is suggested to be solvent accessible (*25*). Even though Ca^{2+} enhances the stability of CBM4-1, other roles cannot be ruled out, especially bearing in mind that the homologous CBM4-2 (termed CBD_{N2}) of the same enzyme which binds no calcium has a higher unfolding temperature than both the apo and calcium forms of CBM4-1 (CBD_{N1}). CBM9-2 of Xyn10A from *Thermotoga maritima* is the only module besides the investigated *R. marinus* modules that binds more than one Ca^{2+}. Three octahedrally coordinated metal ions in loop regions on the sides and on the opposite face of a shallow substrate binding groove were identified as calcium ions. None of these metal binding sites resembles the ones in CBM4-2. The CBM22 from the thermostable *Clostridium thermocellum* xylanase Xyn10B has a calcium site very similar to the high-affinity site of CBM4-2 (Figure 5). This module belongs to a group that has been termed the thermostabilizing domains (TD), as their sole function was believed to be to enhance the thermostability of the catalytic modules to which they are fused. However, some members of this group have recently been shown to bind xylan (*18*, *46*). Recently, CBM families 4, 16, 17, and 22, and the newly identified family 27, were proposed to constitute a superfamily on the basis of a common motif, and other structural similarities (*36*). The results presented here support claims for similarity within the superfamily.

A striking observation is that the moderate-affinity site is located at the edge of the substrate binding groove (Figure 4), which is unique among calcium binding CBMs with known structures. This location for calcium may explain the observation (*27*) that addition of calcium causes a large increase in the level of adsorption of the module to insoluble xylan. Further work is necessary to reach a detailed understanding of this phenomenon.

REFERENCES

1. Evenäs, J., Malmendal, A., and Forsen, S. (1998) *Curr. Opin. Chem. Biol. 2*, 293−302.
2. Buisson, G., Duee, E., Haser, R., and Payan, F. (1987) *EMBO J. 6*, 3909−3916.
3. Violet, M., and Meunier, J.-C. (1989) *Biochem. J. 263*, 665−670.
4. Declerck, N., Machius, M., Wiegand, G., Huber, R., and Gaillardin, C. (2000) *J. Mol. Biol. 301*, 1041−1057.
5. Choi, S. K., and Ljungdahl, L. G. (1996) *Biochemistry 35*, 4897−4905.
6. Choi, S. K., and Ljungdahl, L. G. (1996) *Biochemistry 35*, 4906−4910.
7. Lytle, B. L., Volkman, B. F., Westler, W. M., Heckman, M. P., and Wu, J. H. (2001) *J. Mol. Biol. 307*, 745−753.
8. Lytle, B. L., Volkman, B. F., Westler, W. M., and Wu, J. H. (2000) *Arch. Biochem. Biophys. 379*, 237−244.
9. Chauvaux, S., Beguin, P., Aubert, J. P., Bhat, K. M., Gow, L. A., Wood, T. M., and Bairoch, A. (1990) *Biochem. J. 265*, 261−265.
10. Chauvaux, S., Souchon, H., Alzari, P. M., Chariot, P., and Beguin, P. (1995) *J. Biol. Chem. 270*, 9757−9762.
11. Charnock, S. J., Spurway, T. D., Xie, H., Beylot, M. H., Virden, R., Warren, R. A., Hazlewood, G. P., and Gilbert, H. J. (1998) *J. Biol. Chem. 273*, 32187−32199.
12. Spurway, T. D., Morland, C., Cooper, A., Sumner, I., Hazlewood, G. P., O'Donnell, A. G., Pickersgill, R. W., and Gilbert, H. J. (1997) *J. Biol. Chem. 272*, 17523−17530.
13. Juy, M., Amit, A. G., Alzari, P. M., Poljak, R. J., Claeyssens, M., Beguin, P., and Aubert, J. P. (1990) *Nature 357*, 89−91.
14. Parsiegla, G., Juy, M., Reverbel-Leroy, C., Tardif, C., Belaich, J. P., Driguez, H., and Haser, R. (1998) *EMBO J. 17*, 5551−5562.
15. Varrot, A., Schülein, M., and Davies, G. J. (2000) *J. Mol. Biol. 297*, 819−828.
16. Hahn, M., Pons, J., Planas, A., Querol, E., and Heinemann, U. (1995) *FEBS Lett. 30*, 221−224.
17. Keitel, T., Simon, O., Borriss, R., and Heinemann, U. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 5287−5291.
18. Charnock, S. J., Bolam, D. N., Turkenburg, J. P., Gilbert, H. J., Ferreira, L. M., Davies, G. J., and Fontes, C. M. (2000) *Biochemistry 39*, 5013−5021.
19. Tomme, P., Warren, R. A. J., and Gilkes, N. R. (1995) in *Advances in Microbial Physiology* (Poole, R. K., Ed.) pp 2−81, Academic Press, London.
20. Warren, R. A. (1996) *Annu. Rev. Microbiol. 50*, 183−212.
21. Bolam, D. N., Ciruela, A., McQueen-Mason, S., Simpson, P., Williamson, M. P., Rixon, J. E., Boraston, A., Hazlewood, G. P., and Gilbert, H. J. (1998) *Biochem. J. 331*, 775−781.
22. Coutinho, P. M., and Henrissat, B. (1999) in *Carbohydrate binding modules: Diversity of structure and function* (Gilbert, H. J., Davies, G. J., Henrissat, B., and Svensson, B., Eds.) pp 3−12, The Royal Society of Chemistry, Cambridge, U.K.
23. Sakon, J., Irwin, D., Wilson, D. B., and Karplus, P. A. (1997) *Nat. Struct. Biol. 4*, 810−818.
24. Tormo, J., Lamed, R., Chirino, A. J., Morag, E., Bayer, E. A., Shoham, Y., and Steitz, T. A. (1996) *EMBO J. 15*, 5739−5751.
25. Johnson, P. E., Creagh, A. L., Brun, E., Joe, K., Tomme, P., Haynes, C. A., and McIntosh, L. P. (1998) *Biochemistry 37*, 12772−12781.
26. Nordberg Karlsson, E., Bartonek-Roxå, E., and Holst, O. (1997) *Biochim. Biophys. Acta 1353*, 118−124.
27. Abou Hachem, M., Nordberg Karlsson, E., Bartonek-Roxa, E., Raghothama, S., Simpson, P. J., Gilbert, H. J., Williamson, M. P., and Holst, O. (2000) *Biochem. J. 345*, 53−60.
28. Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E., Markley, J. L., and Sykes, B. D. (1995) *J. Biomol. NMR 6*, 135−140.
29. André, I., and Linse, S. (2002) *Anal. Biochem.* (in press).
30. Råde, L., and Westergren, B. (1993) *Beta Mathematics Handbook*, Student litteratur, Lund, Sweden.
31. Bryant, D. T. (1985) *Biochem. J. 226*, 613−616.
32. Linse, S., Brodin, P., Drakenberg, T., Thulin, E., Sellers, P., Elmden, K., Grundstrom, T., and Forsen, S. (1987) *Biochemistry 26*, 6723−6735.
33. Linse, S. (2002) in *Calcium-Binding Protein Protocols* (Vogel, H. J., Ed.) pp 15−24, Humana Press, Inc., Totowa, NJ.

34. Linse, S., Johansson, C., Brodin, P., Grundstrom, T., Drakenberg, T., and Forsen, S. (1991) *Biochemistry 30*, 154−162.
35. Simpson, P. J., Jamieson, S. J., Abou-Hachem, M., Nordberg Karlsson, E., Gilbert, H. J., and Williamson, M. P. (2002) *Biochemistry 41*, 5712−5719.
36. Sunna, A., Gibbs, M. D., and Bergquist, P. L. (2001) *Biochem. J. 356*, 791−798.
37. Linse, S., Brodin, P., Johansson, C., Thulin, E., Grundstrom, T., and Forsen, S. (1988) *Nature 335*, 651−652.
38. Vallee, B. L., Stein, E. A., Sumerwell, W. N., and Fisher, E. H. (1959) *J. Biol. Chem. 234*, 2901−2929.
39. Linse, S., and Forsen, S. (1995) *Adv. Second Messenger Phosphoprotein Res. 30*, 89−152.
40. Smith, C. A., Toogood, H. S., Baker, H. M., Daniel, R. M., and Baker, E. N. (1999) *J. Mol. Biol. 294*, 1027−1040.
41. Zverlov, V. V., Volkov, I. Y., Velikodvorskaya, G. A., and Schwarz, W. H. (2001) *Microbiology 147*, 621−629.
42. Welfle, K., Misselwitz, R., Welfle, H., Politz, O., and Borriss, R. (1995) *Eur. J. Biochem. 229*, 726−735.
43. Nordberg Karlsson, E., Dahlberg, L., Torto, N., Gorton, L., and Holst, O. (1998) *J. Biotechnol. 60*, 23−35.
44. Notenboom, V., Boraston, A. B., Kilburn, D. G., and Rose, D. R. (2001) *Biochemistry 40*, 6248−6256.
45. Shimon, L. J. W., Pages, S., Belaich, A., Belaich, J.-P., Bayer, E. A., Lamed, R., Shoham, Y., and Frolow, F. (2000) *Acta Crystallogr. D56*, 1560−1568.
46. Sunna, A., Gibbs, M. D., and Bergquist, P. L. (2000) *Biochem. J. 346*, 583−600.

BI012094A