Structure and Binding Specificity of the Second N-Terminal Cellulose-Binding Domain from *Cellulomonas fimi* Endoglucanase C[†,‡]

Emmanuel Brun,[§,||] Philip E. Johnson,[§,⊥,#] A. Louise Creagh,[§,▽,*] Peter Tomme,[§,▽,×,◇] Philip Webster,[||] Charles A. Haynes,[§,▽,*] and Lawrence P. McIntosh*[,§,||,⊥,▽]

Protein Engineering Network of Centres of Excellence, Department of Biochemistry and Molecular Biology, Department of Chemistry, The Biotechnology Laboratory, Department of Chemical Engineering, and Department of Microbiology and Immunology, The University of British Columbia, Vancouver, BC, Canada, V6T 1Z3

Received September 3, 1999; Revised Manuscript Received December 15, 1999

ABSTRACT: The 1,4-β-glucanase CenC from *Cellulomonas fimi* contains two cellulose-binding domains, CBD_{N1} and CBD_{N2}, arranged in tandem at its N-terminus. These homologous CBDs are distinct in their selectivity for binding amorphous and not crystalline cellulose. Multidimensional heteronuclear nuclear magnetic resonance (NMR) spectroscopy was used to determine the tertiary structure of CBD_{N2} in the presence of saturating amounts of cellopentaose. A total of 1996 experimental restraints were used to calculate an ensemble of 21 final structures for the protein. CBD_{N2} is composed of 11 β-strands, folded into two antiparallel β-sheets, with a topology of a jellyroll β-sandwich. On the basis of patterns of chemical shift perturbations accompanying the addition of cellooligosaccharides, as well as the observation of intermolecular protein-sugar NOE interactions, the cellulose-binding site of CBD_{N2} was identified as a cleft that lies across one face of the β-sandwich. The thermodynamic basis for the binding of cellooligosaccharides was investigated using isothermal titration calorimetry and NMR spectroscopy. Binding is enthalpically driven and consistent with a structural model involving hydrogen bonding between the equatorial hydroxyls of the glucopyranosyl rings and polar amino acid side chains lining the CBD_{N2} cleft. Affinity electrophoresis was used to determine that CBD_{N2} also binds soluble β-1,4-linked polymers of glucose, including hydroxyethylcellulose and β-1,3−1,4-glucans. This study complements a previous analysis of CBD_{N1} [Johnson, P. E., Joshi, M. D., Tomme, P., Kilburn, D. G., and McIntosh, L. P. (1996) *Biochemistry 35*, 14381−14394] and demonstrates that the homologous CBDs from CenC share very similar structures and sugar binding properties.

Cellulolytic bacteria and fungi produce a range of enzymes to efficiently degrade the various allomorphs of cellulose, a β-1,4-linked polymer of glucose (*1*−*3*). Generally, cellulases and related β-1,4-glucanases have a modular architecture consisting of a catalytic domain and one or more cellulose-binding domains (CBDs).[1] CBDs facilitate the binding of cellulases to their substrate and possibly aid in the disruption of this insoluble polymer. Over 200 putative CBDs have been identified and classified into 13 families on the basis of their sequence similarity (*4*). In parallel with the structural heterogeneity of naturally occurring cellulose, CBDs from these different families exhibit a wide range of specificities and affinities toward the various allomorphs of cellulose.

The family IV CBDs from the *Cellulomonas fimi* 1,4-β-glucanase CenC are unusual in two major respects. First,

[†] This work was funded by the Government of Canada's Network of Centres of Excellence Program supported by the Medical Research Council of Canada and the Natural Science and Engineering Research Council through PENCE Inc. (L.P.M. and C.A.H.).

[‡] The coordinates of the minimized average structure and the ensemble of 21 structures (1CX1), as well as the X-PLOR restraint tables have been deposited in the Brookhaven Protein Data Bank. The NMR chemical shifts have been deposited to the BioMagResBank.

* To whom correspondence should be addressed: Department of Biochemistry and Molecular Biology, 2146 Health Sciences Mall, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z3. Phone: (604) 822-3341. E-mail: mcintosh@otter.biochem.ubc.ca.

[§] Protein Engineering Network of Centres of Excellence.

[||] Department of Biochemistry and Molecular Biology.

[⊥] Department of Chemistry.

[#] Current address: Howard Hughes Medical Institute, Department of Chemistry and Biochemistry, University of Maryland Baltimore County, Baltimore, MD 21250.

[▽] The Biotechnology Laboratory.

[○] Department of Chemical Engineering.

[×] Department of Microbiology and Immunology.

[◇] Current address: Tibotec Generall de Wittelaan, L-11, 2800 Mechelen, Belgium.

[1] Abbreviations: CBD, cellulose-binding domain; CBD_{CBH1}, the cellulose-binding domain from *T. reesei* cellobiohydrolase 1; CBD_{Cex} the cellulose-binding domain from *Cellulomonas fimi* xylanase-glucanase Cex; CBD_{N1}, the N-terminal cellulose-binding domain from *Cellulomonas fimi* β-1,4-glucanase CenC; CBD_{N2}, the cellulose-binding domain from *Cellulomonas fimi* β-1,4-glucanase CenC following CBD_{N1} in sequence; CBD_{N2}·Glc_5, the 1:1 complex of CBD_{N2} and cellopentaose; CBD_{N1N2}, the tandem cellulose-binding domains from *Cellulomonas fimi* β-1,4-glucanase CenC; CD, circular dichroism; CT-HSQC, constant time heteronuclear single quantum correlation DQF-COSY, double quantum filtered correlation spectroscopy; DSC, differential scanning calorimetry; DSS, 2,2-dimethyl-2-silapentane-5-sulphonic acid, sodium salt; DTNB, 5,5′dithio-bis(2-nitrobenzoic acid); FT-IR, Fourier transform infrared; HMQC, heteronuclear multiple quantum correlation; HSQC, heteronuclear single quantum correlation; IPTG, isopropyl β-D-thiogalactopyranoside; ITC, isothermal titration calorimetry; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; pH*, the observed pH meter reading without correction for isotope effects; rms, root-mean-square; rmsd, root-mean-square deviation; sw, spectral width; TEMPO, 2,2,6,6-tetramethylpiperidine-1-oxyl-4-yl; TOCSY, total correlation spectroscopy.

10.1021/bi992079u CCC: $19.00
Published on Web 02/16/2000

although repeated CBDs do occur in several glucanases, CenC contains two homologous binding domains, named CBD_{N1} and CBD_{N2}, located in tandem at its N-terminus (*5*). Second, CBD_{N1} and CBD_{N2} are distinct in that they bind amorphous cellulose and soluble cellulose derivatives, but not crystalline cellulose (*5*−*7*).

The tertiary structure of the first N-terminal CBD from CenC was determined using NMR spectroscopy. Isolated CBD_{N1} is composed of 10 β-strands, folded into a jellyroll β-sandwich (*8*). The cellulose-binding site of CBD_{N1} is a cleft or groove located across one β-sheet face of the protein (*6*−*8*). The presence of a binding cleft, rather than a flat binding face as encountered with CBDs that bind crystalline cellulose (*9*−*12*), readily explains the specificity of CBD_{N1} for single strands of cellulose. This cleft, which spans approximately the length of five glycosyl units, contains a central strip of hydrophobic residues flanked on both sides by polar, hydrogen-bonding groups. On the basis of thermodynamic studies, which revealed that the association of CBD_{N1} with soluble sugars is enthalpically driven (*7*), a mechanism of binding was proposed that is supported by the structural features of this protein domain. Specifically, the pyranose rings stack against the hydrophobic strip, while the flanking hydrophilic residues provide hydrogen bonds to the equatorial hydroxyl groups of the bound cellulose chain. Subsequent studies using nitroxide spin-labeled cellotetraose demonstrated that oligosaccharides, and by inference single strands of amorphous cellulose, bind in multiple orientations across the β-sheet binding cleft of CBD_{N1} (*13*). This finding is consistent with the approximate symmetry of the hydrogen-bonding groups on both the cellooligosaccharides and the residues forming the binding cleft of the CBD.

We have extended our analysis of *C. fimi* CenC by characterizing the second N-terminal binding domain, CBD_{N2}. As expected from their 36.5% sequence identity (*8*), the NMR-derived tertiary structure of CBD_{N2} closely resembles that of CBD_{N1}, namely that of a jellyroll β-sandwich with a pronounced cleft. Chemical shift perturbations accompanying the addition of cellotetraose and cellopentaose confirm that this cleft is also the carbohydrate-binding site of CBD_{N2}. NMR spectroscopy and isothermal titration calorimetry reveal that the binding of cellooligosaccharides to CBD_{N2} is enthalpically driven and occurs with approximately the same affinity as measured previously for CBD_{N1}. Therefore, CBD_{N1} and CBD_{N2} share the same mechanism of cellulose recognition. These results provide the foundation for understanding the role of the two CBDs when located in tandem within their native context in *C. fimi* CenC.

MATERIALS AND METHODS

Cloning of CBD_{N2}. The gene fragment encoding CBD_{N2} (residues 146−296 of CenC) was cloned in the high expression pTug vector downstream of the Cex leader peptide using PCR based methods (*6*, *14*). DNA was amplified from a template gene in the pTugN1N2 plasmid encoding the tandem CBD_{N1N2} (*5*). An *Nhe*1 restriction site (underlined), corresponding to an additional two residues (Ala-Ser) at the amino terminus of the mature protein, was introduced at the 5′ end of the CBD_{N2} gene using the oligonucleotide 5′-CATATGGCTAGCCTCGACTCCGAGG TCGAGCTC-3′ as a primer. A *Hin*dIII restriction site (underlined) was introduced at the 3′ end of CBD_{N2} gene using the oligonucleotide 5′-AGAATGAATTCAAGCTTACGTCGCCGAG-GTGGTGAGCGA-3′. To obtain the final pTugN2 vector, the resulting 0.38 kb PCR fragment was cloned in-frame and downstream of the gene encoding the Cex leader peptide in the pTugEO7K3 plasmid (*14*). The identity of the final construct was verified by sequencing of the double-stranded DNA using the deoxy chain termination method with modified T7 DNA polymerase (*15*).

Expression, Labeling, and Purification of CBD_{N2}. The CBD_{N2} gene in the pTugN2 vector was expressed in *Escherichia coli* JM101 cells (*16*). Unlabeled protein was produced in TYP rich media (*17*). Uniformly ^{15}N-labeled protein was prepared using M9 media containing 1 g/L of 99% $^{15}NH_4Cl$ and 1 g/L of 99% ^{15}N-Isogro (Isotec Inc.) as the sole sources of nitrogen (*18*). Uniformly $^{13}C/^{15}N$-labeled protein was produced using M9 media containing 1 g/L of 99% $^{15}NH_4Cl$, 3 g/L of 99% $^{13}C_6$-D-glucose (Cambridge Isotope Laboratories), and 1 g/L of 99% $^{15}N/^{13}C$-Isogro. Nonrandomly fractionally ^{13}C-labeled protein was prepared using M9 media containing 0.3 g/L of $^{13}C_6$-D-glucose and 2.7 g/L of unlabeled D-glucose as the sole carbon sources (*19*). CBD_{N2} with uniformly deuterated aromatic rings was obtained from a synthetic medium containing 50 mg/L each of L-δ_1,ϵ_2,$\zeta_{1,2}$,η_2-[$^2H_5$]tryptophan, L-$\delta_{1,2}$,$\epsilon_{1,2}$,ζ-[$^2H_5$]phenylalanine, and L-$\delta_{1,2}$,$\epsilon_{1,2}$-[$^2H_4$]tyrosine (Cambridge Isotope Laboratories and Isotec Inc.) (*18*, *20*). The expression and purification of CBD_{N2} was accomplished using the procedure described previously for CBD_{N1} (*6*). The final yields were ∼70 or ∼20 mg of pure unlabeled or labeled CBD_{N2}, respectively, per liter of growth media.

Characterization of CBD_{N2}. Protein concentrations were measured using ultraviolet absorption spectroscopy. The molar absorptivity ϵ_{280} of CBD_{N2} was determined to be 20 500 M^{-1} cm^{-1} by the method of Edelhoch (*21*, *22*). The molecular mass of the unlabeled CBD_{N2} was 15 868 ± 3 Da, as measured by electrospray ionization mass spectroscopy. This is in excellent agreement with the expected molecular mass of 15 868 Da deduced from the amino acid composition of the protein after posttranslational cleavage of the secretory leader peptide and the formation of a single disulfide bond. The N-terminal sequence of the bacterially expressed CBD_{N2} was confirmed to be ASLDSE by Edman degradation.

NMR Spectroscopy. Samples of CBD_{N2} for NMR analysis were exchanged into 50 mM sodium chloride, 50 mM potassium phosphate buffer (pH* 5.9), 0.02% sodium azide, and 10% D_2O/90% H_2O using ultrafiltration through a cellulose-free membrane (Filtron). CBD_{N2} in deuterated buffer was obtained by twice lyophilizing the protein and redissolving after each freeze-drying step in an equivalent volume of 99.9% D_2O. Typical samples contained 1.5−2 mM protein and, unless stated otherwise, up to a 22-fold molar excess of cellopentaose (Seikagaku Corp.).

NMR spectra were recorded at 35 °C on a Varian Unity 500 MHz spectrometer equipped with a gradient triple resonance probe. Experiments with $^1H^N$ detection were recorded using the enhanced sensitivity pulsed-field gradient approach of Kay et al. (*23*) and Muhandiram and Kay (*24*). Selective water flip back pulse was incorporated to minimize the perturbation of the bulk water magnetization (*25*, *26*).

^{1}H chemical shifts were referenced to an external standard of DSS at 0.00 ppm, ^{13}C chemical shifts were referenced to an external DSS standard at 0.00 ppm, and ^{15}N chemical shifts were referenced to external 2.9 M $^{15}NH_4Cl$ in 1 M HCl at 24.93 ppm (*27*). This latter reference yields ^{15}N chemical shifts 1.6 ppm greater than those obtained using liquid NH_3 (*28*). Spectra were processed and analyzed using FELIX software (Biosym Technologies; San Diego, CA).

Amide Hydrogen Exchange. Amide hydrogen exchange rates were estimated by recording a series of sensitivity-enhanced gradient ^{1}H-^{15}N HSQC spectra at 0.5, 1, and 24 h after lyophilized CBD_{N2} was dissolved in 99.9% D_2O. The buffer concentration and pH* were held constant by maintaining the sample volume.

Histidine pK_a Measurements. The apparent pK_a values of the two histidine imidazole side chains in CBD_{N2} were determined by measuring the one-dimensional ^{1}H NMR spectra of the [$^2H_5$]tryptophan/[$^2H_5$]phenylalanine/[$^2H_4$]-tyrosine-labeled protein as a function of pH* at 35 °C. The sample was transferred to 99% D_2O buffer (25 mM potassium phosphate, 25 mM sodium d_3-acetate, and 50 mM sodium chloride) by lyophilization and pH* changed by the addition of microliter aliquots of ~0.1 M HCl or NaOH. A second titration was also carried out with protein (~0.5 mM) in the presence of saturating quantities (~4 mM) of cellohexaose. The pK_a values were obtained from nonlinear least-squares fitting of the observed $^1H^{\delta 2}$ and $^1H^{\epsilon 1}$ chemical shifts as a function of pH* to the equation describing the titration of a single ionizable group (*29*) using the program Kaleidagraph (Abelbeck Software).

Titration of CBD_{N2} with Cellooligosaccharides Monitored by NMR Spectroscopy. The binding of soluble cellooligosaccharides to CBD_{N2} at 35 °C and pH* 5.9 was measured quantitatively using ^{1}H-^{15}N NMR spectroscopy as described by Johnson et al. (*6*). Fifteen ^{1}H-^{15}N HSQC spectra of CBD_{N2}, in the presence of increasing amounts of cellotetraose or cellopentaose, were recorded for these binding-constant measurements. The protein concentration, initially at 0.4 mM, was corrected for the change in sample volume upon addition of cellotetraose and cellopentaose up to a final sugar concentration of 16 and 12.4 mM, respectively. Equilibrium association constants were determined by nonlinear least-squares fitting of the chemical shift perturbation data versus sugar concentration to the Langmuir isotherm describing the binding of one ligand molecule to a single protein site (*6*).

Determination of Binding Constants by Isothermal Titration Calorimetry. Isothermal titration calorimetry (ITC) was performed using a Microcal MCS ITC. All samples were pH 7.0 in 50 mM potassium phosphate. Titrations were performed by injecting 26 consecutive 10 μL aliquots of a cellooligosaccharide solution into the ITC cell (volume = 1.3528 mL), initially containing sugar-free CBD_{N2}. The data were corrected for the heat of dilution of the titrant by subtracting mixing enthalpies for 10 μL injections of cellooligosaccharide solution into protein-free buffer. At least two independent titration experiments were performed at 35 °C to determine the binding constant of CBD_{N2}. Binding stoichiometry, enthalpy, and equilibrium association constants were determined by fitting to the 1:1 bimolecular interaction binding model, determined by optimizing the binding polynomial, to the corrected data.

Structure Calculations. All structure calculations were performed using a simulated annealing protocol in X-PLOR 3.8 (*30*). The initial model was an extended chain, and the target function contained only potential terms for covalent geometry, experimental distance, and dihedral angle restraints, and a van der Waals repulsion term for nonbonded contacts. NOE-derived distance restraints were identified from a 3D ^{15}N-editted NOESY-HSQC spectrum, a simultaneous 3D ^{15}N/^{13}C NOESY-HSQC spectrum, and, for aromatic residues, a 2D ^{1}H-^{1}H NOESY spectrum acquired in D_2O. A 4D ^{13}C-^{13}C resolved NOESY experiment was used to help resolve ambiguous NOE interactions. Intermolecular NOE interactions between ^{15}N/^{13}C-labeled CBD_{N2} and bound, unlabeled cellopentaose were identified through a 3D ^{13}C ω1-edited, ω3-filtered NOESY experiment (*31*), but were not used in the structure calculations. In all cases, a mixing time of 140 ms was employed. Interproton distances were assigned to three ranges (weak, 1.8−5.0 Å; medium, 1.8−3.8 Å; and strong, 1.8−2.9 Å), following a square-well potential energy function. These distance ranges were calibrated as described previously (*8*). A correction of 0.5 Å was added to the upper bounds of restraints involving methyl groups (*32*). Hydrogen bonds, identified from patterns of amide hydrogen exchange kinetics and NOE interactions involving main-chain protons, were restrained to 2.5−3.5 Å between O and N atoms and 1.5−2.5 Å between H^N and O atoms. The ϕ torsion angles were restrained to 60 ± 30° for J < 6 Hz, −140 ± 40° for 8 Hz < J < 9 Hz, and −140 ± 30° for J > 9 Hz, using $^3J_{HN-H\alpha}$ couplings determined from a ^{1}H-^{15}N HNHA spectrum (*33*). The ψ angles were restrained to 120 ± 100° or −30 ± 110° based on the ratio of $^1H^{\alpha}_{i-1}$-$^1H^N_i$ and $^1H^{\alpha}_i$-$^1H^N_i$ NOE intensities (*34*). The χ^1 angles restrained to ±30° from their assigned rotamer values. These data were obtained for Thr, Ile, and Val residues from $^{13}C^{\gamma}$-{^{15}N} and $^{13}C^{\gamma}$-{$^{13}C'$} spin−echo difference CT-HSQC spectra (*33*, *35*), and for residues with two H^{β}'s from HNHB (*36*) and 40 ms mixing time ^{15}N TOCSY-HSQC (*37*). With the exception of the cis peptide bond between Thr75 and Pro76, all amides were constrained in a trans conformation. The imidazole rings of both histidine residues were in the neutral $N^{\epsilon 2}H$ tautomeric state, as determined from an ^{1}H-^{15}N HMBC spectrum (*38*).

A preliminary fold of CBD_{N2} was determined using unambiguous NOE-derived distance restraints, hydrogen bonds involved in pairing β-strands, and dihedral angle restraints. Additional NOE interactions were assigned in reiterative fashion based on successive structures of increasing precision. A total of 1510 NOE-derived distance restraints were used in the final calculation of an ensemble of 60 structures. This data set was comprised of 700 nontrivial intraresidue, 332 sequential, 83 short range ($1 < |i - j| \leq 4$), and 394 long range ($|i - j| > 4$) distance restraints. In addition, 86 hydrogen bond restraints (43 hydrogen bonds), one disulfide bond distance restraint between the C^{γ} of Cys38 and Cys142, 53 ϕ-angle restraints, 70 ψ-angle restraints, and 48 χ^1-angle restraints were utilized for the structure calculations. The stereochemical quality of the structural ensemble was analyzed using X-PLOR, Procheck-NMR and Aqua (*39*). Angular order parameters, Sϕ, Sψ, and Sχ1, were calculated according to ref *40*.

RESULTS

Expression of CBD_{N2}. The gene encoding CBD_{N2} was cloned into the high expression pTug vector (*14*). The boundaries of the isolated domain were chosen to span residues 146−296 of CenC based on the sequence alignment of four known family IV CBDs (*1*, *8*). Without an obvious linker between CBD_{N1} and CBD_{N2}, the N-terminus boundary of the subcloned CBD_{N2} was selected to ensure the expression of the full domain binding. Specifically, it includes seven residues of the disordered C-terminus of the isolated CBD_{N1} (see Figure 6) (*8*). Furthermore, the introduction of an *Nhe*I restriction site by PCR at the 5′ end of the CBD_{N2} gene for cloning purposes resulted in the addition of the dipeptide Ala-Ser at the N-terminus of this protein. The C-terminal boundary of CBD_{N2} was chosen to precede three consecutive proline residues (297−299 of CenC) that form part of a short sequence postulated to serve as a flexible linker between the binding and catalytic domains of this endoglucanase. For simplicity, residues in the isolated CBD_{N2} are numbered from Ala1 to Thr153.

As expected based on previous studies of CBD_{N1} and CBD_{N1N2} (*7*), isolated CBD_{N2} folds into a native structure that is functional for cellulose binding. This is evident, for example, from the use of an Avicel affinity chromatography step to purify the isolated domain. Differential scanning calorimetric studies indicate that CBD_{N2} is very stable, with a midpoint denaturation temperature of 67 °C at pH 7.0 (not shown). A correlation time of 7.5 ns for the global tumbling of CBD_{N2}, as determined from ^{15}N relaxation measurements (not shown), confirms that the protein is monomeric under the conditions used for this NMR study.

Resonance Assignments for CBD_{N2}. To help characterize the mechanism of cellulose binding and to facilitate comparison with CBD_{N1}, which was studied in the presence of cellooligosacharides (*8*), all NMR spectra used for the structure determination of CBD_{N2} were obtained with protein in the presence of ∼10 mM cellopentaose. As discussed below, the isolated binding domain forms a 1:1 complex with the cellooligosaccharide. Unfortunately, due to severe spectral overlap, the resonances of the bound cellopentaose could not be unambiguously assigned, and thus, the sugar was not included in the structure calculations.

The $^1H^N$ and ^{15}N resonances from the main-chain amides of CBD_{N2} bound to cellopentaose were assigned using the combination of HNCACB (*41*) and CBCACONH (*42*) experiments to correlate the $^{15}NH_i$ and $^{13}C^{\alpha}_i/^{13}C^{\beta}_i$, $^{13}C^{\alpha}_{i-1}/^{13}C^{\beta}_{i-1}$, and $^{15}NH_i$ and $^{13}C^{\alpha}_{i-1}/^{13}C^{\beta}_{i-1}$ resonances, respectively. Due to their distinctive chemical shift patterns, several peptide segments containing primarily glycine, serine, or threonine residues were used as specific starting points in the sequential assignment procedure. A modified version of the CBCACO(CA)HA (*43*) experiment that serves to detect only the $^1H^{\alpha}$ and $^{13}C^{\alpha}/^{13}C^{\beta}$ resonances of residues preceding a proline was also extremely useful for providing unambiguous reference points (*44*) (L. Kay, personal communication). Sequential assignments of proline residues were achieved using ^{15}N NOESY-HSQC (*45*, *46*) and $^{13}C/^{15}N$ NOESY-HSQC (*47*) experiments. Of the 13 prolines in CBD_{N2}, only Pro76 exhibits a cis conformation as indicated by a strong $^1H^{\alpha}-^1H^{\alpha}$ NOE correlation with Thr75 (*48*). Resonances from the single pair of neighboring prolines were identified based on the observation of a strong NOE interaction between Pro120 $^1H^{\alpha}$ and Pro121 $^1H^{\delta}$ in a ^{13}C-resolved NOESY spectrum. With the exception of the first two residues, the resonances from the main-chain $^1H^N$ and ^{15}N nuclei of CBD_{N2} were fully assigned by this strategy (Figure 1).

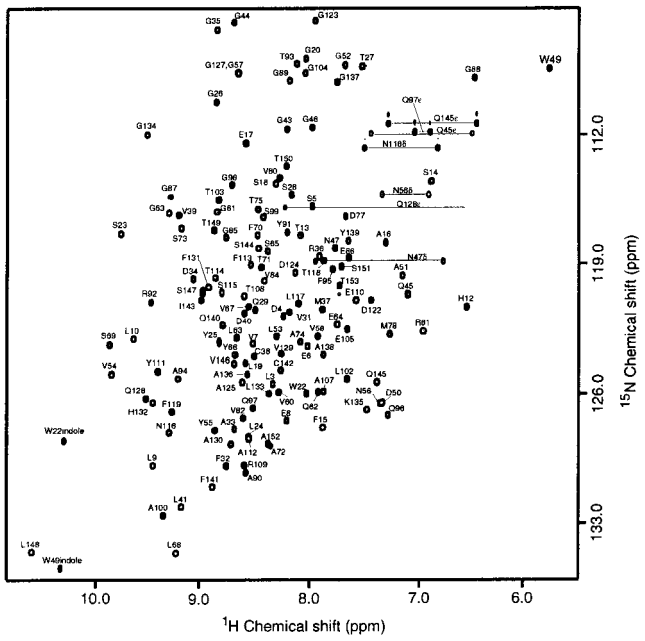

FIGURE 1: ^1H-^{15}N HSQC spectrum of CBD_{N2} at 35 °C and pH* 5.9, showing the assignments of the resonances from backbone amide and side-chain asparagine, glutamine and tryptophan ^{15}N-^1H groups.

Using the ^1H-^{15}N HSQC spectrum as a reference point, virtually complete assignments of the resonances from the side-chain ^1H, ^{13}C, and ^{15}N nuclei in CBD_{N2} were obtained readily from a combination of ^{15}N TOCSY-HSQC (*37*), H(CCO)NH, and C(CO)NH spectra (*49*, *50*) recorded on the labeled protein in H_2O buffer. These spectra were augmented with an HCCH-TOCSY (*51*, *52*) experiment, measured with CBD_{N2} in D_2O buffer. The resonances from the $^1H^{\delta}$ and $^1H^{\epsilon}$ of the aromatic rings were directly correlated to those of the side-chain $^{13}C^{\beta}$ nuclei using the (Hβ)Cβ(CγCδ)Hδ and (Hβ)Cβ(CγCδCϵ)Hϵ experiments (*53*, *54*). Additional aromatic ring spin system assignments were obtained from ^1H-^{13}C CT-HSQC spectra of $^{13}C/^{15}N$-labeled CBD_{N2} (*55*, *56*) and a combination of homonuclear 2D DQF-COSY, TOCSY, and NOESY spectra recorded with unlabeled protein in D_2O buffer (*54*). With the exception of Phe113, complete ^1H resonance assignments were obtained for all of the aromatic rings in CBD_{N2}. Resonances from the $^{15}N^{\delta 1}$ and $^{15}N^{\epsilon 2}$ of two histidine imidazole rings were identified using a ^1H-^{15}N HMBC experiment. On the basis of the observed patterns of ^{15}N chemical shifts and 2- and 3-bond heteronuclear *J*-couplings, both His12 and His132 were found to exist in the neutral $N^{\epsilon 2}H$ tautomeric form at pH* 5.9 (*38*, *54*).

Stereospecific assignment of 38 of the 79 residues in CBD_{N2} with prochiral $^1H^{\beta}$ protons were obtained through an analysis of HNHB (*36*), 40 ms mixing time ^{15}N TOCSY (*37*), and 140 ms $^{13}C/^{15}N$ NOESY-HSQC (*47*) spectra. These assignments, and the corresponding $\chi 1$ dihedral angles, were determined on the basis of a staggered rotamer model, as outlined by Powers et al. (*57*). Stereospecific assignments of the resonances from 22 of the 25 diastereotopic methyls

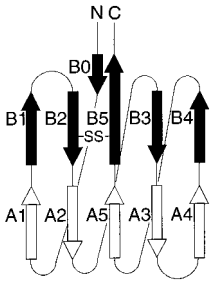

FIGURE 2: Schematic diagram showing the jellyroll β-sandwich topology of CBD_{N2}. Sheets A and B are indicated by open and solid arrows, respectively, and the position of the disulfide between Cys38 and Cys142 is indicated. Strands A2−A5 and B2−B5 comprise the jellyroll motif, with the two short strands A1 and B1 appended along one of this core structure. The tertiary structure of CBD_{N2} can be envisioned by folding the figure such that sheet B lies below sheet A. The lengths of the strands and loops are not drawn to scale.

of Val and Leu were obtained from both high-resolution and constant-time ^1H-^{13}C HSQC spectra of 10% nonrandomly fractionally ^{13}C-enriched CBD_{N2} (*8*, *19*). Degeneracy of both carbon and proton frequency prevents the stereospecific assignment of three of the diastereotopic methyls. The NH_2 resonances from the primary amides of asparagine and glutamine side chains were stereospecifically assigned using a combination of an EZ-HMQC-NH_2 (*58*) and a ^{15}N-resolved NOESY-HSQC (*26*) experiment, as described by McIntosh et al. (*58*). The resonance assignments of CBD_{N2} have been submitted to the BioMagResBank.

Secondary Structure Determination. Initial studies using CD spectroscopy indicated that CBD_{N2} is composed of β-strands and is devoid of α-helices (data not shown). This global analysis was confirmed when the regular secondary structural elements of CBD_{N2} were determined on the basis of patterns of sequential and cross-stranded NOEs, $^3J_{HN-H\alpha}$ coupling constants, $^{13}C^{\alpha}$, $^{13}C^{\beta}$, and $^1H^{\alpha}$ secondary chemical shifts, and amide hydrogen exchange rates (*48*, *59*, *60*). The information defining the 11 β-strands in CBD_{N2} is summarized in Figures S1 and S2 of the Supporting Information. On the basis of the patterns of cross-strand backbone NOEs, CBD_{N2} folds into two β-sheets (Figure 2). Following the nomenclature utilized previously for CBD_{N1}, these sheets are denoted as A and B. Sheet A is composed of five antiparallel β-strands (A1−A5) and contain two bulges (Supporting Information, Figure S2). One, comprising Ala94 and Phe95 in strand A4 and Val84 in strand A3, is classified as "anti-parallel classic" by Promotif (*61*, *62*). The second bulge, formed by Ser99 and Ala100 in strand A4 and Arg81 in strand A3, is "anti-parallel wide". Sheet B is composed of six antiparallel β-strands (B0−B5) and also contains two bulges. Strand B0, named due to the lack of a counterpart in CBD_{N1}, is very short and defined primarily by three interstrand NOEs to residues in strand B5. On the basis of the tertiary structure of CBD_{N2}, the polypeptide backbone turns sharply in the region of Glu62−Ser65 in strand B3 and Thr114−Asn116 in strand B4 to link sheets A and B. This sharp turn corresponds to a bulge, comprising Glu64 in strand B3 and Thr114 and Ser115 in strand B4, which is classified as "anti-parallel G1" by Promotif. The second anti-parallel classic bulge of β-sheet B is formed by Ser144 and Gln145 in strand B5 and Thr71 in strand B3.

From the topological arrangement of the β-strands of CBD_{N2}, it is evident that this protein adopts a jellyroll β-sandwich structure (Figure 2) (*8*, *63*). The jellyroll is formed by strands A2−A5 and B2−B5, with the two short strands A1 and B1 appended along one side of this core motif. Strands A1/B1 and A4/B4 are not connected by hydrogen bonding, nor are amides on outer edges of the β-sheets A and B protected from hydrogen−deuterium exchange. This defines the structure of CBD_{N2} as a β-sandwich rather than a continuous β-barrel. The poorly defined strand B0 is not a component of the jellyroll motif. Note that the N-terminal region of CBD_{N2} (Leu3−Leu10), which includes strand B0, overlaps with the C-terminal region of CBD_{N1} in the full-length enzyme (see Figure 6). It is therefore possible that this strand is a prolongation of the final β-strand B5 found in CBD_{N1}.

A disulfide bond between the only two cysteine residues (Cys38 and 142) serves to bridge strands B2 and B5. The presence of this disulfide was initially detected by the difference in the migration rates of the oxidized and reduced forms of CBD_{N2} on native gel electrophoresis (data not shown). Further support was provided by the observation of $^{13}C^{\beta}$ chemical shifts indicative of a cystine group involving residues 38 and 142 (*60*), as well as cross-strand NOEs between these two amino acids. As also seen with CBD_{N1} (*64*), reduction of this single disulfide leads to the complete unfolding of CBD_{N2} under otherwise benign experimental conditions (data not shown).

Tertiary Structure Determination. A total of 1996 distance, hydrogen bond, and dihedral restraints were used to calculate 60 structures following a hybrid distance geometry/simulated annealing protocol in X-PLOR 3.8 (*30*, *65*). NOE-derived distance restraints were identified from a 2D ^1H-^1H NOESY acquired in D_2O and a simultaneous 3D $^{15}N/^{13}C$ NOESY-HSQC experiments. A 4D ^{13}C-^{13}C resolved NOESY experiment was exploited only to resolve ambiguous NOE interactions. A 3D ^{12}C-^{13}C filtered NOESY experiment was used to identify specific NOE interactions between the ^{13}C-labeled CBD_{N2} and the unlabeled cellopentaose. However, due to the severe degeneracy of the cellooligosaccharide resonances, these intermolecular NOEs could not be assigned to specific protons of cellopentaose and, therefore, were not used in the structure calculations. Of the 60 structures calculated, the 21 with the lowest total energy and fewest restraint violations were selected for comparison. None of the latter had NOE violations greater than 0.4 Å, nor dihedral violations greater than 5.0°. Statistics for the 21 accepted structures are listed in Table 1, and the superimposition of the final ensemble of structures calculated for CBD_{N2} is shown in Figures 3 and 4.

The structural ensemble is clearly consistent with the jellyroll β-sandwich topology of CBD_{N2} deduced at the level of secondary structure analysis. The strands forming the two sheets are well-determined, having an rms deviation of 0.39 ± 0.07 Å for backbone heavy atoms with respect to the average structure (Figures 3 and 5). Most of the side chains that make up the hydrophobic core of CBD_{N2} are also well-defined structurally. This is reflected by both low rms deviations in the structural ensemble and high angular order parameters S(χ1) (Figures 4 and 5). Of the 33 side chains that comprise the hydrophobic core of CBD_{N2}, the positions of Leu53, Val58, Leu82, and Val129 are the least well-defined,

Table 1: Structural Statistics and Atomic RMS Differences[a]

	⟨SA⟩	⟨SA⟩$_{av}$
rmsd from experimental distance restraints[b] (Å) (1597)	0.16 ± 0.001	0.012
rmsd from experimental dihedral restraints[c] (deg) (171)	0.522 ± 0.043	0.44
deviations from idealized geometry		
bonds (Å) (2196)	0.001 ± 0.000	0.0009
angles (deg) (3957)	0.322 ± 0.007	0.3
impropers[d] (deg) (1140)	0.321 ± 0.014	0.29
XPLOR energies[e] (kcal mol^{-1})		
E_{NOE}	21.06 ± 3.325	13.39
E_{cdih}	2.86 ± 0.47	2.03
E_{imp}	18.47 ± 1.69	15.33
E_{angle}	62.84 ± 2.79	54.44
E_{bond}	4.3 ± 0.3	2.86
E_{vdw}	3.61 ± 0.45	20.15
E_{L-J}[f]	−711.78 ± 14.8	−709.3

atomic rms differences[g] (Å)	backbone[h]	all heavy atoms
residues 8−148	0.83 ± 0.09	1.60 ± 0.14
β-sheets regions[i]	0.39 ± 0.07	1.73 ± 0.18

[a] ⟨SA⟩ represents the final ensemble of 21 simulated annealing structures; ⟨SA⟩ $_{av}$ is the restrained minimized average structure obtained by averaging the 21 final structures over residues 8−148. Errors reported are ±1 standard deviation. The number of restraints is given in parentheses. [b] This includes 1510 NOE-derived distance restraints, 1 disulfide bond distance restraint, and 86 hydrogen-bond restraints (43 hydrogen bonds). [c] Dihedral angle restraints includes 53 ϕ-angle restraints, 70 ψ-angle restraints, and 48 χ^1-angle restraints. [d] Improper torsion angle restraints maintain peptide planarity and chirality. [e] The square-well NOE (E_{NOE}) using r^{-6} summing averaging, the restrained dihedral (E_{cdih}), the improper torsion angles (Ei_{mp}), the angle (E_{angle}), the bond (E_{bond}), and the quartic van der Waals repulsion energies (E_{vdw}) were calculated using force constants of 50 kcal mol^{-1}, 200 kcal mol^{-1} rad^{-2}, 500 kcal mol^{-1} rad^{-2}, 500 kcal mol^{-1} rad^{-2}, 1500 kcal mol^{-1} Å^{-1}, and 4 kcal mol^{-1} Å^{-1}. [f] E_{L-J} is the Lennard-Jones van der Waals energy. This term was not included in any of the structure-generating steps but was calculated for the final 21 structures and the restrained minimized average structure. [g] Atomic rms differences were calculated using the average structure before restrained minimization. [h] Atoms used were N, Ca, and C′. [i] This comprised residues 22−25, 30−34, 36−41, 52−56, 60−75, 77−87, 91−103, 106−119, 126−135, and 139−150 inclusive.

with S(χ1) values lower than 0.8. The Lennard-Jones energy of each of the accepted structures is large and negative (Table 1), indicating that no unfavorable van der Waals contacts exist. The stereochemical quality of the backbone coordinates for the ensemble of 21 structures was also checked using the program Procheck-NMR (*66*). For this ensemble, 98% of the residues lies in the allowed regions of the Ramachandran plot. Apart from the N- and C-termini, the regions which have the highest rms deviation from the average structure and lowest angular order parameters, $S(\phi)$ and $S(\psi)$, are the loops between residues 41−51, 86−92, 121−126, and 134−138. These regions also contain the few residues with main-chain dihedral angles that often fell outside of the allowed regions of the Ramachandran plot, namely, Ala90, Tyr91, Ala138, and Tyr139. The apparent disorder of these four segments of CBD_{N2} is not readily evident in ^{15}N relaxation studies (data not shown) and is thus attributed primarily to the detection of an insufficient number of long-range NOE distance restraints necessary for a precise and accurate structural determination.

Binding of CBD_{N2} to Cellooligosaccharides Measured by NMR Spectroscopy. The binding of cellotetraose and cellopentaose to CBD_{N2} was detected and quantified by the observation of ^1H and ^{15}N chemical shift changes in a series of ^1H-^{15}N HSQC spectra acquired in the presence of increasing amount of each cellooligosaccharide. Figure 6 summarizes the changes of the $^1H^N$ and ^{15}N chemical shifts of each amide in CBD_{N2} due to the binding of cellotetraose. Similar shift perturbations are seen with cellopentaose (not shown). The average change in the absolute value of the H^N chemical shifts of all resolved amides in CBD_{N2} due to cellotetraose binding was 0.025 ppm with a standard deviation of 0.03 ppm. In the ^{15}N dimension, the average absolute change in shift was 0.16 ppm with a standard deviation of 0.2 ppm. From Figure 6, it is evident that the majority of residues with chemical shift changes greater than these average values lie within or immediately adjacent to the β-sheet cleft of CBD_{N2} composed of strands A1−A5. This strongly indicates that the binding site for the two cellooligosaccharides lies on the face of the CBD formed by these strands. Furthermore, residues showing pronounced chemical shift changes upon the addition of sugar are located in all five β-strands that form this sheet. Therefore, the cellooligosaccharides are likely to bind across, and not parallel to, strands A1−A5. This conclusion is supported further by the observation of intermolecular proton NOE interactions between bound cellooligosaccharide and protein side chains located within β-sheet A in CBD_{N2} (data not shown). A very similar pattern of chemical shift changes was also observed in the cellooligosaccharide titrations of CBD_{N1} (Figure 6) (*6*).

Cellotetraose and cellopentaose bind to CBD_{N2} with a stoichiometry of one sugar molecule per protein molecule. This conclusion is supported by the following evidence. First, as exemplified in Figure 7A, all amides with H^N and ^{15}N chemical shifts that are perturbed upon cellotetraose binding show coincident titration curves. That discounts the possibility of multiple binding sites with differing affinities on the CBD. Furthermore, given that CBD_{N2} is monomeric in solution, it is unlikely that this protein domain has two or more distinct binding sites with equal affinities for the sugar ligands. Therefore, the simplest interpretation for the coincident titration curves is that cellooligosaccharides bind to CBD_{N2} at a single site and that each ^1H-^{15}N group reports the same association event. Second, plots of CBD_{N2} amide chemical shifts versus added cellopentaose show a plateau at approximately equal concentrations of total sugar and protein, indicating a binding stoichiometry of 1:1 (Figure 7B). Cellotetraose does not exhibit such a pronounced

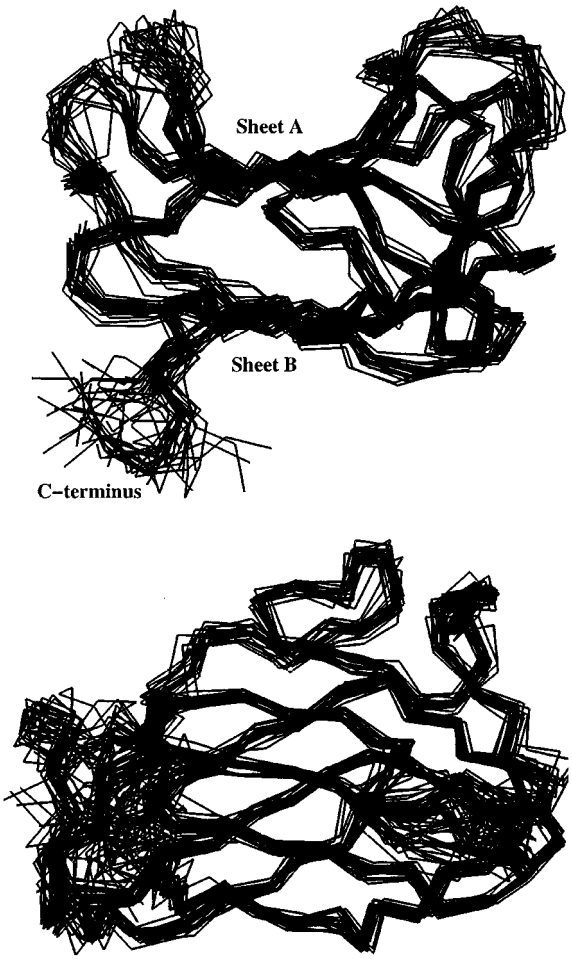

FIGURE 3: Two views of the C^{α} traces of the final ensemble of 21 structures calculated for CBD_{N2}, superimposed using the backbone atoms from residues 8 to 148. The presence of a binding cleft formed by β-sheet A is clearly evident in the top view, whereas the jellyroll β-sandwich topology is apparent in the lower view. The top view looks across the binding cleft of CBD_{N2}, whereas the lower view is rotate by approximately 90° to look down on β-sheet A.

FIGURE 4: Superimposition of the side chains that make up the hydrophobic core of CBD_{N2} on the C^{α} trace of the minimized average structure of this protein. All heavy atoms between residues 8 and 148 were used to superimpose the 21 structures. The CBD is oriented with the binding cleft, formed by β-sheet A, toward the page and β-sheet B closest to the reader.

titration end point due to its lower affinity for CBD_{N2}. However, cellotetraose and cellopentaose cause similar changes in the ^{1}H-^{15}N HSQC spectrum of the labeled protein, strongly suggesting that each bind CBD_{N2} with the same stoichiometry and at the same site. Third, the titration curves

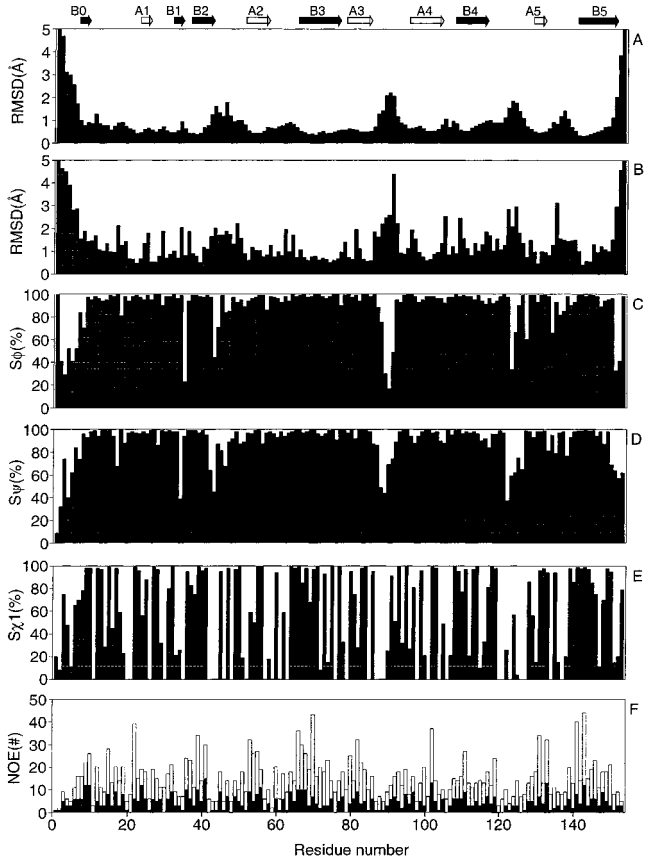

FIGURE 5: The rms deviation for all heavy atoms (A) and main-chain atoms (B) for the ensemble of 21 accepted structures with respect to the average structure obtained by superimposing residues 8−148. Angular order parameters Sϕ (C), Sψ (D), and Sχ1 (E) for the ϕ and ψ main chain, and the χ1 dihedral angles, respectively, observed in the final ensemble of 21 structures (*40*). Distribution of NOE restraints per residue (F). Filled bars represent the number of nontrivial intraresidue NOEs and open bars represent interresidue NOEs. For every restraint, the originating and destination residues were each counted once. The locations of the 11 β-strands are indicated on the top of the figure.

for the association of cellotetraose and cellopentaose with CBD_{N2}, measured by both NMR spectroscopy and isothermal titration microcalorimetry (ITC), are adequately fit to the binding isotherm describing the formation of a 1:1 ligand−protein complex (Table 2).

In the case of each cellooligosaccharide investigated, the free and bound forms of the protein are in fast exchange on the NMR time scale, resulting in the observation of population-weighted average chemical shifts throughout the titration series. This allows us to use the chemical shift changes as a direct measure of the bound fraction of the protein. The association constants (K_a) describing the interaction of cellotetraose and cellopentaose to CBD_{N2} were determined by nonlinear least-squares fitting of the chemical shift titration data to the binding isotherm for a protein with a single ligand recognition site (Figure 7) (*6*). The average association constants, measured with both the H^N and ^{15}N chemical shifts of residues Ser23, Tyr25, Gly52, Leu53, Tyr55, Val84, Gly85, Glu86, Gly87, Gly88, Thr93, Gln128, and Ala130, are 8500 ± 200 M^{-1} for cellotetraose and 30 700 ± 2000 M^{-1} for cellopentaose (Table 2). For comparison, the association constants of CBD_{N1} were 4200 ± 720 M^{-1} for cellotetraose and 34 000 ± 7600 M^{-1} for cellopentaose, as determined by NMR spectroscopy (*6*).

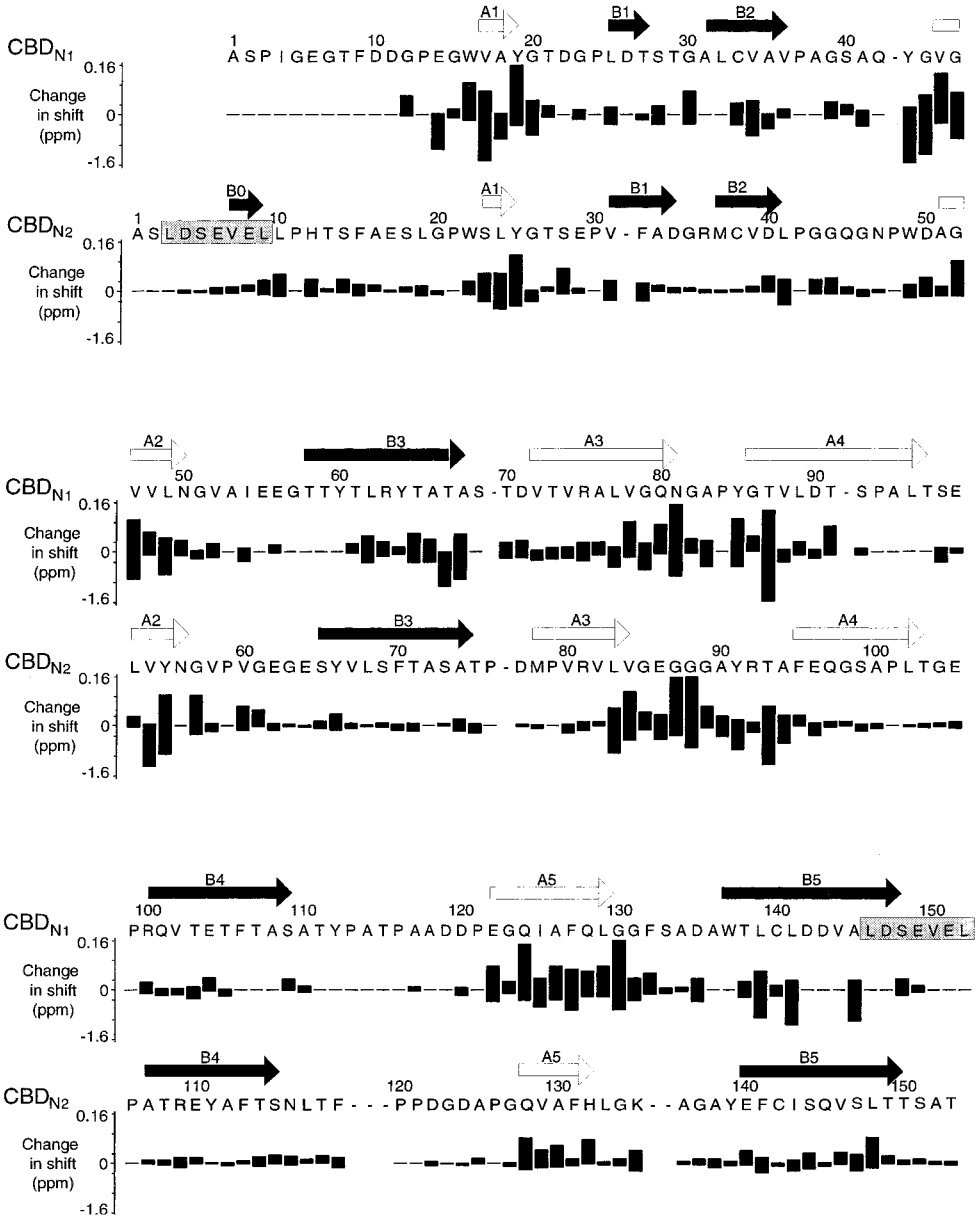

FIGURE 6: Patterns of NMR chemical shift perturbations demonstrate that cellotetraose binds to β-sheet A of both CBD_{N1} (*6*) and CBD_{N2}. The absolute values of the differences between the $^1H^N$ and ^{15}N chemical shifts of the main-chain amides in the free and cellotetraose-bound forms of the CBDs are indicated as positive and negative numbers, respectively. Similar patterns are observed with cellopentaose. The sequences of the two CBDs are aligned based on Tomme et al. (*1*, *6*) and the observed secondary structures of the two protein domains. The locations of the 10 β-strands in CBD_{N1} and 11 β-strands in CBD_{N2} are shown by arrows. The two β-sheets identified in both CBD_{N1} and CBD_{N2} are composed of strands A1-A4 (open arrows) and B1−B5 or B0−B5 (solid arrows). Within their native context in *C. fimi* CenC, the boxed residues at the C-terminus of isolated CBD_{N1} and the N-terminus of isolated CBD_{N2} are the same.

Binding of CBD_{N2} to Cellooligosaccharides Measured by ITC. The thermodynamic parameters governing the binding of cellooligosaccharides to CBD_{N2} were determined independently using ITC. As summarized in Table 2, this CBD forms 1:1 complexes with cellotetraose and cellopentaose, with measured association constants of $4500 \pm 100\ M^{-1}$ and $13\,000 \pm 100\ M^{-1}$, respectively. Furthermore, binding is controlled by a favorable standard enthalpy change, compensated in part by a significant decrease in entropy. Similar behavior was observed with the association of soluble sugar ligands to CBD_{N1} and indicates that binding is driven predominantly by a favorable combination of hydrogen-bonding and van der Waals interactions (*7*). The binding constants measured by NMR and calorimetry agree reasonably well, although clearly differ by more than the sums of their standard errors. This may result from the slightly different experimental conditions of pH and ionic strength used for the two analyses. Alternatively, note that these standard errors reflect the quality of the data fitting and do not include possible systematic differences such as variations in the purity of the samples of commercially available cellooligosaccharides used for the various studies (*73*) and difficulties in preparing, by weight, independent stock solutions of these sugars at defined concentrations.

Binding of CBD_{N2} to Polymeric Ligands. The binding of CBD_{N2} to various polymeric glucans was assayed using either affinity gel electrophoresis for soluble substrates (*7*) or sedimentation assays for insoluble substrates. Isolated CBD_{N2} binds Avicel, phosphoric acid swollen cellulose, carboxymethylcellulose, hydroxyethylcellulose, and barley β-glucan,

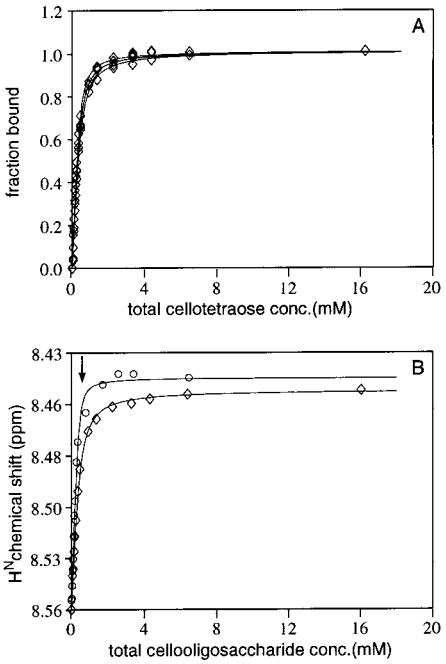

FIGURE 7: The association constants of CBD_{N2} for soluble cellooligosaccharides were determined from titration curves monitored by ^1H-^{15}N NMR spectroscopy. (A) Cellotetraose binds to CBD_{N2} at a single site. The coincident plots of the normalized H^N chemical shift changes for residues Tyr25, Gly57, Val84, and Gln128 versus total added cellotetraose demonstrate that each amide group in CBD_{N2} monitors the same binding event. The solid lines represent the titration isotherms obtained by fitting the observed data point (◇) to the equation describing the association of cellotetraose and CBD_{N2} to form a 1:1 protein-sugar complex. (B) The amide H^N chemical shift of Val84 is plotted as a function of the total concentration of added cellotetraose (◇) and cellopentaose (○). The solid line represents the best fit of the experimental data to the equilibrium equation describing binding to a single protein site. The arrow marking the plateau in the titration curve for cellopentaose fall at the point where the total sugar concentration equals the total protein concentration (~0.4 mM CBD_{N2}), indicating a 1:1 binding stoichiometry.

but not crystalline cellulose xylan or arabinan (data not shown). For all substrates tested thus far, CBD_{N2} and CBD_{N1} exhibit the same binding specificities and approximate affinities.

Histidine pK_a Values. CBD_{N2} contains two histidine residues, one of which (His132) lies within the binding cleft and the other (His12) near the edge of β-sheet B. On the basis of the ^1H and ^{15}N chemical shifts observed in a ^1H-^{15}H HMBC spectra, both residues are unambiguously in the neutral $N^{\epsilon 2}H$ tautomeric state at pH* 5.9. To determine the pK_a values of the imidazole side chains, pH titrations were carried out using a sample of CBD_{N2} in which all tryptophan, tyrosine, and phenylalanine aromatic rings were biosynthetically deuterated. After exchange of the amide protons for deuterons by reversible thermal unfolding and refolding in D_2O buffer, the downfield region of the ^1H NMR spectrum of this labeled protein contained signals from only the carbon-bonded $^1H^{\delta 2}$ and $^1H^{\epsilon 1}$ of the two histidine residues. This permitted the facile measurement of the pH dependence of the chemical shifts of these imidazole protons by one-dimensional ^1H NMR spectroscopy (Figure 8).

The chemical shifts of the $^1H^{\delta 2}$ and $^1H^{\epsilon 1}$ resonances of His12 move downfield by 0.29 and 1.00 ppm, respectively, upon titration from basic to acidic pH. This is clearly indicative of the protonation of the imidazole ring (*67*). On the basis of these data, the measured pK_a value of His12, without correction for isotope effects, is 5.1 ± 0.1 in the absence or presence of saturating quantities of cellohexaose.

In the absence of added sugar, the $^1H^{\delta 2}$ and $^1H^{\epsilon 1}$ resonances of His132 shift downfield by only 0.13 and 0.20 ppm upon reduction in pH and follow an apparent pK_a of 4.1 ± 0.2. Under these conditions, the $^1H^{\epsilon 1}$ line width of His132 is also relatively large (12 Hz versus 4 Hz for His12), indicative of conformational exchange broadening. This suggests that the small chemical shift changes observed in the titration measurements arise from an unusual feature of environment of His132 in CBD_{N2}, such as its partial burial along the binding cleft. Unfortunately, due to the limited solubility of this CBD under acidic conditions, we were unable to measure the ^{15}N shifts of His132 at low pHs with a ^1H-^{15}N HMBC experiment in order to verify that the observed spectral changes indeed arise from the protonation of the imidazole ring, rather than from indirect effect such as the titration of adjacent ionizable groups in the protein. Upon addition of cellohexaose, the $^1H^{\delta 2}$ and $^1H^{\epsilon 1}$ chemical shifts of His132 change and the $^1H^{\epsilon 1}$ line width sharpens to 4 Hz. This indicates that His132 is involved in the binding of cellooligosaccharides and suggests that any conformational mobility is restricted in the presence of added sugar. Furthermore, the apparent pK_a of His132 drops to 2.45 ± 0.1, with extrapolated chemical shift changes of 0.14 and 0.62 for $^1H^{\delta 2}$ and $^1H^{\epsilon 1}$, respectively.

DISCUSSION

Structure of CBD_{N2}. Using NMR methods, we have determined the structure of isolated CBD_{N2} in the presence of saturating concentrations of cellopentaose. As illustrated in Figure 9, CBD_{N2} is a monomeric protein domain, comprised of two antiparallel β-sheets (A and B) folded into a jellyroll β-sandwich topology. The overall structure is stabilized by a large hydrophobic core between the two β-sheets (Figure 4), as well as by a disulfide bond bridging strands B2 and B5. The strands forming β-sheet A are shorter than those forming sheet B, thus giving the molecule a "bean" shape with a distinct cleft or groove into which sugars bind. Although intermolecular NOEs between the residues within this cleft and bound cellopentaose were detected, these could not be unambiguously assigned due to severe overlap of the resonances from the sugar molecule. Accordingly, the cellooligosaccharide was not included in the structure calculations.

The individual β-strands in CBD_{N2} are well ordered as judged by low rmsd values in the ensemble of calculated structures and by ^{15}N relaxation measurements (not shown). In contrast, the exposed loops formed by residues 41−51, 86−92, 121−126, and 134−138 are less well-defined, with high rmsds and low angular order parameters (Figures 3 and 5). Because these loops form the extreme edges of β-sheet A, this results in a range of depths and widths calculated for the cellulose-binding cleft within the set of 21 accepted structures. As an estimation of this range, the distance between C^α atoms of Trp49 and Tyr91, located on opposite sides of the cleft, varies from 12 to 20 Å (the mean being 15.5 ± 2.4 Å) in the ensemble of structures (Figure 3). Preliminary ^{15}N relaxation measurements do not reveal

Table 2: Thermodynamic Parameters for the Binding of Cellooligosaccharides to CBD_{N2} and CBD_{N1}

	CBD_{N2}		CBD_{N1}[a]	
	cellotetraose	cellopentaose	cellotetraose	cellopentaose
NMR[b]				
K_a (M^{-1})	8.5 ($\pm$0.2) $\times 10^3$	3.07 ($\pm$.2) $\times 10^4$	4.2 ($\pm$0.7) $\times 10^3$	3.4 ($\pm$0.8) $\times 10^4$
ITC[c]				
K_a (M^{-1})	4.5 ($\pm$0.1) $\times 10^3$	1.30 ($\pm$0.01) $\times 10^4$	3.2 ($\pm$0.5) $\times 10^3$	2.1 ($\pm$0.3) $\times 10^4$
$\Delta G°$(kJ/mol)	$-$21.55 ($\pm$0.06)	$-$24.27 ($\pm$0.02)	$-$20.7 ($\pm$3.5)	$-$25.5 ($\pm$2.7)
$\Delta H°$(kJ/mol)	$-$37.2 ($\pm$0.2)	$-$42.1 ($\pm$0.1)	$-$40.4 ($\pm$1.7)	$-$53.0 ($\pm$1.3)
$T\Delta S°$(kJ/mol)	$-$9.7 ($\pm$0.2)	$-$17.8 ($\pm$0.1)	$-$19.7 ($\pm$3.8)	$-$27.6 ($\pm$3.5)
n	0.97	0.93	0.92	0.92

[a] Results for CBD_{N1} are taken from refs 6 and *7*. [b] Data obtained at 35 °C and pH* 5.9 in 50 mM sodium chloride, 50 mM potassium phosphate buffer, 0.02% sodium azide, and 10% D_2O/90% H_2O. Values in parentheses indicate the standard deviation of the fit. [c] Data obtained at 35 °C in potassium phosphate 50 mM, pH 7.0.

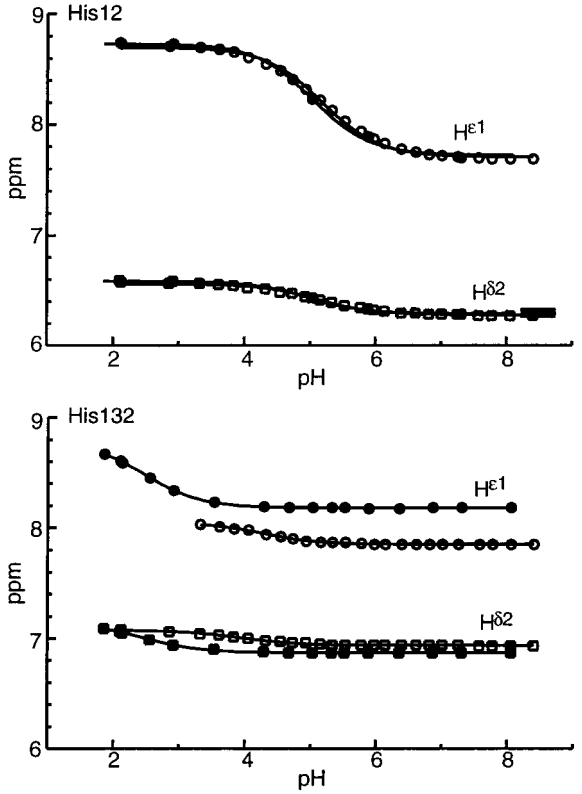

FIGURE 8: The apparent pK_a values of His12 (top) and His132 (bottom) in the absence (○, □) and presence (●, ■) of saturating quantities of cellohexaose at 35 °C were determined from pH-dependence of the $^1H^{\delta 2}$ (□, ■) and $^1H^{\epsilon 1}$ (○, ●) chemical shifts. Consistent with its location in the binding cleft of CBD_{N2}, the addition of cellohexaose perturbs both the chemical shifts and pK_a values of His132. The solid lines represent the best fit of the data points to the equation describing a single ionization monitored by NMR spectroscopy. The measured pH* values were not corrected for isotope effects due to the use of 99% D_2O buffer.

significant mobility on a subnanosecond time scale for the loop residues of CBD_{N2} in either the presence or the absence of cellopentaose (not shown). This implies that the apparent disorder of these residues, and hence the variation in the width of the binding cleft, results largely from the detection of an insufficient number of long-range NOE to adequately define the positions of the loops in CBD_{N2} (Figure 5).

Cellulose-Binding Mechanism of CBD_{N2}. CBD_{N2} binds soluble cellooligosaccharides with a stoichiometry of one sugar molecule per CBD molecule and in the order of increasing affinity cellotetraose < cellopentaose. On the basis of the patterns of amide ^{15}N and $^1H^N$ chemical shift perturbations resulting from the addition of cellooligosaccharides to CBD_{N2}, we demonstrate that exposed residues in β-sheet A (strands A1−A5) all interact with these soluble ligands (Figure 6). The identification of this β-sheet and the surrounding loops as the ligand-binding site of CBD_{N2} is confirmed by studies using a nitroxide spin-label derivative of cellotetraose (*13*), as well as by the observation of intermolecular NOE's between bound cellopentaose and residues located within this region of the protein. Although the ^1H resonances of the cellopentaose are currently unassigned, NOEs from unlabeled sugar protons to ^{13}C-labeled Ser23 $^1H^{\beta}$, Tyr25 $^1H^{\beta}$, $^1H^{\delta}$ and $^1H^{\epsilon}$, Trp49 $^1H^{\beta}$ and aromatics ^1Hs, Val54 $^1H^{\gamma 1}$ and $^1H^{\gamma 2}$, Asn56 $^1H^{\delta 1}$ and $^1H^{\delta 2}$, Leu83 $^1H^{\delta 1}$ and $^1H^{\delta 2}$, Tyr91 $^1H^{\beta}$, $^1H^{\delta}$, and $^1H^{\epsilon}$, and Gln128 $^1H^{\epsilon 1}$ and $^1H^{\epsilon 2}$ were detected using a ^{13}C ω_1-edited, ω_3-filtered HMQC-NOESY experiment (data not shown) (*31*).

Perhaps the most important feature of the structure of CBD_{N2} is that the cellulose-binding site is a groove or a cleft. This provides an immediate explanation for the specificity of CBD_{N2} toward amorphous, but not crystalline, forms of cellulose. That is, single strands of cellooligosaccharides and, by inference, soluble β-1,4-glucans and insoluble amorphous cellulose can bind within this cleft, whereas sugars within an ordered, crystalline array are excluded. On the basis of the pattern of amide chemical shift perturbations due to the addition of cellopentaose to CBD_{N2} (Figure 6), as well as the above listed intermolecular NOEs, these strands of cellulose must lie perpendicular to β-strands A1−A5 and thus lengthwise along the binding cleft. Consistent with the observed dependence of binding affinity on the degree of oligosaccharide polymerization (Table 2) (*6*), the length of the CBD_{N2} binding groove (~22 Å) is approximately equal to that of an extended cellopentaose chain (~23.5 Å).

The positions of the exposed side chains present on the binding face of the minimized structure of CBD_{N2} are illustrated in Figure 9. These amino acids can be classified into three distinct groups. First, a strip of hydrophobic residues composed of Val54, Ala130, and Leu83 lies along the center of the binding cleft. Second is a set of hydrophilic residues, including Ser23, Asn56, Gln128, Thr83, Glu96, Arg81, and His132, that flank this nonpolar strip. Finally, the aromatic rings of Tyr25, Tyr91, and Trp49 are located at the edges and top of the binding site. This distinctive pattern of amino acids leads to a structural model of cellulose binding to CBD_{N2} in which the pyranose rings of a single strand of β-1,4-linked glucose stack against the hydrophobic and aromatic residues, while the equatorial hydroxyls form

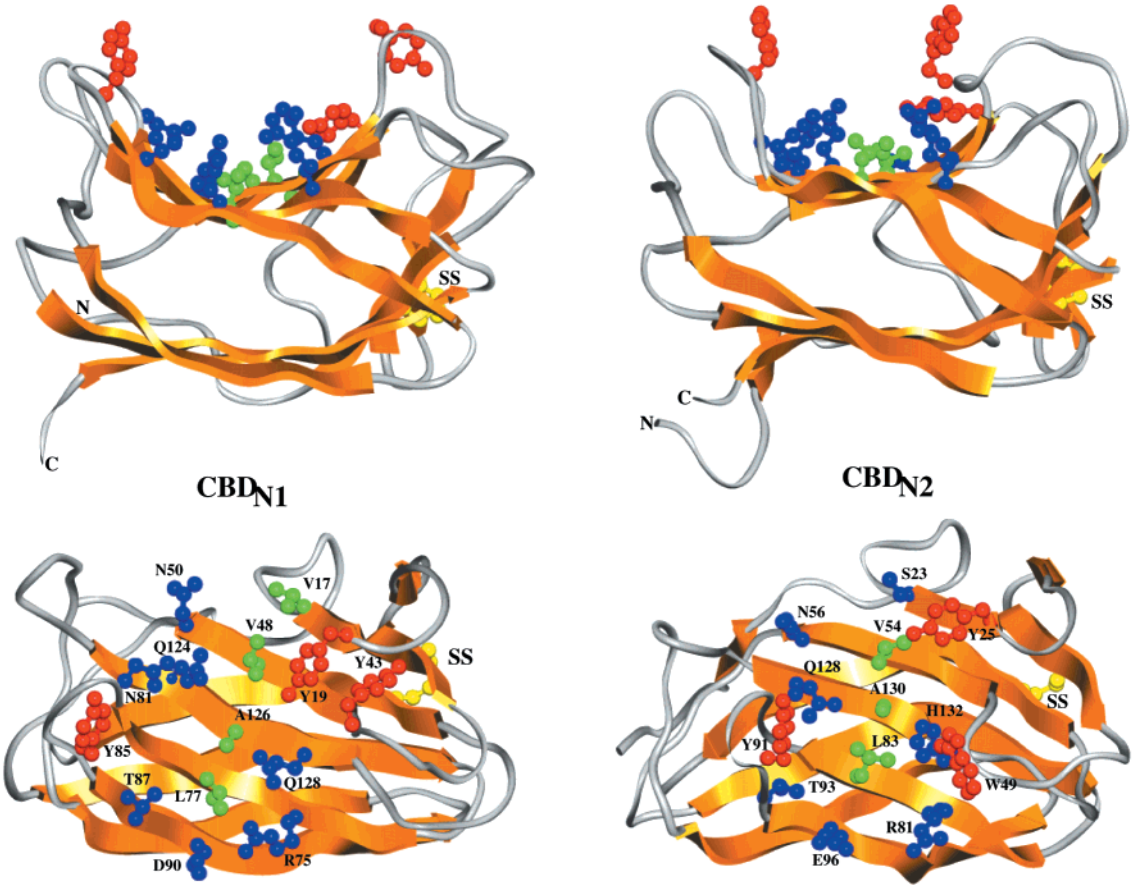

FIGURE 9: Ribbon diagrams of the minimized average structures of CBD_{N1} (left) and CBD_{N2} (right). Residues from β-sheet A of both CBDs implicated in oligosaccharide binding are shown in ball-and-stick format. Hydrophilic side chains are in blue, the hydrophobic side chains in green, and aromatic side chains in red. The disulfide bond between β-strands B2 and B5 of both CBDs is in yellow. The bottom view looks directly down on the binding face of each CBD, whereas the top view is rotated by 90° to emphasize the binding cleft. These two proteins share common jellyroll β-sandwich topology and can be aligned with an rmsd of 2.0 Å (β-strands) or 2.6 Å (central 141 residues).

hydrogen bonds with the flanking polar side chains. On the basis of their positions within the loop regions at the ends of the binding clefts, the two tyrosines may shift to cover the rings of the bound sugar. As reviewed by Vyas (*68*) and Quicho (*69*, *70*), such interactions are commonly observed at protein−sugar binding interfaces.

The proposed structural mechanism of cellulose binding is entirely consistent with the thermodynamic parameters governing the association of soluble cellooligosaccharides with CBD_{N2}. As summarized in Table 2, isothermal titration calorimetry reveals that sugar binding is enthalpically driven. The release of heat upon complex formation can be attributed to hydrogen bonding and van der Waals interactions between the sugar and protein. These favorable interactions are, however, partially offset by an overall decrease in entropy due to net loss in the conformational freedom of the oligosaccharide, protein side chains, and hydrating water upon binding.

An interesting feature of the cellulose-binding mechanism of CBD_{N2} is revealed by NMR studies of His132. In the absence of added cellopentaose, the imidazole $^1H^{\delta 2}$ and $^1H^{\epsilon 1}$ resonances of this histidine are unusually broad and have an apparent acid dissociation constant of 4.1 ± 0.2. This depressed pK_a value, relative to that of ~6.5 for a histidine in a random coil polypeptide (*71*), is attributed to the partial burial of the side chain of His132 in a nonpolar environment and to its separation from any other charged residues in CBD_{N2}. Specifically, the average accessible-surface-area (ASA) of the His132 side chain is 29.8 ± 7 Å^2 in the ensemble of 21 accepted structure or 15% of the total ASA of an isolated histidine side chain, while the average distance between His132 $^1H^{\epsilon 1}$ and its closest ionizable neighbors is 8.0 ± 1.7 Å to Arg81 C^{ζ} and 10.3 ± 1.6 Å to Glu96 C^{δ}. (Note, however, that these values were measured from the ensemble of structures determined for CBD_{N2} in the presence of cellopentaose.) In addition, the anomalous NMR signals from the ring protons of His132 suggest that its imidazole side chain is undergoing conformational motions on a millisecond time scale, which leads to exchange broadening (*72*). In the presence of saturating concentrations of cellohexaose, this motion appears suppressed as evident by a sharpening of the resonances of $^1H^{\delta 2}$ and $^1H^{\epsilon 1}$. Furthermore, the pK_a of His132 drops to a surprisingly low value of 2.45 ± 0.1. In combination with structural studies, which reveal that His132 lies along the wall of the CBD_{N2} binding cleft (Figure 9), these data suggest that cellooligosaccharides hydrogen bond directly to the imidazole ring, which is preferentially in its neutral $N^{\epsilon 2}H$ tautomeric form. In the structural ensemble, $N^{\epsilon 2}H$ is always exposed to the cleft and, thus, must serve as a hydrogen bond donor to a bound

oligosaccharide. Accordingly, protonation at the buried $N^{\delta 1}$ position of His132 will introduce a positive charge in the hydrophobic core of CBD_{N2}; this appears possible, albeit very unfavorable, as evident by the different NMR chemical shifts of this residue in the absence and presence at cellohexaose at both high and low pH values (Figure 8). Although not measured experimentally, the linkage of the deprotonation of His132 and the binding of cellooligosaccharides implies that the affinity of CBD_{N2} for cellohexaose will decrease by ~45-fold at 30 °C as the solution pH drops from neutral to highly acidic. This corresponds to a $\Delta\Delta G^{\circ}_{bind} = 2.303RT\Delta pK_a = 9.6$ kJ mol^{-1} for cellohexaose to CBD with His132 in its protonated versus deprotonated states. This prediction is consistent with a recent study by Linder et al. (*75*), who demonstrated that histidine-containing mutants of *Trichoderma reesei* CBD_{CBH1} bound to crystalline cellulose with pH-dependent affinity. Finally, by way of comparison, we note that the pK_a value of His12, which lies distal to the binding cleft of the CBD, is 5.1 ± 0.1 in the absence or presence of cellohexaose. The marginally depressed pK_a value is also attributed to the partial burial of the imidazole ring of this histidine along the edge of β-sheet B. The average side chain ASA of His12 is 45.4 ± 11.3 Å^2 over the 21 accepted structures or 23% of that of an isolated histidine side chain. However, in contrast to His132, $N^{\epsilon 2}H$ of His12 is buried while $N^{\delta 1}$ is exposed to the solvent and thus accessible for protonation.

At the present time, we have not confirmed the proposed binding model at an atomic level through a detailed structural analysis of the CBD_{N2}-cellopentaose complex due to the incomplete NMR spectral assignment of the sugar. Moreover, based on a quantitative analysis of the paramagnetic relaxation of the amide $^1H^N$ resonances of this CBD due to the nitroxide moiety of a bound TEMPO-cellotetraose molecule, we have shown that cellooligosaccharides associate with CBD_{N2}, as well as CBD_{N1}, in multiple orientations or polarities (*13*). That is, the reducing end of the sugar can be positioned near either edge (e.g., β-strand A1 or A4) of the CBD's binding cleft. The relative affinities for association in these two possible orientations are estimated to be within a factor of 5−10-fold. As discussed by Johnson et al. (*13*), this mode of protein−sugar complex formation can be attributed to an approximate symmetry in the positions of hydrogen-bonding groups both along a strand of cellulose and within the binding cleft of the CBD. Parenthetically, NMR relaxation studies of CBD_{N1} also suggest that the binding clefts of the CenC CBDs are relatively dynamic or plastic (in preparation). This dynamic behavior is exemplified by the conformational exchange broadening of His132 described for sugar-free CBD_{N2}. We speculate that the conformational flexibility of the CBDs may lead to their promiscuous association with a variety of β-1,4-glucans.

Global Comparison of CBD_{N1} and CBD_{N2}. CBD_{N2} is the second Family IV binding domain, after CBD_{N1} (*8*), for which a solution structure has been determined. At the level of their primary structures, these two CenC CBDs are clearly homologous, sharing 36.5% sequence identity and over 50% sequence similarity (*4*). The two proteins also adopt the same secondary structure, consisting of 10 major β-strands arranged into two β-sheets. Residues forming these β-strands are well conserved between CBD_{N1} and CBD_{N2}, with deletions or insertions found in the intervening loop regions (Figure 6). As discussed below, CBD_{N2} contains an additional small N-terminal β-sheet B0, which may represent a continuation of the C-terminal β-strand B5 from CBD_{N1} within the context of native CenC. At the level of their tertiary structures, the two β-sheets in each CBD adopt the identical jellyroll β-sandwich fold (Figure 9). Accordingly, the coordinates of the backbone atoms of the minimized average structures of CBD_{N1} and CBD_{N2} superimpose with a rms deviation of 2.0 Å (β-strands) or 2.6 Å (central 141 residues). As noted previously for CBD_{N1}, CBD_{N2} also shows structural similarity to the bacterial 1,3-1,4-β-glucanases (*8*).

A major structural difference between the two CenC domains is that CBD_{N1}, but not CBD_{N2}, binds calcium (*73*). Of the three postulated metal-chelating residues in CBD_{N1} (Thr8, Gly30, and Asp142), the threonine and glycine are conserved in CBD_{N2}, whereas the aspartic acid is replaced by a serine. The lack of this negatively charged carboxylate side chain undoubtedly contributes to the inability of CBD_{N2} to bind divalent metal ions. Detailed calorimetric studies have shown that calcium stabilizes CBD_{N1} against thermal denaturation, but does not alter its affinity for cellooligosaccharides (*73*). This is consistent with the location of the metal-binding site on the opposite face of the protein (β-strand B) as the sugar-binding cleft. Intriguingly, the midpoint unfolding temperatures for CBD_{N1} in its calcium-free and -loaded states are 46.7 and 55.2 °C, respectively (*73*), whereas that of CBD_{N2} is 67 °C. Therefore, CBD_{N2} is significantly more stable than CBD_{N1}, even though it does not bind a metal ion. However, reduction of the single disulfide bond in either CBD_{N1} or CBD_{N2} leads to complete unfolding under otherwise benign solution conditions (*64* and unpublished results). A comparison of the sequences and structures of the two CenC CBDs does not yield an obvious explanation for their pronounced difference in stability.

Comparison of the Cellulose-Binding Properties of CBD_{N1} and CBD_{N2}. As expected based on their conserved sequences and structures, the cellulose-binding clefts of CBD_{N1} and CBD_{N2} are very similar. In particular, both contain a central strip of hydrophobic residues, flanked by nonpolar side chains and two tyrosine rings. Accordingly, it is not unexpected that these two Family IV domains also exhibit the same binding specificity for cellooligosaccharides, amorphous cellulose, and a variety of soluble β-1,4-glucans (*7* and herein). Furthermore, as summarized in Table 2, the thermodynamic parameters governing the binding of CBD_{N1} and CBD_{N2} to cellotetraose and cellopentaose are very comparable. Thus, it is certain that the two CenC CBDs share the same basic mechanism of cellulose binding.

A detailed inspection of Figure 9, however, reveals that the β-sheet clefts of CBD_{N1} and CBD_{N2} are not identical. Specifically, the side chain of Asn81 of CBD_{N1} is absent in CBD_{N2} as the corresponding residue is a glycine. In addition, Val17 of CBD_{N1} is replaced by Ser23 in CBD_{N2}, Tyr43 by Trp49, and Gln128 by His132. Using site-directed mutagenesis, Kormos et al. (submitted for publication) demonstrated that Asn81 and Gln128 contribute approximately 3.6 and 4.3 kJ mol^{-1} toward the binding of PASA by CBD_{N1}. Note that a neutral histidine side chain with $N^{\epsilon 2}H$ exposed to the binding cleft can donate, but not accept, a hydrogen bond, whereas a glutamine could do both. Therefore, the expected reduction in the binding affinity of CBD_{N2} for cellooligosaccharides relative to that of CBD_{N1} due to the absence of the

asparagine and the replacement of glutamine by histidine could in part be compensated for by the valine to serine and/or tyrosine to tryptophan substitutions. That is, the serine could provide alternative hydrogen-bonding interactions, whereas the tryptophan could enhance aromatic-sugar stacking. Note that the smaller magnitude of the favorable ΔH°_{bind} measured for the binding of CBD_{N2} to cellooligosaccharides as compared to CBD_{N1} suggests that the former process indeed involves fewer hydrogen-bonding interactions (Table 2).

In addition to the obvious difference in their complement of binding cleft residues, a second more subtle difference between the structures of the two CenC CBDs is seen in the width of the cellulose-binding groove. Using the distance between the C^{α} atoms of Tyr43/Trp49 and Tyr85/Tyr91 as a measure of this width, we find values of 22.8 ± 1.2 Å for the ensemble of CBD_{N1} structures and 15.5 ± 2.4 Å for CBD_{N2}. This difference may simply reflect our inability to precisely define the conformations of the loop regions bordering the binding clefts of the two CBDs. However, it is noteworthy that β-strands A3, A4, and A5 are 2, 2, and 3 residues, respectively, shorter in CBD_{N2} than in CBD_{N1}. Such a variation in secondary structure would at least in part lead to the observed difference in the curvature of the sugar-binding sites in these two protein domains. On the basis of manual docking experiments, it follows that a cellooligosaccharide or single chain of cellulose may lie in a more tilted, rather than a flat, orientation across the binding cleft of CBD_{N2} relative to CBD_{N1}. The potential for subtle differences in the exact positioning of the sugars within the binding grooves of these proteins may also compensate for any differences in the measured binding affinity of CBD_{N1} and CBD_{N2} for cellooligosaccharides due to sequence variations. Clearly, detailed structural studies of the two isolated CenC CBDs complexed with selectively ^{13}C-labeled cellooligosaccharides will be required in order to compare and contrast their exact mechanisms of sugar binding.

Tandem CBD_{N1N2} of Native CenC. CBD_{N1} and CBD_{N2} are arranged in tandem at the N-terminus of native CenC to form a double binding domain, CBD_{N1N2}. Preliminary studies reveal that the $^{1}H/^{13}C/^{15}N$ NMR spectra of this 30 kDa protein is essentially the summation of the spectra measured for CBD_{N1} and CBD_{N2} in isolation. Thus, to the first approximation, the structures determined for the separate CBDs remain unperturbed when they are joined in tandem. However, as outlined in Figure 6, there is no discrete linker region between the two domains within their native context. Due to the exact boundaries chosen for the cloning and expression of the individual proteins, the seven C-terminal residues of CBD_{N1} (Leu146−Leu152) correspond to the seven N-terminal residues of CBD_{N2} (Leu3−Leu9). Moreover, these overlapping residues are in part structured as β-strands in both of the isolated domains. Specifically, Leu146−Ser148 form the end of β-strand B5 in CBD_{N1}, while residues Val7−Leu9 form the small, albeit poorly defined, β-strand B0 of CBD_{N2}. It is therefore plausible that B0 of CBD_{N2} is actually a continuation of B5 in CBD_{N1}, such that the domains of CBD_{N1N2} are joined in tandem by a continuous β-strand. If so, this could certainly restrict, in a structural and dynamic sense, the relative orientations of CBD_{N1} and CBD_{N2} such that their binding clefts would be positioned roughly side-by-side and not end-to-end (see Figure 9). Such an alignment of the individual domains could explain the observation that CBD_{N1N2} binds phosphoric acid swollen cellulose with a relative affinity that is only approximately 2-fold greater than that of CBD_{N1} alone (*7*). That is, the individual CBDs appear to bind amorphous cellulose in an additive, rather than cooperative fashion. This apparent lack of cooperativity suggests that the CBD_{N1} and CBD_{N2} are restrained such that they cannot bind simultaneously to adjacent sites on the same polymeric sugar chain. Further structural and thermodynamic studies of the two CenC CBDs, both in tandem and in isolation, are underway in order to understand the role played by these two Family IV binding domains in facilitating the hydrolysis of cellulose by this *C. fimi* endoglucanase.

ACKNOWLEDGMENT

We are indebted to Lewis Kay for providing NMR pulse sequences and advice. We also thank Douglas Kilburn and Antony Warren for insightful discussions, Koman Joe and Sarah Everts-Hammond for help with protein preparation, and L. Nina for inspiration.

SUPPORTING INFORMATION AVAILABLE

Two figures, S1 and S2, summarizing the definition of the secondary structure of CBD_{N2} based on amide hydrogen exchange rates, $^{3}J_{HN-H\alpha}$ coupling constants, secondary chemical shifts, and NOE interactions, are presented. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Tomme, P., Warren, R. A. J., Miller, R. C., Jr., Kilburn, D. G., and Gilkes, N. R. (1995) in *Enzymatic Degradation of Insoluble Polysaccharides* (Saddler, J. N., and Penner, M., Eds.) pp 142−161, American Chemical Society, Washington DC.
2. Beguin, P., and Aubert, J.-P. (1994) *FEMS Microbiol. Rev. 13*, 25−58.
3. Henrissat, B. (1994) *Cellulose 1*, 169−196.
4. Tomme, P., Warren, R. A. J., and Gilkes, N. R. (1995) *Adv. Microbial Physiol. 37*, 1−81.
5. Coutinho, J. B., Gilkes, N. R., Warren, R. A. J., Kilburn, D. G., and Miller, R. C., Jr. (1992) *Mol. Microbiol. 6*, 1243−1252.
6. Johnson, P. E., Tomme, P., Joshi, M. D., and McIntosh, L. P. (1996) *Biochemistry 35*, 13895−13906.
7. Tomme, P., A. L., C., Kilburn, D. G., and Haynes, C. A. (1996) *Biochemistry 35*, 13885−13894.
8. Johnson, P. E., Joshi, M. D., Tomme, P., Kilburn, D. G., and McIntosh, L. P. (1996) *Biochemistry 35*, 14381−14394.
9. Kraulis, P. J., Clore, G. M., Nilges, M., Jones, T. A., Petterson, G., Knowles, J., and Gronenborn, A. M. (1989) *Biochemistry 28*, 7241−7257.
10. Xu, G.-Y., Ong, E., Gilkes, N. R., Kilburn, D. G., Muhandiram, D. R., Harris-Brandts, M., Carver, J. P., Kay, L. E., and Harvey, T. S. (1995) *Biochemistry 34*, 6993−7009.
11. Tormo, J., Lamed, R., Chirino, A. J., Morag, E., Bayer, E. A., Shoham, Y., and Steitz, T. A. (1996) *EMBO J. 15*, 5739−5751.
12. Brun, E., Morland, F., Gans, P., Blackledge, M. J., Barras, F., and Marion, D. (1997) *Biochemistry 36*, 16074−16086.
13. Johnson, P. E., Brun, E., Mackenzie, L., Withers, S. G., and McIntosh, L. P. (1999) *J. Mol. Biol. 287* (2), 609−625.
14. Graham, R. W., Greenwood, J. M., Warren, R. A. J., Kilburn, D. G., and Trimbur, D. E. (1995) *Gene 158*, 51−54.
15. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463−5467.

16. Yanisch-Perron, C., Vieira, J., and Messing, J. (1985) *Gene 33*, 103−119.
17. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: a Laboratory Manual*, 2nd ed., Cold Spring Harbour Laboratory Press, Plainview, NY.
18. McIntosh, L. P., and Dahlquist, F. W. (1990) *Q. Rev. Biophys. 23*, 1−38.
19. Neri, D., Szyperski, T., Otting, G., Senn, H., and Wüthrich, K. (1989) *Biochemistry 28*, 7510−7516.
20. McIntosh, L. P., Wand, A. J., Lowry, D. F., Redfield, A. G., and Dahlquist, F. D. (1990) *Biochemistry 29*, 6341−6362.
21. Edelhoch, H. (1967) *Biochemistry 6*, 1948−1954.
22. Pace, C., Vajdos, F., Fee, L., Grimsley, G., and Gray, M. (1995) *Protein Sci. 4*, 2411−2432.
23. Kay, L., Keifer, P., and Saarinen, T. (1992) *J. Am. Chem. Soc. 114*, 10663−10665.
24. Muhandiram, D., and Kay, L. (1994) *J. Magn. Reson., Ser. B 103*, 203−216.
25. Grzesiek, S., and Bax, A. (1993) *J. Am. Chem. Soc. 115*, 12593−12594.
26. Zhang, O., Kay, L. E., Olivier, J. P., and Foreman-Kay, J. D. (1994) *J. Biomol. NMR 4*, 845−858.
27. Levy, G. C., and Lichter, R. L. (1979) *Nitrogen-15 Nuclear Magnetic Resonance Spectroscopy*, J. Wiley & Sons, New York.
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. Shrager, R. I., Cohen, J. S., Heller, S. R., Sachs, D. H., and Schechter, A. N. (1972) *Biochemistry 11*, 541−547.
30. Brünger, A. T. (1992) *X-PLOR: A System for X-ray Crystallography and NMR*, Yale University Press: New Haven, CT.
31. Lee, W., Revington, M. J., Arrowsmith, C., and Kay, L. E. (1994) *FEBS Lett. 350*, 87−90.
32. Clore, G., Gronenborn, A., Nilges, M., and Ryan, C. (1987) *Biochemistry 26*, 8012−23.
33. Vuister, G. W., Wang, A. C., and Bax, A. (1993) *J. Am. Chem. Soc. 115*, 5334−5335.
34. Gagne, S., Tsuda, S., Li, M. X., Chandra, M., Smillie, L. B., and Sykes, B. D. (1994) *Protein Sci. 3*, 1961−1974.
35. Grzesiek, S., Vuister, G. W., and Bax, A. (1993b) *J. Biomol. NMR 3*, 487−493.
36. Archer, S. J., Ikura, M., Sporn, M. B., Torcia, D. A., and Bax, A. (1991) *J. Magn. Reson. 95*, 636−641.
37. Marion, D., Driscoll, P., Kay, L., Wingfield, P., Bax, A., Gronenborn, A., and Clore, G. (1989a) *Biochemistry 28*, 6150.
38. Pelton, J. G., Torchia, D. A., Meadow, N. D., and Roseman, S. (1993) *Protein Sci. 2*, 543−558.
39. Laskowski, R. A., Rullma, J. A. C., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) *J. Biomol. NMR 8*, 477−486.
40. Hyberts, S. G., Goldberg, M. S., Havel, T. F., and Wagner, G. (1992) *Protein Sci 1*, 736−751.
41. Wittekind, M., and Müeller, L. (1993) *J. Magn. Reson., Ser. B 101*, 201−205.
42. Grzesiek, S., and Bax, A. (1992) *J. Am. Chem. Soc. 114*, 6291−6293.
43. Kay, L. (1993) *J. Am. Chem. Soc. 115*, 2055−2057.
44. Olejniczak, E. T., and Fesik, S. W. (1994) *J. Am. Chem. Soc. 116*, 2215−2216.
45. Marion, D., Kay, L. E., Sparks, S. W., Torcia, D. A., and Bax, A. (1989) *J. Am. Chem. Soc. 111*, 1515−1517.
46. Zuiderweg, E. R., and Fesik, S. W. (1989) *Biochemistry 28*, 2387−2391.
47. Pascal, S. M., Muhandiram, D. R., Yamazaki, T., Kay, J. D. F., and Kay, L. E. (1994) *J. Magn. Reson. 103*, 197−201.
48. Wüthrich, K. (1986) *NMR of Proteins and Amino Acids*, Wiley, New York.
49. Montelione, G. T., Lyons, B. A., Emerson, S. D., and Tashiro, M. (1992) *J. Am. Chem. Soc. 114*, 10974−10975.
50. Grzesiek, S., Anglister, J., and Bax, A. (1993) *J. Magn. Reson., Ser. B 101*, 114−119.
51. Bax, A., Clore, M., and Gronenborn, A. M. (1990) *J. Magn. Reson. 88*, 425−431.
52. Kay, L., Xu, G.-Y., Singer, A., Muhnadiram, D., and Forman-Kay, J. (1993) *J. Magn. Reson. 101*, 333−337.
53. Yamazaki, T., Foreman-Kay, J. D., and Kay, L. (1993) *J. Am. Chem. Soc. 115*, 11054−11055.
54. Slupsky, C. M., Gentile, L. N., and McIntosh, L. P. (1998) *Biochem. Cell Biol. 76*, 379−390.
55. Santoro, J., and King, G. C. (1992) *J. Magn. Reson. 97*, 202−207.
56. Vuister, G. W., and Bax, A. (1992) *J. Magn. Reson. 98*, 428−435.
57. Powers, R., Garrett, D., March, C., Frieden, E., Gronenborn, A., and Clore, G. (1993) *Biochemistry 32*, 6744−6762.
58. McIntosh, L. P., Brun, E., and Kay, L. E. (1997) *J. Biomol. NMR 9*, 306−312.
59. Wishart, D., Richards, F., and Sykes, B. (1992) *Biochemistry 31*, 1647−1651.
60. Wishart, D., and Sykes, B. (1994) *J. Biomol. NMR 4*, 171−180.
61. Hutchinson, E. G., and Thornton, J. M. (1996) *Protein Sci. 5*, 212−220.
62. Chan, A. W. E., Hutchnison, E. G., Harris, D., and Thornton, J. M. (1993) *Protein Sci. 2*, 1574−1590.
63. Brandon, C., and Tooze, J. (1991) *Introduction to Protein Structure*, New York.
64. Creagh, A. L., Koska, J., Johnson, P. E., Tomme, P., Joshi, M. D., McIntosh, L. P., Kilburn, D. G., and Haynes, C. A. (1998) *Biochemistry 37*, 3529−3537.
65. Nilges, M., Clore, G. M., and Gronenborn, A. M. (1988) *FEBS Lett. 229*, 317−324.
66. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
67. Markley, J. L. (1975) *Acc. Chem. Res. 8*, 70−80.
68. Vyas, N. K. (1991) *Curr. Opin. Struct. Biol. 1*, 732−740.
69. Quiocho, F. A. (1986) *Annual Review of Biochemistry 55*, 287−315.
70. Quiocho, F. (1989) *Pure Appl. Chem. 61*, 1293−1306.
71. Creighton, T. E. (1993) *Proteins: Structures and Molecular Principles*, 2nd ed., W. H. Freeman, New York.
72. Sandström, J. (1982) *Dynamic NMR Spectroscopy*, Academic Press, New York.
73. Johnson, P. E., Creagh, L. A., Brun, E., Joe, K., Tomme, P., Haynes, C. A., and McIntosh, L. P. (1998) *Biochemistry 37*, 12772−12781.
74. Linder, M., Nevanen, T., and Teeri T. T. (1999) *FEBS Lett. 447*, 13−16.

BI992079U