Structure of the N-Terminal Cellulose-Binding Domain of *Cellulomonas fimi* CenC Determined by Nuclear Magnetic Resonance Spectroscopy[†,‡]

Philip E. Johnson,[§] Manish D. Joshi,[§] Peter Tomme,[∥] Douglas G. Kilburn,[∥] and and Lawrence P. McIntosh*[,§]

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

Received July 3, 1996; Revised Manuscript Received August 24, 1996[⊗]

ABSTRACT: Multidimensional heteronuclear nuclear magnetic resonance (NMR) spectroscopy was used to determine the tertiary structure of the 152 amino acid N-terminal cellulose-binding domain from *Cellulomonas fimi* 1,4-β-glucanase CenC (CBD_{N1}). CBD_{N1} was studied in the presence of saturating concentrations of cellotetraose, but due to spectral overlap, the oligosaccharide was not included in the structure calculations. A total of 1705 interproton nuclear Overhauser effect (NOE), 56 ϕ, 88 ψ, 42 χ1, 9 χ2 dihedral angle, and 88 hydrogen-bond restraints were used to calculate 25 final structures. These structures have a rmsd from the average of 0.79 ± 0.11 Å for all backbone atoms excluding disordered termini and 0.44 ± 0.05 Å for residues with regular secondary structures. CBD_{N1} is composed of 10 β-strands, folded into two antiparallel β-sheets with the topology of a jelly-roll β-sandwich. The strands forming the face of the protein previously determined by chemical shift perturbations to be responsible for cellooligosaccharide binding [Johnson, P. E., Tomme, P., Joshi, M. D., & McIntosh, L. P. (1996) *Biochemistry 35*, 13895−13906] are shorter than those forming the opposite side of the protein. This results in a 5-stranded binding cleft, containing a central strip of hydrophobic residues that is flanked on both sides by polar hydrogen-bonding groups. The presence of this cleft provides a structural explanation for the unique selectivity of CBD_{N1} for amorphous cellulose and other soluble oligosaccharides and the lack of binding to crystalline cellulose. The tertiary structure of CBD_{N1} is strikingly similar to that of the bacterial 1,3−1,4-β-glucanases, as well as other sugar-binding proteins with jelly-roll folds.

Proteins and enzymes that bind oligosaccharides are ubiquitous in nature. Carbohydrates are used for fuel storage for cells, as well as for forming the structural components of bacterial and plant cell walls. As a result, carbohydrate−protein interactions are involved in numerous diverse functions ranging from adhesion at the cell surface, mediated by proteins such as lectins, to the degradation of oligosaccharides by enzymes such as lysozyme. These interactions generally involve hydrogen-bond formation, coupled with van der Waals interactions, such as stacking of aromatic side chains of the protein against the pyranosyl rings of the sugar (Quiocho, 1986, 1989).

One class of enzymes involved in oligosaccharide degradation are the cellulases. These β-1,4-glycosidases are found in many species of bacteria and fungi and are characterized by a modular design with binding to the substrate, cellulose, often mediated by one or more cellulose-binding domains (CBDs;[1] Beguin & Aubert, 1994; Henrissat, 1994; Tomme et al., 1995a). Over 120 different CBDs have been identified and classified into 10 families on the basis of their sequence similarities (Tomme et al., 1995b). These binding domains, which are found at either terminus of the enzymes as well as internally, range in size from the small (36 residues) fungal CBDs to the larger bacterial CBDs that can be up to 200 residues long. In addition, CBDs display a spectrum of specificities for the crystalline and amorphous forms of cellulose and related soluble polysaccharides.

Currently, there are two published CBD structures, the first of which is for the family I CBD found in cellobiohydrolase 1 from the fungus *Trichoderma reesei* (CBD_{CBH1}; Kraulis et

[†] This work was supported by a grant from the Protein Engineering Network of Centres of Excellence (L.P.M.). P.E.J. is the recipient of a MacMillan-Bloedel Limited Roger and Wiewel Fellowship in Wood Chemistry.

[‡] The coordinates of the minimized average structure (1ULO) and the ensemble of 25 structures (1ULP) and the X-PLOR restraint tables (1ULOMR) have been deposited in the Brookhaven Protein Data Bank.

* Author to whom correspondence should be addressed at Department of Biochemistry and Molecular Biology, 2146 Health Sciences Mall, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z3.

[§] Protein Engineering Network of Centres of Excellence and Department of Chemistry and Department of Biochemistry and Molecular Biology.

[∥] Protein Engineering Network of Centres of Excellence, Biotechnology Laboratory, and Department of Microbiology and Immunology.

[⊗] Abstract published in *Advance ACS Abstracts,* October 15, 1996.

[1] Abbreviations: CBD, cellulose-binding domain; CBD_{CBH1}, the cellulose-binding domain from *Trichoderma 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_{N1N2}, the tandem cellulose-binding domains from *Cellulomonas fimi* β-1,4-glucanase CenC; CT-HSQC, constant-time heteronuclear single quantum correlation; DQF-COSY, double-quantum-filtered correlation spectroscopy; DSS, 2,2-dimethyl-2-silapentane-5-sulfonic acid, sodium salt; DTNB, 5,5′-dithiobis(2-nitrobenzoic acid); FT-IR, Fourier transform infrared; HMQC, heteronuclear multiple quantum correlation; HSQC, heteronuclear single quantum correlation; INEPT, insensitive nuclei enhanced by polarization transfer; IPTG, isopropyl β-D-thiogalactopyranoside; 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; TOCSY, total correlation spectroscopy.

S0006-2960(96)01612-1 CCC: $12.00

al., 1989). The second is that of the family II CBD from the mixed function xylanase−glucanase Cex secreted by the bacterium *Cellulomonas fimi* (CBD_{Cex}; Xu et al., 1995). Both of these proteins are composed of antiparallel β-sheets that form a flat binding "surface" containing three exposed tyrosine or tryptophan rings, respectively. These aromatic rings are essential for the ability of the CBDs to bind cellulose (Reinikainen et al., 1992, 1995; Poole et al., 1993; Din et al., 1994; Linder et al., 1995; Bray et al., 1996).

The family IV CBDs from the *C. fimi* 1,4-β-glucanase CenC are unusual in two major respects. First, although repeated CBDs do occur in several glucanases, either internally (family IIIa) or at their C-termini (family VI, IX), CenC contains two binding domains (CBD_{N1} and CBD_{N2}) located in tandem at its N-terminus (Coutinho et al., 1992). Second, CBD_{N1} and CBD_{N1N2} are distinct among CBDs in that both bind phosphoric acid-swollen cellulose and soluble oligosaccharides but not crystalline cellulose (Coutinho et al., 1992; Johnson et al., 1996; Tomme et al., 1996). In an effort to provide an explanation for this binding selectivity, conformational studies of CBD_{N1} were carried out using NMR spectroscopy.

In this report, we describe the three-dimensional structure of CBD_{N1}. The tertiary fold of this CBD is strikingly different from those of CBD_{Cex} and CBD_{CBH1} in that it is a jelly-roll β-sandwich containing a binding cleft, not a flat surface. Furthermore, this cleft contains a central strip of nonpolar residues, flanked on both sides by polar hydrogen-bonding side chains. The nature of these amino acid residues is consistent with data from calorimetric measurements, which revealed that association with soluble sugars is enthalpically driven (Tomme et al., 1996). The presence of a binding groove also explains the selectivity of CBD_{N1} for single strands of cellooligosaccharides and its lack of affinity toward the flat surface of crystalline cellulose. Finally, we note that the tertiary structure of CBD_{N1} is similar to that of the bacterial 1,3−1,4-β-glucanases, suggesting an evolutionary relationship between the family IV CBDs and this class of enzymes.

EXPERIMENTAL PROCEDURES

Sample Preparation. Samples of uniformly (∼99%) ^{15}N- and $^{13}C/^{15}N$-labeled CBD_{N1} were produced by expression of the plasmid pTugN1n in *Escherichia coli* JM101 cells, as described previously (Johnson et al., 1996). In the case of the $^{13}C/^{15}N$-labeled protein, the growth medium contained 2 g/L [$^{13}C_6$]glucose and 1 g/L $^{13}C/^{15}N$ Isogro algal extract (Isotec, Inc.). Since the JM101 cells containing pTugN1n grew poorly in minimal media, CBD_{N1} nonrandomly fractionally ^{13}C-labeled at a level of 10% (Neri et al., 1989) was produced by initially growing bacteria in M9 medium (Miller, 1972) supplemented with 1 g/L unlabeled Isogro. At an OD_{600} of 1.0, the cells were spun down, washed with M9 medium, and resuspended in M9 medium containing 0.3 g of 99% [$^{13}C_6$]glucose and 2.7 g of unlabeled glucose as the sole carbon sources. IPTG (0.5 mM) was then added to induce expression of the gene encoding CBD_{N1}. The cells were grown at 30 °C for an additional 16 h, and the secreted protein was purified as outlined previously (Johnson et al., 1996). The amides of Tyr, Leu, Asp/Asn, and Val were selectively α-^{15}N-enriched using the protocol of McIntosh and Dahlquist (1990). CBD_{N1} samples with selectively deuterated aromatic rings were obtained from a synthetic medium containing 100 mg/L L-[δ_1,ϵ_2,$\zeta_{1,2}$,η_2-$^2H_5$]tryptophan and 100 mg/L of either L-[$\delta_{1,2}$,$\epsilon_{1,2}$,ζ-$^2H_5$]phenylalanine or [$\delta_{1,2}$,$\epsilon_{1,2}$-$^2H_4$]tyrosine (Cambridge Isotope Laboratories and Isotec Inc.) (McIntosh et al., 1990; McIntosh & Dahlquist, 1990). Characterization of the expressed CBD_{N1} by N-terminal amino acid sequencing, ultracentrifugation, mass spectrometry, and ligand-binding assays was described by Tomme et al. (1996) and Johnson et al. (1996). The CBD is a monomer under the experimental conditions of this study.

Samples of CBD_{N1} for NMR analysis were exchanged into 50 mM sodium chloride, 50 mM potassium phosphate (pH* 5.9), 0.02% sodium azide, and 10% D_2O/90% H_2O using ultrafiltration through a cellulose-free membrane (Filtron). Samples in deuterated buffer were obtained by twice lyophilizing the CBD_{N1} and redissolving in an equivalent amount of 99.9% D_2O. Typical protein concentrations were 2 mM as determined by $\epsilon_{280} = 21\,370\ M^{-1}\ cm^{-1}$. With the exception of the selectively ^{15}N-labeled proteins, CBD_{N1} samples contained up to a 40-fold molar excess of cellotetraose (Seikagaku Corp.) to prevent exchange broadening of several amide resonances (Johnson et al., 1996).

NMR Spectroscopy. NMR spectra were recorded on a Varian Unity 500 MHz spectrometer equipped with a triple resonance probe and a pulsed field gradient accessory. ^1H chemical shifts were referenced to an internal 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 was referenced to external 2.9 M $^{15}NH_4Cl$ in 1 M HCl at 24.93 ppm (Levy & Lichter, 1979). This latter reference yields ^{15}N chemical shifts 1.6 ppm greater than those obtained using liquid NH_3 (Wishart et al., 1995).

All spectra were collected at 35 °C and analyzed using a combination of FELIX v2.30 (Biosym Technologies, San Diego, CA), NMRPipe (Delaglio et al., 1995), and PIPP (Garrett et al., 1991). Experiments with $^1H^N$ detection were recorded using the enhanced-sensitivity pulsed field gradient approach of Kay et al. (1992) and Muhandiram and Kay (1994). Selective water flip-back pulses were incorporated to minimize the perturbation of the bulk water magnetization (Grzesiek & Bax, 1993; Zhang et al., 1994). Quadrature detection was accomplished using the States−TPPI method (Marion et al., 1989a). The initial delays in most of the indirectly detected dimensions were set to 1/(2*sw), resulting in a 180° first-order phase shift across the transformed spectrum and the inversion of aliased peaks (Bax et al., 1991). A summary of the data collection and processing parameters for the NMR experiments used to determine the structure of CBD_{N1} is given in Table S1 of the supporting information.

Amide Hydrogen Exchange. Amide hydrogen exchange rates were determined by recording a series of sensitivity-enhanced gradient ^1H-^{15}N HSQC spectra at 10, 29, 59, 95, 131, 266, 451, 688, 962, 5377, 10 790, and 25 211 min after lyophilized CBD_{N1} was dissolved in D_2O. To minimize the quantity of residual H_2O, the uniformly ^{15}N-labeled protein was lyophilized twice, being resuspended in D_2O after each freeze-drying step. The buffer concentration and pH were held constant by maintaining the sample volume.

Structure Calculations. All structure calculations were performed using X-PLOR 3.1 (Brünger, 1992). Distance restraints from NOE experiments were tabulated with in-

house programs available from http://otter.biochem.ubc.ca/www/nmrtools.html. Initially, a preliminary fold for the CBD_{N1} was calculated following the hybrid distance geometry/simulated annealing protocol using NOE restraints involved in β-strand pairing, unambiguous NOE restraints identified from 3 and 4D heteronuclear NOESY spectra, and dihedral angle restraints. This preliminary structure was used in a reiterative fashion to assign additional NOE interactions. A total of 1705 NOE-derived distance restraints were used in the final calculation of an ensemble of 60 structures. This data set was comprised of 711 nontrivial intraresidue, 411 sequential, 90 short-range ($1 < |i - j| \leq 4$), and 463 long-range ($|i - j| > 4$) distance restraints. Interproton distances were assigned to three strengths following a square-well potential energy function: weak 1.8−5.0 Å, medium 1.8−3.5 Å, and strong 1.8−2.9 Å. A correction of 0.5 Å was added to the upper bounds of restraints involving methyls (Clore et al., 1987). The distance ranges for the ^{15}N NOESY-HSQC were calibrated using the intensities of the $^1H^N_i$−$^1H^\alpha_{i-1}$ NOEs, which should be strong (∼2.2 Å) in regions of β-strand conformation, and cross-strand $^1H^\alpha_i$−$^1H^N_j$ NOEs, which in antiparallel β-sheets should have a medium intensity (∼3.2 Å). For the simultaneous 3D ^{13}C/^{15}N NOESY-HSQC spectra (Pascal et al., 1994), these sequential NOEs, as well as the strong cross-strand $^1H^\alpha_i$−$^1H^\alpha_j$ NOEs (∼2.3 Å), were used to calibrate the intensity ranges. For the 2D NOESY spectra involving Phe and Tyr aromatic rings, the $^1H^\delta$−$^1H^\epsilon$ NOEs were used for calibration. In the case of the 4D ^{13}C−^{13}C NOESY, a 150-ms mixing time was used, resulting in extensive spin diffusion. As a result of this complication, all NOEs extracted from this experiment were classified as weak. In addition, 88 hydrogen-bond restraints (44 hydrogen bonds), 56 ϕ-angle restraints, 88 ψ-angle restraints, 42 χ1-angle restraints, and 9 χ2-angle restraints were included in structure calculations. Hydrogen bonds, deduced from patterns of amide hydrogen exchange 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 $60° \pm 30°$ for $J < 5.5$ Hz, $-120° \pm 30°$ for 8 Hz $< J <$ 9 Hz, and $-140° \pm 20°$ for $J > 9$ Hz. These $^3J_{HN-H\alpha}$ couplings were determined from a ^{1}H−^{15}N HMQC-J spectrum (Kay & Bax, 1990) using software provided by Lewis Kay, as described by Foreman-Kay et al. (1990). The ψ angles were restrained to $120° \pm 100°$ or $-30° \pm 110°$ based on the ratio of $^1H^\alpha_{i-1}$−$^1H^N_i$ and $^1H^\alpha_i$−$^1H^N_i$ NOE intensities (Gagné et al., 1994). The χ1 and χ2 angles were restrained to $\pm 30°$ from their assigned rotamer values.

RESULTS

In the absence of high concentrations of cellotetraose, resonances from 13 amides near the N-terminus and adjacent to the disulfide bond of CBD_{N1} are not observed in the ^{1}H−^{15}N HSQC spectrum of this protein (Johnson et al., 1996). It is unlikely that the lack of these signals results from rapid hydrogen exchange with the solvent as the gradient-enhanced HSQC spectrum was recorded at pH* 5.9 using selective flip-back pulses to minimize any perturbation of the bulk water magnetization. Therefore, we tentatively attribute this to line broadening due to conformational averaging of these amides on a millisecond time scale. Consistent with this explanation, some of these peaks are observed in HSQC spectra of the protein recorded at elevated temperatures. Inspection of the structure of CBD_{N1}, determined in the presence of cellotetraose, reveals no obvious reason for this anomalous behavior as these residues do not reside in the ligand-binding groove (*vide infra*). Two possible explanations are that a second, low-affinity binding site for cellotetraose is located near the N-terminal region of CBD_{N1} or that, in the absence of cellotetraose, this region adopts one or more possible interconverting conformations. Further analysis of the structure of the free CBD_{N1} and characterization of the bound sugar will test these possibilities.

For this current study, all spectra of CBD_{N1} were obtained in the presence of a 40-fold excess of cellotetraose (typically 80 mM) in order to obtain complete resonance assignments, as well as to investigate the effects of ligand binding. At 35 °C, the free and cellotetraose-bound forms of CBD_{N1} are in fast exchange on the NMR time scale, resulting in the observation of population-weighted average chemical shifts. With the concentration of cellotetraose used, CBD_{N1} is fully saturated with this ligand ($K_a = 4200 \pm 720$ M^{-1}; Johnson et al., 1996). However, the cellotetraose was not specifically included in the structure calculations due to the degeneracy of its NMR spectrum, which prevented the unambiguous assignment of intermolecular NOEs.

Main-Chain Resonance Assignments. The $^1H^N$ and ^{15}N resonances from the main-chain amides of CBD_{N1} were assigned using the combination of HNCACB (Wittekind & Mueller, 1993) and CBCACONH (Grzesiek & Bax, 1992) 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. Selectively α-^{15}N-labeled Tyr, Leu, Val, and Asp/Asn CBD_{N1} provided amino acid-specific starting points for this assignment procedure. A modified version of the CBCACO(CA)HA experiment that detects only the $^1H^\alpha$ and $^{13}C^\alpha$/$^{13}C^\beta$ resonances of residues preceding prolines was also extremely useful for providing unambiguous reference points (Olejniczak & Fesik, 1994; L. Kay, personal communication). The assignments of the resonances from the remaining backbone residues were obtained using the HNCO (Ikura et al., 1990; Muhandiram & Kay, 1994) and CBCACO(CA)HA experiments (Kay, 1993). Figure 1 shows the assigned ^{1}H−^{15}N HSQC spectrum of CBD_{N1}. With the exception of the 13 amides not detected in the absence of added cellotetraose, near-complete assignments of the resonances from the main-chain $^1H^N$ and ^{15}N nuclei of uncomplexed CBD_{N1} were also obtained by following the progressive titration of the protein with cellotetraose (Johnson et al., 1996).

Aliphatic Side-Chain Resonance Assignments. Virtually complete side-chain assignments of the resonances from ^{1}H, ^{13}C, and ^{15}N nuclei in CBD_{N1} were obtained using a combination of ^{15}N TOCSY-HSQC (Marion et al., 1989c), H(CCO)NH, C(CO)NH (Logan et al., 1992; Montelione et al., 1992; Grzesiek et al., 1993a), HCCH-TOCSY (Bax et al., 1990; Kay et al., 1993), and HCCH-COSY (Kay et al., 1990; Ikura et al., 1991) experiments. The HCCH-COSY and HCCH-TOCSY experiments were recorded with the protein in D_2O buffer. These assignments are reported in Table S2 of the supporting information.

Initial inspection of the ^{1}H−^{15}N HSQC spectrum of CBD_{N1} revealed that four of seven expected peaks from the side-chain $^{15}NH_2$ resonances were either missing, very weak, or degenerate. The strong resonances from the $^{15}N^{\epsilon 2}H_2$ groups of Gln42, Gln80, and Gln101 were readily assigned from the HNCACB experiment and confirmed using the ^{15}N

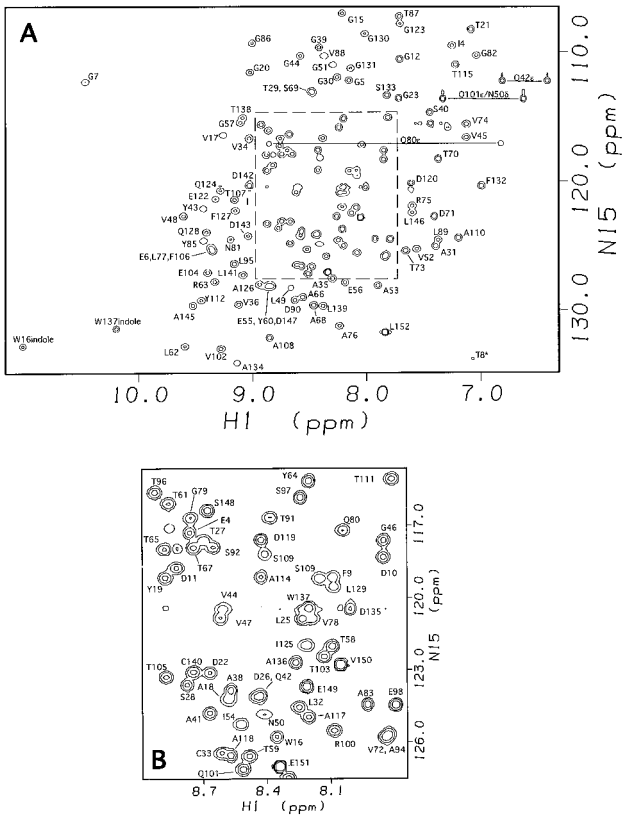

FIGURE 1: Sensitivity-enhanced gradient $^1H-^{15}N$ HSQC spectrum of CBD_{N1} at 35 °C and pH* 5.9, showing the assignments of the resonances from backbone amide and tryptophan indole $^{15}N^{\epsilon 1}H$ groups and those of the observable signals from the side-chain $^{15}N^{\delta}H_2$ and $^{15}N^{\epsilon}H_2$ of asparagine and glutamine, respectively. The aliased peak from Thr8 is denoted by an asterisk. The spectrum in (B) is the expansion of the boxed central region of (A).

NOESY-HSQC spectrum. Therefore, the unidentified side-chain amide resonances were those of Asn50, Asn81, Gln124, and Gln128. All four of these residues lie within the oligosaccharide-binding cleft of CBD_{N1} and could possibly play an important role in complexation with sugar. As a result of their potential importance, a special effort was made to assign them.

First, a modified version of the CBCACO(CA)HA experiment that links the side-chain carbonyls to the previously assigned $^{13}C^{\alpha}/^{13}C^{\beta}$ and $^1H^{\beta 2,\beta 3}$ of Asp/Asn and $^{13}C^{\beta}/^{13}C^{\gamma}$ and $^1H^{\gamma 2,\gamma 3}$ of Glu/Gln was recorded (Kay, 1993). Next an HNCO experiment, tuned using a total delay of $1/(4J_{NH})$ during the first reverse INEPT sequence to favor AX_2 spin systems (Schleucher et al., 1994), was used to correlate the resonances of these side-chain carbonyls to the corresponding $^{15}NH_2$ groups. Finally, an HSQC spectrum, with similar delays to enhance AX_2 spin systems, was recorded overnight to help identify the $^{15}NH_2$ resonances. This strategy yielded assignments for all seven side-chain $^{15}NH_2$ groups, as well as those of the carbonyls of almost all Asp/Asn and Glu/Gln residues (Table S2, supporting information).

Both side-chain amide $^1H^{\delta 2}$ protons as well as the $^{15}N^{\delta 2}$ of Asn50 are completely degenerate with those of Gln101. Fortunately, the difference in the chemical shift of the side-chain carbonyl resonances of these two residues allowed them to be distinguished. The chemical shift of the $^{13}C^{\gamma}$ carbonyl of Asn50 is unusually upfield shifted to 173.5 ppm. In the absence of structural information for the bound cellotetraose, the reason for this perturbed shift is not immediately apparent. The side-chain $^{15}NH_2$ groups of Asn81, Gln124, and Gln128, all of which are likely to participate in hydrogen bonding to the cellotetraose, exhibit very weak resonances. Only one $^1H^{\delta 2}$ for the amide of Asn81 was found. It is not known if the resonances from the two amide protons are degenerate or if the second is very weak. It is likely that the anomalous behavior of these side-chain amide resonances results from conformational exchange broadening, possibly due to unfavorable kinetics of cellotetraose association and dissociation from CBD_{N1} or due to mobility of the sugar ligand within the binding cleft. In support of this suggestion, we note that the side chain of Gln80, which lies on a β-strand that is part of the binding face but does not point into the binding cleft, has a strong $^{15}N^{\epsilon}H_2$ signal. In contrast, the side chain of the adjacent Asn81, which points into the binding cleft, has a very weak $^{15}N^{\delta}H_2$ signal.

Aromatic Side-Chain Assignments. Due to spectral overlap, the use of protein samples with selectively deuterated aromatic rings was an extremely useful tactic for the assignment of these side-chain spin systems (Johnson et al., 1996). The ^1H resonances from the aromatic rings were identified from homonuclear DQF-COSY, 70-ms mixing time TOCSY, and 150-ms mixing time NOESY spectra. These were recorded with samples of 1.9 mM unlabeled, 2.0 mM ([$^2H_4$]Tyr and [$^2H_5$]Trp)-labeled, and 1.9 mM ([$^2H_5$]-Phe and [$^2H_5$]Trp)-labeled CBD_{N1}, each in the presence of 80 mM cellotetraose. The unlabeled CBD_{N1} sample was recorded in both H_2O and D_2O buffers, while the two deuterated CBD_{N1} samples were recorded solely in D_2O buffer. The $^1H^{\delta}$ and $^1H^{\epsilon}$ of the aromatic ring spin systems were then directly connected to the previously assigned $^{13}C^{\beta}$ nuclei using the (Hβ)Cβ(CγCδ)Hδ and (Hβ)Cβ(CγCδCϵ)Hϵ experiments (Yamakazi et al., 1993). Finally the assignment of the ^{13}C resonances of the aromatic rings were obtained from $^1H-^{13}C$ HSQC and CT-HSQC spectra acquired using the $^{13}C/^{15}N$-labeled CBD_{N1} (Santoro & King, 1992; Vuister & Bax, 1992). The CT-HSQC was particularly useful for distinguishing the $^{13}C^{\delta 1}$ of tryptophan residues due to their inverted signals relative to those of other ^{13}C nuclei with an even number of neighboring carbons.

Stereospecific Assignments and Side-Chain Torsion Angle Restraints. Stereospecific assignment of 23 of the 67 residues in CBD_{N1} with prochiral $^1H^{\beta}$ protons was obtained from a conservative analysis of the HNHB (Archer et al., 1991), 40- and 72-ms mixing time ^{15}N TOCSY-HSQC (Marion et al., 1989c), and 50-ms $^{13}C/^{15}N$ NOESY-HSQC (Pascal et al., 1994) spectra of CBD_{N1}. These assignments and the corresponding $\chi 1$ restraints were determined on the basis of staggered rotamer model, as outlined by Powers et al. (1993).

Near-complete stereospecific assignments of the diastereotopic methyls of Val and Leu were obtained using the elegant approach of Neri et al. (1989). As demonstrated by these authors, in high-resolution HSQC spectra of 10% nonrandomly fractionally ^{13}C-enriched proteins, the *pro-R* ($Leu^{\delta 1}$, $Val^{\gamma 1}$) methyls are doublets due to $^{13}C-^{13}C$ couplings, while the *pro-S* ($Leu^{\delta 2}$, $Val^{\gamma 2}$) methyls are singlets. An additional level of discrimination is provided by the use of a constant-time $^1H-^{13}C$ HSQC experiment with a total evolution delay of $1/J_{CC} = 1/34$ Hz (L. Kay, personal communication). In the resulting CT-HSQC spectrum, the signals due to $Val^{\gamma 1}$, $Leu^{\delta 1}$, and $Ile^{\gamma 2}$ and the signals due to $Val^{\gamma 2}$ and $Leu^{\delta 2}$ all

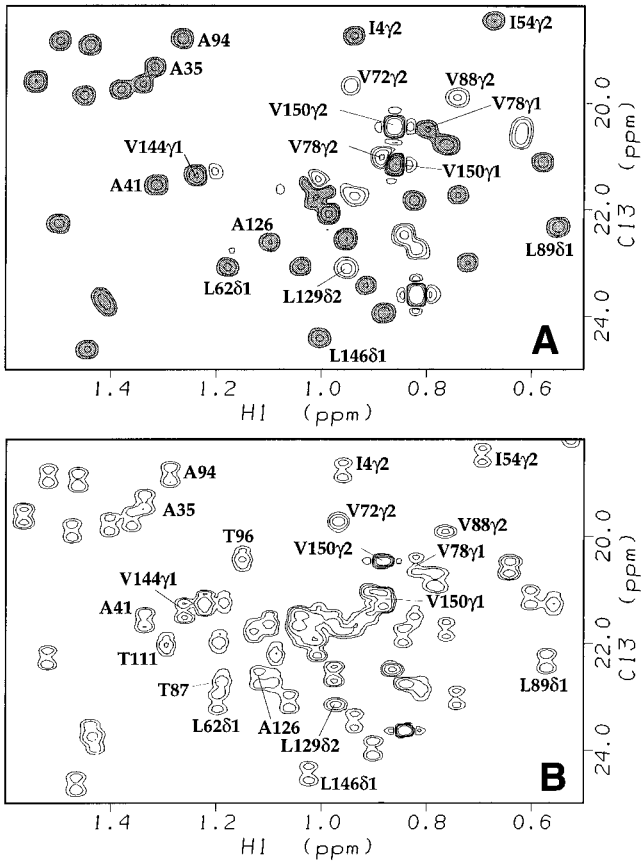

FIGURE 2: A portion of the (A) constant-time $^1H-^{13}C$ HSQC and (B) $^1H-^{13}C$ HSQC spectra of 10% fractionally ^{13}C-enriched CBD_{N1}. Due to the biosynthetic labeling pathways, the *pro-R* ($Leu^{\delta 1}$, $Val^{\gamma 1}$) methyls are doublets and *pro-S* ($Leu^{\delta 2}$, $Val^{\gamma 2}$) methyls are singlets in the regular HSQC (B). In the CT-HSQC (A), all peaks are singlets but the *pro-R* ($Leu^{\delta 1}$, $Val^{\gamma 1}$) methyls (shaded) have the opposite sign compared to the *pro-S* ($Leu^{\delta 2}$, $Val^{\gamma 2}$) methyls (open). The use of constant-time results in further simplification of the spectra by the elimination of $Thr^{\gamma 2}$ and $Ile^{\delta 1}$ methyls (not shown in the spectral window) that are apparent triplets in the non-constant-time HSQC. For clarity, assignments are indicated only for selected peaks.

appear as singlets but with opposite signs (Figure 2). Also, the peaks due to the $Thr^{\gamma 2}$ and $Ile^{\delta 1}$ methyls, which are apparent triplets in the ^{13}C HSQC due to approximately equal levels of $^{13}C-^{13}C$ and $^{13}C-^{12}C$ labeling (Szyperski et al., 1992), are nulled in the constant-time ^{13}C HSQC. Together, these two factors help simplify crowded regions of the spectrum, allowing previously overlapped resonances to be assigned.

$\chi 1$ restraints for 9 of the 14 valines, 6 of the 20 threonines, and all 3 isoleucines were determined on the basis of $^3J_{NC\gamma}$ and $^3J_{C'C\gamma}$ coupling constants (Tables 1 and 2) and intraresidue NOE interactions, according to the staggered rotamer model. The coupling constants were determined quantitatively from $^{13}C-\{^{15}N\}$ and $^{13}C-\{^{13}C'\}$ spin echo difference CT-HSQC spectra (Grzesiek et al., 1993b; Vuister et al., 1993). Peak volumes were used in the calculation of the coupling constants using in-house written programs, as described by Grzesiek et al. (1993b) and Vuister et al. (1993). The $\chi 1$ analysis of valine was aided by the previously determined stereospecific assignments, particularly in the case of resonances obscured by spectral overlap.

Qualitative analysis of the 3D $^{13}C-^{13}C$ long-range correlation experiment (Bax et al., 1992) provided $\chi 2$ restraints

Table 1: Coupling Constants[a] and $\chi 1$ Assignments[b] for Valine Residues in CBD_{N1}

	$\gamma 1$		$\gamma 2$		
residue	$^3J_{NC\gamma}$ (Hz)	$^3J_{C'C\gamma}$ (Hz)	$^3J_{NC\gamma}$ (Hz)	$^3J_{C'C\gamma}$ (Hz)	χ_1 (deg)
V17	*c*	*c*	0.7	1.7	*d*
V34	0.4	3.4	0.6	0.8	−60
V36	*c*	*c*	0.6	3.4	180
V45	*c*	*c*	0.3	3.0	180
V47	1.5	1.1	0.7	2.9	180
V48	0.7	2.7	*c*	*c*	−60
V52	*e*	*e*	*e*	*e*	*e*
V72	0.2	1.8	1.5	0.5	*d*
V74	*f*	3.3	0.8	1.2	−60
V78	2.1	0.7	0.7	1.7	180
V88	*c*	*c*	0.9	0.9	*d*
V102	1.9	0.7	0.5	3.6	180
V144	1.7	1.1	*c*	*c*	180
V150	1.6	2.7	0.8	2.8	*d*

[a] Coupling constants were determined from $^{13}C-\{^{15}N\}$ and $^{13}C-\{^{13}C'\}$ spin echo CT-HSQC spectra (Grzesiek et al., 1993b; Vuister et al., 1993). The coupling constants were corrected for a systematic underestimation in the determined coupling constant value inherent in the method by multiplication of the measured value by 1.06 ($^3J_{C'C\gamma}$) and 1.08 ($^3J_{NC\gamma}$) (Damberger et al., 1994). [b] $\chi 1$ angles were determined on the basis of $^3J_{C'C\gamma}$ and $^3J_{NC\gamma}$ values, as well as intraresidue NOEs determined with a 50-ms mixing time $^{13}C/^{15}N$ NOESY−HSQC spectrum and a 125-ms ^{15}N NOESY−HSQC spectrum. [c] Not determined due to spectral overlap. [d] Not assigned due to conflicting evidence for the presence of a single $\chi 1$ angle among $^3J_{C'C\gamma}$ values and/or $^3J_{NC\gamma}$ values and/or intraresidue NOEs. [e] Not determined due to degeneracy of the two V52 methyls in both ^{13}C and ^1H dimensions. [f] Coupling constant value too small to accurately measure.

Table 2: Coupling Constants[a] and $\chi 1$ Assignments[b] for the Ile and Thr Residues of CBD_{N1}

residue	$^3J_{NC\gamma 2}$ (Hz)	$^3J_{C'C\gamma 2}$ (Hz)	χ_1 (deg)	residue	$^3J_{NC\gamma 2}$ (Hz)	$^3J_{C'C\gamma 2}$ (Hz)	χ_1 (deg)
T8	0.7	2.9	60	T73	*d*	*d*	*e*
T21	0.9	2.4	60	T87	0.7	3.2	60
T27	*c*	3.3	*e*	T91	0.3	2.4	60
T29	1.0	2.1	*e*	T96	*c*	2.4	60
T58	2.0	0.7	*e*	T103	*d*	*d*	*e*
T59	*d*	*d*	*e*	T105	*d*	*d*	*e*
T61	1.7	*c*	*e*	T107	*d*	*d*	*e*
T65	*d*	*d*	*e*	T111	1.6	0.6	−60
T67	*d*	*d*	*e*	T115	1.0	2.6	*e*
T70	*c*	1.8	*e*	T138	*d*	*d*	*e*
I4	0.8	3.3	60	I125	2.2	0.6	−60
I54	2.0	0.8	−60				

[a] Coupling constants were determined from a $^{13}C-\{^{15}N\}$ and $^{13}C-\{^{13}C'\}$ spin echo CT-HSQC spectra (Grzesiek et al., 1993b; Vuister et al., 1993). The coupling constants were corrected for a systematic underestimation in the determined coupling constant value inherent in the method by multiplication of the measured value by 1.06 ($^3J_{C'C\gamma}$) and 1.08 ($^3J_{NC\gamma}$) (Damberger et al., 1994). [b] $\chi 1$ angles were determined on the basis of $^3J_{C'C\gamma}$ and $^3J_{NC\gamma}$ values, as well as intraresidue NOEs using a 50-ms mixing time $^{13}C/^{15}N$ NOESY−HSQC spectrum and a 125-ms ^{15}N NOESY−HSQC spectrum. [c] Coupling constant value too small to accurately measure. [d] Not determined due to spectral overlap. [e] Not assigned due to conflicting evidence for the presence of a single $\chi 1$ angle among $^3J_{C'C\gamma}$ values and/or $^3J_{NC\gamma}$ values and/or intraresidue NOEs.

for 7 of the 12 leucine residues and 2 of the 3 isoleucine residues. As shown in Figure 3, a strong $^1H^{\delta 1}-^{13}C^{\alpha}$ cross peak was observed for Ile4 and Ile125, reflecting a large $^3J_{C\delta 1C\alpha}$ coupling indicative of a *trans* conformation ($\chi 2 = 180°$). In the case of Ile54, there was no $^1H^{\delta 1}-^{13}C^{\alpha}$ peak and a weak $^1H^{\delta 1}-^{13}C^{\gamma 2}$ peak. This indicates that both the $^3J_{C\delta 1C\alpha}$ and the $^3J_{C\delta 1C\gamma 2}$ couplings are small, suggesting the

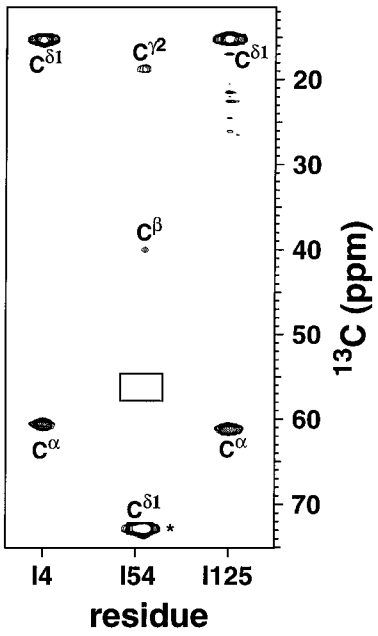

FIGURE 3: Strip plot of a portion of the long-range $^{13}C-^{13}C$ coupling constant experiment recorded for CBD_{N1} (Bax et al., 1992). Shown are $\omega1(^{13}C)-\omega3(^{1}H)$ strips for Ile4 and Ile125, which exhibit strong $^{1}H^{\delta1}-^{13}C^{\alpha}$ cross peaks relative to the $^{1}H^{\delta1}-^{13}C^{\delta1}$ autocorrelation peaks. This reflects large $^{3}J_{C\delta1C\alpha}$ couplings, indicative of *trans* (180°) $\chi2$ dihedral angles. In contrast, no $^{1}H^{\delta1}-^{13}C^{\alpha}$ cross peak is observed for Ile54 (a box indicates the position at which it is expected). Also, only a weak $^{1}H^{\delta1}-^{13}C^{\gamma2}$ cross peak is observed. Together, this reflects both small $^{3}J_{C\delta1C\gamma2}$ and $^{3}J_{C\delta1C\alpha}$ coupling constants, indicative of a *gauche*$^{+}$ conformation ($\chi2 = -60°$). The $^{13}C^{\delta1}$ of Ile54 is marked with an asterisk indicating it is aliased in the ^{13}C dimension shown in this plot. Cross peaks and autocorrelation peaks are shown without distinction of their opposite signs.

side chain adopts a *gauche*$^{+}$ conformation ($\chi2 = -60°$). In the absence of positive evidence for this conformation, no $\chi2$ torsion restraint was included for Ile54 during the structure calculations. However, in the final ensemble of structures, this $\chi2$ angle was indeed well defined at $-65°$ with an order parameter, $S(\chi2)$, of 0.99 (Hyberts et al., 1992).

Secondary Structure Determination. Initial studies using CD and FT-IR spectroscopy indicated that CBD_{N1} is composed of β-strands and devoid of helices (data not shown). This global analysis was confirmed when the regular secondary structural elements of CBD_{N1} were determined on the basis of patterns of amide hydrogen exchange rates, sequential and cross-strand NOEs, $^{3}J_{HN-H\alpha}$ coupling constants, and $^{13}C^{\alpha}$, $^{13}C^{\beta}$, $^{13}C'$, and $^{1}H^{\alpha}$ chemical shifts. The information defining the β-strands in CBD_{N1} is summarized in Figure 4. In general, residues in β-strand conformations are characterized by $^{3}J_{HN-H\alpha} \geq 8$ Hz, positive $^{1}H^{\alpha}$ and $^{13}C^{\beta}$ secondary chemical shifts ($\delta_{observed} - \delta_{random\ coil}$; Wishart et al., 1992; Wishart & Sykes, 1994), negative $^{13}C^{\alpha}$ and $^{13}C'$ secondary chemical shifts, strong sequential $^{1}H^{\alpha}_{i}-^{1}H^{N}_{i+1}$ NOEs, cross-strand $^{1}H^{\alpha}_{i}-^{1}H^{\alpha}_{j}$, $^{1}H^{\alpha}_{i}-^{1}H^{N}_{j}$, and $^{1}H^{N}_{i}-^{1}H^{N}_{j}$ NOEs, and protection from exchange due to cross-strand hydrogen bonding (Wüthrich, 1986).

Figure 5 summarizes the interresidue NOEs observed in the NOESY spectra of CBD_{N1} (Figure 6), as well as the hydrogen bonds used in the structure calculations. On the basis of the patterns of cross-strand $^{1}H^{\alpha}_{i}-^{1}H^{\alpha}_{j}$, $^{1}H^{N}_{i}-^{1}H^{N}_{j}$, and $^{1}H^{\alpha}_{i}-^{1}H^{N}_{j}$ NOEs, CBD_{N1} contains two β-sheets, denoted A and B, each composed of five antiparallel β-strands. Defining the exact boundaries of the regular elements of secondary structure in a protein is often difficult. For example, in the case of CBD_{N1}, there is some ambiguity in identifying the ends of strands B3 and B4. As seen in Figure 5, interstrand NOEs indicate that strand B3 could be defined to start at Ile54. Correspondingly strand B4 would end at Tyr112. The lack of cross-strand $^{1}H^{\alpha}_{i}-^{1}H^{\alpha}_{j}$ NOEs is a result of extensive degeneracy of the $^{1}H^{\alpha}$ resonances in this region. However, the classification of these residues as having a β-sheet conformation is not supported by the secondary chemical shifts of these residues (Figure 4). From the tertiary structure of CBD_{N1}, it is found that the polypeptide backbone in these regions turn sharply, linking sheets A and B. Therefore, these sequences were not defined as part of the β-strands in CBD_{N1}.

Two strands of CBD_{N1} are broken by β-bulges. One bulge, classified by Promotif (Hutchinson & Thornton, 1996) as being classical (Chan et al., 1993), begins at residue Thr87, following which Val88 and Leu89 both lie in the hydrophobic core of the protein. The second bulge starts at Leu141, after which Asp142 and Asp143 both have their side chains on the exterior of the protein. The presence of both of these bulges is evident from the secondary chemical shifts shown in Figure 4. This emphasizes the potential wealth of structural information contained in NMR chemical shifts.

From the topological arrangement of the β-strands of CBD_{N1}, it is evident that this protein adopts a jelly-roll β-sandwich structure (Figure 7; Brandon & Tooze, 1991). The jelly-roll is comprised of 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 together by hydrogen bonding, thus defining the structure as a β-sandwich as opposed to a continuous β-barrel. This is evident from the protection patterns of the backbone amide $^{1}H^{N}$ protons in the hydrogen−deuterium exchange experiments (Figures 4 and 5). The outer edges of the β-sheets do not show protection for the backbone amide $^{1}H^{N}$ protons.

Tertiary Structure. A total of 1988 distance, hydrogen-bond, and dihedral restraints were used to calculate 60 structures following the hybrid distance geometry/simulated annealing protocol (Nilges et al., 1988) with X-PLOR 3.1 (Brünger, 1992). The 25 structures with the lowest total energy and fewest NOE violations were selected for comparison. None of these had NOE violations greater than 0.4 Å, and, except for the two $\chi1$ restraints involving Cys33 and Cys140, none of the 25 structures had dihedral violations greater than 4.3°. Statistics for the 25 accepted structures are listed in Table 3.

The superimposition of the final ensemble of structures calculated for CBD_{N1} is shown in Figure 8. The structural ensemble is clearly consistent with the jelly-roll β-sandwich topology deduced at the level of secondary structure analysis. The backbone conformation of the strands in each of the sheets is well determined, having an rmsd of 0.44 ± 0.05 Å with respect to the average structure. Apart from the N- and C-termini, the regions which have the highest rms deviation from the average structure are between residues 20 and 30, which contains the short β-strand B1 involving residues 25−27, and in the loops between residues 37−44, 81−85, and 113−121. All these stretches contain no, or few, long-distance restraints (Figure S1A of the supporting information). These stretches are also regions where the

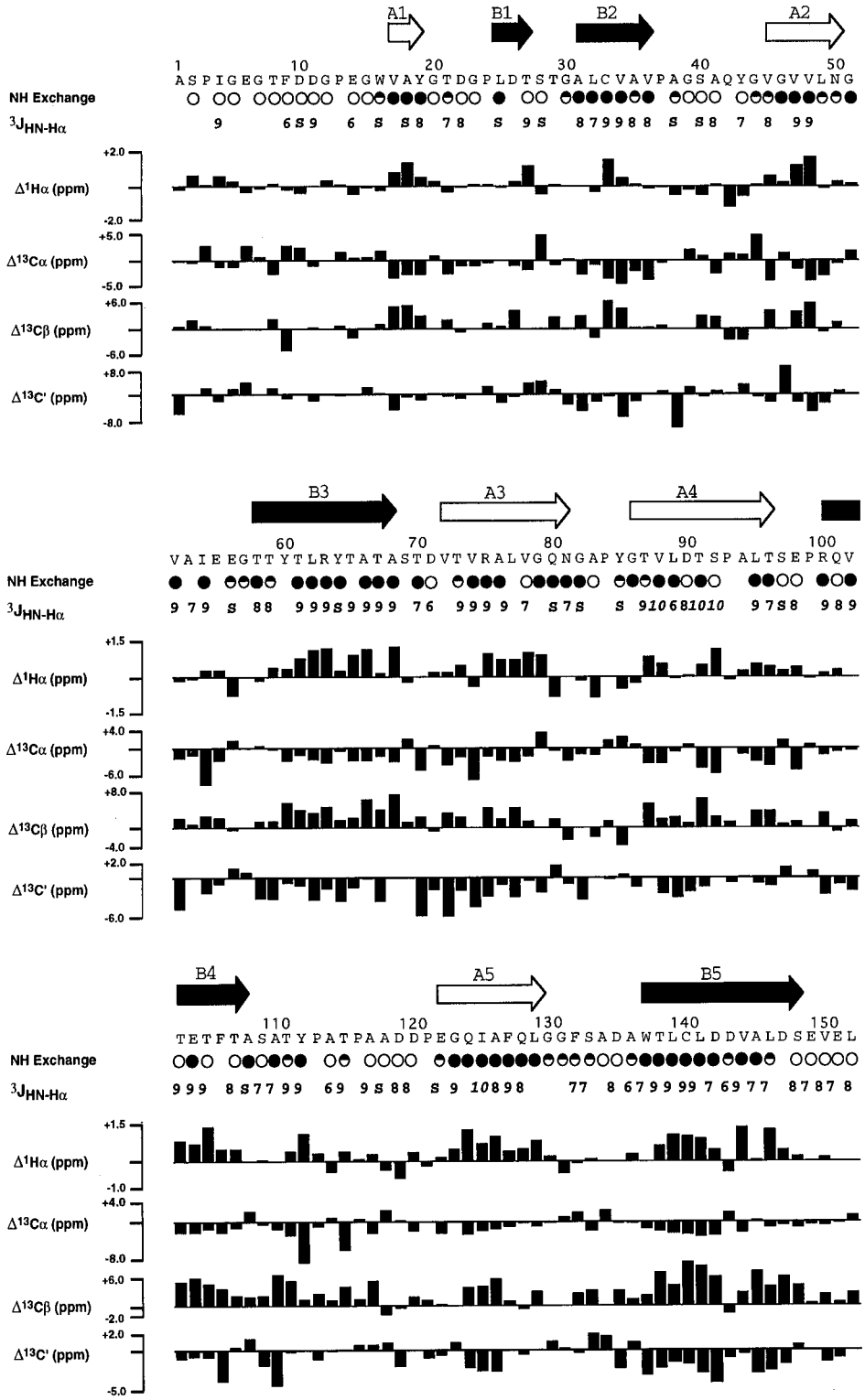

FIGURE 4: Summary of amide hydrogen exchange rates, $^3J_{HN-H\alpha}$ coupling constants, and $^1H^\alpha$, $^{13}C^\alpha$, $^{13}C^\beta$, and $^{13}C'$ chemical shifts used to deduce the β-strand secondary structure of CBD_{N1}. The locations of the five β-strands of sheet A are indicated by open arrows, and those of the five β-strands of sheet B are indicated by solid arrows. (i) Hydrogen exchange: Filled circles indicate residues with slow hydrogen−deuterium exchange kinetics ($t_{1/2}$ > 1000 min), half-filled circles indicate those with intermediate hydrogen−deuterium exchange kinetics (10 min < $t_{1/2}$ < 1000 min), and open circles indicate those with fast hydrogen−deuterium exchange kinetics ($t_{1/2}$ < 10 min) at 35 °C and pH* 5.90. (ii) $^3J_{HN-H\alpha}$: measured values are reported in hertz. S denotes couplings that are too small to be determined reliably using the HMQC-J experiment. (iii) Main-chain $^1H^\alpha$, $^{13}C^\alpha$, $^{13}C^\beta$, and $^{13}C'$ chemical shifts are plotted as the difference from the random coil values ($\delta_{observed} - \delta_{random\ coil}$). Residues in β-strands have a positive change in $^1H^\alpha$ and $^{13}C^\beta$ shift and a negative change in $^{13}C^\alpha$ and $^{13}C'$ chemical shift (Wishart et al., 1992; Wishart & Sykes, 1994).

angular order parameters S(ϕ) and S(ψ) are the lowest, indicative of local disorder (Figure S1B, C; Hyberts et al., 1992).

The stereochemical quality of the backbone coordinates for the ensemble of 25 structures was checked using the programs Procheck and Procheck-NMR (Laskowski et al., 1993). For this ensemble, 98% of the residues lie in the allowed regions of the Ramachandran plot. The few residues with main-chain dihedral angles that often fall outside the allowed regions, namely, Glu14, Ala83, and Tyr85, are all found in the parts of CBD_{N1} that exhibit high rms deviations,

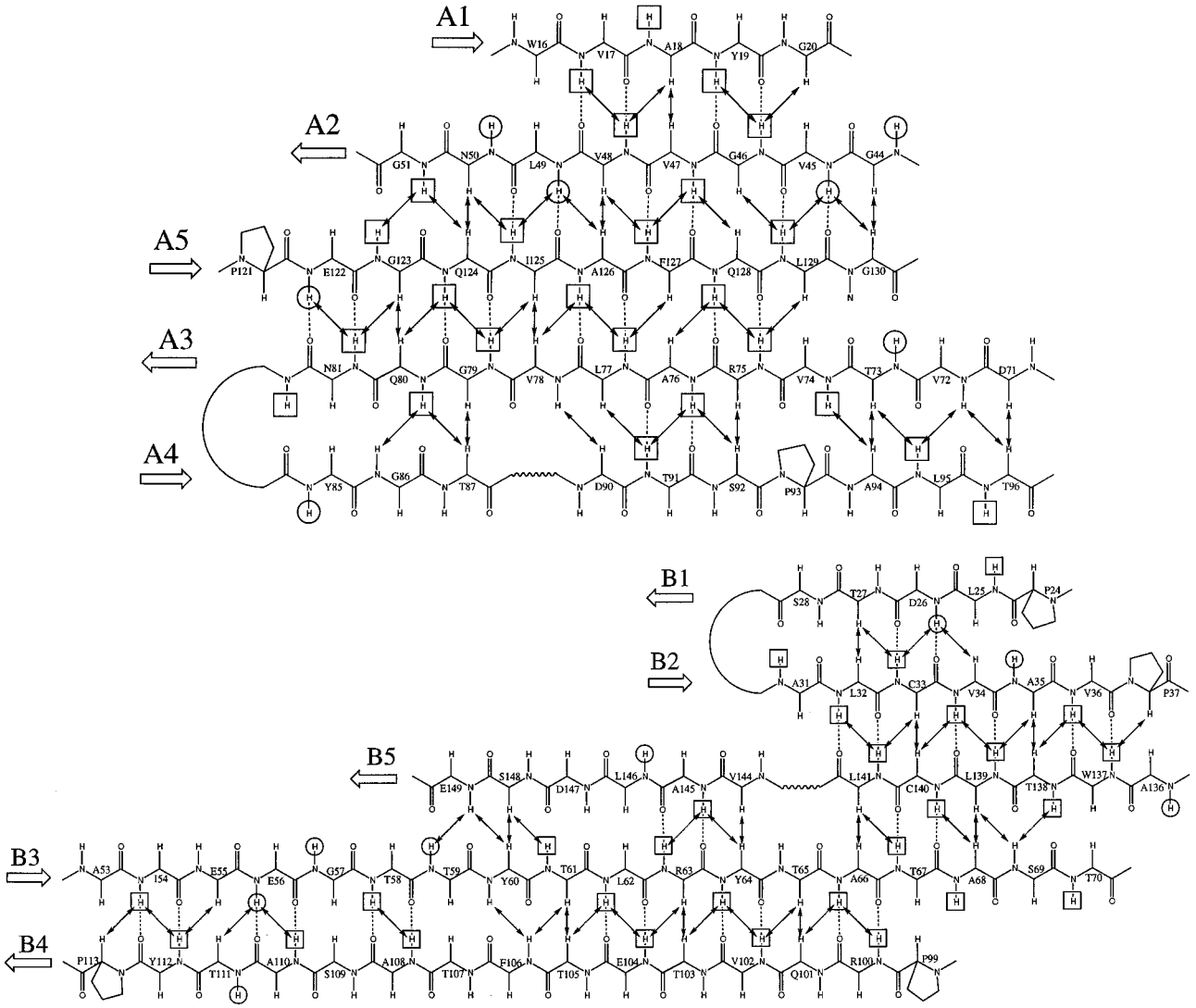

FIGURE 5: Alignment of the β-strands to form sheets A and B present in CBD_{N1}. The NOEs used to deduce these alignments are shown by arrows. Dotted lines indicate the hydrogen bonds included in the structure calculations. Boxed amide hydrogens have slow hydrogen−deuterium exchange kinetics (filled circles in Figure 4), while circled amide hydrogens have intermediate hydrogen−deuterium exchange kinetics (half-filled circles in Figure 4). The positions of β-bulges are indicated by jagged lines.

low angular order parameters, and the lack of regular secondary structure.

As shown in Figure 9, the side chains that make up the hydrophobic core of CBD_{N1} are well-defined structurally. This is reflected by both low rms deviations and high angular order parameters $S(\chi 1)$ (Figure S1D,E). Of the 33 side chains that comprise the hydrophobic core of CBD_{N1}, 32 have values of $S(\chi 1)$ greater than 0.98. The one exception is for Leu25, which has a value of 0.62. This residue lies in the strand that is at one edge of face B, in an area that is not well-defined as discussed above. In addition, the Lennard−Jones energy of each of the accepted structures is large and negative (Table 3) indicating that no unfavorable van der Waals contacts exist.

The presence of a cis-peptide bond for Ala83−Pro84 in CBD_{N1} was initially identified on the basis of the $^{13}C^{\gamma}$ chemical shift of the proline at 24.9 ppm. This is approximately 3 ppm upfield from the value found for prolines with trans-peptide linkages (Stanczyk et al., 1989). The cis-peptide bond was further confirmed by the observation of a strong $^{1}H^{\alpha}_{A83}-{}^{1}H^{\alpha}_{P84}$ NOE, and the lack of an $^{1}H^{\alpha}_{A83}-{}^{1}H^{\delta}_{P84}$ NOE in the 50-ms mixing time $^{13}C/^{15}N$ NOESY-HSQC of CBD_{N1} (Wüthrich, 1986). This NOE pattern is reversed for every other proline in CBD_{N1}, all of which are trans-linked. Pro84 lies in a turn between strands A3 and A4, classified as type VIa1 by Promotif (Hutchinson & Thornton, 1996). This type of turn requires a cis-linked proline at position $i+2$.

A disulfide bond between Cys33 and Cys140 was identified initially from the fact that the $^{13}C^{\beta}$ chemical shift of a cysteine is indicative of its oxidation state. The $^{13}C^{\beta}$ shift of Cys33 is 47.4 ppm and that of Cys140 is 47.5 ppm. These are close to the expected range for an oxidized Cys, which has a random coil value of 41.8 ppm, as opposed to 28.6 ppm for the reduced form (Wishart & Sykes, 1994). This conclusion is consistent with the results of a DTNB titration of CBD_{N1}, demonstrating that there are no free thiols present in the protein (data not shown). Cross-strand NOEs, illustrated in Figure 5, also confirm the pairing of these two cysteine residues.

The disulfide bond connecting Cys33 and Cys140 is unusual in that it bridges two β-strands (Figures 7 and 10). Such covalent bridges between paired β-strands are not observed commonly, although examples are found in azurin and chymotrypsin (Thornton, 1981). The disulfide in CBD_{N1} is classified as a short right-handed hook by Promotif

Table 3: Structural Statistics and Atomic RMS Differences[a]

	⟨SA⟩	⟨SA⟩$_{av}$
rmsd from experimental distance restraints[b] (Å) (1793)	0.013 ± 0.002	0.011
rmsd from experimental dihedral restraints[c] (deg) (195)	0.45 ± 0.13	0.29
deviations from idealized geometry		
bonds (Å) (2129)	0.0019 ± 0.0002	0.0016
angles (deg) (3841)	0.54 ± 0.01	0.52
impropers[d] (deg) (1102)	0.41 ± 0.02	0.39
XPLOR energies[e] (kcal mol^{-1})		
E_{NOE}	16.53 ± 3.9	10.72
E_{cdih}	2.64 ± 1.4	1.01
E_{imp}	28.2 ± 2.3	25.5
E_{angle}	168.2 ± 6.6	158.6
E_{bond}	7.5 ± 1.3	5.7
E_{vdw}	13.9 ± 3.4	10.3
E_{L-J}[f]	−569.2 ± 19.8	−612.5

atomic rms differences[g] (Å)	backbone[h]	all heavy atoms
residues 4−148	0.79 ± 0.14	1.60 ± 0.20
β-sheet regions[i]	0.44 ± 0.05	1.61± 0.20

[a] ⟨SA⟩ represents the final ensemble of 25 simulated annealing structures; ⟨SA⟩$_{av}$ is the restrained minimized average structure obtained by averaging the 25 structures over residues 4−148. Errors reported are ± 1 standard deviation. The number of restraints is given in parentheses. [b] This includes 1705 NOE-derived distance restraints and 88 hydrogen-bond restraints (44 hydrogen bonds). [c] Torsion angle restraints include 56 ϕ-angle restraints, 88 ψ-angle restraints, 42 χ1-angle restraints, and 9 χ2-angle restraints based on $^3J_{HN-H\alpha}$, $^3J_{NC\gamma}$, and $^3J_{C'C\gamma}$ measurements, qualitative analysis of $^3J_{C\delta C\alpha}$ values, and intraresidue and sequential NOEs from NOESY spectra. [d] Improper torsion angle restraints maintain peptide planarity and chirality. [e] The square-well NOE (E_{NOE}) using center averaging, the restrained dihedral (E_{cdih}), the improper torsion angle (E_{imp}), 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} Å^{-2}, 200 kcal mol^{-1} rad^{-2}, 500 kcal mol^{-1} rad^{-2}, 500 kcal mol^{-1} rad^{-2}, 1000 kcal mol^{-1} Å^{-2}, and 4 kcal mol^{-1} Å^{-4}. The van der Waals repulsion energy was calculated using sphere radii set to 0.75 times that supplied with XPLOR (parallhdg.pro). [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 25 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, C$^\alpha$, and C′. [i] This comprised residues 17−19, 25−27, 31−36, 45−50, 58−68, 72−81, 86−96, 100−108, 122−129, and 137−148 inclusive.

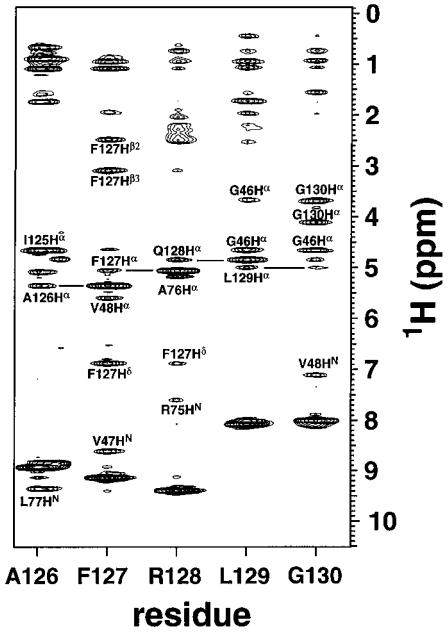

FIGURE 6: ω1(^{1}H)-ω3(^{1}H) strip plots of a portion of the ^{15}N HSQC−NOESY spectrum of CBD$_{N1}$, recorded with a mixing time of 125 ms. Selected NOE interactions are labeled, and solid lines connect the strong NOEs of a $^1H^N$ to the $^1H^\alpha$ of the previous residue, indicative of an extended β-strand conformation. This is an example of the data obtained to derive distance restraints for the structure generating process using the hybrid distance geometry/simulated annealing protocols in XPLOR 3.1.

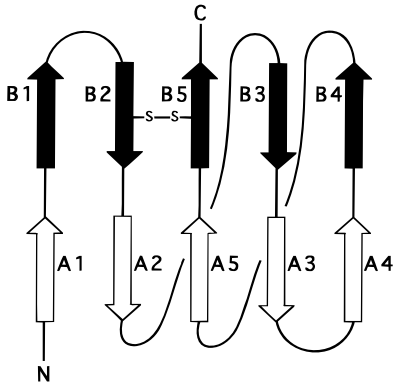

FIGURE 7: Schematic diagram showing the jelly-roll β-sandwich topology of CBD$_{N1}$. Sheets A and B are indicated by open and solid arrows, respectively, and the position of the disulfide between Cys33 and Cys140 is indicated. Strands A2−A5 and B2−B5 comprise the jelly-roll motif, with the two short strands A1 and B1 appended along one side. The global structure of CBD$_{N1}$ can be envisioned by folding the Figure such that sheet B lies below sheet A and sheet A is concave. The lengths of the strands and loops are not drawn to scale.

(Hutchinson & Thornton, 1996). Although the (ϕ, ψ) angles of the cysteines and their neighboring amino acids do deviate slightly from those expected for residues in ideal antiparallel β-strands, the presence of the disulfide bond does not produce any pronounced distortions in strands B2 and B5. The backbone dihedral angles of the cysteines and adjacent residues still lie within the most favored regions of the Ramachandran plot and are well-defined. In contrast, we observed that the χ1 restraints for Cys33 and Cys140 were often violated during the structural calculations. This may reflect conformational isomerization of the disulfide, particularly since the side-chain dihedral angles were determined from a staggered rotamer model using qualitatively estimated coupling constants. Consistent with this suggestion, several of the amides that are not detected in the absence of high concentrations of cellotetraose lie within the β-strands connected by this disulfide (Johnson et al., 1996).

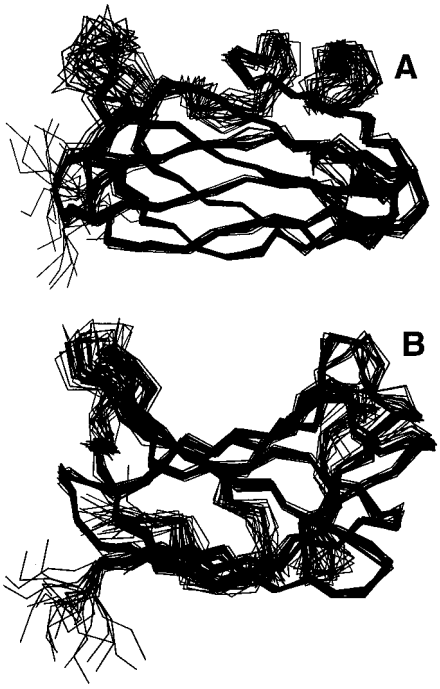

FIGURE 8: Two views of the C^{α} traces of the final ensemble of 25 structures calculated for CBD_{N1}. The coordinates were superimposed using the backbone atoms from residues 4 to 148. The jelly-roll β-sandwich topology is apparent in view A, whereas the presence of a binding cleft formed by β-sheet A is clearly evident in view B. View A looks down on the binding cleft, with view B rotated by approximately 90° compared to view A. The Figure was produced using the program Molscript (Kraulis, 1991).

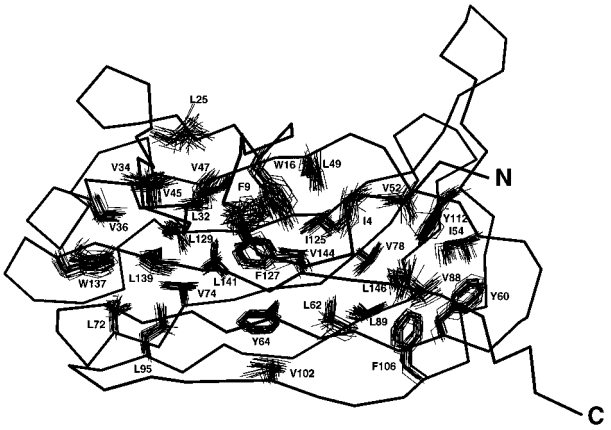

FIGURE 9: Superimposition of the coordinates of the side chains that make up the hydrophobic core of CBD_{N1} on the C^{α} trace of the minimized average structure (thick line) of this protein. All heavy atoms between residues 4 and 148 were used to superimpose the 25 structures. The CBD is oriented with the binding cleft toward the page and β-sheet B closest to the reader. This Figure was produced using the program Molscript (Kraulis, 1991).

DISCUSSION

Binding Mechanism. The structure of CBD_{N1} in the presence of saturating concentrations of cellotetraose was determined in this study using NMR methods. Previously, we reported that this CBD binds soluble cellooligosaccharides in order of increasing affinity cellotriose < cellotetraose < cellopentaose ~ cellohexaose (Johnson et al., 1996; Tomme et al., 1996). In addition, CBD_{N1} associates with oat and barley β-glucans with affinities approximately equal to that of cellopentaose (Tomme et al., 1996). Together, these results implied that the binding site spans approximately

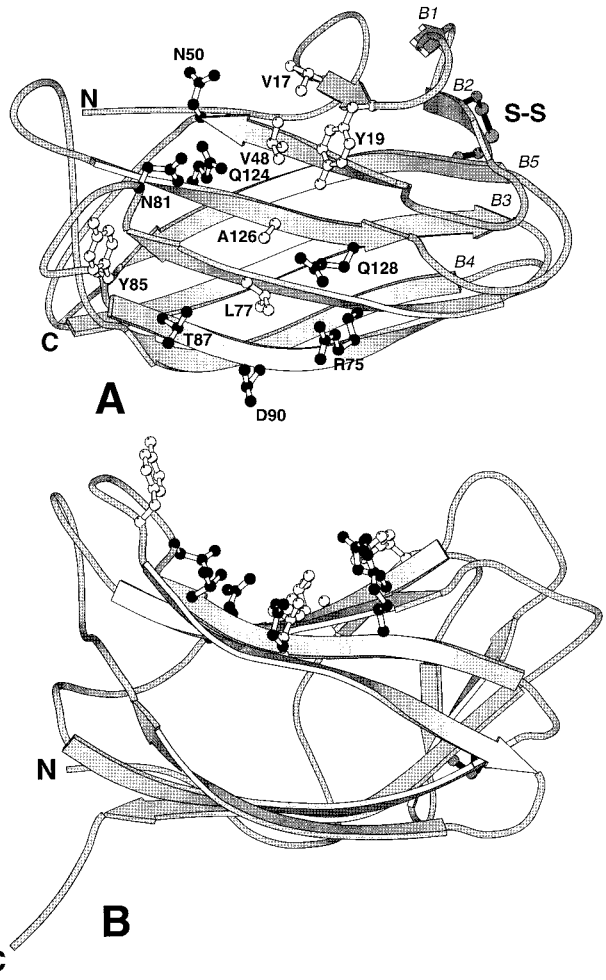

FIGURE 10: Molscript ribbon diagram of the minimized average structure of CBD_{N1} with residues from β-sheet A that are implicated in oligosaccharide binding shown in ball-and-stick format (Johnson et al., 1996). Hydrophilic residues are identified with black atoms and hydrophobic residues with white atoms. The disulfide bond between β-strands B2 and B5 is also presented using gray atoms and dark gray bonds. View A looks directly down on the binding face, whereas view B is rotated by 90° to emphasize the binding cleft. These are the same molecular orientations used in Figure 8. Strands B1−B5 are labeled in lightface italic type.

the length of a cellopentaose molecule. From the patterns of amide ^{15}N and $^{1}H^{N}$ chemical shift perturbations resulting from the addition of cellooligosaccharides to CBD_{N1}, it was demonstrated that the residues in strands A1−A5 (β-sheet A) all interact with these soluble ligands (Johnson et al., 1996). The involvement of the aromatic side chains of Tyr19 and Tyr85 in sugar binding was also suggested on the basis of changes in the NMR chemical shifts and line shapes of these residues and on perturbations of the near-ultraviolet absorption spectrum of CBD_{N1} due to addition of cellooligosaccharides (Johnson et al., 1996).

Expanding upon this initial work, we have now shown that the binding face of CBD_{N1} is a groove or cleft that extends across β-sheet A. The identification of this β-sheet as the ligand binding cleft of CBD_{N1} is confirmed by the observation of intermolecular NOEs between bound cellotetraose and residues located within this region of the protein. Although the ^{1}H resonances of the cellotetraose are currently unassigned, NOEs from unlabeled sugar protons to ^{13}C-labeled Tyr19 $^{1}H^{\delta}$ and $^{1}H^{\epsilon}$, Val48 $^{1}H^{\gamma 1}$ and $^{1}H^{\gamma 2}$, Leu77 $^{1}H^{\delta 2}$, Tyr85 $^{1}H^{\delta}$ and $^{1}H^{\epsilon}$, and Ala126 $^{1}H^{\beta}$ were

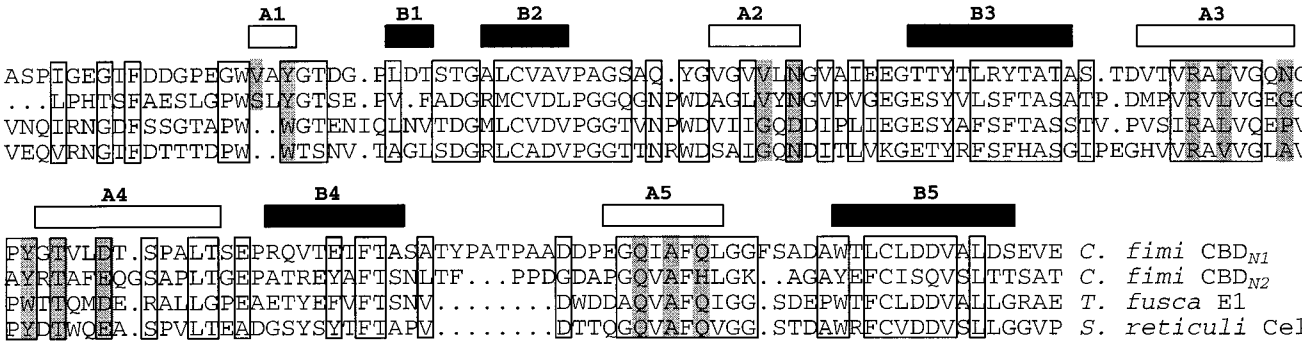

FIGURE 11: Alignment of the amino acid sequences of four family IV CBDs (Tomme et al., 1995b), *C. fimi* CenC (CBD_{N1} and CBD_{N2}), *Thermomonospora fusca* E1, and *Streptomyces reticuli* Cel1. The alignment was obtained using the program PHD (Rost, 1996). Boxes highlight positions where residues are conserved in three or more family members. The secondary structure of CBD_{N1} is shown as open and filled boxes to represent the β-strands of sheets A and B, respectively. The residues in sheet A that are implicated in cellooligosaccharide binding are shaded.

detected using a ^{13}C ω1-edited, ω3-filtered HMQC-NOESY experiment (unpublished observations; Lee et al., 1994). Consistent with the reported dependence of binding affinity on the degree of oligosaccharide polymerization, the length of the CBD_{N1} binding groove is approximately equal to that of an extended cellopentaose chain.

The individual β-strands in CBD_{N1} that comprise this binding face are very well ordered, with the loops regions less defined. In particular, the loops formed by residues 37−44, 82−84, and 113−121 are the most disordered, as evident by their high rms deviations and low angular order parameters (Figures 8 and S1). These loops form the extreme edges of the binding face, and thus a range of depths and widths are found for the cleft within the set of 25 accepted structures. An estimate of the variation in groove width is provided by the distances between the C^{α} atoms of Tyr43 and Ala118, which range from 24 to 31 Å in the ensemble of structures. It is not known whether this is a function of true mobility or simply a result of having few long-range restraints for residues in these loop regions of CBD_{N1}. This important issue is currently under investigation using NMR methods to characterize the dynamic behavior of this binding domain in the presence and absence of bound oligosaccharide.

Using the NMR-derived tertiary structure of CBD_{N1}, it is now possible to look more closely at the residues involved in oligosaccharide binding, particularly in light of recent thermodynamic studies of the interactions between CBD_{N1} and various sugars (Tomme et al., 1996). Figure 10 shows the exposed side chains present on the binding face of the minimized average structure of CBD_{N1}. Two distinct features of CBD_{N1} are evident from this figure. First, a strip of hydrophobic residues, composed of Val17, Tyr19, Val48, Leu77, and Ala126, runs along the center of the binding cleft. Tyr 85, which may also be involved in binding, is located in a loop region near this hydrophobic strip. These nonpolar residues may contact the pyranose rings of the oligosaccharide, providing favorable hydrophobic and van der Waals interactions. Second, there are numerous hydrophilic groups, including Asn50, Arg75, Asn81, Thr87, Asp90, Gln124, and Gln128, that flank the hydrophobic strip by lining the sides of the binding cleft. The polar residues likely provide hydrogen bonds to the equatorial hydroxyl groups of the sugar rings. In accordance with the relative insensitivity of sugar binding to pH and ionic strength (Tomme et al., 1996), only two of these seven polar groups are ionizable. Although CBD_{N1} was investigated in the presence of saturating amounts of cellotetraose, the sugar was not included in the structural calculations. Thus, it is not possible to state confidently that all of these polar and nonpolar residues are involved directly in ligand binding. To address this issue, the specific roles played by these amino acids will be investigated by studying complexes of CBD_{N1} with selectively labeled cellooligosaccharides and by site-directed mutagenesis.

The proposed structural mechanism of oligosaccharide binding is entirely consistent with the thermodynamic parameters characterizing the association of soluble oligosaccharides to CBD_{N1} (Tomme et al., 1996). On the basis of detailed calorimetric analyses, it was reported that sugar binding results in a favorable enthalpic change, compensated in part by a decrease in entropy. This implies that a predominance of polar interactions, such as hydrogen bonding, provides the primary driving force for binding. The dominant role of hydrogen bonding is observed with most other carbohydrate-binding proteins (Quiocho, 1986, 1989).

Finally, and perhaps most importantly, the presence of a binding cleft in the structure of CBD_{N1} provides a simple explanation as to why this protein has the ability to bind soluble oligosaccharides and amorphous cellulose but not crystalline cellulose. The residues that mediate binding are all located within this cleft and, as a result, are unable to interact with the flat surface presented by crystalline cellulose. In contrast, soluble sugars and single strands of amorphous cellulose can bind to CBD_{N1} by lying within this groove.

Comparison to Other Family IV CBDs. CBD_{N1} is the first family IV CBD for which a structure has been determined. Figure 11 presents the sequences of the four known members of this family (Tomme et al., 1995b), aligned with the secondary structural elements identified in CBD_{N1}. This alignment shows that the residues forming the eight strands of the jelly-roll β-sandwich motif of CBD_{N1} are in general well conserved, with deletions/insertions found in the intervening loop regions and the small β-strands A1 and B1. The two cysteines that form the disulfide bridge are invariant, and most of the residues involved in ligand binding by CBD_{N1} are conserved. Accordingly, it is reasonable to postulate that CBD_{N2} from CenC, as well as the related CBDs from *Thermomonospora fusca* E1 and *Streptomyces reticuli* Cel1, adopt secondary and tertiary structures similar to those determined for CBD_{N1}. Although it remains to be demonstrated if the *T. fusca* E1 and *S. reticuli* Cel1 CBDs exhibit

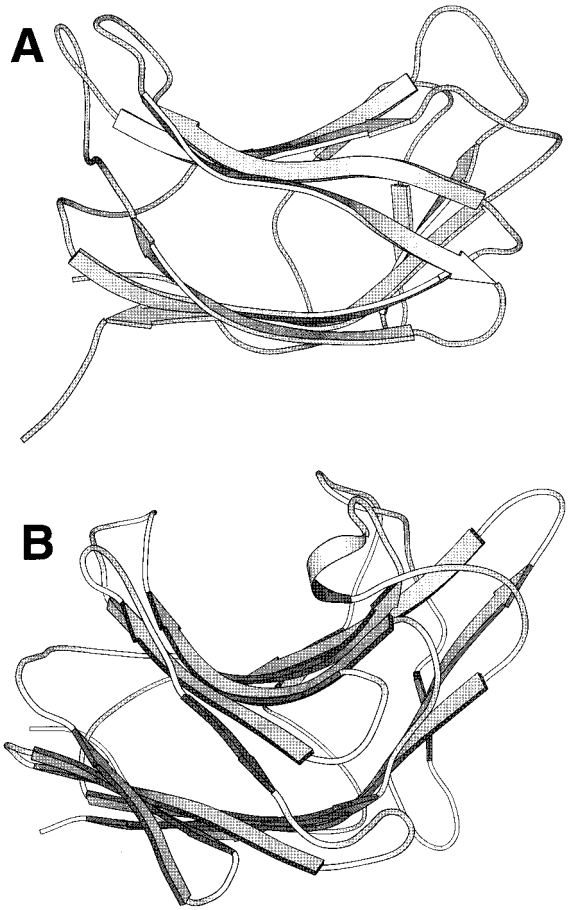

FIGURE 12: Molscript ribbon diagrams of (A) the minimized average structure of *C. fimi* CBD_{N1} and (B) the crystal structure of the hybrid *Bacillus* 1,3−1,4-β-glucanase (Keitel et al., 1993; Hahn et al., 1995b). These two proteins share a common jelly-roll β-sandwich topology. Despite only 8% sequence identity, 118 residues can be aligned with a rmsd of 3.7 Å for C^{α} coordinates.

the same specificity for soluble forms of cellulose, preliminary NMR studies of isolated CBD_{N2} confirms that the second CBD from CenC does indeed bind cellooligosaccharides (E. Brun, personal communication). This leads to the more tantalizing questions as to whether or not interdomain interactions exist between CBD_{N1} and CBD_{N2} within the native enzyme and why, when linked together in CBD_{N1N2}, do these two domains appear to bind phosphoric acid-swollen cellulose in an independent and not cooperative manner (Tomme et al., 1996).

Comparison to CBD_{CBH1} and CBD_{Cex}. The dominant structural feature of CBD_{N1} is the presence of a binding cleft. This distinguishes CBD_{N1} from the previously characterized CBD_{CBH1} (Kraulis et al., 1989) and CBD_{Cex} (Xu et al., 1995). Although all three of these protein domains are composed of antiparallel β-strands and contain at least one disulfide bridge, CBD_{CBH1} and CBD_{Cex} have flat binding surfaces characterized by three exposed aromatic rings. Paralleling this structural difference, CBD_{CBH1} and CBD_{Cex} also stand apart from CBD_{N1} in their affinity for crystalline cellulose and in the different thermodynamic forces that lead to carbohydrate binding by these protein domains. Whereas the association of CBD_{N1} with soluble oligosaccharides and phosphoric acid-swollen cellulose is enthalpically driven (Tomme et al., 1996), the affinity of CBD_{Cex} for insoluble bacterial microcrystalline cellulose results primarily from a favorable increase in entropy, indicative of a hydrophobic interaction (Creagh et al., 1996). Furthermore, it is known that the exposed tryptophans of CBD_{Cex} are involved in ligand binding (Poole et al., 1993; Din et al., 1994; Bray et al., 1996). We postulate that CBD_{Cex} relies on hydrophobic stacking of these aromatic rings with the flat surface of crystalline cellulose, as the hydroxyl groups of the glucosyl residues are involved primarily in interactions with adjacent polysaccharide chains. In contrast, CBD_{N1} associates with a single strand of cellulose by exploiting a binding cleft in which polar residues are positioned to hydrogen-bond to the exposed equatorial hydroxyl groups of the glucopyranose rings. These distinct structural and thermodynamic mechanisms highlight the complexity of cellulose as a substrate for enzymatic recognition and degradation.

Structural Similarity with 1,3−1,4-β-Glucanase. The tertiary structure of CBD_{N1} closely resembles those of *endo*-1,3−1,4-β-glucanase from *Bacillus macerans* (Hahn et al., 1995a) and a hybrid *Bacillus* 1,3−1,4-β-glucanase (Keitel et al., 1993; Hahn et al., 1995b). Although the larger 1,3−1,4-β-glucanases are composed of two β-sheets of seven strands each, as opposed to five strands each for CBD_{N1}, all three of these proteins share a common jelly-roll β-sandwich fold (Figure 12). According to the program DALI (Holm & Sanders, 1995), *C. fimi* CBD_{N1} and the hybrid *Bacillus* 1,3−1,4-β-glucanase were found to have 118 residues aligned with an rmsd of 3.7 Å based on superposition of C^{α} coordinates. This alignment occurs despite only 8% sequence identity of these residues. As pointed out by Hahn et al. (1995b), the prokaryotic 1,3−1,4-β-glucanases are topologically related to other polysaccharide degrading enzymes such as cellobiohydrolase I (Devine et al., 1994) and 1,4-β-xylanase II (Törrönen et al., 1994), both from *T. reesei.* Thus, it is not surprising that the program DALI identified these, as well as *Bacillus circulans* xylanase (Campbell et al., 1993) and an S-lectin (Liao et al., 1994) as having tertiary structures similar to that of CBD_{N1}. The lectin also has a jelly-roll fold. It is intriguing that these polysaccharide-binding domains resemble structurally the catalytic domains of several polysaccharide-degrading enzymes.

The observation that CBD_{N1} appears structurally related to these β-glucanases is particularly interesting given that this CBD also binds the mixed 1,3- and 1,4-linked oat and barley β-glucans (Tomme et al., 1996). These two soluble oligosaccharides, along with lichenan from the Icelandic moss *Cetraria islandica*, are natural substrates for the 1,3−1,4-β-glucanses. Furthermore, the catalytic residues Glu105 and Glu109 of the hybrid 1,3−1,4-β-glucanases are located on opposite sides of the active site of the enzyme, whereas the similarly spaced Gln124 and Gln128 lie on the two sides of the CBD_{N1} binding cleft. The structural similarities between CBD_{N1} and the 1,3−1,4-β-glucanases, combined with their binding to a common class of polysaccharides, suggests that these proteins are evolutionarily related, perhaps through divergence from a common ancestor with a jelly-roll sugar-binding motif. The unusual function of CBD_{N1} in binding soluble or amorphous, but not crystalline, glucans may stem from an evolutionary relationship with the 1,3−1,4-β-glucanases, and not an ancestral binding domain shared by the members of other CBD families.

SUPPORTING INFORMATION AVAILABLE

Two tables, summarizing the data collection and processing parameters for the NMR experiments used in this study (Table S1) and providing ^{1}H, ^{13}C, and ^{15}N NMR assignments for CBD_{N1} (Table S2), and one figure showing the per-residue distribution of NOE restraints, angular order parameters $S(\phi)$, $S(\psi)$, and $S(\chi 1)$, and rms deviations for all heavy atoms and main-chain atoms (Figure S1) (17 pages). Ordering information is given on any current masthead page.

ACKNOWLEDGMENT

We are indebted to Lewis Kay for providing NMR pulse sequences and advice. We also thank Anne Johnson for enthusiastic help in the assignment of CBD_{N1}, Logan Donaldson for writing numerous data conversion programs, Emmanuel Brun, Rick Dahlquist, Neil Gilkes, Charles Haynes, Jack Skalicky, Carolyn Slupsky, and Tony Warren for insightful discussions, and Nirvana for many mornings of pick-picking musical accompaniment.

REFERENCES

Archer, S. J., Ikura, M., Torcia, D. A., & Bax, A. (1991) *J. Magn. Reson. 95*, 636−641.
Bax, A., Clore, M., & Gronenborn, A. M. (1990) *J. Magn. Reson. 88*, 425−431.
Bax, A., Ikura, M., Kay, L. E., & Zhu, G. (1991) *J. Magn. Reson. 91*, 174−178.
Bax, A., Max, D., & Zax, D. (1992) *J. Am. Chem. Soc. 114*, 6923−6925.
Beguin, P., & Aubert, J.-P. (1994) *FEMS Microbiol. Rev. 13*, 25−58.
Bodenhausen, G., & Ruben, D. J. (1980) *Chem. Phys. Lett. 69*, 185−189.
Brandon, C., & Tooze, J. (1991) *Introduction to Protein Structure*, Garland Publishing Inc., New York.
Bray, M. R., Johnson, P. E., Gilkes, N. R., McIntosh, L. P., Kilburn, D. G., & Warren, R. A. J. (1996) *Protein Sci. (*in press).
Brünger, A. T. (1992) *X-PLOR Version 3.1: A System for X-ray Crystallography and NMR*, Yale University Press, New Haven, CT.
Campbell, R., Rose, D., Wakarchuk, W., To, R., Sung, W., & Yaguchi, M. (1993) in *Proceedings of the second TRICEL symposium on Trichoderma reesei cellulases and other hydrolases.* (Suominen, P., & Reinikainen, T., Eds.) pp 63−77, Foundation for Biotechnological and Industrial Fermentation Research, Helsinki, Finland.
Chan, A. W. E., Hutchinson, E. G., Harris, D., & Thornton, J. M. (1993) *Protein Sci. 2*, 1574−1590.
Clore, G., Gronenborn, A., Nilges, M., & Ryan, C. (1987) *Biochemistry 26*, 8012−8023.
Clore, G. M., Kay, L. E., Bax, A., & Gronenborn, A. M. (1991) *Biochemistry 30*, 12−18.
Coutinho, J. B., Gilkes, N. R., Warren, R. A. J., Kilburn, D. G., & Miller, R. C., Jr. (1992) *Mol. Microbiol. 6*, 1243−1252.
Creagh, A. L., Ong, E., Jervis, E., Kilburn, D. G., & Haynes, C. A. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 12229−12234.
Damberger, F. F., Pelton, J. G., Harrison, C. J., Nelson, H. C. M., & Wemmer, D. E. (1994) *Protein Sci. 3*, 1806−1821.
Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., & Bax, A. (1995) *J. Biomol. NMR 6*, 277−293.
Devine, C., Ståhlberg, J., Reinikainen, T., Ruohonen, L., Pettersson, G., Knowles, J. K. C., Teeri, T. T., & Jones, T. A. (1994) *Science 265*, 524−528.
Din, N., Forsythe, I. J., Burtnick, L. D., Gilkes, N. R., Miller, R. C. J., Warren, R. A. J., & Kilburn, D. G. (1994) *Mol. Microbiol. 11*, 747−755.
Forman-Kay, J. D., Gronenborn, A. M., Kay, L. E., Wingfield, P. T., & Clore, G. M. (1990) *Biochemistry 29*, 1566−1572.
Gagné, S. M., Tsuda, S., Li, M., Chandra, L. B., Smillie, L. B., & Sykes, B. D. (1994) *Protein Sci. 3*, 1961−1974.
Garrett, D. S., Powers, R., Gronenborn, A. M., & Clore, G. M. (1991) *J. Magn. Reson. 95*, 214−220.
Griesenger, C., Otting, G., Wüthrich, K., & Ernst, R. R. (1988) *J. Am. Chem. Soc. 110*, 7870−7872.
Grzesiek, S., & Bax, A. (1992) *J. Am. Chem. Soc. 114*, 6291−6293.
Grzesiek, S., & Bax, A. (1993) *J. Am. Chem. Soc. 115*, 12593−12594.
Grzesiek, S., Anglister, J., & Bax, A. (1993a) *J. Magn. Reson. Ser. B 101*, 114−119.
Grzesiek, S., Vuister, G. W., & Bax, A. (1993b) *J. Biomol. NMR 3*, 487−493.
Hahn, M., Olsen, O., Politz, O., Borriss, R., & Heinemann, U. (1995a) *J. Biol. Chem. 270*, 3081−3088.
Hahn, M., Keitel, T., & Heinemann, U. (1995b) *Eur. J. Biochem. 232*, 849−858.
Henrissat, B. (1994) *Cellulose 1*, 169−196.
Holm, L., & Sanders, C. (1995) *J. Mol. Biol.* 123−138.
Hutchinson, E. G., & Thornton, J. M. (1996) *Protein Sci. 5*, 212−220.
Hyberts, S. G., Goldberg, M. S., Havel, T. F., & Wagner, G. (1992) *Protein Sci. 1*, 736−751.
Ikura, M., Kay, L. E., & Bax, A. (1990) *Biochemistry 29*, 4659−4667.
Ikura, M., Kay, L. E., & Bax, A. (1991) *J. Biomol. NMR 1*, 299.
Johnson, P. E., Tomme, P., Joshi, M. D., & McIntosh, L. P. (1996) *Biochemistry 35*, 13895−13906.
Kay, L. (1993) *J. Am. Chem. Soc. 115*, 2055−2057.
Kay, L., & Bax, A. (1990) *J. Magn. Reson. 86*, 110−126.
Kay, L. E., Ikura, M., & Bax, A. (1990) *J. Am. Chem. Soc. 112*, 888.
Kay, L., Keifer, P., & Saarinen, T. (1992) *J. Am. Chem. Soc. 114*, 10663−10665.
Kay, L., Xu, G.-Y., Singer, A., Muhnadiram, D., & Forman-Kay, J. (1993) *J. Magn. Reson. 101*, 333−337.
Keitel, T., Simon, O., Borriss, R., & Heinemann, U. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 5287−5291.
Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
Kraulis, P. J., Clore, G. M., Nilges, M., Jones, T. A., Petterson, G., Knowles, J., & Gronenborn, A. M. (1989) *Biochemistry 28*, 7241−7257.
Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
Lee, W., Revington, M. J., Arrowsmith, C., & Kay, L. E. (1994) *FEBS Lett. 350*, 87−90.
Levy, G. C., & Lichter, R. L. (1979) *Nitrogen-15 Nuclear Magnetic Resonance Spectroscopy*, J. Wiley & Sons, New York.
Liao, D., Kapadia, G., Ahmed, H., Vasta, G. R., & Herzberg, O. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 1428−1432.
Linder, M., Mattinen, M.-L., Kontteli, M., Lindberg, G., Ståhlberg, J., Drakenberg, T., Reinikainen, T., Pettersson, G., & Annila, A. (1995) *Protein Sci. 4*, 1056−1064.
Logan, T. M., Olejniczak, E. T., Xu, E. T., & Fesik, S. W. (1992) *FEBS Lett. 314*, 413−418.
Marion, D., Ikura, M., Tschudin, R., & Bax, A. (1989a) *J. Magn. Reson. 85*, 393−399.
Marion, D., Kay, L. E., Sparks, S. W., Torcia, D. A., & Bax, A. (1989b) *J. Am. Chem. Soc. 111*, 1515−1517.
Marion, D., Driscoll, P., Kay, L., Wingfield, P., Bax, A., Gronenborn, A., & Clore, G. (1989c) *Biochemistry 28*, 6150−6156.
Macura, S., Huang, Y., Suter, D., & Ernst, R. R. (1981) *J. Magn. Reson. 85*, 259−281.
McIntosh, L. P., & Dahlquist, F. W. (1990) *Q. Rev. Biophys. 23*, 1−38.
McIntosh, L. P., Wand, A. J., Lowry, D. F., Redfield, A. G., & Dahlquist, F. D. (1990) *Biochemistry 29*, 6341−6362.
Miller, J. (1972) *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Montelione, G. T., Lyons, B. A., Emerson, S. D., & Tashiro, M. (1992) *J. Am. Chem. Soc. 114*, 10974−10975.
Muhandiram, D., & Kay, L. (1994) *J. Magn. Reson. Ser. B 103*, 203−216.
Neri, D., Szyperski, T., Otting, G., Senn, H., & Wüthrich, K. (1989) *Biochemistry 28*, 7510−7516.
Nilges, M., Clore, G. M., & Gronenborn, A. M. (1988) *FEBS Lett. 229*, 317−324.

Olejniczak, E. T., & Fesik, S. W. (1994) *J. Am. Chem. Soc. 116*, 2215−2216.
Pascal, S. M., Muhandiram, D. R., Yamazaki, T., Kay, J. D. F., & Kay, L. E. (1994) *J. Magn. Reson. 103*, 197−201.
Poole, D. B., Hazlewood, G. P., Huskisson, N. S., Virden, R., & Gilbert, H. J. (1993) *FEMS Microbiol. Lett. 106*, 77−84.
Powers, R., Garrett, D., March, C., Frieden, E., Gronenborn, A., & Clore, G. (1993) *Biochemistry 32*, 6744−6762.
Quiocho, F. A. (1986) *Ann. Rev. Biochem. 55*, 287−315.
Quiocho, F. (1989) *Pure Appl. Chem. 61*, 1293−1306.
Reinikainen, T., Ruohonen, L., Nevanen, T., Laaksonen, L., Kraulis, P., Jones, T. A., Knowles, J. K. C., & Teeri, T. T. (1992) *Proteins: Struct., Funct., Genet. 14*, 475−482.
Reinikainen, T., Teleman, O., & Teeri, T. T. (1995) *Proteins: Struct., Funct., Genet. 22*, 392−403.
Rost, B. (1996) *Method Enzymol. 266*, 525−539.
Santoro, J., & King, G. C. (1992) *J. Magn. Reson. 97*, 202−207.
Schleucher, J., Schwendinger, M., Sattler, M., Schmidt, P., Schedletzky, O., Glase, S. J., Sørensen, O. W., & Griesinger, C. (1994) *J. Biomol. NMR 4*, 301−306.
Stanczyk, S. M., Bolton, P. H., Dell'Acqua, M., & Gerlt, J. A. (1989) *J. Am. Chem. Soc. 111*, 8317−8318.
Szyperski, T., Neri, D., Leiting, B., Otting, G., & Wüthrich, K. (1992) *J. Biomol. NMR 2*, 323−334.
Thornton, J. (1981) *J. Mol. Biol. 151*, 261−287.
Tomme, P., Warren, R. A. J., & Gilkes, N. R. (1995a) *Adv. Microbial Physiol. 37*, 1−81.
Tomme, P., Warren, R. A. J., Miller, R. C., Jr., Kilburn, D. G., & Gilkes, N. R. (1995b) in *Enzymatic Degradation of Insoluble Polysaccharides* (Saddler, J. M., & Penner, M., Eds.) pp 142−161, American Chemical Society, Washington, DC.
Tomme, P., Creagh, L., Kilburn, D., & Haynes, C. (1996) *Biochemistry 35*, 13885−13894.
Törrönen, A., Harkki, A., & Rouvinen, J., (1994) *EMBO J. 13*, 2493−2501.
Vuister, G. W., & Bax, A. (1992) *J. Magn. Reson. 98*, 428−435.
Vuister, G. W., Wang, A. C., & Bax, A. (1993) *J. Am. Chem. Soc. 115*, 5334−5335.
Wishart, D., & Sykes, B. (1994) *J. Biomol. NMR 4*, 171−180.
Wishart, D., Richards, F., & Sykes, B. (1992) *Biochemistry 31*, 1647−1651.
Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E., Markley, J. L., & Sykes, B. D. (1995) *J. Biomol. NMR 6*, 135−140.
Wittekind, M., & Mueller, L. (1993) *J. Magn. Reson. Ser. B 101*, 201−205.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, New York.
Xu, G.-Y., Ong, E., Gilkes, N. R., Kilburn, D. G., Muhandiram, D. R., Harris-Brandts, M., Carver, J. P., Kay, L. E., & Harvey, T. S. (1995) *Biochemistry 34*, 6993−7009.
Yamazaki, T., Foreman-Kay, J. D., & Kay, L. (1993) *J. Am. Chem. Soc. 115*, 11054−11055.
Zhang, O., Kay, L. E., Olivier, J. P., & Foreman-Kay, J. D. (1994) *J. Biomol. NMR 4*, 845−858.
Zuiderweg, E. R., & Fesik, S. W. (1989) *Biochemistry 28*, 2387−2391.

BI961612S