Determination of the Three-Dimensional Solution Structure of the C-Terminal Domain of Cellobiohydrolase I from *Trichoderma reesei*. A Study Using Nuclear Magnetic Resonance and Hybrid Distance Geometry–Dynamical Simulated Annealing†

Per J. Kraulis,‡,§ G. Marius Clore,*,‡ Michael Nilges,‡ T. Alwyn Jones,§ Göran Pettersson,∥ Jonathan Knowles,⊥ and Angela M. Gronenborn*,‡

Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892, Department of Molecular Biology, Biomedical Center, University of Uppsala, P.O. Box 590, S-751 24 Uppsala, Sweden, Department of Biochemistry, Biomedical Center, University of Uppsala, P.O. Box 576, S-751 23 Uppsala, Sweden, and Biotechnical Laboratory, VTT, Tietotie 2, SF-02150 Espoo, Finland

Received March 10, 1989; Revised Manuscript Received May 12, 1989

ABSTRACT: The solution structure of a synthetic 36-residue polypeptide comprising the C-terminal cellulose binding domain of cellobiohydrolase I (CT-CBH I) from *Trichoderma reesei* was investigated by nuclear magnetic resonance (NMR) spectroscopy. The ^{1}H NMR spectrum was completely assigned in a sequential manner by two-dimensional NMR techniques. A large number of stereospecific assignments for β-methylene protons, as well as ranges for the ϕ, ψ, and χ_1 torsion angles, were obtained on the basis of sequential and intraresidue nuclear Overhauser enhancement (NOE) and coupling constant data in combination with a conformational data base search. The structure calculations were carried out in an iterative manner by using the hybrid distance geometry–dynamical simulated annealing method. This involved computing a series of initial structures from a subset of the experimental data in order to resolve ambiguities in the assignments of some NOE cross-peaks arising from chemical shift degeneracy. Additionally, this permitted us to extend the stereospecific assignments to the α-methylene protons of glycine using information on ϕ torsion angles derived from the initial structure calculations. The final experimental data set consisted of 554 interproton distance restraints, 24 restraints for 12 hydrogen bonds, and 33 ϕ, 24 ψ, and 25 χ_1 torsion angle restraints. CT-CBH I has two disulfide bridges whose pairing was previously unknown. Analysis of structures calculated with all three possible combinations of disulfide bonds, as well as without disulfide bonds, indicated that the correct disulfide bridge pairing was 8–25 and 19–35. Forty-one structures were computed with the 8–25 and 19–35 disulfide bridges, and the average atomic rms difference between the individual structures and the mean structure obtained by averaging their coordinates was 0.33 ± 0.04 Å for the backbone atoms and 0.52 ± 0.06 Å for all atoms. The protein has a wedgelike shape with an amphiphilic character, one face being predominantly hydrophilic and the other mainly hydrophobic. The principal element of secondary structure is made up of an irregular triple-stranded antiparallel β-sheet composed of residues 5–9 (β1), 24–28 (β2), and 33–36 (β3) in which strand β3 is hydrogen bonded to the other two strands. Strand β1 is preceded by and overlaps with a type II turn (residues 3–6), and strands β2 and β3 are connected by a type I turn (residues 29–32). Strands β1 and β2 are connected by a loop (residues 14–19) via type II turns (residues 10–13 and 20–23). A number of side-chain interactions are discussed in the light of the structure.

Cellulose constitutes the major component of the cell walls of plants and is degraded in nature by a wide variety of microorganisms that make use of a family of enzymes, generally referred to as cellulases. The cellulolytic enzymes of one such microorganism, *Trichoderma reesei*, have been extensively studied (Eriksson, 1969; Gritzali & Brown, 1979; Gong et al., 1979; Montenecourt, 1983; Henrissat et al., 1985; Knowles et al., 1987; Enari & Niku-Paavola, 1987). The cellulases of *T. reesei* degrade cellulose in a cooperative manner and consist of two cellobiohydrolases (CBH I[1] and CBH II) and two endoglucanases (EG I and EG III), the former releasing cellobiose from the nonreducing ends of the cellulose chain and the latter cleaving internal glucosidic bonds in native cellulose (Enari & Niku-Paavola, 1987). The amino acid sequences of all four *T. reesei* cellulases have been determined

† This work was supported by the Intramural AIDS Targeted Antiviral Program of the Office of the Director of the National Institutes of Health (G.M.C. and A.M.G.) and by the Swedish Natural Science Research Council and the Swedish Board for Technical Development (T.A.J. and G.P.).

* Address correspondence to either author.
‡ National Institutes of Health.
§ Department of Molecular Biology, University of Uppsala.
∥ Department of Biochemistry, University of Uppsala.
⊥ Biotechnical Laboratory, VTT.

[1] Abbreviations: CBH, cellobiohydrolase (EC 3.2.1.91) from *Trichoderma reesei*; CT-CBH I, synthetic C-terminal domain comprising residues 462–497 of CBH I; EG, endoglucanase (EC 3.2.1.4) from *T. reesei*; BPTI, basic pancreatic trypsin inhibitor; CPI, potato carboxypeptidase inhibitor; CMTI-I, *Cucurbita maxima* trypsin inhibitor I; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, two-dimensional NOE spectroscopy; E-COSY, two-dimensional exclusive correlated spectroscopy; PE-COSY; two-dimensional primitive exclusive correlated spectroscopy; HOHAHA, two-dimensional homonuclear Hartmann–Hahn spectroscopy; rms, root mean square; SA, simulated annealing. The residue numbering scheme used for CT-CBH I in this paper corresponds to the peptide sequence; the sequence number in CBH I is obtained by adding 461 to the peptide number.

0006-2960/89/0428-7241$01.50/0

(Shoemaker et al., 1983; Fägerstam et al., 1984; Bhikhabhai & Pettersson, 1984; Penttilä et al., 1986; Teeri et al., 1987; Chen et al., 1987; Saloheimi et al., 1988) and reveal a common structural organization consisting of a catalytically active domain (~400–500 residues), a highly conserved (~70%) terminal domain (~40 residues) that may occur either at the C-terminus (cf. CBH I and EG I) or at the N-terminus (cf. CBH II and EG III), and a heavily glycosylated linker region (~30 residues) connecting the two domains. The domain structure of the cellulases has been confirmed by physicochemical (Schmuck & Pilz, 1986; Abuja et al., 1988) and functional (van Tilbeurgh et al., 1986; Langsford et al., 1987; Ståhlberg et al., 1988; Tomme et al., 1988; Johansson et al., 1989) studies. Further, these studies have revealed that the role of the terminal domain and linker region is to enhance the activity of the enzymes toward solid substrates and that the terminal domain directly interacts and binds to cellulose independently of the presence of the catalytic domain. This activity requires a unique disulfide-bridged conformation as reduction of the terminal domain eliminates its ability to bind cellulose (Johansson et al., 1989).

A more detailed understanding of the mechanism of action of these enzymes clearly requires a knowledge of their three-dimensional structures. The crystallization of an endoglucanase from *Clostridium hemocellum* has been reported (Jöliff et al., 1986). To date, there has been no success in the crystallization of intact cellulases from *T. reesei*. The domain architecture of the *T. reesei* cellulases and the observation that the domains of the CBH I retain their respective activities after cleavage suggest a dual approach to this problem involving the application of both X-ray crystallography and NMR spectroscopy. Crystals have been obtained of the catalytic domains of both CBH I (Westergren et al., 1989) and CBH II (Bergfors et al., 1989). The polypeptide produced by enzymatic cleavage of CBH I comprises both the heavily glycosylated linker region and the terminal domain. The NMR spectrum of this polypeptide is complicated by the presence of carbohydrate, which cannot be readily removed (unpublished data). Fortunately, however, it has been shown that the cellulose binding properties of a synthetic 36-residue polypeptide comprising only the C-terminal domain of CBH I (residues 462–497) are identical to those of the proteolytic fragment (Johansson et al., 1989). In this paper we present the determination of the three-dimensional solution structure of this 36-residue polypeptide using NMR spectroscopy.

In recent years NMR has been used to determine a number of protein structures in solution [see Wüthrich (1986) and Clore and Gronenborn (1987, 1989) for reviews]. Comparisons with structures determined by X-ray crystallography show that the overall polypeptide fold can be reliably defined (Williamson et al., 1985; Kline et al., 1986, 1988; Clore et al., 1986b, 1987a–c; Wagner et al., 1987). The precision, however, with which most of the NMR structures published to date have been determined has been of the order of 1–2 Å for the backbone atoms and 1.5–3 Å for all atoms, as measured by the atomic rms difference between the individual structures and the mean structure obtained by averaging their coordinates. Recently, it has been shown that the precision of an NMR structure determination can be significantly improved by incorporating restraints to stereospecifically assigned β-methylene protons (Kline et al., 1988; Driscoll et al., 1989b,c; Folkers et al., 1989). One goal of the present work was to attempt a further increase in the precision of the structure determination by stereospecifically assigning as many of the chiral methylene protons as possible.

The spectrum of CT-CBH I was completely assigned by applying a combination of two-dimensional NMR techniques (Ernst et al., 1987) to demonstrate through-bond and through-space connectivities. Stereospecific assignments of β-methylene protons, as well as ranges for ϕ, ψ and χ_1 torsion angles, were obtained on the basis of the experimental intraresidue and sequential NOEs and the $^3J_{HN\alpha}$ and $^3J_{\alpha\beta}$ coupling constants in combination with a conformational data base search (Nilges et al., 1989). By carrying out the structure determination in an iterative manner, it was possible to extend this procedure to the stereospecific assignment of glycine $C^{\alpha}H$ protons and of additional β-methylene protons that could not be unambiguously assigned initially. The final structure calculations, using the hybrid distance geometry–dynamical simulated annealing method (Nilges et al., 1988a), were based on a total of 660 experimental restraints made up of 554 interproton distance restraints, including a large number to stereospecifically assigned protons, 82 torsion angle restraints involving the ϕ, ψ, and χ_1 torsion angles, and 24 hydrogen-bonding restraints for 12 hydrogen bonds identified on the basis of NOE and amide-exchange data. A total of 41 final structures were obtained that had an atomic rms distribution about the mean coordinate positions of 0.33 ± 0.04 Å for the backbone atoms and 0.52 ± 0.06 Å for all atoms.

Experimental Procedures

Sample Preparation. The 36-residue C-terminal domain of CBH I, CT-CBH I, was synthesized by automated solid-phase synthesis (Merrifield, 1963) on an Applied Biosystems (Foster City, CA) Model 430A peptide synthesizer and purified as described previously (Johansson et al., 1989). The samples for NMR contained ~2 mM peptide, pH 3.9, in either 90% H_2O/10% D_2O or 99.996% D_2O.

NMR Spectroscopy. All NMR spectra were recorded at 600 MHz on a Bruker AM600 spectrometer. All two-dimensional spectra were recorded in the pure-phase absorption mode by using the time-proportional incrementation method (Redfield & Kunz, 1975) as described by Marion & Wüthrich (1983). NOESY (Jeener et al., 1979; Macura et al., 1981) and HOHAHA (Braunschweiler & Ernst, 1983; Davis & Bax, 1985; Bax & Davis, 1985) spectra were recorded at 15 and 27 °C in both D_2O and H_2O. Mixing times of 50, 75, 150, and 200 ms were used for the NOESY spectra. In the case of the 50-ms NOESY spectrum a 5% random variation of the mixing time was used to eliminate zero quantum coherence transfer (Macura et al., 1981). HOHAHA spectra were recorded with a WALTZ17$_y$ mixing sequence (Bax, 1989) of 35–45-ms duration sandwiched between 1.5-ms trim pulses. Typically 800–1024 t_1 increments of 2K data points were collected per experiment yielding, after appropriate zero filling, spectra with a digital resolution of 6–8 Hz in each dimension. Lorentz-to-Gaussian transformation was applied in F2 and a $\pi/6$ phase-shifted sine-square bell window function in F1.

For NOESY spectra recorded in H_2O the water resonance was suppressed either by using a semiselective "jump-return" pulse (Plateau & Guéron, 1982) in place of the last 90° pulse in the sequence or by presaturating the water resonance during the relaxation delay and the mixing time. For the HOHAHA spectra recorded in H_2O water suppression was achieved by using a 90° "flip-back" pulse, a 200-μs recovery delay, and the jump-return sequence after the WALTZ17$_y$ mixing sequence (Bax et al., 1987). In addition, the adverse effects of radiation damping were partially avoided by cycling the phase of the preparation pulse 45° out of register with those of the preparation and detection periods (Bax, Clore, Driscoll, and Gronenborn, unpublished data; Driscoll et al., 1989a). This

ensures that radiation damping starts immediately after the detection pulse and the water resonance is never fully inverted along the $-z$ axis, thereby reducing the magnitude of the radiation damping.

To eliminate base-line distortions for all NOESY and HOHAHA spectra, the acquisition parameters, in particular the relative receiver phase and the delay between the last pulse and the start of acquisition, were adjusted such that the zero- and first-order phase corrections in F2 were 90° and 180°, respectively (Marion & Bax, 1988a). Additionally, for all NOESY and HOHAHA spectra recorded in H_2O with a semiselective excitation pulse, undesirable t_2 ridges arising from the strong residual H_2O resonance were suppressed by zeroing the first point of each FID and by linear base-line correction of both the initial FIDs (which reduces the line width of the residual H_2O signal) and the F2 cross sections prior to Fourier transformation in F1.

$^3J_{HN\alpha}$ coupling constants, excluding those of glycine residues, were measured from DQF-COSY spectra (Rance et al., 1983) recorded in H_2O at 27 and 40 °C. Water suppression was achieved by selective irradiation during the relaxation delay. A total of 1024 increments of 4K data points were acquired, with a sweep width of 7353 Hz. For analysis of coupling constants, zero filling was employed to increase the digital resolution to 0.9 Hz/point in F2 and 3.68 Hz/point in F1.

Identification of direct scalar connectivities in D_2O as well as the accurate measurement of $^3J_{\alpha\beta}$ coupling constants was carried out with a PE-COSY spectrum recorded at 36 °C to obtain reduced multiplet patterns. For the same reason, a PE-COSY spectrum recorded in H_2O at 27 °C was used to measure the $^3J_{HN\alpha}$ coupling constants of the glycine residues. The PE-COSY experiment (Mueller, 1987) is a variant of the small flip angle mixing pulse COSY experiment (Bax & Freeman, 1981) (also known as β-COSY) in which the dispersive character of the diagonal in a regular β-COSY spectrum is purged. This was achieved by using the method of Marion and Bax (1988b) as follows. A reference one-dimensional FID of 8K data points was recorded with a 0° mixing pulse, $t_1 = 0$ μs, and 64 scans, eight times the number of scans per increment in the β-COSY experiment recorded with a 35° mixing pulse. By left shifting the data of this single reference FID, the time domain data for successive values of the 0° COSY experiment were obtained. The 0° COSY time domain data were then subtracted from the 35° COSY time domain data recorded with 4K data points per FID to generate the PE-COSY time domain data. The sensitivity of the PE-COSY experiment is only slightly reduced relative to that of a regular COSY experiment and a factor of at least four times higher than that of an E-COSY (Griesinger et al., 1982) experiment. A total of 1024 increments were recorded with a sweep width of 5814 Hz, yielding after appropriate zero filling a spectrum with a digital resolution of 0.73 Hz/point in F2 and 2.91 Hz/point in F1.

Spectral Assignment. Spectral assignment was aided by the use of the interactive computer graphics program ANSIG (P. J. Kraulis, T. A. Jones, G. M. Clore, and A. M. Gronenborn, unpublished data) operating on an Evans & Sutherland PS390, which enables one to assign and connect cross-peaks interactively and to maintain a list of the assignments made. In addition, it permits one to generate a list of distance restraints directly from the intensity of the assigned NOESY cross-peaks. The conformational data base search used to obtain stereospecific assignments of β-methylene protons on the basis of intraresidue and sequential interresidue NOEs and coupling constant data was carried out with the program STEREOSEARCH (Nilges et al., 1989).

Structure Computations. Three-dimensional structures were calculated on the basis of the NMR data by using the hybrid metric matrix distance geometry–dynamical simulated annealing method (Nilges et al., 1988a). The metric matrix distance geometry part of the calculations was carried out with the program DISGEO (Kuntz et al., 1979; Havel et al., 1979, 1983; Havel & Wüthrich, 1985; Havel, 1986). The dynamical simulated annealing calculations were carried out with the program XPLOR (Brünger, 1988a) which is derived originally from the program CHARMM (Brooks et al., 1983) and has been specially adapted for restrained molecular dynamics and dynamical simulated annealing calculations with NMR (Clore et al., 1985, 1986a,b; Brünger et al., 1986; Nilges et al., 1988a–c) as well as X-ray (Brünger et al., 1987; Brünger, 1988b) data. All computations were carried out on either MicroVAX III, VAX 8530, or Stellar GS1000 computers. The computed structures were inspected on an Evans & Sutherland PS390 color graphics system with the program FRODO (Jones, 1978) and its function networks interfaced to XPLOR.

A detailed description of the hybrid distance geometry–dynamical simulated annealing calculations is given by Nilges et al. (1988a), so only a brief description follows.

A set of substructures, comprising the N, C, C^α, $C^\alpha H$, C^β, nonterminal C^γ and C^δ atoms, and a pseudo atom for the Tyr rings, is computed with the metric matrix distance geometry program DISGEO without checking the triangle inequalities. This procedure takes ~5 min on a MicroVAX III computer. The coordinates of the substructures are an approximation of the complete structures and are used as a starting point for the subsequent calculations. The substructures, however, have very poor nonbonded contacts and fail to satisfy a large number of experimental interproton distance restraints with an average of more than 100 violations greater than 0.5 Å.

The starting structures for the dynamical simulated annealing calculations are obtained by best fitting one residue at a time to the subset of atoms present in the substructures. The protocol of dynamical simulated annealing involves solving Newton's equations of motion to locate the global minimum region of a target function made up of covalent, nonbonded, and experimental NMR restraints. The covalent component is made up of quadratic terms for bonds, angles, and improper torsions with force constants of 600 kcal·mol^{-1}·Å^{-2}, 500 kcal·mol^{-1}·rad^{-2}, and 500 kcal·mol^{-1}·rad^{-2}, respectively. The improper torsions serve to maintain planarity and appropriate chirality, and the conformation about the peptide bond is assumed to be planar and trans. The nonbonded component is described entirely by a van der Waals repulsion term, F_{repel}

$$\begin{aligned} F_{repel} &= 0 && \text{if } r \geq sr_{min} \\ &= k_{vdw}(s^2 r_{min}^2 - r^2)^2 && \text{if } r < sr_{min} \end{aligned} \quad (1)$$

where k_{vdw} is the force constant, r_{min} are the standard values for the sum of the van der Waals radii of two neighboring atoms as represented by the Lennard–Jones van der Waals radii used in the CHARMM empirical energy function (Brooks et al., 1983), and s is a van der Waals radius scale factor which is set to 0.8 in the present calculations. The experimental interproton distance and torsion angle restraints are described by square-well potentials F_{NOE} and F_{tor}, respectively, of the form (Clore et al., 1986b)

$$\begin{aligned} F_{NOE} &= k_{NOE}(r_{ij} - r_{ij}^u)^2 && \text{if } r_{ij} > r_{ij}^u \\ &= 0 && \text{if } r_{ij}^l \leq r_{ij} \leq r_{ij}^u \\ &= k_{NOE}(r_{ij} - r_{ij}^l)^2 && \text{if } r_{ij} < r_{ij}^l \end{aligned} \quad (2a)$$

and

$$\begin{aligned} F_{tor} &= k_{tor}(\varphi_i - \varphi_i^u)^2 && \text{if } \varphi_i > \varphi_i^u \\ &= 0 && \text{if } \varphi_i^l \leq \varphi_i \leq \varphi_i^u \\ &= k_{tor}(\varphi_i - \varphi_i^l)^2 && \text{if } \varphi_i < \varphi_i^l \end{aligned} \quad (2b)$$

where r_{ij}^u and ψ_i^u are the upper limits for the distance and torsion angle restraints, respectively, and r_{ij}^l and φ_i^l are the corresponding lower limits. The lower limits r_{ij}^l for all interproton distance restraints were set to 1.8 Å, equal to the sum of the van der Waals radii of two protons.

The dynamical simulated annealing protocol, which takes ~2 h on a MicroVax III, proceeds in four phases: (i) 200 cycles of Powell minimization with the force constants k_{NOE} and k_{tor} set to zero and k_{vdw} set to 0.1 kcal·mol^{-1}·Å^{-1} to regularize the peptide bond; (ii) 3.75 ps of dynamics at 1000 K, during which time k_{NOE} and k_{tor} are increased from 1 to 50 kcal·mol^{-1}·Å^{-2} and from 5 to 200 kcal·mol^{-1}·rad^{-2}, respectively, by doubling their values every 75 fs, k_{rep} is increased from 0.01 to 4.0 kcal·mol^{-1}·Å^{-4} by multiplying its value by $400^{1/50}$ every 75 fs, and the velocities are rescaled every 75 fs to 1000 K; (iii) 1.5 ps at 300 K rescaling the velocities every 100 fs; and (iv) 200 cycles of Powell minimization. The final values of the various force constants have been chosen empirically to ensure that the experimental restraints are satisfied within the errors of the data, that the deviations from idealized covalent geometry are small (≤0.01 Å for bonds, ≤3° for angles, and ≤2° for planes and chirality), and that the nonbonded contacts are good as judged by Lennard–Jones van der Waals energies smaller than −50 kcal·mol^{-1}.

In addition to structure calculations employing the hybrid distance geometry–dynamical simulated method, a number of calculations were carried out with a dynamical simulated annealing protocol starting from a completely random array of atoms (Nilges et al., 1988c). This particular protocol circumvents the folding problem generally associated with real space methods. The key to its success lies in setting the force constants of all terms in the target function to very low values during the early stages of the simulation. As a result, the atoms are initially only weakly coupled and can move essentially independently of each other to satisfy the restraints. This protocol is computationally more intensive than the hybrid distance geometry–dynamical simulated annealing method and takes ~20 h per structure on a MicroVAX III.

As described previously, the individual calculated structures were best fitted to each other and their coordinates averaged to yield a mean structure referred to as $\overline{SA}$ (Clore et al., 1986a). The mean structure displays poor stereochemistry and nonbonded contacts which could easily be relieved by 1000 cycles of Powell restrained minimization with only minor atomic rms shifts to yield a restrained minimized mean structure referred to as $(\overline{SA})r$.

Ring Current Shift Calculations. Ring current shift calculations were carried out with the program VNMR (Hoch et al., 1983; Redfield et al., 1983) in which ring current fields are calculated by the method of Johnson and Bovey (1958).

Results and Discussion

Sequential Assignment. The sequential assignment was carried out by identifying direct and relayed through-bond connectivities and through-space connectivities (Billeter et al., 1982; Wüthrich, 1986; Clore & Gronenborn, 1987). Both direct and relayed through-bond connectivities were detected by HOHAHA spectroscopy, and direct scalar connectivities were unambiguously identified in the DQF-COSY and PE-COSY spectra, recorded in H_2O and D_2O, respectively. Examples of the NH–aliphatic and $C^\alpha H$–aliphatic regions of the HOHAHA spectra in H_2O and D_2O, respectively, are shown in Figure 1. Through-space connectivities were established by means of NOESY spectra, an example of which is shown in Figure 2. A summary of short-range NOE connectivities involving the NH, $C^\alpha H$, and $C^\beta H$ protons, together with the slowly exchanging backbone amide protons whose resonances were still visible 24 h after the protein was dissolved in D_2O, is shown in Figure 3. The complete list of assignments at 27 °C is given in Table I. (The list of assignments at 15 °C is provided; see paragraph at end of paper regarding supplementary material.)

Stereospecific Assignment of β-Methylene Protons and Determination of Torsion Angle Restraints. Stereospecific assignments were based on intraresidue and sequential (i, i + 1) NOEs involving the NH, $C^\alpha H$, and $C^\beta H$ protons, and $^3J_{HN\alpha}$ and $^3J_{\alpha\beta}$ coupling constants. The $^3J_{HN\alpha}$ coupling constants, excluding those of the glycine residues, were measured from two DQF-COSY spectra recorded at 27 and 40 °C. No measurable difference could be detected in the values of these coupling constants at the two different temperatures. The $^3J_{HN\alpha}$ coupling constants of the glycine residues were measured from a PE-COSY spectrum recorded in H_2O at 27 °C. $^3J_{\alpha\beta}$ coupling constants were measured from a PE-COSY spectrum recorded in D_2O at 36 °C, a part of which is shown in Figure 4. (A listing of the values of the measured coupling constants is provided in the supplementary material.) The intensities of the NOEs were measured from 50-ms NOESY spectra recorded at 15 °C. The strategy employed for obtaining both stereospecific assignments as well as restraints on the values of the ϕ and ψ backbone and χ_1 side-chain torsion angles involved the use of a conformational data base search (Nilges et al., 1989). Two data bases were employed. The first was a systematic one containing interproton distances and coupling constant values for a truncated tripeptide segment with idealized geometry consisting of different ϕ, ψ, and χ_1 values, systematically varied at 10-deg intervals. The tripeptide segment comprises a central serine residue, the carbonyl group of the first residue, and the amide group of the third one. Of the 36^3 possible conformations, a small number with very severe nonbonded contacts were excluded (i.e., those centered around $\phi = \psi = 0°$) such that the resulting data base contained 37 649 conformations. In the second data base the interproton distance and coupling constant values were derived from 34 high-resolution X-ray structures with a nominal resolution of ≤2 Å and a crystallographic R factor of <20%. The ϕ and χ_1 dihedral angles were related to $^3J_{HN\alpha}$ and $^3J_{\alpha\beta}$ coupling constants, respectively, by means of empirical Karplus (1963) relationships (Pardi et al., 1984; DeMarco et al., 1978). The experimental coupling constant data were assumed to have an accuracy of better or equal to ±2 Hz. The NOE data were classified into three distance ranges, ≤2.7, ≤3.3, and ≤5 Å, corresponding to strong, medium, and weak NOEs. Additionally, relative NOE intensities for pairs of NOEs between a proton x and the two β-methylene protons of a single β-methylene group were also specified in terms of distance differences. Thus, for example, we assume in these cases that a strong NOE always represents a shorter distance than a medium or weak NOE, the approximate difference being calculated from the ratio of the NOE intensities. In general, therefore, the difference in distance between a strong and weak NOE is assumed to be ≥0.5 Å. The distances used involve the following proton pairs: $C^\alpha H(i)$–$C^{\beta 2}H(i)$, $C^\alpha H(i)$–$C^{\beta 3}H(i)$, $NH(i)$–$C^{\beta 2}H(i)$, $NH(i)$–$C^{\beta 3}H(i)$, $C^\alpha H(i)$–$NH(i+1)$, $C^{\beta 2}H(i)$–$NH(i+1)$, $C^{\beta 3}H(i)$–$NH(i+1)$, and $NH(i)$–$NH(i$

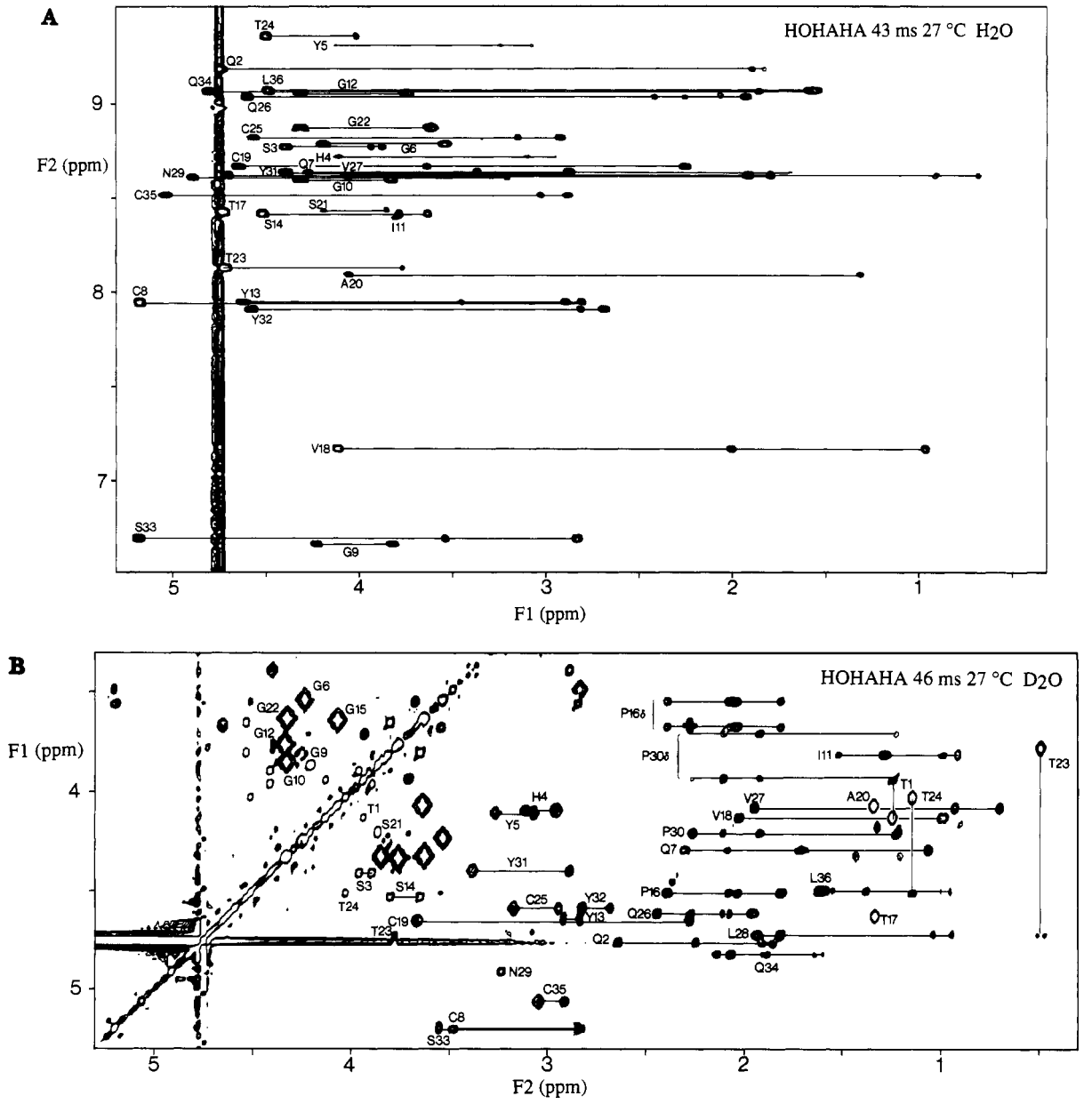

FIGURE 1: (A) NH (F2 axis)–aliphatic (F1 axis) and (B) $C^{\alpha}H$ (F1 axis)–aliphatic (F2 axis) regions of HOHAHA spectra recorded in H_2O and D_2O at 27 °C, respectively. Both direct and relayed through-bond connectivities are apparent. The labels in (A) are at the direct NH–$C^{\alpha}H$ cross-peaks.

+ 1). The stereospecific assignment is obtained by searching the two data bases and comparing the experimental restraints with the corresponding values in the data bases for both possible assignments. Whenever all the experimental restraints are matched within the specified errors by an entry in the data base, the corresponding ϕ, ψ, and χ_1 torsion angles are stored separately for each assignment. The correct stereospecific assignment is determined if the data bases contain only conformations that are consistent with one of the two possible assignments. Even if a stereospecific assignment is not obtained, the results of the search provide ranges of the three torsion angles that are consistent with the experimental data. For the special case of proline, the search is carried out by restricting the range of χ_1 to the region around 0 ± 30°. Additionally, the search can be used to obtain ranges of torsion angles for residues with only a single β-proton as the β-proton of Thr and Ile is equivalent to β_2 and that of Val to β_3. The rationale behind the use of the two data bases is that while the systematic data base contains many conformations that are not found in crystal structures, the crystallographic data base tends to contain a larger variation of conformations within particular torsion angle ranges. The latter arises from the fact that the systematic data base does not take into account small deviations from idealized geometry.

Using this approach, we were able to stereospecifically assign 15 out of 22 β-methylene groups and to obtain 27 ϕ, 22 ψ, and 25 χ_1 torsion angle restraints. Of the latter, 20 could be defined within a range of $\leq\pm 50°$. For the purposes of the calculations, the minimum deviations for the ϕ, ψ, and χ_1 angles were set to ±30°, ±50°, and ±20°, respectively, although in some cases the ranges derived from the data base searches were smaller. Additionally, the two methyl groups of Val-27 could be stereospecifically assigned on the basis of the criteria laid out by Zuiderweg et al. (1985). This was not possible for Val-18 as the chemical shifts of the methyl groups were almost degenerate.

After the initial structure calculations (see below), it became apparent that the χ_1 angles of Tyr-5, Tyr-32, and Leu-36 were all clustered around a single rotamer population, thereby enabling both stereospecific assignments and a concomitant

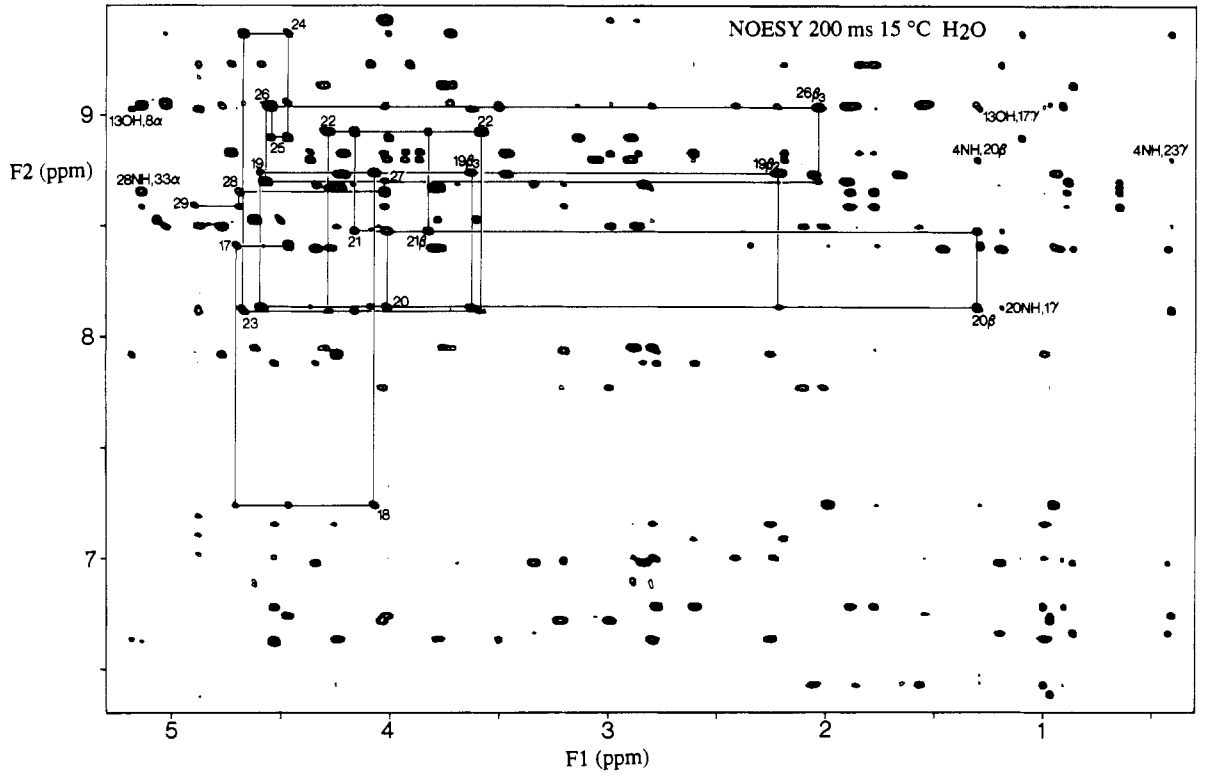

FIGURE 2: NH (F2 axis)–aliphatic (F1 axis) region of the NOESY spectrum in H_2O at 15 °C. A number of sequential $C^{\alpha}H(i)$–$NH(i+1)$ and $C^{\beta}H(i)$–$NH(i+1)$ connectivities are indicated with the labels at the intraresidue $NH(i)$–$C^{\alpha}H(i)$ or $NH(i)$–$C^{\beta}H(i)$ cross-peaks. Some long-range NOEs as well as two NOEs involving the $O^{\eta}H$ proton of Tyr-13 are also labeled.

Table I: Proton Resonance Assignments of CT-CBH I at 27 °C and pH 3.9[a]

residue	NH	$C^{\alpha}H$	$C^{\beta}H$	others
T1		3.93	4.12	$C^{\gamma}H_3$ 1.23
Q2	9.19	4.76	1.90, 1.83	$C^{\gamma}H$ 2.63, 2.23; $N^{\epsilon}H_2$[b] 8.82, 7.10*
S3	8.78	4.40	3.94, 3.88	
H4	8.73	4.08	3.10, 2.95*	$C^{\delta 2}H$ 6.82; $C^{\epsilon 1}H$ 8.56
Y5	9.31	4.09	3.25*, 3.06	$C^{\delta}H$ 6.80; $C^{\epsilon}H$ 6.47
G6	8.80	4.23*, 3.52		
Q7	8.65	4.28	2.07*, 1.69	$C^{\gamma}H$ 2.28, 1.05; $N^{\epsilon}H_2$ 7.16*, 6.98
C8	7.95	5.20	3.47*, 2.82	
G9	6.67	4.24, 3.83		
G10	8.61	4.32, 3.83*		
I11	8.40	3.80	1.27	$C^{\gamma}H$ 1.50, 0.97; $C^{\gamma}H_3$ 0.90; $C^{\delta}H_3$ 0.48
G12	9.06	4.32, 3.76*		
Y13	7.96	4.64	2.91, 2.82*	$C^{\delta}H$ 6.92; $C^{\epsilon}H$ 6.49; $O^{\delta}H$ 9.01[c]
S14	8.43	4.52	3.79*, 3.64	
G15	5.16	4.06*, 3.63		
P16		4.50	2.37, 1.80*	$C^{\gamma}H$ 2.06, 2.02; $C^{\delta}H$ 3.66, 3.53
T17	8.43	4.74	4.62	$C^{\gamma}H_3$ 1.32
V18	7.18	4.12	2.01	$C^{\gamma}H_3$ 0.99, 0.97
C19	8.68	4.65	3.65, 2.26*	
A20	8.10	4.06	1.32	
S21	8.44	4.20	3.86, 3.86	
G22	8.88	4.31, 3.62*		
T23	8.14	4.72	3.77	$C^{\gamma}H_3$ 0.48
T24	9.36	4.50	4.02	$C^{\gamma}H_3$ 1.13
C25	8.83	4.58	3.16*, 2.93	
Q26	9.04	4.61	2.07, 1.94*	$C^{\gamma}H$ 2.43, 2.26; $N^{\epsilon}H_2$ 7.04*, 6.74
V27	8.63	4.07	1.93	$C^{\gamma 1}H3$ 0.68; $C^{\gamma 2}H3$ 0.91
L28	8.63	4.72	1.91, 1.80*	$C^{\gamma}H$ 1.80; $C^{\delta}H_3$ 1.03, 0.93
N29	8.62	4.90	3.22, 3.22	$N^{\delta}H2$ 7.85*, 7.01
P30		4.20	2.24, 1.22*	$C^{\gamma}H$ 2.09, 1.90; $C^{\delta}H$ 3.92, 3.69
Y31	8.65	4.39	3.37, 2.87*	$C^{\delta}H$ 7.03; $C^{\epsilon}H$ 6.72
Y32	7.92	4.58	2.81*, 2.67	$C^{\delta}H$ 6.85; $C^{\epsilon}H$ 6.49
S33	6.70	5.19	3.54, 2.83*	
Q34	9.07	4.81	1.88, 1.61*	$C^{\gamma}H$ 2.13, 2.05; $N^{\epsilon}H_2$ 7.68*, 6.66
C35	8.52	5.05	3.03, 2.89*	
L36	9.08	4.49	1.60, 1.56*	$C^{\gamma}H$ 1.36; $C^{\delta}H_3$ 0.99, 0.94

[a] Stereospecific assignments are denoted as follows: for the C^{β} methylene protons, the asterisk indicates the $H^{\beta 2}$ proton; for the Gly C^{α} methylene protons, the asterisk indicates the $H^{\alpha 2}$ proton; for the NH_2 protons of Asn and Gln, the asterisk indicates the proton ($H^{\delta 21}$ for Asn, $H^{\epsilon 21}$ for Gln) cis to C^{β} and C^{γ}, respectively. The notation used is that of IUPAC–IUB (1970). [b] The NH_2 protons of Gln-2 were only observed at 15 °C. [c] The $O^{\eta}H$ proton of Tyr-13 was identified through an intraresidue NOE to the $C^{\epsilon}H$ ring proton.

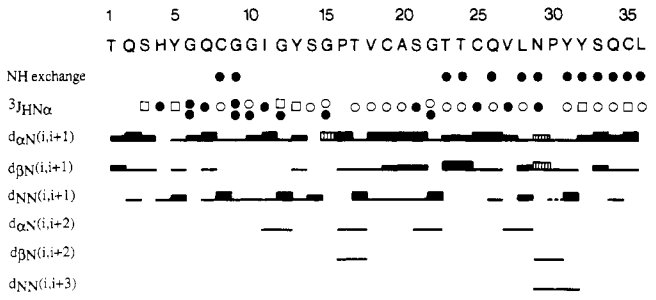

FIGURE 3: Summary of the short-range NOE data involving the NH, $C^{\alpha}H$, and $C^{\beta}H$ protons, as well as the $C^{\delta}H$ protons of proline, together with the NH exchange and $^3J_{HN\alpha}$ coupling constant data. The relative strengths of the NOEs are indicated by the thickness of the lines. NOEs involving the $C^{\delta}H$ protons of proline are given on the same line as those involving backbone NH protons and are distinguished by a hatched box. Slowly exchanging amide protons whose resonances were still visible in a HOHAHA spectrum recorded 24 h after the protein was dissolved in D_2O are indicated by solid circles (●). $^3J_{HN\alpha}$ coupling constants ≤6, 6–9, and ≥9 Hz are indicated by a solid circle (●), and open square (□), and an open circle (○), respectively.

reduction in the χ_1 angle ranges to <50° for these three residues.

The data base search on its own cannot be used to stereospecifically assign Gly residues as these are achiral within a truncated tripeptide segment. However, the conformation of the Gly residues may be forced into either the positive (30–160°) or negative region (−30° to −160°) of ϕ conformational space by nonsequential NOE restraints. Analysis of the structures produced in the early calculations (see below) indicated that the ϕ angles of each Gly residue were clustered in only one region of conformational space. With this knowledge at hand, the data base search can then be used to obtain stereospecific assignments as well as ϕ and ψ torsion angle restraints. By this means, we were able to stereospecifically assign the methylene protons of five of the six glycine residues, to obtain ϕ torsion angle restraints for all six residues and ψ torsion angle restraints for two of them (Gly-6 and Gly-10). In the case of Gly-9 no stereospecific assignment was possible as both $^3J_{HN\alpha}$ coupling constants were small (3.5 Hz) and both NH(i)–$C^{\alpha}H(i)$ NOEs weak.

Interproton Distance Restraints. NOEs were interpreted in terms of distance ranges of 1.8–2.7, 1.8–3.3, and 1.8–5.0 Å corresponding to strong, medium, and weak NOEs respectively, on the basis of 50-, 75-, and 150-ms NOESY spectra. In the case of distances involving methyl protons an additional 0.5 Å was added to the upper distance limit to account for the higher apparent intensity of methyl resonances (Clore et al., 1987a; Wagner et al., 1987). As no difference in NOE intensities within the limits of accuracy of this classification range could be detected between the data recorded at 15 and 27 °C, both data sets were used. In the case of NOEs involving nonstereospecifically assigned methylene protons, methyl groups of leucine and valine, and $C^{\delta}H$ and $C^{\epsilon}H$ protons of tyrosine, $\langle r^{-6}\rangle^{-1/6}$ averaging was used (Clore et al., 1986a). This quantity is heavily weighted toward the distance with the shorter value and thus enables the proton of closest approach to another proton to be automatically selected during the course of the dynamical simulated annealing calculations. This avoids the loss of information entailed by using a pseudo-atom representation or center averaging combined with appropriate distance corrections (Wüthrich et al., 1983). In those cases, however, where the NOE intensities from a resonance x to the two resonances of a methylene group were the same, explicit restraints to both protons were included, and similarly for the methyl groups of the leucine residues.

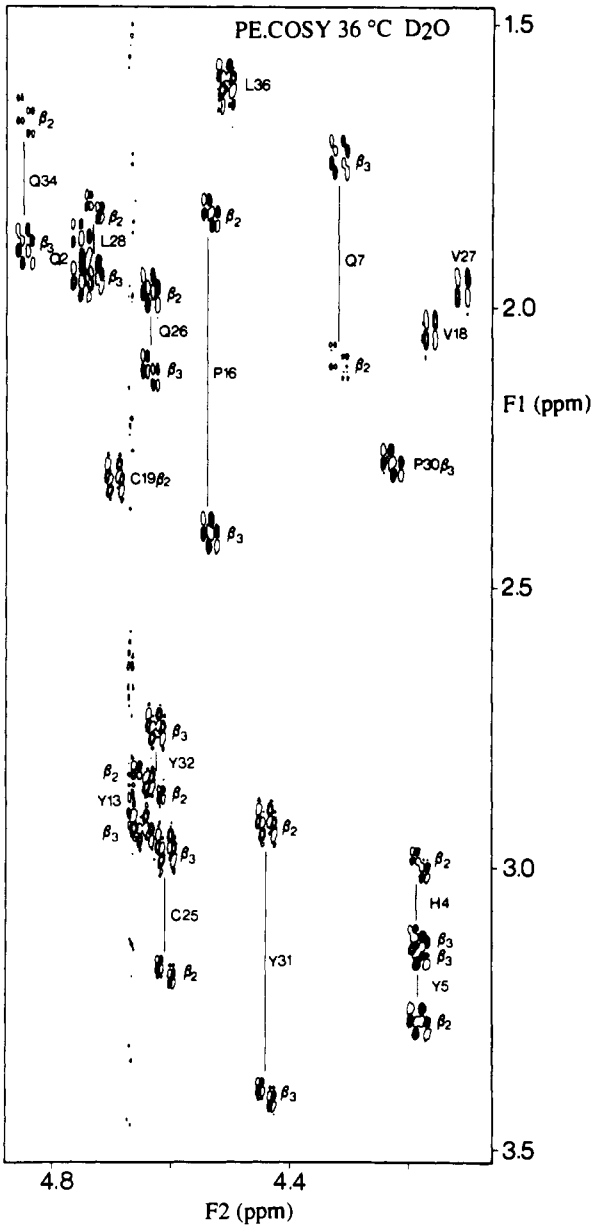

FIGURE 4: $C^{\alpha}H$ (F2 axis)–aliphatic (F1 axis) region of the PE-COSY spectrum used to measure $^3J_{\alpha\beta}$ coupling constants. Positive and negative contours are distinguished by filled and open contour levels. Stereospecific assignments of a number of β-methylene protons are indicated.

Hydrogen-Bonding Restraints. Backbone hydrogen bonds within β-sheets and β-turns can potentially be identified from a qualitative interpretation of interstrand NOEs, $^3J_{HN\alpha}$ coupling constants, and slowly exchanging amide protons, coupled with model building (Wüthrich et al., 1984; Wüthrich, 1986; Kline & Wüthrich, 1985; Wagner et al., 1986, 1987). In general, a slowly exchanging NH proton seems to be a reliable indication that this NH proton is involved in a hydrogen bond (Wüthrich, 1986), and in the case of CT-CBH I there are 13 such backbone NH protons (Figure 3). Initial model building studies on the basis of the NOE and coupling constant data suggested that two of the slowly exchanging NH protons, namely, those of Thr-23 and Asn-29, could be accounted for by β-turns from residues 20–23 and 29–32, respectively, and that 10 others could be explained by backbone NH(i)–CO(j) hydrogen bonding within a triple-stranded antiparallel β-sheet formed by residues 6–9 (strand a), 24–28 (strand b), and 33–36 (strand c). Strand c is hydrogen bonded to the other two strands, and strand b is connected to strand c by a β-turn

(residues 29–32). In addition, model building suggested the presence of a β-bulge at residues 28–29 and possibly at residues 8–9 as well. The β-bulges are indicated by the presence of medium to strong sequential NH(i)-NH(i + 1) NOEs between residues 8 and 9 and between residues 28 and 29, coupled with a number of interstrand NOEs. For the bulge formed by residues 28–29, these NOEs are NH(Asn-29)–NH(Tyr-32), NH(Asn-29)–C$^{\alpha}$H(Ser-33), and NH(Leu-28)–C$^{\alpha}$H(Ser-33), while for the 8–9 bulge they are NH(Cys-8)–NH(Cys-35), NH(Cys-8)–C$^{\alpha}$H(Gln-34), and C$^{\alpha}$H(Gly-9)–C$^{\beta}$H(Ser-33). [Note that the expected NH(Gly-9)–NH(Ser-33) NOE could not be observed as the NH proton chemical shifts of Gly-9 and Ser-33 are virtually degenerate.]

In a classical β-sheet, an interstrand backbone hydrogen bond between NH(i) on strand i and CO(j) on strand j should be manifested by three NOEs: NH(i)–NH(j), NH(i)–C$^{\alpha}$H(j + 1), and C$^{\alpha}$H(i)–C$^{\alpha}$H(j + 1) (Wagner et al., 1987). In general, however, it is often the case that not all three NOEs are observed. Indeed only two NH(i)-CO(j) hydrogen bonds display all three NOEs, namely, (i, j) = (34, 26) and (35, 6). For this reason we carried out a series of 10 initial structure calculations without any hydrogen-bonding restraints using a restraints lists that included stereospecific and torsion angle restraints obtained from the data base search described above. These confirmed the hydrogen-bonding pattern obtained from the model building studies. In these initial structures, 10 of the postulated NH(i)–CO(j) backbone hydrogen bonds consistently had N(i)–O(j) distances less than 3.3 Å and N(i)–NH(i)–O(j) angles greater than 120° and were as follows: (i, j) = (8, 33), (23, 20), (24, 36), (26, 34), (28, 32), (29, 32), (32, 29), (34, 26), (35, 6), and (36, 24). In the case of the NH protons of Gly-9 and Ser-33, however, the N(i)–O(j) distances to their respective hydrogen-bond acceptors, the carbonyl oxygen atoms of Ser-33 and Gly-9, were less than 3.3 Å in only a minority of the initial structures and tended to lie between 3.5 and 4 Å in the majority of initial structures. Analysis, however, of all the initial structures indicated that the only available hydrogen-bond acceptors for the NH protons of Gly-9 and Ser-33 were the carbonyl oxygen atoms of Ser-33 and Gly-9, respectively.

These 12 backbone hydrogen bonds account for all the slowly exchanging backbone amide protons (Figure 3), with the exception of that of Tyr-31 (see below). Two restraints were used for each NH(i)–CO(j) backbone hydrogen bond in subsequent calculations with r_{N-O} restricted to 2.4–3.3 Å and r_{NH-O} to 1.7–2.3 Å.

As a final check on the correctness of the assignment of these 12 backbone hydrogen bonds, we also carried out a series of eight structure calculations using the final set of restraints but excluding hydrogen-bonding restraints. The 12 previously explicitly incorporated hydrogen bonds were preserved in these structures, and the average atomic rms difference between these structures and the average coordinate positions of the structures calculated with the hydrogen-bonding restraints was 0.37 ± 0.04 Å for the backbone atoms and 0.55 ± 0.07 Å for all atoms. It is also worth pointing out that the inclusion of any alternative hydrogen-bonding scheme for these 12 slowly exchanging NH protons would be inconsistent with the NOE and torsion angle restraints and would lead to large violations with respect to the experimental data.

Structure Calculations. Although the assignment of the majority of NOEs could be interpreted unambiguously, a number could potentially be attributed to several alternative pairwise interactions due to chemical shift degeneracies. Such ambiguities can in most cases be resolved by inspection of low-resolution structures. For this reason we used an iterative approach to the structure determination. This involved carrying out a series of calculations with more and more restraints being incorporated at each successive stage. The large majority of structures were computed by using the hybrid distance geometry–dynamical simulated annealing protocol (Nilges et al., 1988a). As the pattern of disulfide bonding was unknown, all distance geometry substructures were calculated without incorporating disulfide restraints. These substructures were then subjected to the protocol of dynamical simulated annealing described under Experimental Procedures in which the disulfide bridges were explicitly included in the covalent terms of the target function. Dynamical simulated annealing calculations were carried out for all three possible disulfide bridge pairings as well as without disulfide bridge restraints. In addition, a number of structures were computed by using the dynamical simulated annealing protocol starting from a completely random array of atoms (Nilges et al., 1988c) in order to verify the sampling properties of the hybrid method and to ensure that stereospecific assignments derived from inspection of the calculated structures were correct. The resulting structures converged to the same global minimum region of conformational space, were distributed about the same mean, and had the same atomic rms distribution as the structures calculated from the faster hybrid method.

The final set of structures was computed with a total of 660 experimental restraints. There were 554 interproton distance restraints comprising 206 short-range ($|i - j| \leq 5$) and 137 long-range ($|i - j| > 5$) interresidue restraints and 211 intraresidue restraints, 24 distance restraints for 12 clearly identified hydrogen bonds, and 82 torsion angle restraints comprising 33 ϕ, 24 ψ, and 25 χ_1 restraints. A complete listing of all the experimental restraints is given in the supplementary material. Forty-one structures (SA) were calculated for the disulfide bridges 8–25 and 19–35, 11 (SA$_x$) without disulfide bridges, 10 (SA$_y$) for the disulfide bridges 8–19 and 25–35, and 12 (SA$_z$) for the disulfide bridges 8–35 and 19–25. Structural statistics for the structures and atomic rms differences are given in Tables II and III, respectively.

The overall polypeptide fold is not affected by the introduction of different disulfide bridge restraints (Table III). This result is clearly unusual and is due to the fact that the cysteine residues are tightly clustered in space, thereby enabling different combinations of disulfide bridges to be accommodated within a single polypeptide fold. The data in Table II indicate that the structures calculated with disulfides 8–25 and 19–35 and those calculated without disulfides satisfy the experimental restraints equally well, display similar very small deviations from idealized covalent geometry, and have good nonbonded contacts as evidenced not only by low values of the van der Waals repulsion term but, more importantly, by negative Lennard–Jones van der Waals energies of around −118 kcal·mol^{-1} (Table II). At the same time, it is clear that all the structural statistics, with the exception of those relating to the intraresidue interproton distance restraints, are worse for the other two disulfide pairing combinations and that this is particularly marked for the long-range interresidue interproton distance restraints and the torsion angle restraints (Table II). The atomic rms distribution for the four sets of structures about their respective means is similar (~0.3–0.4 Å for the backbone atoms and ~0.5–0.6 Å for all atoms). However, the structures calculated without disulfide restraints are significantly closer to those calculated with the 8–25 and 19–35 disulfide pairing than to the other two pairing combinations (Table III). Analysis of the S$^{\gamma}$–S$^{\gamma}$ distances in the

Table II: Structural Statistics[a]

	disulfide bridges				
	8–25, 19–35		none	8–19, 25–35	8–35, 19–25
	⟨SA⟩	$(\overline{SA})$r	⟨SA_x⟩	⟨SA_y⟩	⟨SA_z⟩
rms deviations from exptl distance restraints (Å)[b]					
all (578)	0.034 ± 0.002	0.024	0.033 ± 0.003	0.046 ± 0.001	0.043 ± 0.003
interresidue short range ($\|i - j\| \leq 5$) (206)	0.042 ± 0.006	0.030	0.040 ± 0.008	0.047 ± 0.005	0.046 ± 0.008
interresidue long range ($\|i - j\| > 5$) (137)	0.027 ± 0.006	0.017	0.021 ± 0.006	0.046 ± 0.006	0.050 ± 0.004
intraresidue (211)	0.029 ± 0.005	0.021	0.030 ± 0.004	0.034 ± 0.005	0.030 ± 0.005
H-bond (24)[c]	0.033 ± 0.011	0.019	0.032 ± 0.010	0.099 ± 0.009	0.061 ± 0.010
rms deviations from exptl dihedral restraints (82)(deg)[b]	1.06 ± 0.22	0.001	1.08 ± 0.35	1.99 ± 0.33	3.05 ± 0.27
F_{NOE} (kcal·mol^{-1})[d]	34 ± 5	17	33 ± 5	61 ± 4	54 ± 8
F_{tor} (kcal·mol^{-1})[e]	5 ± 2	0	3 ± 2	14 ± 4	36 ± 9
F_{repel} (kcal·mol^{-1})[f]	24 ± 6	34	20 ± 2	51 ± 5	27 ± 3
E_{L-J} (kcal·mol^{-1})[g]	−118 ± 6	−118	−117 ± 7	−105 ± 6	−110 ± 7
deviations from idealized geometry[h]					
bonds (Å) (503)	0.008 ± 0.0003	0.010	0.008 ± 0	0.011 ± 0.0004	0.010 ± 0.0004
angles (deg) (896)	2.153 ± 0.010	2.170	2.156 ± 0.017	2.219 ± 0.006	2.221 ± 0.014
impropers (deg) (227)	0.818 ± 0.027	0.911	0.570 ± 0.044	1.017 ± 0.043	0.984 ± 0.037

[a] The notation of the structures is as follows: ⟨SA⟩ are the 41 final dynamical simulated annealing structures with disulfide bridges between Cys-8 and Cys-25 and between Cys-19 and Cys-35; $\overline{SA}$ is the mean structure obtained by averaging the coordinates of the 41 individual SA structures best fitted to each other; $(\overline{SA})$r is the structure obtained by restrained minimization of $\overline{SA}$; ⟨SA_x⟩ are the 11 dynamical simulated annealing structures calculated without any disulfide bridges; ⟨SA_y⟩ are the 10 dynamical simulated annealing structures with disulfide bridges between Cys-8 and Cys-19 and between Cys-25 and Cys-35; and ⟨SA_z⟩ are the 12 dynamical simulated annealing structures with disulfide bridges between Cys-9 and Cys-35 and between Cys-19 an Cys-25. [b] The rms deviations from the experimental restraints are calculated with respect to the upper and lower limits of the distance and dihedral restraints (Clore et al., 1986b). None of the structures exhibited distance violations greater than 0.5 Å. The number of restraints in each category is given in parentheses next to the category name. [c] For each backbone hydrogen bond there are two restraints: r_{NH-O} 1.7–2.3 Å and r_{N-O} 2.4–3.3 Å. Twelve backbone hydrogen bonds within regular elements of secondary structure were identified on the basis of the NOE and NH exchange data (see text). [d] The values of the square-well NOE potential F_{NOE} (eq 2) are calculated with a force constant of 50 kcal·mol^{-1}·Å^{-2}. [e] The values of F_φ are calculated with a force constant of 200 kcal·mol^{-1}·rad^{-2}. F_φ is a square-well dihedral potential (eq 3) that is used to restrict the ranges of 33 ϕ, 24 ψ, and 25 χ_1 torsion angles (see text). [f] The values of the van der Waals repulsion term F_{repel} (cf. eq 1) are calculated with a force constant of 4 kcal·mol^{-1}·Å^{-4} with the hard-sphere van der Waals radii set to 0.8 times the standard values used in the CHARMM empirical energy function (Brooks et al., 1983). [g] E_{L-J} is the Lennard–Jones van der Waals energy calculated by using the CHARMM empirical energy function (Brooks et al., 1983). [h] The number of bond, angle, and improper terms is given in parentheses. The improper torsion terms serve to maintain planarity and appropriate chirality; they also maintain the peptide bond of all residues in the trans conformation. In the dynamical simulated annealing calculations, the restraints for the disulfide bridges are included in the bond, angle, and improper torsion terms.

structures calculated without the disulfide restraints yields values of 4.1 ± 0.6, 3.6 ± 0.5, 3.0 ± 0.06, 5.0 ± 0.4, 3.2 ± 0.3, and 5.9 ± 0.7 Å for the distances between residues 8–19, 8–25, 8–35, 19–25, 19–35, and 25–35, respectively. The long distance between the S^γ atoms of residues 19 and 25 would exclude the 8–35 and 19–25 disulfide pairing, whereas the long distance between the S^γ atoms of residues 25 and 35 would exclude the 8–19 and 25–35 disulfide pairing. Finally, the only NOEs between cysteine residues that are observed are those between the β-methylene protons of Cys-19 and Cys-35. Taking all these facts into account leaves little doubt that the correct disulfide pairing is 8–25 and 19–35. Consequently, all further discussion is restricted to the structures calculated with the 8–25 and 19–35 disulfide bridges.

As an additional qualitative measure of the consistency of the converged SA structures with the experimental NOESY spectra, two simulated NOESY spectra were computed from the interproton distances in the structures and the proton chemical shift values. One simulated spectrum was calculated solely on distance criteria in which interproton distances between 1.8–2.7, 2.7–3.3, and 3.3–4 Å were classified as strong, medium, and weak NOEs, respectively. The other simulated spectrum was computed by complete relaxation matrix analysis (Keepers & James, 1984; Borgias & James, 1988) and by plotting NOE intensities at a cutoff corresponding to an interproton distance of ~4 Å. A comparison between the 200-ms NOESY spectrum in H_2O and the two simulated spectra for the restrained minimized mean structure $(\overline{SA})$r is included as supplementary material. There is good qualitative agreement between the experimental and simulated spectra. Nearly all the peaks present in the experimental spectrum are present in the simulated ones. In addition, only three to four weak cross-peaks (corresponding to distances larger than 3.3 Å) are present in the simulated spectra but absent in the experimental one. These results indicate that the calculated structures provide a good representation of the experimental NOE data.

The Converged Structures. The polypeptide fold and structural details of the elements of regular secondary structure of CT-CBH I are shown by the backbone atom superpositions of all 41 SA structures in Figure 5. Best fit superpositions of the side chains of some selected regions of the protein are illustrated in Figure 6. The atomic rms distribution of the individual SA structure about the mean coordinate positions as a function of residue number is shown in Figure 7, and the angular rms distribution of ϕ and ψ torsion angles is available as supplementary material. Both the backbone conformation and most of the side-chain positions are very well defined with an average atomic rms difference between the individual SA structures and the mean structure of 0.33 ± 0.04 Å for the backbone atoms and 0.52 ± 0.06 Å for all atoms and an average angular rms difference for the ϕ and ψ backbone torsion angles of 11 ± 8°. The largest deviations in both atomic and backbone angular rms differences occur at the N-terminus (residues 1 and 2) and in the segment from residues 13 to 15. With respect to the side chains, only seven residues have atomic rms distributions about the mean coordinate positions of ≥1 Å. For three of these, Gln-2, Ser-21, and Asn-29, stereospecific assignments of the β-methylene protons could not be obtained. In the case of Ser-21 and Asn-29, this was due to $C^\beta H$ chemical shift degeneracy, while the $^3J_{\alpha\beta}$ coupling constant data for Gln-2 ($^3J_{\alpha\beta a} = {}^3J_{\alpha\beta b} = 7.5$ Hz) are indicative of multiple χ_1 side-chain conformations. For the other four residues, Gln-26, Leu-28, Gln-34, and

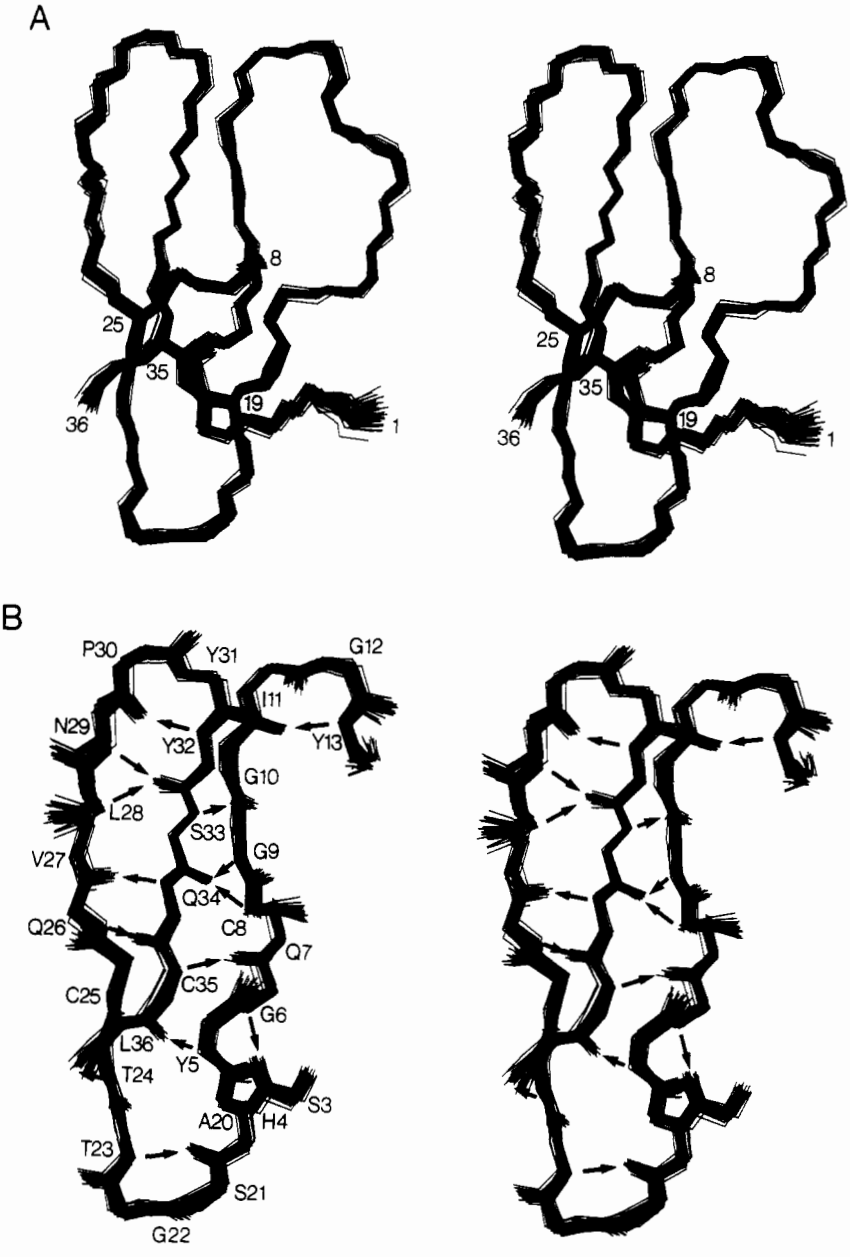

FIGURE 5: (A) Stereoview showing a best fit superposition of the backbone (N, C^{α}, C) atoms of the 41 SA structures. The disulfide bridges are also shown. (B) Stereoview showing a best fit superposition of the backbone (N, C^{α}, C, O) atoms of the 41 SA structures for the regions of regular secondary structure comprising residues 3–13 and 20–36. Backbone hydrogen bonds are shown as arrows with the direction of the arrow from the N to the O atom. There are three type II turns formed by residues 3–6, 10–13, and 20–23 and one type I turn formed by residues 29–32.

Leu-36, the side chains are highly solvent accessible and the density of interproton distances constraining the side-chain positions is low.

All the non-glycine residues with the exception of Tyr-5 lie densely clustered in the allowed regions (Ramachandran & Sasisekharan, 1968; Burgess et al., 1974) of the Ramachandran ϕ, ψ plot (Figure 8).

CT-CBH I has a wedgelike shape with overall dimensions of approximately 30 × 18 × 10 Å (Figure 5A). The hydrogen-bonding pattern is shown schematically in Figure 9. The main secondary structure element consists of an irregular triple-stranded antiparallel β-sheet in which strand β3 (residues 33–36) hydrogen bonds to strand β1 (residues 5–9) and strand β2 (residues 24–28). It is distorted by two β-bulges, the first comprises Cys-8 and Gly-9 whose NH protons hydrogen bond to the carbonyl oxygen atom of Ser-33, and the second Leu-28 and Asn-29 whose NH protons hydrogen bond to the carbonyl oxygen atom of Tyr-32 (Figures 5B and 9). Of the 17 backbone hydrogen bonds shown in Figure 9, 5 involve NH protons that are not slowly exchanging and were identified by inspection of the final converged structures. These five NH-(i)–CO(j) backbone hydrogen bonds are (5, 35), (6, 3), (13, 10), (20, 2), and (2, 18). All of these occur in solvent-exposed positions in the molecule and either are located at the end of a β-sheet structure (e.g., 5N, 35O) or are relatively isolated.

Strands β2 and β3 are connected by an approximate type I turn (residues 29–32), according to the classification of Venkatachalam (1968), which is characterized by ϕ, ψ angles of $(-69 \pm 3°, 25 \pm 2°)_{\text{Pro-30}}$ and $(-152 \pm 4°, 12 \pm 8°)_{\text{Tyr-31}}$ and a hydrogen bond between Asn-29 CO and Tyr-32 NH. The NH of Tyr-31 is slowly exchanging, and the only available hydrogen-bonding partner is the $O^{\delta 1}$ atom of Asn-29. The χ_2 torsion angle of Asn-29 is not very well defined as it is restricted by only a few NOEs from the $N^{\delta 2}H_2$ group of Asn-29 to the $C^{\beta}H$ and methyl protons of Leu-28. Nevertheless, in about half the SA structures, the $O^{\delta 1}$ atom of Asn-29 is within

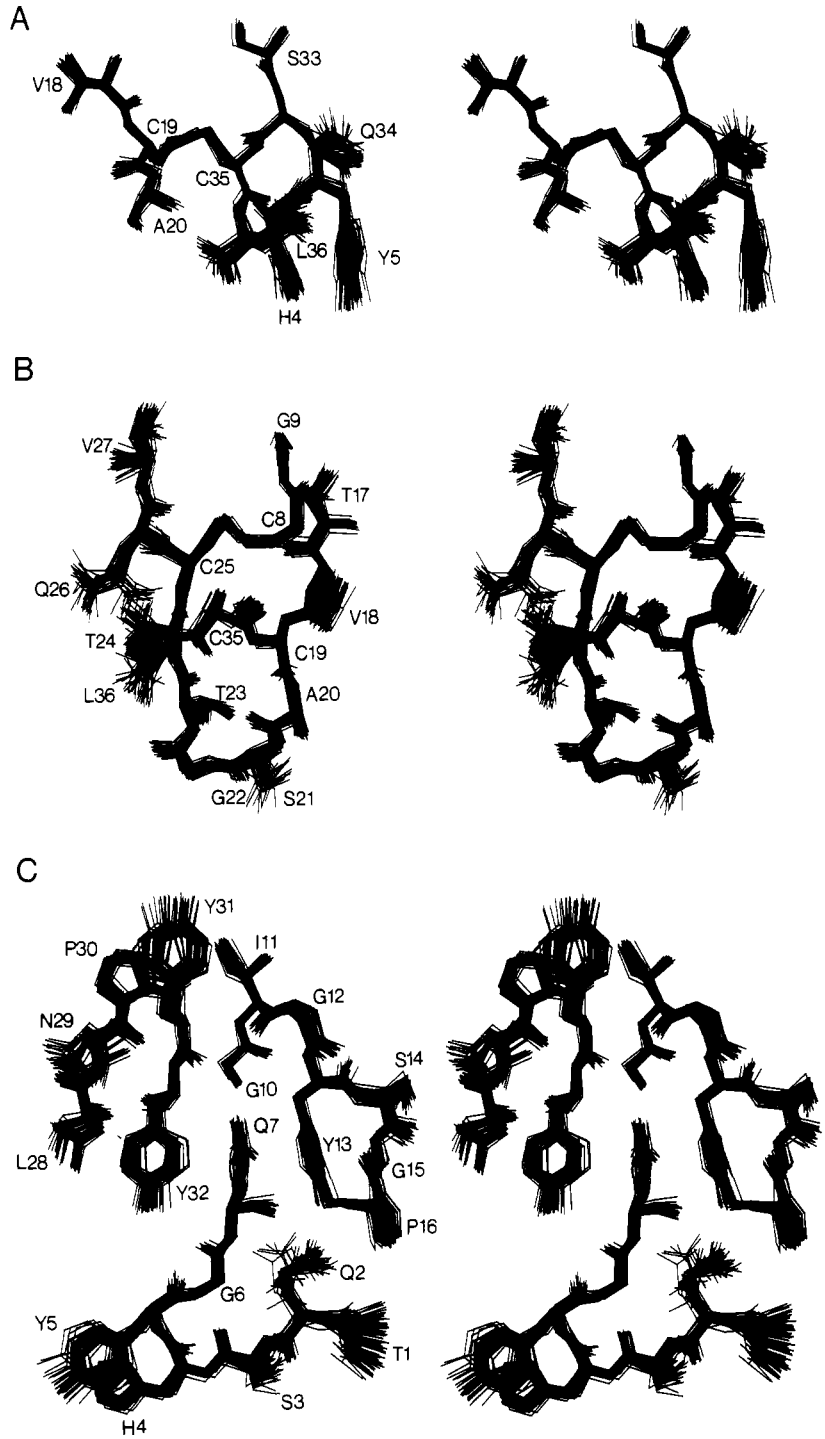

FIGURE 6: Stereoviews showing best fit superposition of all atoms (excluding protons) of the 41 SA structures for three selected segments of the protein. (A) Residues 4–5, 18–20, and 33–36, (B) residues 8–9, 17–27, and 35–36, and (C) residues 1–7, 10–16, and 28–32.

hydrogen-bonding distance of the NH of Tyr-31 (Figure 6C). This is consistent with the frequent occurrence of an asparagine residue in the first position of a type I turn with a hydrogen bond from its side-chain CO to the backbone NH proton of the $n + 2$ residue (Richardson, 1981; Baker & Hubbard, 1984). It would also account for the conservation of Asn-29 in the homologous sequences of the terminal domains of the four *T. reesei* cellulases (Figure 10).

The type I turn at residues 29–32 is immediately preceded by the β-bulge involving the amide protons of Leu-28 and Asn-29 and the carbonyl group of Tyr-32. Such a β-hairpin formed by this combination of β-bulge and type I turn is a structural feature that can be found in other proteins. A search of a data base of crystallographically determined protein structures for fragments that satisfied the NOE distances involving the NH, $C^{\alpha}H$, and $C^{\beta}H$ protons (Kraulis & Jones, 1987) within the sequence 27–33 resulted in four examples of this particular β-hairpin in the trypsin-like serine proteases. Residues 46–52 (using the residue numbering of chymotrypsinogen) of trypsin (Marquart et al., 1983), rat mast cell protease II (Reynolds et al., 1985), kallikrein (Bode et al., 1983), and chymotrypsin (Blevins & Tulinsky, 1985) superimpose within a C^{α} atomic rms deviation of 0.5 Å onto residues 27–33 of CT-CBH I. Three of these β-hairpins have an asparagine residue in position 1 of the turn, while the fourth has a serine, and in three out of the four cases there is a hydrogen bond between a side-chain atom of this residue and the backbone amide proton of the residue in position 3 of the turn.

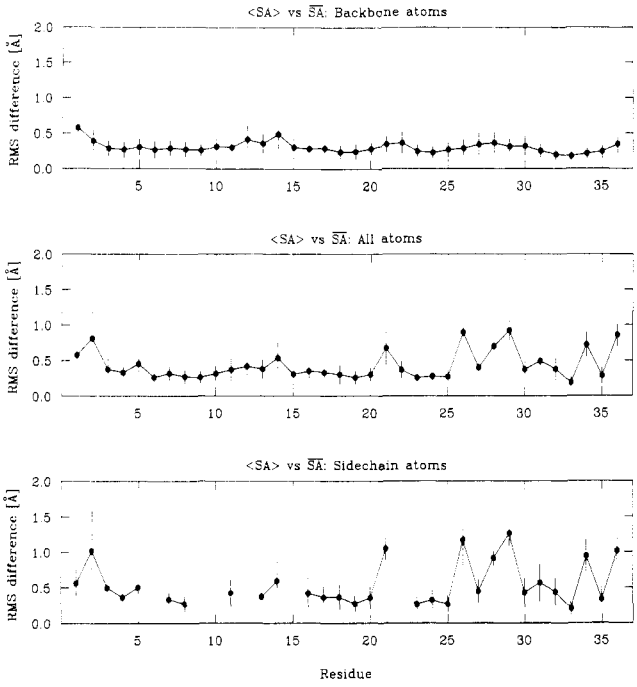

FIGURE 7: Atomic rms distribution of the 41 individual SA structures about the mean structure $\overline{SA}$ best fitted to residues 1–36. The filled-in circles (●) represent the average rms difference at each residue between the individual structures and the mean structure, and the bars represent the standard deviations in these values.

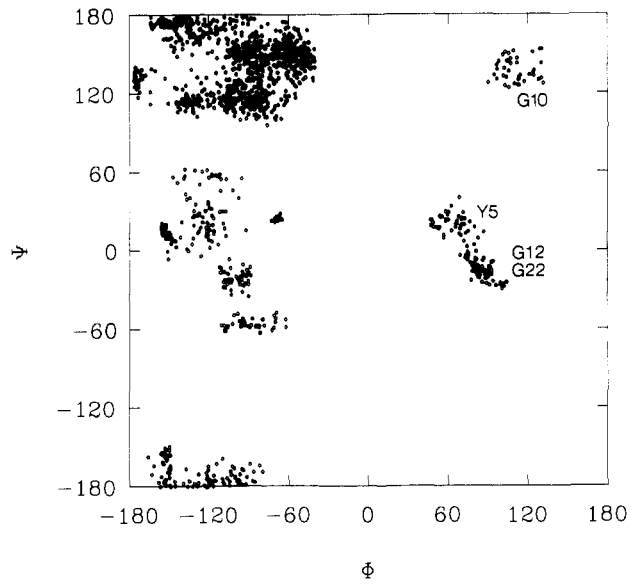

FIGURE 8: Ramachandran ϕ, ψ plot for the 41 SA structures. All non-glycine residues with the exception of Tyr-5 lie within the allowed region of ϕ, ψ phase space (Ramachandran & Sasisekharan, 1968; Burgess et al., 1974; Richardson, 1981). The third residue position in the three type II turns present in CT-CBH I is occupied by Tyr-5, Gly-12, and Gly-22.

The β1 strand is overlapped and preceded by a type II turn (residues 3–6) with ϕ, ψ angles of $(-62 \pm 12°, 14 \pm 14°)_{His\text{-}4}$ and $(64 \pm 13°, 22 \pm 10°)_{Tyr\text{-}5}$ and a hydrogen bond between the CO of Ser-3 and the NH of Gly-6. This type II turn is unusual insofar as it has a tyrosine residue rather than a glycine at position 3 in a left-handed helical conformation (Figure 9). Such a conformation for a non-glycine residue is generally disfavored due to steric clash between the C^β atom of residue $n + 2$ (Tyr-5) and the carbonyl oxygen atom of residue $n + 1$ (His-4). This distance has a value of 2.9 Å in CT-CBH I, slightly larger than the sum of the van der Waals

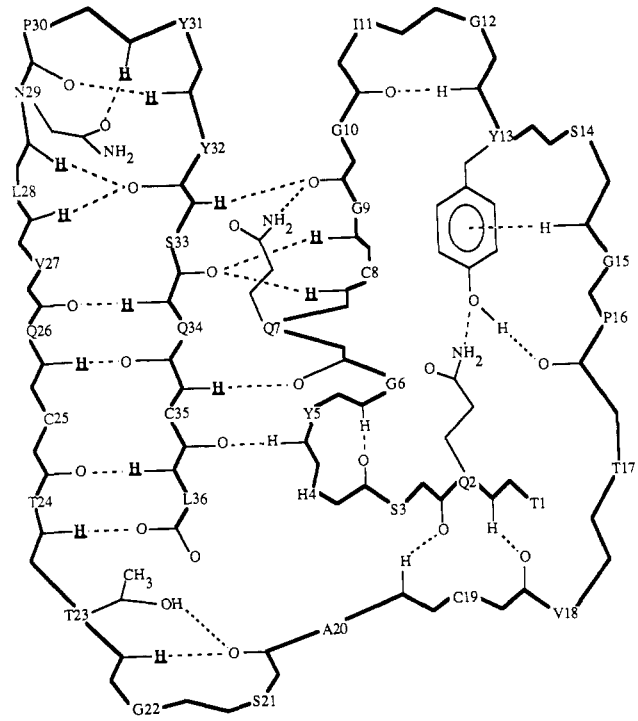

FIGURE 9: Schematic representation of the hydrogen-bond pattern in the CT-CBH I structure. Slowly exchanging amide protons are in bold letters and are underlined. Hydrogen bonds involving backbone NH and CO groups were deduced from the observation of slowly exchanging amide protons and a qualitative interpretation of the NOEs involving the NH, $C^\alpha H$, and $C^\beta H$ protons (see text). The hydrogen bonds involving other protons were derived on the basis of the calculated structures.

		0	1	5	10	15	20	25	30	35	36
CBH I	461	P	TQSHY	GQCGG	IGYSG	PTVCA	SGTTC	QVLNP	YYSQC	L	
CBH II	3	A	CSSVW	GQCGG	QNWSG	PTCCA	SGSTC	VYSND	YYSQC	L	
EG I	401	C	TQTHW	GQCGG	IGYSG	CKTCT	SGTTC	QYSND	YYSQC	L	
EG III	0	-	QQTVW	GQCGG	IGWSG	PTNCA	PGSAC	STLNP	YYAQC	I	

FIGURE 10: Sequence alignment of the homologous terminal domains of the four *T. reesei* cellulases. The numbering above the sequences (from residue 1) is that of the synthetic polypeptide CT-CBH I used in this work. The column 0 is included to show the preceding cysteine residue in EG I (see text). The numbers at the left of the figure are the residue numbers in the native enzyme for the column 0 residues. The domain is C-terminal in CBH I and EG I and N-terminal in CBH II and EG III. Residues conserved in all four sequences are boxed.

radii of carbon and oxygen atoms (2.8 Å). This relatively unfavorable steric interaction can potentially be relieved by flipping the ϕ backbone torsion angle of Tyr-5 by 180°, resulting in a type I turn. To investigate the factors favoring the type II over the type I turn in this case, we therefore performed a number of structure calculations in which the ϕ torsion angle of Tyr-5 was restrained to $-120 \pm 30°$, thereby forcing a type I turn. A comparison of the two sets of structures revealed that in those with the type II turn there is a hydrogen bond between the NH of Tyr-5 and the carbonyl oxygen atom of Cys-35 (Figures 5A and 9), while those with the type I turn have a very short distance of ~2.3 Å between the carbonyl oxygens of His-4 and Cys-35. The latter interaction is unfavorable not only from a steric viewpoint (as the sum of the van der waals radii of two oxygen atoms is 2.6 Å) but also from an electrostatic one as well. Interestingly, a very similar turn has recently been found in the X-ray structure of glycolate oxidase in which the NH of the tyrosine residue at position 3 is also involved in hydrogen bonding to another residue, in that case the $O^{\gamma 1}$ atom of a threonine (Y. Lindquist, personal communication).

Table III: Atomic rms Differences[a]

structures	atomic rms difference (Å) backbone atoms	all atoms
Disulfides 8–25 and 19–35		
$\langle SA \rangle$ vs $\overline{SA}$[b]	0.33 ± 0.04	0.52 ± 0.06
$\langle SA \rangle$ vs $(\overline{SA})r$	0.36 ± 0.07	0.56 ± 0.08
$(\overline{SA})r$ vs $\overline{SA}$	0.06	0.22
No Disulfides		
$\langle SA_x \rangle$ vs $\overline{SA_x}$	0.36 ± 0.07	0.58 ± 0.09
$\langle SA_x \rangle$ vs $\overline{SA}$	0.45 ± 0.07	0.67 ± 0.10
Disulfides 8–19 and 25–35		
$\langle SA_y \rangle$ vs $\overline{SA_y}$	0.36 ± 0.06	0.58 ± 0.08
$\langle SA_y \rangle$ vs $\overline{SA}$	0.72 ± 0.06	0.91 ± 0.04
Disulfides 8–35 and 19–25		
$\langle SA_z \rangle$ vs $\overline{SA_z}$	0.39 ± 0.10	0.58 ± 0.13
$\langle SA_z \rangle$ vs $\overline{SA}$	0.63 ± 0.05	0.81 ± 0.07
Comparison of Mean Structures		
$\overline{SA_x}$ vs $\overline{SA}$	0.26	0.33
$\overline{SA_y}$ vs $\overline{SA}$	0.62	0.71
$\overline{SA_z}$ vs $\overline{SA}$	0.48	0.55
$\overline{SA_y}$ vs $\overline{SA_x}$	0.58	0.68
$\overline{SA_z}$ vs $\overline{SA_x}$	0.45	0.53
$\overline{SA_z}$ vs $\overline{SA_y}$	0.65	0.74

[a] The notation of the structures is the same as that in Table II. In addition, $\overline{SA_x}$, $\overline{SA_y}$, and $\overline{SA_z}$ are the mean coordinates obtained by averaging the coordinates of the individual $\langle SA_x \rangle$ (no disulfides), $\langle SA_y \rangle$ (disulfides 8–19 and 25–35), and $\langle SA_z \rangle$ (disulfides 8–35 and 19–25) structures, respectively. There are 41 $\langle SA \rangle$ (disulfides 8–25 and 19–35), 11 $\langle SA_x \rangle$, 10 $\langle SA_y \rangle$, and 12 $\langle SA_z \rangle$ structures. [b] By standard statistical theory it is easily shown that the average rms difference between all pairs of SA structures is related to the average rms difference between the individual SA structures and the mean $\overline{SA}$ structure by a factor of $\sim[2n/(n-1)^{1/2}$, which in this case is equal to 1.43. Further, the standard atomic rms error s_{mean} in the coordinates of the average mean structure $\overline{SA}$ is given by $[\sum(\text{rmsd}_i)^2/n(n-1)]^{1/2}$, where rmsd_i is the atomic rms difference between the ith SA structure and the mean SA structure and n is the number of structures.

There are two other type II turns in the structure, one (residues 10–13) connects β-strand 1 to the loop comprising residues 14–19; the other (residues 20–23) connects the loop to β-strand 2. Both these turns have characteristic CO(i)–NH(i + 3) hydrogen bonds with a glycine residue at position 3. The ϕ, ψ angles for these two turns are (−49 ± 11°, 150 ± 6°)$_{\text{Ile-11}}$ and (80 ± 6°, −9 ± 14°)$_{\text{Gly-12}}$, and (−57 ± 11°, 148 ± 8°)$_{\text{Ser-21}}$ and (89 ± 11°, −17 ± 10°)$_{\text{Gly-22}}$. Gly-22 is conserved in the four homologous terminal domain sequences (Figure 10). Gly-12, on the other hand, is present in three of the sequences but is substituted by asparagine in CBH II (Figure 10). This substitution, however, can be regarded as conservative as asparagine is frequently found to adopt a glycine conformation in type II turns (Richardson, 1981). The conformation of the 20–23 turn is further stabilized by a side-chain–backbone hydrogen bond between the $O^{\gamma 1}H$ of Thr-23 and the carbonyl oxygen atom of Ala-20 (Figures 6B and 9). The presence of an $O^{\gamma 1}H$ proton at position 23 is conserved in the four sequences by the presence of either a threonine (CBH I and EG I) or a serine (CBH II and EG III).

Examination of the SA structures suggests that the conformation of the central portion of the loop is principally stabilized by hydrogen bonds between the $O^{\eta}H$ of Tyr-13 and the carbonyl oxygen atom of Pro-16 with $r_{O^{\eta}H(13)-O(16)} \sim 3.1$ Å and the angle $O^{\eta}(13)$–$O^{\eta}H(13)$–O(16) ~140–150° (Figure 6C). Further, in most of the SA structures there appears to be a hydrogen bond between the $N^{\epsilon 22}H$ of Gln-2 and the O^{η} of Tyr-13 (Figure 6C). This, together with its position in the plane of the aromatic ring, would account for the observation that it is highly downfield shifted, resonating at 8.82 ppm. In contrast, the NH proton of Gly-15 lies directly on top of the plane of aromatic ring, accounting for its large upfield-shifted resonance position at 5.16 ppm. This is consistent with the results of ring current shift calculations that predict an upfield shift of ~2 ppm. The distal end of the loop is stabilized by two backbone hydrogen bonds between the NH of Gln-2 and the CO of Val-18, and between the NH of Ala-20 and the CO of Gln-2.

The distances between the center of the aromatic ring of Tyr-13 and the NH and N atoms of Gly-15 are ~2.6 and 3.4 Å, respectively. This suggests the possibility of a NH–aromatic ring hydrogen bond (Perutz et al., 1986; Levitt & Perutz, 1988). We note, however, that the position of the NH group of Gly-15 is not absolutely optimal for such an interaction as, although it is equidistant from the C^{ζ} and C^{γ} atoms of Tyr-13, it is closer to one edge of the ring than to the other ($r_{NH(15)-C^{\delta 1}/C^{\epsilon 1}(13)}$ ~2.5 Å, $r_{NH(15)-C^{\delta 2}/C^{\epsilon 2}(13)}$ ~3.4 Å; $r_{N(15)-C^{\delta 1}/C^{\epsilon 1}(13)}$ ~3.4 Å, $r_{N(15)-C^{\delta 2}/C^{\epsilon 2}(13)}$ ~4.0 Å). Interestingly, there is a very similar interaction in basic pancreatic trypsin inhibitor (BPTI) where the structure determined by joint refinement of X-ray and neutron diffraction data (Wlodawer et al., 1984) shows that the NH proton of Gly-37 is located almost exactly on the symmetry axis perpendicular to the plane of the Tyr-35 ring at a distance of 2.8 Å, consistent with the observation that its resonance is highly upfield shifted with a chemical shift of 4.3 ppm (Tüchsen & Woodward, 1987). In BPTI there is also a hydrogen bond from the side-chain group of Asn-44 to the other side of the ring of Tyr-35. In CT-CBH I the side chain of Gln-7 is located on the other side of the Tyr-13 ring, but neither the computed structures nor the chemical shift values of the side-chain amide protons provide any evidence for a comparable hydrogen-bonding interaction. In this respect it is interesting to note that the aromatic ring of Tyr-35 in BPTI flips slowly on the NMR time scale, giving rise to four separate aromatic resonances (Wüthrich & Wagner, 1978), while that of Tyr-13 of CT-CBH I flips rapidly so that only two aromatic resonances are seen. A possible explanation for this difference may be afforded by the hydrogen-bonding pattern to the respective aromatic rings, with the two hydrogen bonds in the case of BPTI hindering mobility to a greater extent than the single one in CT-CBH I.

The tyrosine at position 13 is not absolutely conserved in the four *T. reesei* cellulases (Figure 10). While the terminal domains of CBH I and EG I have a tyrosine in this position, those of CBH II and EG II have a tryptophan. The aromatic ring of tryptophan could act as a hydrogen-bond acceptor for the NH of Gly-15, and a hydrogen bond between the CO side-chain group of Gln-2 and the $N^{\epsilon}H$ atom of tryptophan could substitute for the hydrogen bond between the side-chain amide group of Gln-2 and the Tyr-13 O^{η} hydroxyl oxygen in the CT-CBH I structure.

A number of other side-chain interactions are of interest. Gln-7 is conserved in the terminal domains of the four *T. reesei* cellulases and is located between the aromatic rings of Tyr-32 and Tyr-13, whose planes are approximately orthogonal to each other. The resonance of one of the $C^{\gamma}H$ protons of Gln-7 is highly upfield shifted with a chemical shift of 1.05 ppm. This is consistent with the results of ring current shift calculations which indicate that the $C^{\gamma 2}H$ proton of Gln-7 should be upfield shifted by ~1 ppm. In addition, the $N^{\epsilon 21}H$ side-chain amide proton of Gln-7 forms a hydrogen bond to the

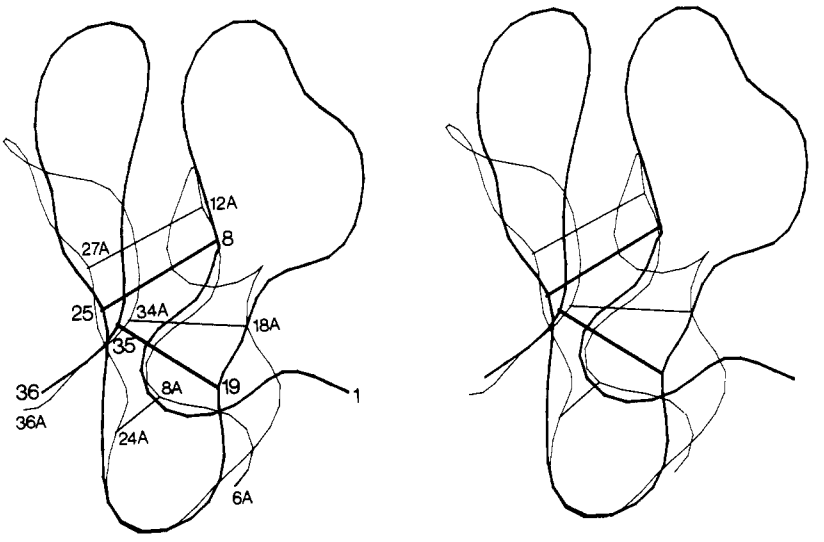

FIGURE 11: Best fit superposition of a smoothed backbone (N, C^{α}, C) atom representation of the restrained minimized mean structure ($\overline{SA}$)r of CT-CBH I (thick line) and the X-ray structure of residues 6–36 of potato carboxypeptidase inhibitor (thin line). The X-ray structure of CPI was solved by Rees and Lipscomb (1983). The larger numbers relate to the CT-CBH I residues, while the smaller numbers followed by a capital A refer to those in CPI. The alignment is as follows: residues 3–9, 20–28, and 33–36 of CT-CBH I superimpose on residues 7–13, 21–29, and 32–35 of CPI with an atomic rms difference of ~1.6 Å. The disulfide bridges are represented by straight lines connecting the C^{α} atoms. The Cys-8–Cys-25 and Cys-19–Cys-35 disulfide bridges in CT-CBH I correspond to the Cys-12–Cys-27 and Cys-18–Cys-34 disulfide bridges in CPI. In addition, CPI has a third disulfide bridge between Cys-8 and Cys-24.

carbonyl oxygen atom of Gly-9 [$r_{N^{\epsilon 21}(7)-O(9)}$ ~3 Å and angle $N^{\epsilon 21}(7)$–$N^{\epsilon 21}H(7)$–O(9) ~150–160°], thereby stabilizing the β-bulge at Gly-9 in strand $\beta 1$. The $C^{\delta}H3$ group of Ile-11 and the $C^{\beta 2}H$ proton of Pro-30 are located on opposite sides of the plane of the aromatic ring of Tyr-31 (Figure 6C), and their upfield-shifted resonances (cf. Table I) can also be attributed to ring current shifts. Likewise, the methyl group of Thr-23 is situated on top of the plane of the aromatic ring of His-4 and thus experiences an upfield shift.

The disulfide bridge between Cys-8 and Cys-25 has a right-handed spiral conformation (χ_1 = 37 ± 3°, χ_2 = 118 ± 6°, χ_3 = 101 ± 5°, $\chi_{2'}$ = 55 ± 4°, $\chi_{1'}$ = −166 ± 3°), while that between Cys-19 and Cys-35 has a left-handed one (χ_1 = −58 ± 12°, χ_2 = −70 ± 13°, χ_3 = −82 ± 11°, $\chi_{2'}$ = −80 ± 16°, $\chi_{1'}$ = −79 ± 9°). The sulfur atoms of the four cysteines are closely clustered in space. The largest S^{γ}–S^{γ} separation is ~4.8 Å between Cys-25 and Cys-35, while the S^{γ} atom of Cys-8 is in approximate van der Waals contact with the S^{γ} atoms of both Cys-19 and Cys-35 [$r_{S^{\gamma}(8)-S^{\gamma}(19)}$ ~3.7 Å, $r_{S^{\gamma}(8)-S^{\gamma}(35)}$ ~3.2 Å]. However, only two $C^{\beta}H(Cys_i)$–$C^{\beta}H(Cys_j)$ distances are less than 5 Å, namely, between Cys-19 and Cys-35 (~3 Å) and between Cys-8 and Cys-25 (~3.8 Å). Of these two interproton distances only the former is observed in the NOESY spectra.

Cysteines 8, 19, 25, and 35 are absolutely conserved in all four homologous terminal domains (Figure 10), and the "staggered" pairing of the disulfide bonds may account for the unusually well defined structure for a polypeptide of this size, as the cysteines form a network cross-linking the different parts of the peptide chain. The terminal domains of CBH II and EG III, however, have two additional cysteine residues. In the case of CBH II these are at positions 1 and 18 (Figure 10). The distance between the C^{α} atoms of residues 1 and 18 in CT-CBH I is ~5.3 Å, which is fully consistent with the formation of an extra disulfide bridge in the terminal domain of CBH II. In the case of EG III the extra cysteines are at positions 0 and 16 (Figure 10). Obviously, we cannot locate the position of residue 0 in CT-CBH I. Nevertheless, the C^{α}–C^{α} distance between residues 1 and 16 in CT-CBH I is ~6.2 Å, which would also be compatible with the formation of an extra disulfide bridge between residues 0 and 16 in EG III.

Examination of the structure of CT-CBH I revealed a similar topological fold to that found in potato carboxypeptidase inhibitor (CPI), a small disulfide-bridged 39-residue protein whose structure has been solved by both X-ray crystallography (Rees & Lipscomb, 1982) and NMR (Clore et al., 1987a). Despite the fact that there is no significant sequence homology between the two proteins, the C^{α} atoms of residues 3–9, 20–28, and 33–36 of CT-CBH I can be superimposed on those of residues 7–13, 21–29, and 32–35 of CPI with an atomic rms deviation of ~1.6 Å. This is illustrated in Figure 11. In this alignment there are three conserved residues in the two sequences with Cys-8, Ala-20, and Cys-35 in CT-CBH I corresponding to Cys-12, Ala-21, and Cys-34 in CPI. Further, the pattern, sequential spacing, and spatial orientation of two of the three disulfide bridges in CPI, namely, between Cys-12 and -27 and between Cys-18 and -34, correspond to disulfide bridges 8–25 and 19–35, respectively, in CT-CBH I. Recently, Bode et al. (1989) have reported the crystal structure of a 29-residue trypsin inhibitor from squash seeds known as CMTI-I and have shown that it also has a similar fold to CPI. We hope to make a more detailed comparison of the structures at a later date. It is not clear whether these similarities can be attributed to some evolutionary relationship or whether they are simply the consequence of the presumably limited number of stable three-dimensional folds that a small cysteine-rich peptide can assume.

The distribution of surface residues is noteworthy insofar as it is amphiphilic in character. The van der Waals surface of one face of the molecule (the front face of Figures 5A,B and 6C) is remarkably flat and predominantly hydrophilic, being formed by residues Gln-2, Ser-3, Tyr-5, Gln-7, Tyr-13, Ser-14, Asn-29, Tyr-31, Tyr-32, and Gln-36. The other face is hydrophobic in character, being formed principally by residues Gly-9, Ile-11, Thr-17, Val-18, Cys-19, Thr-24, Cys-25, and Val-27. Its van der Waals surface is also flat but has a slight indentation in the middle of it lined by Thr-17 and Val-18 on one side and by Thr-24, Cys-25, and Val-27 on the other. Gly-9 lies in the indentation with Thr-17 and Val-27

on either side of it and Ile-11 above it. It is interesting to note that Gly-9 is conserved (Figure 10) and that the main effect of a mutation at this site would be to disrupt both the small indentation and flat surface of the hydrophobic face. Due to the nonglobular form and small size of CT-CBH I, only two residues, namely, Cys-8 and Cys-35, are completely buried.

At present we are unsure of how the structure interacts with the cellulase catalytic domain or with cellulose. We note, however, that crystalline native cellulose has a flat, layered structure with hydrogen bonding in the plane of the layer and van der Waals interactions holding the layers together (Gardner & Blackwell, 1974). Both the hydrophobic and hydrophilic faces of CT-CBH I present flat surfaces that could potentially interact with cellulose.

Conclusions

In this paper we have determined the three-dimensional structure of CT-CBH I, the C-terminal cellulose binding domain of cellobiohydrolase I. The structure is exceptionally well determined with an average atomic rms distribution about the mean coordinate positions of 0.33 Å for the backbone atoms and 0.52 Å for all atoms (Figures 5–8 and Table III). This is due not only to the large number of loose interproton distance restraints acting in a cooperative manner but also to the large number of stereospecific assignments of prochiral protons and torsion angle restraints derived from the NOE and coupling constant data in combination with a conformational data base search. Further, the use of an iterative strategy in calculating the structures enabled us to resolve ambiguities relating not only to the assignment of some NOE cross-peaks but also to stereospecific assignments (e.g., the α-methylene protons of glycine). This overall approach leads to yet a further qualitative and quantitative improvement in the precision of the structure determination compared to previous recent studies both by ourselves (Driscoll et al., 1989b,c; Folkers et al., 1989) and others (Wagner et al., 1987; Arseniev et al., 1988; Kline et al., 1988) in which stereospecific assignments were obtained from a purely qualitative analysis of intraresidue NOE data and $^3J_{\alpha\beta}$ coupling constants. This high-resolution NMR structure has permitted us to assign the disulfide bridges with confidence, to postulate the presence of hydrogen bonds involving fast-exchanging NH protons and hydroxyl and amino protons of side chains, and to analyze certain side-chain interactions in detail. With this structure in hand, the relationship between the structure and function of CT-CBH I, as well as that of the homologous terminal domains of the other *T. reesei* cellulases, can now be investigated by techniques such as site-directed mutagenesis.

Acknowledgments

We thank the group of Dr. Hans Bennich for the gift of CT-CBH I, Drs. Ad Bax and Paul Driscoll for useful discussions, Dr. R. Huber for pointing out the similarity of the folds of CPI and CMTI, and Dr. Jeff Hoch for providing us with his ring current shift program.

Supplementary Material Available

Two figures giving (1) a comparison between an experimental NOESY spectrum and simulated NOESY spectra and (2) the angular rms distributions of the backbone ϕ and ψ torsion angles and three tables giving the chemical shifts of the resonances of CT-CBH I at 15 °C, the values of the $^3J_{HN\alpha}$ and $^3J_{\alpha\beta}$ coupling constants, and the complete list of experimental restraints (i.e., NOE interproton and hydrogen-bonding distance restraints and ϕ backbone, ψ backbone, and χ_1 side-chain torsion angle restraints) used in the computation of the three-dimensional structure of CT-CBH I (23 pages). Ordering information is given on any current masthead page. In addition, the coordinates of the 41 SA structures of CT-CBH I, as well as of the restrained minimized mean structure $(\overline{SA})r$, together with the complete list of experimental restraints, have been deposited in the Brookhaven Protein Data Bank.

References

Abuja, P. M., Schmuck, M., Pilz, I., Tomme, P., Claeyssens, M., & Esterbauer, H. (1988) *Eur. Biophys. J. 15*, 339–342.

Arseniev, S. A., Schultze, P., Wörgötter, E., Braun, W., Wagner, G., Vasák, M., Kägi, J. H. R., & Wüthrich, K. (1988) *J. Mol. Biol. 201*, 637–657.

Baker, E. N., & Hubbard, R. E. (1984) *Prog. Biophys. Mol. Biol. 44*, 97–179.

Bax, A. (1989) *Methods Enzymol. 176*, 151–168.

Bax, A., & Freeman, R. (1981) *J. Magn. Reson. 45*, 177–181.

Bax, A., & Davis D. G. (1985) *J. Magn. Reson. 65*, 355–366.

Bax, A., Sklenar, V., Clore, G. M., & Gronenborn, A. M. (1987) *J. Am. Chem. Soc. 109*, 6511–6513.

Bergfors, T., Rouvinen, J., Lehtovaara, P., Caldentey, X., Tomme, P., Claeyssens, M., Pettersson, G., Teeri, T., Knowles, J., & Jones, T. A. (1989) *J. Mol. Biol.* (in press).

Bhikhabhai, R., & Pettersson, G. (1984) *Biochem. J. 222*, 729–736.

Billeter, M., Braun, W., & Wüthrich, K. (1982) *J. Mol. Biol. 155*, 321–345.

Blevins, R. A., & Tulinsky, A. (1985) *J. Biol. Chem. 260*, 4264–4275.

Bode, W., Chen, Z., Bartels, K., Kutzbach, C., Schmidt-Kastner, G., & Bartunik, H. (1983) *J. Mol. Biol. 164*, 237–257.

Bode, W., Greyling, H. J., Huber, R., Otlewski, J., & Wilusz, T. (1989) *FEBS Lett. 242*, 285–292.

Borgias, B. A., & James, T. L. (1988) *J. Magn. Reson. 79*, 493–512.

Braunschweiler, L., & Ernst, R. R. (1983) *J. Magn. Reson. 53*, 521–528.

Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., & Karplus, M. (1983) *J. Comput. Chem. 4*, 187–217.

Brünger, A. T. (1988a) *X-PLOR Manual*, Yale University, New Haven, CT.

Brünger, A. T. (1988b) *J. Mol. Biol. 203*, 803–816.

Brünger, A. T., Clore, G. M., Gronenborn, A. M., & Karplus, M. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 3801–3805.

Brünger, A. T., Kuryan, J., & Karplus, M. (1987) *Science 235*, 458–460.

Burgess, A. W., Ponnuswamy, P. K., & Scheraga, H. A. (1974) *Isr. J. Chem. 12*, 239–286.

Chen, C. M., Gritzali, M., & Stafford, D. W. (1987) *Biotechnology 5*, 274–278.

Clore, G. M., & Gronenborn, A. M. (1987) *Protein Eng. 1*, 275–288.

Clore, G. M., & Gronenborn, A. M. (1989) *CRC Crit. Rev. Biochem. Mol. Biol.* (in press).

Clore, G. M., Gronenborn, A. M., Brünger, A. T., & Karplus, M. (1985) *J. Mol. Biol. 186*, 435–455.

Clore, G. M., Brünger, A. T., Karplus, M., & Gronenborn, A. M. (1986a) *J. Mol. Biol. 191*, 523–551.

Clore, G. M., Nilges, M., Sukumaran, D. K., Brünger, A. T., Karplus, M., & Gronenborn, A. M. (1986b) *EMBO J. 5*, 2729–2735.

Clore, G. M., Gronenborn, A. M., Nilges, M., & Ryan, C. A. (1987a) *Biochemistry 26*, 8012–8023.

Clore, G. M., Gronenborn, A. M., Kjaer, M., & Poulsen, F. M. (1987b) *Protein Eng. 1*, 305–311.

Clore, G. M., Gronenborn, A. M., James, M. N. G., Kjaer, M., McPhalen, C. A., & Poulsen, F. M. (1987c) *Protein Eng. 1*, 313–318.

Davis, D. G., & Bax, A. (1985) *J. Am. Chem. Soc. 107*, 2820–2821.

DeMarco, A., Llinas, M., & Wüthrich, K. (1978) *Biopolymers 17*, 617–636.

Driscoll, P. C., Clore, G. M., Berêss, L., & Gronenborn, A. M. (1989a) *Biochemistry 28*, 2178–2187.

Driscoll, P. C., Gronenborn, A. M., Berêss, L., & Clore, G. M. (1989b) *Biochemistry 28*, 2188–2198.

Driscoll, P. C., Gronenborn, A. M., & Clore, G. M. (1989c) *FEBS Lett. 243*, 223–233.

Enari, T. M., & Niku-Paavola, M.-L. (1987) *CRC Crit. Rev. Biotechnol. 5*, 67–87.

Ernst, R. R., Bodenhausen, G., & Wokaun, A. (1987) *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*, Clarendon Press, Oxford.

Fägertam, L. G., Pettersson, G., & Engström, J. Å. (1984) *FEBS Lett. 167*, 309–315.

Folkers, P. J. M., Clore, G. M., Dodt, J., Köhler, S., & Gronenborn, A. M. (1989) *Biochemistry 28*, 2601–2617.

Gardner, H., & Blackwell, J. (1974) *Biopolymers 13*, 1975–2001.

Gong, C.-S., Ladish, M. R., & Tsao, G. T. (1979) *Adv. Chem. Ser. No. 181*, 261–287.

Griesinger, C., Sørensen, O. W., & Ernst, R. R. (1982) *J. Am. Chem. Soc. 104*, 6800–6802.

Gritzali, M., & Brown, J. D. (1979) *Adv. Chem. Ser. No. 181*, 237–260.

Havel, T. F. (1986) DISGEO, *Quantum Chemistry Program Exchange No. 507*, Indiana University, Bloomington, IN.

Havel, T. F., & Wüthrich, K. (1985) *J. Mol. Biol. 182*, 381–394.

Havel, T. F., Crippen, G. M., & Kuntz, I. D. (1979) *Biopolymers 18*, 73–81.

Havel, T. F., Kuntz, I. D., & Crippen, G. M. (1983) *Bull. Math. Biol. 45*, 665–720.

Henrissat, B., Driquez, H., Viet, C., & Schulein, M. (1985) *Bio/Technology 3*, 722–726.

Hoch, J. C. (1983) Ph.D. Thesis, Harvard University, Cambridge, MA.

Hoch, J. C., Dobson, C. M., & Karplus, M. (1982) *Biochemistry 21*, 1118–1125.

IUPAC–IUB (1970) *Biochemistry 9*, 3471–3479.

Jeener, J., Meier, B. H., Bachmann, P., & Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546–4553.

Johansson, J., Ståhlberg, J., Lindeberg, G., Engström, Å, & Pettersson, G. (1989) *FEBS Lett. 243*, 381–393.

Johnson, C. E., & Bovey, F. A. (1958) *J. Chem. Phys. 29*, 1012–1014.

Jöliff, G., Beguin, P., Millet, J., Aubert, J.-P., Alzari, P., Juy, M., & Poljak, R. J. (1986) *J. Mol. Biol. 189*, 249–250.

Jones, T. A. (1978) *J. Appl. Crystallogr. 11*, 268–272.

Karplus, M. (1963) *J. Am. Chem. Soc. 85*, 2870–2871.

Keepers, J. W., & James, T. L. (1984) *J. Magn. Reson. 57*, 404–426.

Kline, A. D., & Wüthrich, K. (1985) *J. Mol. Biol. 183*, 503–507.

Kline, A. D., Braun, W., & Wüthrich, K. (1986) *J. Mol. Biol. 183*, 503–507.

Kline, A. D., Braun, W., & Wüthrich, K. (1988) *J. Mol. Biol. 204*, 675–724.

Knowles, J., Lehtovaara, P., Teeri, T., Pentillaä, M., Salovuori, I., & Andre, L. (1987) *Philos. Trans. R. Soc. London A231*, 449–454.

Kraulis, P. J., & Jones, T. A. (1987) *Proteins 2*, 188–201.

Kuntz, I. D., Crippen, G. M., & Kollman, P. A. (1979) *Biopolymers 18*, 939–957.

Langsford, M. L., Gilkes, N. R., Singh, B., Moser, B., Miller, R. C., Warren, R. A. J., & Kilburn, D. G. (1987) *FEBS Lett. 225*, 163–167.

Levitt, M., & Perutz, M. F. (1988) *J. Mol. Biol. 201*, 751–754.

Macura, C., Huang, Y., Suter, D., & Ernst, R. R. (1981) *J. Magn. Reson. 43*, 259–282.

Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967–974.

Marion, D., & Bax, A. (1988a) *J. Magn. Reson. 79*, 352–356.

Marion, D., & Bax, A. (1988b) *J. Magn. Reson. 80*, 528–533.

Marquart, M., Walter, J., Deisenhofer, J., Bode, W., & Huber, R. (1983) *Acta Crystallogr., Sect. B 39*, 480–490.

Merrifield, R. B. (1963) *J. Am. Chem. Soc. 85*, 304–305.

Montenecourt, B. S. (1983) *Trends Biotechnol. 1*, 156–161.

Mueller, L. (1987) *J. Magn. Reson. 72*, 191–196.

Nilges, M., Clore, G. M., & Gronenborn, A. M. (1988a) *FEBS Lett. 229*, 317–324.

Nilges, M., Gronenborn, A. M., Brünger, A. T., & Clore, G. M. (1988b) *Protein Eng. 2*, 27–38.

Nilges, M., Clore, G. M., & Gronenborn, A. M. (1988c) *FEBS Lett. 239*, 129–136.

Nilges, M., Clore, G. M., & Gronenborn, A. M. (1989) *Biopolymers* (in press).

Pardi, A., Billeter, M., & Wüthrich, K. (1984) *J. Mol. Biol. 180*, 741–751.

Penttilä, M., Lehtovaara, P., Nevalainen, H., Bhikhabhai, R., & Knowles, J. (1986) *Gene 45*, 253–263.

Perutz, M. F., Fermi, G., Abraham, D., Poyart, C., & Bursaux, E. (1986) *J. Am. Chem. Soc. 108*, 1064–1078.

Plateau, P., & Guêron, M. (1982) *J. Am. Chem. Soc. 104*, 7310–7311.

Ramachandran, G. M., & Sasisekharan, V. (1968) *Adv. Protein Chem. 23*, 283–437.

Rance, M., Sørensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 117*, 479–485.

Redfield, A. G., & Kunz, S. D. (1975) *J. Magn. Reson. 19*, 250–254.

Redfield, C., Hoch, J. C., & Dobson, C. M. (1983) *FEBS Lett. 159*, 132–136.

Rees, D. C., & Lipscomb, W. N. (1982) *J. Mol. Biol. 160*, 475–498.

Reynolds, R. A., Remington, S. J., Weaver, L. H., Fischer, R. G., Andersen, W. F., Ammon, H. L., & Matthews, B. W. (1985) *Acta Crystallogr., Sect. B 41*, 139–147.

Richardson, J. S. (1981) *Adv. Protein Chem. 34*, 167–339.

Saloheimi, M., Lehtovaara, P., Penttilä, M., Teeri, T., Ståhlberg, J., Johansson, G., Pettersson, G., Claeyssens, M., & Knowles, J. (1988) *Gene 63*, 11–21.

Schmuck, M., & Pilz, I. (1986) *Biotechnol. Lett. 8*, 397–402.

Shoemaker, S., Schweckart, V., Ladner, M., Gelfand, D., Kwok, S., Myambo, K., & Innis, M. (1983) *Bio/Technology 1*, 691–695.

Ståhlberg, J., Johansson, G., & Pettersson, G. (1988) *Eur. J. Biochem. 173*, 179–183.

Teeri, T., Lehtovaara, P., Kaupinen, S., Salovouri, I., & Knowles, J. (1987) *Gene 51*, 43–52.
Tomme, P., Van Tilbeurgh, H., Pettersson, G., Vandekerckhove, J., Knowles, J., & Teeri, T. (1988) *Eur. J. Biochem. 170*, 575–581.
Tüchsen, E., & Woodward, C. (1987) *Biochemistry 26*, 1918–1925.
Van Tilbeurgh, H., Tomme, P., Caleyssens, M., Bhikhabhai, R., & Pettersson, G. (1986) *FEBS Lett. 204*, 223–227.
Venkatachalam, C. M. (1968) *Biopolymers 6*, 425–436.
Wagner, G., Neuhaus, D., Wörgötter, E., Vasák, M., Kägi, J. H. R., & Wüthrich, K. (1986) *J. Mol. Biol. 187*, 131–135.
Wagner, G., Braun, W., Havel, T. F., Schaumann, T., Go, N., & Wüthrich, K. (1987) *J. Mol. Biol. 196*, 611–639.
Westergren, C., Pettersson, G., Knowles, J., & Jones, T. A. (1989) *J. Mol. Biol.* (in press).
Williamson, M. P., Havel, T. F., & Wüthrich, K. (1985) *J. Mol. Biol. 182*, 295–315.
Wlodawer, A., Walter, J., Huber, R., & Sjölin, L. (1984) *J. Mol. Biol. 180*, 301–329.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.
Wüthrich, K., Billeter, M., & Braun, W. (1983) *J. Mol. Biol. 169*, 949–961.
Wüthrich, K., Billeter, M., & Braun, W. (1984) *J. Mol. Biol. 180*, 715–740.
Zuiderweg, E. R. P., Boelens, R., & Kaptein, R. (1985) *Biopolymers 24*, 601–611.

Zinc Environment in Sheep Liver Sorbitol Dehydrogenase†

Martinus C. Feiters*,‡ and Jonathan Jeffery§

Department of Bioorganic Chemistry, State University of Utrecht, Padualaan 8, P.O. Box 80.075, NL-3508 TB Utrecht, The Netherlands, and Department of Biochemistry, University of Aberdeen, Marischal College, Aberdeen AB9 1AS, Scotland, United Kingdom

Received February 6, 1989; Revised Manuscript Received May 19, 1989

ABSTRACT: The extended X-ray absorption fine structure (EXAFS) associated with the zinc K-absorption edge has been recorded for sorbitol dehydrogenase. It is interpreted in terms of one cysteine sulfur among the ligands to the active site zinc atom. Simulations of the EXAFS based on the presence of two such sulfurs are less satisfactory, and comparison with the EXAFS of such systems points to the presence of only one sulfur ligand in sorbitol dehydrogenase. These results provide evidence that sorbitol dehydrogenase does not have the characteristic one water, one His, two Cys arrangement of ligands to the active site zinc found in the homologous alcohol dehydrogenases and are consistent with the one water, one His, one Cys, one Glu ligand arrangement of the proposed model of sorbitol dehydrogenase [Eklund, H., Horjales, E., Jörnvall, H., Brändén, C.-I., & Jeffery J. (1985) *Biochemistry 24*, 8005–8012]. Evidence for the correctness of the model is also evidence for validity of predictive techniques used in constructing the model, i.e., computer graphics fitting of the amino acid sequence to the crystallographically derived structure of a different but homologous protein.

The amino acid sequence of sheep liver sorbitol dehydrogenase (Jeffery et al., 1984a) has been aligned with those of mammalian, plant, fungal, and yeast alcohol dehydrogenases, clearly establishing that they are homologous proteins (Jörnvall et al., 1984, 1987; Eklund et al., 1985). Sorbitol dehydrogenase is a zinc enzyme (Jeffery et al., 1984b), as are the homologous alcohol dehydrogenases. The protein ligands to the active site zinc atom of these alcohol dehydrogenases are one His and two Cys (His-67, Cys-46, and Cys-174 in the numbering system of the horse liver enzyme; Brändén et al., 1975). The fourth ligand is water, or substrate during the reaction process.

A model has been constructed for the three-dimensional structure of sorbitol dehydrogenase by fitting the amino acid sequence to the crystallographically determined tertiary structure of horse liver alcohol dehydrogenase (Eklund et al., 1985). According to this model, the protein ligands to the active site zinc atom in sorbitol dehydrogenases are one His, one Cys, and one Glu (His-68, Cys-43, and Glu-154 in the continuous numbering system of sorbitol dehydrogenase; Jeffery et al., 1984a; Karlsson et al., unpublished data). In this part of the polypeptide chain, neither the Cys nearest to Glu-154 (Cys-163), which is the Cys that could be considered to be a candidate for coordination to zinc with another alignment (Eklund et al., 1985), nor the next-nearest Cys (Cys-138), which is not conserved in the alcohol dehydrogenases, is positioned in the model in a way that would allow it to serve as a zinc ligand. The fourth ligand in the model is a water molecule (or substrate during the reaction process), and a further water molecule is thought possibly to

† This work was supported by the British Science and Engineering Research Council (SERC), providing funds to the University of Manchester to establish the EXAFS service, with experimental and computational facilities at the SERC Daresbury Laboratory. General financial support by the University of Manchester and the University of Aberdeen is gratefully acknowledged. This work was carried out under SERC Grant GR/D 44188.

* Correspondence should be addressed to this author.
‡ University of Utrecht.
§ University of Aberdeen.

0006-2960/89/0428-7257$01.50/0