Solution Structure of the Active Domain of Tissue Inhibitor of Metalloproteinases-2. A New Member of the OB Fold Protein Family†

Richard A. Williamson,‡ Gabriel Martorell,§,‖ Mark D. Carr,*,§,⊥ Gillian Murphy,▽ Andrew J. P. Docherty,# Robert B. Freedman,‡ and James Feeney*,§

Research School of Biosciences, Biological Laboratory, University of Kent, Canterbury CT2 7NJ, U.K., Laboratory of Molecular Structure, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, U.K., Department of Cell and Molecular Biology, Strangeways Research Laboratory, Cambridge CB1 4RN, U.K., and Celltech Limited, 216 Bath Road, Slough SL1 4EN, U.K.

Received April 20, 1994; Revised Manuscript Received July 13, 1994®

ABSTRACT: Homonuclear two-dimensional and three-dimensional ^{1}H nuclear magnetic resonance spectroscopy has been used to obtain essentially complete sequence-specific assignments for 123 of the 127 amino acid residues present in the truncated form of tissue inhibitor of metalloproteinases-2 (ΔTIMP-2), the active N-terminal domain of the protein. Analysis of the through-space nuclear Overhauser effect data obtained for ΔTIMP-2 allowed determination of both the secondary structure of the domain and also a low-resolution tertiary structure defining the protein backbone topology. The protein contains a five-stranded antiparallel β-sheet that is rolled over on itself to form a closed β-barrel, and two short helices which pack close to one another on the same barrel face. A comparison of the ΔTIMP-2 structure with other known protein folds reveals that the β-barrel topology is homologous to that seen in proteins of the oligosaccharide/oligonucleotide binding (OB) fold family. The common structural features include the number of β-strands and their arrangement, the β-barrel shear number, an interstrand hydrogen bond network, the packing of the hydrophobic core, and a conserved β-bulge. Superpositions of the β-barrels from ΔTIMP-2 and two previously known members of the OB protein fold family (staphylococcal nuclease and *Escherichia coli* heat-labile enterotoxin) confirmed the similarity in β-barrel topology. The three-dimensional structure of ΔTIMP-2 has allowed a more detailed interpretation than was previously possible of the functional significance of available protein sequence and site-directed mutagenesis data for the TIMP family. Furthermore, the structure has revealed conserved surface regions of potential functional importance.

Breakdown of connective tissue is an important event in many normal and pathological processes, such as growth, wound repair, and tumor metastasis. The proteins known as tissue inhibitors of metalloproteinases (TIMPs)[1] are thought to play an important role in the control of this turnover by regulating the activity of a family of matrix-degrading enzymes, the matrix metalloproteinases (MMPs; Woessner, 1991; Docherty et al., 1992). This family includes the collagenases, gelatinases, and stromelysins, and together these zinc-dependent endopeptidases can degrade all the major extracellular matrix components (Matrisian, 1992). The TIMPs are specific inhibitors of the MMPs. They bind to an active MMP to form a tight (K_i < 1 nM), 1:1, non-covalent complex (Cawston et al., 1983; Welgus et al., 1985; Murphy et al., 1989). In vivo, inhibition occurs within the extracellular matrix and probably represents the final mechanism of MMP control before tissue degradation (Matrisian, 1990). TIMPs have also been reported to possess growth factor activity (Hayakawa et al., 1992; Yang & Hawkes, 1992).

Three members of the TIMP family have now been identified: TIMP-1, TIMP-2, and TIMP-3 (Docherty et al., 1985; Boone et al., 1990; Pavloff et al., 1992). The degree of amino acid sequence similarity between the members is high at about 45% (25% identity) and includes total conservation of 12 Cys residues known to form six disulfide bonds (Williamson et al., 1990). The TIMPs studied to date also show little difference in their specificity for the MMPs, each TIMP being able to inhibit equally well all members of the MMP family (Ward et al., 1991). The major structural and functional differences between the TIMPs are in their glycosylation and their specificity of binding the latent proenzyme form of gelatinase A and B. TIMP-1 is a complex glycoprotein with heterogeneous glycan units attached at Asn30 and Asn78 (Murphy & Werb, 1985; Carmichael et al., 1986). This glycosylation accounts for 30% of the molecular mass of this glycoprotein, but its removal has no effect on TIMP-1 inhibitory activity (Stricklin, 1986; Tolley et al., 1993). In contrast, TIMP-2 is nonglycosylated, and first reports on TIMP-3 suggest that this protein similarly lacks carbohydrate, although it does possess a potential N-glycosylation site at the C-terminus (Pavloff et al., 1992;

† This work was funded by the Medical Research Council (U.K.), both directly and by a project grant to the University of Kent, the Ministry of Education and Science (Spain), and the Arthritis and Rheumatism Council (U.K.). G. Martorell acknowledges the award of a research fellowship from the Spanish Ministry of Education and Science.

* To whom correspondence should be addressed.
‡ University of Kent.
§ National Institute for Medical Research.
‖ Present address: Departament de Quimica Organica, Facultat de Ciencies, Universitat de les Illes Balears, Palma de Mallorca 07071, Spain.
⊥ Present address: National Institute for Biological Standards and Control, Blanche Lane, South Mimms, Hertfordshire EN6 3QG, U.K.
▽ Strangeways Research Laboratory.
Celltech Ltd.

® Abstract published in *Advance ACS Abstracts,* September 1, 1994.

[1] Abbreviations: 2D, two-dimensional; 3D, three-dimensional; 2D COSY, two-dimensional correlation spectroscopy; DQF-COSY, double-quantum-filtered correlation spectroscopy; LTB, B subunit of *E. coli* heat-labile enterotoxin; MMP, matrix metalloproteinase; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; OB, oligosaccharide/oligonucleotide binding; SN, staphylococcal nuclease; TIMP, tissue inhibitor of metalloproteinases; ΔTIMP, N-terminal domain of TIMP; TOCSY, two-dimensional total correlation spectroscopy.

0006-2960/94/0433-11745$04.50/0

Apte et al., 1994). Interestingly, TIMP-1 and TIMP-2 show distinct specificity for the proenzyme form of gelatinase A and B. TIMP-2 complexes only with progelatinase A, while TIMP-1 binds only progelatinase B (Goldberg et al., 1989; Wilhelm et al., 1989).

Some insight into the structural organization of the TIMPs has been provided by the construction of truncated forms of the molecules. Residues 1–126 of TIMP-1 and 1–127 of TIMP-2 (referred to as ΔTIMP-1 and ΔTIMP-2, respectively), which contain three of the six disulfide bonds in the full-length molecules (C1–C72, C3–C101, and C13–C126), have been expressed in mammalian cells in the absence of the C-terminal region and are secreted in a soluble form (Murphy et al., 1991). Fluorescence studies show that these fragments adopt a folded stable structure in solution, indicating that residues 1–126/127 form an autonomous protein domain (Williamson et al., 1994). In addition, both ΔTIMP-1 and ΔTIMP-2 effectively inhibit matrilysin and the catalytic domains of stromelysin and gelatinase A, demonstrating that there is a direct interaction between the catalytic domains of the MMPs and the N-terminal domain of the TIMPs (Murphy et al., 1991; Nguyan et al., 1994). Further comparative work on the kinetics of inhibition by the truncated and full-length TIMP molecules showed that the C-terminal domain also makes some binding contribution to the TIMP–MMP complex (Willenbrock et al., 1993). In particular, the C-terminal domain of TIMP-2 appears to be responsible for the specific interaction of this molecule with progelatinase A. Although the emerging picture for the TIMP–MMP interaction is quite complex, it is clear that the N-terminal domain of the TIMPs represents a stable, minimized form of the inhibitor that includes the major site (or sites) necessary for metalloproteinase inhibition. Site-directed mutagenesis studies on TIMP-1 (O'Shea et al., 1992) indicated that no single residue is likely to be responsible for TIMP activity.

A more detailed insight into the interaction between TIMPs and MMPs has been frustrated by the absence of three-dimensional structural information for either protein family. Remarkably, four independent tertiary structures for the catalytic domain of a MMP (collagenase and stromelysin) by X-ray crystallography and NMR spectroscopy were reported in the first month of 1994 (Borkakoti et al., 1994; Gooley et al., 1994; Lovejoy et al., 1994; Stams et al., 1994). However, no structural data is currently available for the TIMPs; furthermore, these proteins show no significant sequence homology to other proteins in the sequence and structural databases. The ΔTIMP-2 protein is a good candidate for detailed NMR-based structural studies since it is stable, soluble at millimolar concentrations, and nonglycosylated. However, this size of protein (14.1 kDa) represents a considerable challenge for complete resonance assignment and subsequent structure determination by ^{1}H NMR analysis alone. The availability of a detailed three-dimensional TIMP structure would represent a significant step toward the goal of fully characterizing the TIMP–metalloproteinase interaction. Such a structure would provide a framework for designing variant proteins with modified inhibitory properties and would assist in the proper interpretation of existing results from site-directed mutagenesis and chemical modification studies which, on their own, have not been able to convincingly identify the inhibitory site for the molecule (O'Shea et al., 1992; Williamson et al., 1993).

In this study we report the sequence-specific ^{1}H resonance assignments for 123 of the 127 residues in ΔTIMP-2 based on 2D and 3D ^{1}H experiments. This information allows the determination of both the secondary structure of the protein and also a low-resolution tertiary structure defining the protein backbone topology. This structure has been compared with other known protein fold families and found to be homologous to the OB fold previously identified only in oligosaccharide or oligonucleotide binding proteins.

MATERIALS AND METHODS

Sample Preparation. ΔTIMP-2 was produced using the mammalian cell expression vector pEE12 transfected into NS0 mouse myeloma cells (Willenbrock et al., 1993). This system results in secretion of the recombinant protein into the cell culture medium, from which the ΔTIMP-2 was purified by a combination of ion-exchange (S-Sepharose) and gel filtration (Sephacryl S100) column chromatography. Initially, MES was added to the cell culture supernatant to a concentration of 30 mM, and the pH was adjusted to 6.0. One-liter batches of this supernatant were then loaded onto a 7.5 cm by 5.3 cm^2 S-Sepharose column, and the weakly bound contaminating proteins were removed by washing with 50 mM sodium phosphate buffer and 120 mM sodium chloride, pH 6.0, prior to elution of the ΔTIMP-2 by increasing the sodium chloride concentration stepwise to 200 mM. The fractions containing ΔTIMP-2 were then pooled and concentrated to 1.5–2.0 mg/mL by ultrafiltration (Amicon YM3 3000 M_r cutoff membrane). Finally, the protein was further purified by loading 10-mL batches onto a 37 cm by 5.3 cm^2 Sephacryl S100 column preequilibrated and eluted with 50 mM sodium phosphate buffer, 100 mM potassium chloride, and 0.05% (w/v) sodium azide, pH 6.0. Typically, 16 mg of purified ΔTIMP-2 was obtained per liter of original cell culture medium and judged to be greater than 97% pure on the basis of SDS–PAGE densitometry. The NMR spectra were recorded from 0.6-mL samples of 1.5–3.0 mM ΔTIMP-2 dissolved in either 25 mM potassium phosphate buffer and 100 mM potassium chloride, pH* 6.7, or 25 mM d_3-acetate buffer and 100 mM potassium chloride, pH* 4.6 (pH* values refer to the actual pH meter readings uncorrected for deuterium isotope effects). The samples were prepared in either 100% D_2O or 90% H_2O/10% D_2O solutions as appropriate.

NMR Measurements. The ^{1}H NMR experiments were carried out on a Varian UNITY 600 spectrometer operating at 600 MHz. All the 2D and 3D spectra were acquired in phase-sensitive mode using the method of States et al. (1982). The NMR measurements were made at 40 °C for pH* 6.7 samples and at 30 °C for samples adjusted to pH* 4.6. The following 2D ^{1}H experiments were carried out on both D_2O and H_2O samples of ΔTIMP-2: DQF-COSY (Rance et al., 1983); TOCSY (Braunschweiler & Ernst, 1983; Davis & Bax, 1985), using isotropic mixing times of 40–60 ms; and NOESY (Jeener et al., 1979; Macura et al., 1981) acquired with NOE buildup periods of 125 ms. The TOCSY spectra were recorded using a spin-lock field of around 10.5 kHz produced by an MLEV17 pulse sequence (Bax & Davis, 1985).

The 2D ^{1}H NMR experiments were typically carried out over a period of 20–48 h, collecting 300 pairs of t_1 increments, 48–128 scans per increment, and 8192 points per scan. The spectra from D_2O samples of ΔTIMP-2 were acquired with a spectral width of either 7000 or 7500 Hz, and the spectra from H_2O samples were acquired with a value of 8000 Hz. The 3D NOESY-TOCSY (Oschkinat et al., 1988) spectrum was recorded from a 3.0 mM sample of ΔTIMP-2 in a 90% H_2O/10% D_2O pH* 6.7 buffer using a NOE mixing time of 125 ms and a 40-ms spin-lock period. The spectra were

acquired over 3.5 days, collecting 128 t_1 by 128 t_2 increments, 8 scans per increment, and 1024 points per scan, with a spectral width of 8000 Hz in all dimensions.

Water suppression in both 2D and 3D spectra recorded from samples of ΔTIMP-2 in 90% H_2O solutions was achieved by the use of selective, on-resonance presaturation at the H_2O frequency. In addition, to prevent the recovery of the solvent signal during the mixing time of NOE-based experiments, a 180° pulse was applied at the center of this period. In order to observe cross peaks involving protons resonating at or very close to the chemical shift of the water, the experiments also included a 60-ms SCUBA sequence immediately after the presaturation period (Brown et al., 1988). The D_2O solutions of ΔTIMP-2 were prepared from protein repeatedly freeze-dried from D_2O, and thus only minimal presaturation was needed to suppress the small residual HOD signal.

The 2D data were processed on a SUN SPARC-330 workstation using the Varian VNMR software package. The original data were zero-filled to 2048 complex points in F_1 prior to Fourier transformation, and mild resolution enhancement was achieved by applying a $\pi/3$ shifted sine-squared apodization function in both dimensions. In order to improve the signal:noise ratio in the spectra, the apodization function employed in F_2 was usually selected to null the latter half of the 8192 complex data points collected. The 3D NOESY-TOCSY experiment was transformed, displayed, and plotted using software written in-house and installed on a SUN SPARCstation 10 (C. J. Bauer, unpublished work). To improve the resolution in the final spectrum, the number of data points in F_1 and F_2 was initially extended by a factor of 1.5 using linear prediction (Press et al., 1988). The time domain matrices were then zero-filled to 512 × 512 × 2048, resulting in a spectrum consisting of 256 × 256 × 1024 real points. Mild resolution enhancement was obtained by applying a $\pi/2.5$ shifted sine-squared apodization function in all dimensions.

Calculation and Analysis of Low-Resolution ΔTIMP-2 Structures. Interresidue ^1H–^1H distance constraints were derived from measurements of the intensities of cross peaks in 2D NOESY and 3D NOESY-TOCSY spectra of ΔTIMP-2. On the basis of the number of exponentially spaced contours observed for peaks in the 125-ms NOESY spectra, the NOEs were divided into five classes, corresponding to upper distance limits of 2.7, 3.3, 4.0, 5.0, and 6.0 Å. In all cases, the lower limits for NOE-based interproton distances were equal to the sum of the atomic radii. Calibrations for the upper distance limits used for the structural calculations were obtained by examining the number of contours observed for sequential (i, $i + 1$) and medium-range (i, $i < 5$) NOEs corresponding to known ^1H–^1H distances in elements of regular secondary structure in ΔTIMP-2; for example, the αCH to NH (i, $i + 3$) NOEs characteristic of helices correspond to distances of 3.3–3.4 Å. Initial structures of ΔTIMP-2 were calculated only on the basis of the NOE constraints, but the final stage in the structural calculations included appropriate upper and lower distance bounds for the three independently determined disulfide bonds (Williamson et al., 1990).

Structural calculations for ΔTIMP-2 were carried out using the variable target function-based distance geometry program DIANA (Güntert et al., 1991; Güntert & Wüthrich, 1991) installed on a SUN SPARCstation 10. In order to obtain a reasonable yield of properly converged structures, it was necessary to use the redundant dihedral angle (REDAC) protocol included as an option in DIANA (Güntert & Wüthrich, 1991). Where appropriate, standard distance corrections were added to constraints involving pseudoatoms (Wüthrich, 1986). The ΔTIMP-2 structures obtained were visualized and analyzed on a Silicon Graphics Indigo workstation using the software packages Insight II from BIOSYM and QUANTA from Molecular Simulations.

RESULTS AND DISCUSSION

Sequence-Specific Assignments. Sequence-specific ^1H resonance assignments for ΔTIMP-2 were obtained using the sequential assignment procedure developed by Wüthrich and co-workers, which relies upon correlating the known amino acid sequence of the protein with the NMR data (Wagner & Wüthrich, 1982; Wüthrich, 1986).

Apart from the aromatic residues, spin system assignments obtained for ΔTIMP-2 were based solely on the through-bond connectivities observed in DQF-COSY or TOCSY spectra and to a lesser extent on the fine structure of DQF-COSY cross-peaks. The signals in ^1H NMR spectra of ΔTIMP-2 are well resolved considering the size of the protein (127 residues), and this is illustrated by the selected regions of DQF-COSY and TOCSY spectra shown in Figures 1 and 2 and supplementary material Figure 1. Consequently, it was possible to establish many direct and relayed through-bond correlations and so group together a large number of signals into amino acid spin systems, which could be assigned to residue type or class. For example, it was possible to assign completely the signals from the nonexchangeable protons of alanine (7) and threonine (6) through the observation of nearly all of the expected αCH to βCH_3, αCH to βCH, αCH to γCH_3, and βCH to γCH_3 through-bond connectivities, as illustrated in Figure 1. Similarly, it was possible to identify spin systems corresponding to all of the valine residues (9) and all but one of the isoleucine residues (10/11) in ΔTIMP-2 from the observation of αCH to βCH, αCH to $\gamma/\gamma' CH_3$ and βCH to $\gamma/\gamma' CH_3$ cross peaks in DQF-COSY and TOCSY spectra, as indicated in Figure 1. It should be noted that even at this early stage in the spectral analysis it was possible to characterize fully 6 of the 11 isoleucine side chains. For the majority of alanine (6/7), valine (8/9), and isoleucine (10/11) residues it proved relatively straightforward to extend the spin system assignments from the side-chain protons to the NH, because of the observation of many NH to αCH, βCH, and $\gamma/\gamma' CH_3$ through-bond correlations (Figure 2). However, only one of the threonine spin systems could be unambiguously associated with its NH proton, due to the lack of relayed NH to βCH and γCH_3 cross peaks for most threonines in the TOCSY spectra.

The amenability of the ΔTIMP-2 ^1H spectra to detailed analysis is further illustrated by the favorable dispersion of the signals in the aromatic region of the DQF-COSY spectrum shown in supplementary material Figure 1. From consideration of this spectrum and the complementary TOCSY spectra it proved relatively straightforward to identify the resonances from the aromatic rings of tyrosine (5), phenylalanine (4), tryptophan (1), and histidine (3) residues in the protein. Spin system assignments for the aromatic side chains were completed by identifying intraresidue NOEs between appropriate aromatic ring protons and their corresponding α and β protons, which had been independently identified as forming αCH–βCH_2 AMX type spin systems (Wüthrich, 1986). It should be noted that ΔTIMP-2 contains 40 residues that give rise to αCH–βCH_2 spin systems, and all but two of these could be clearly traced out using DQF-COSY and TOCSY spectra. It was possible to identify the NH signal for the majority of the tyrosine (4/5), phenylalanine (4),

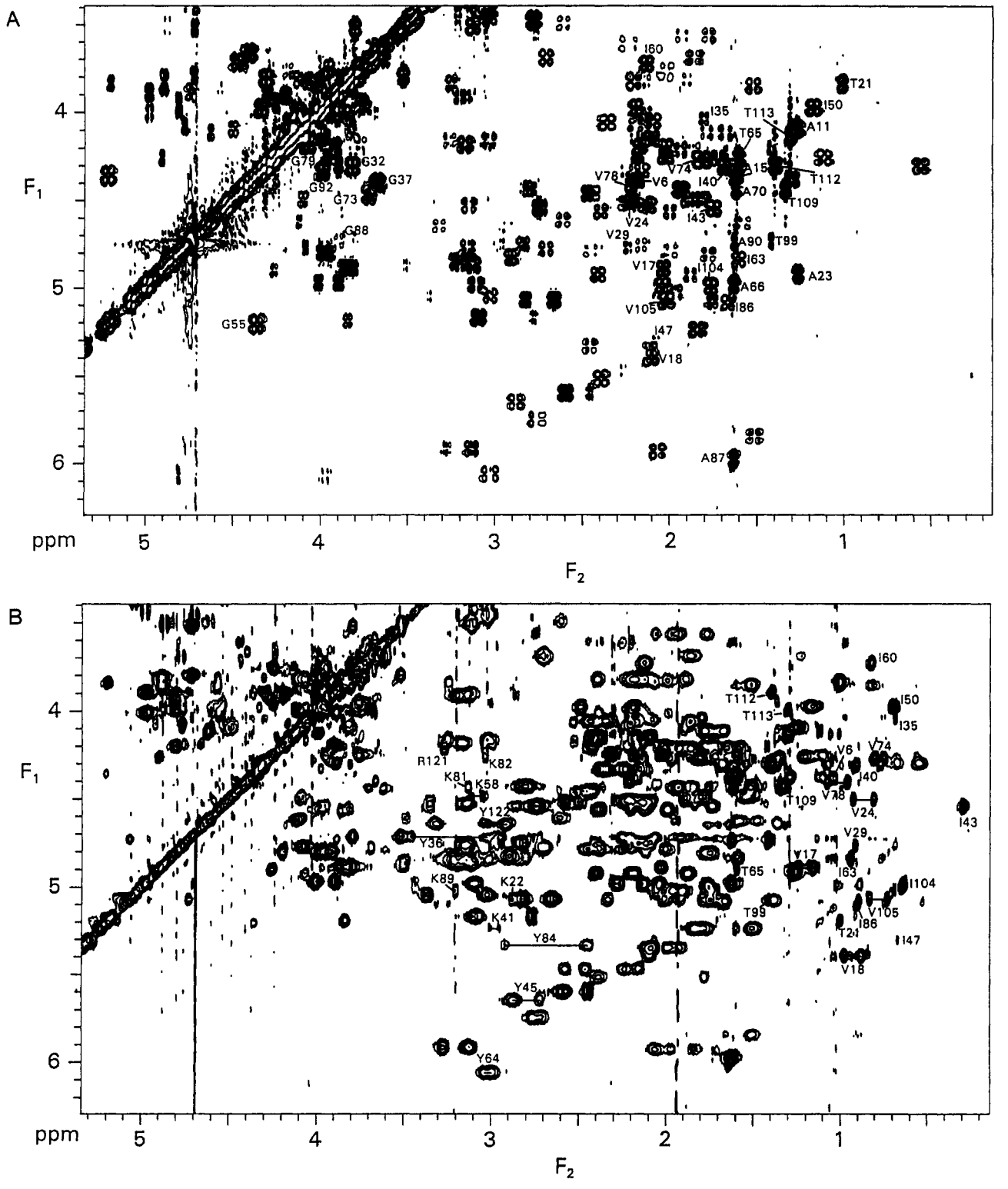

FIGURE 1: Selected portions of the aliphatic regions from DQF-COSY (A) and TOCSY (B) spectra of a D_2O solution of ΔTIMP-2. In the DQF-COSY spectrum the characteristic connectivities involving the methyl groups of alanine and threonine have been labeled, as have the αCH to βCH correlations for the valine and isoleucine residues. Similarly, in the TOCSY spectrum the relayed αCH to γCH_3 cross peaks for threonine, valine, and isoleucine have been indicated to illustrate the way in which the two spectra were used to obtain spin system assignments for the methyl-containing residues. Further assignments are also indicated for a number of glycine αCH to α′CH, tyrosine αCH to βCH, and lysine αCH to εCH correlations in order to illustrate the quality of the NMR data.

tryptophan (1), and histidine (2/3) residues, and thus complete assignment for most of the aromatic residues could be obtained.

From analysis of the DQF-COSY and TOCSY spectra of ΔTIMP-2 the following 82 amino acid spin systems could be characterized, and these were subsequently used as anchor points in the sequential assignment procedure: alanine, 6/7; threonine, 1/6; valine, 8/9; isoleucine, 10/11; leucine, 3/5; tyrosine, 4/5; phenylalanine, 4/4; tryptophan, 1/1; histidine, 2/3; serine, 6/9; half-cystine, aspartate, or asparagine, 11/18; glycine, 4/10; lysine, 3/13; arginine, 1/3; lysine or arginine, 3; methionine, glutamate, or glutamine, 9/17; and methionine, glutamate, glutamine, lysine, or arginine (long chain), 6. In addition, partial spin system assignments originating at an αCH were obtained for 5 threonines; 4 glycines; 10 half-

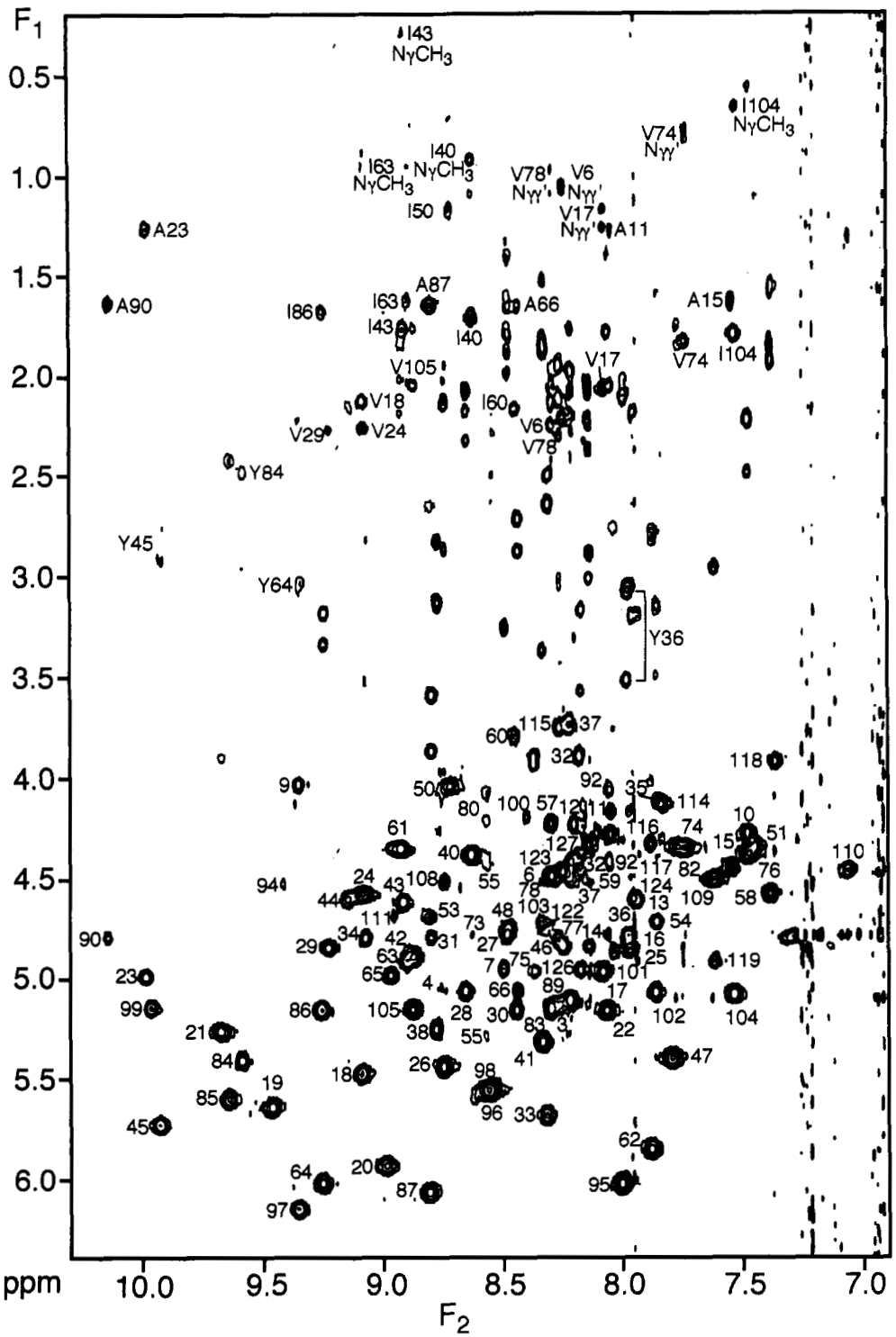

FIGURE 2: The NH to aliphatic region from a TOCSY spectrum of ΔTIMP-2 illustrating the good dispersion of the signals from the protein backbone. The NH to αCH cross peaks have been labeled according to the position of the amino acid in the protein sequence. In addition, relayed correlations to the side-chain protons of alanine, valine, isoleucine, and tyrosine have been highlighted (NH to βCH unless labeled otherwise) to further illustrate the number and type of through-bond connectivities observed for these residues and used to assign their spin systems.

cystine, aspartate, or asparagine residues; 6 long-chain amino acids including proline; and one alanine, valine, tyrosine, histidine, serine, and lysine.

In common with the other signals from ΔTIMP-2, the dispersion of the chemical shifts for the NH and αCH resonances is very good; however, there are still many overlapping signals from backbone protons (Figure 2) simply because of the size of the protein (127 amino acids). The complexity of NOESY spectra recorded from samples of ΔTIMP-2 is illustrated by the αCH to NH and NH to NH regions shown in Figure 3 and supplementary material Figure 2, respectively. Experience gained from NMR-based structural studies of other proteins, ranging in size from 106 to 162 residues (Carr et al., 1991, 1994; Carr, 1992), suggested that it would be impossible to obtain nearly complete sequence-specific resonance assignments for ΔTIMP-2 using only 2D NOESY data, because of the problems of cross-peak overlap and ambiguities in the assignment of sequential NOEs, both

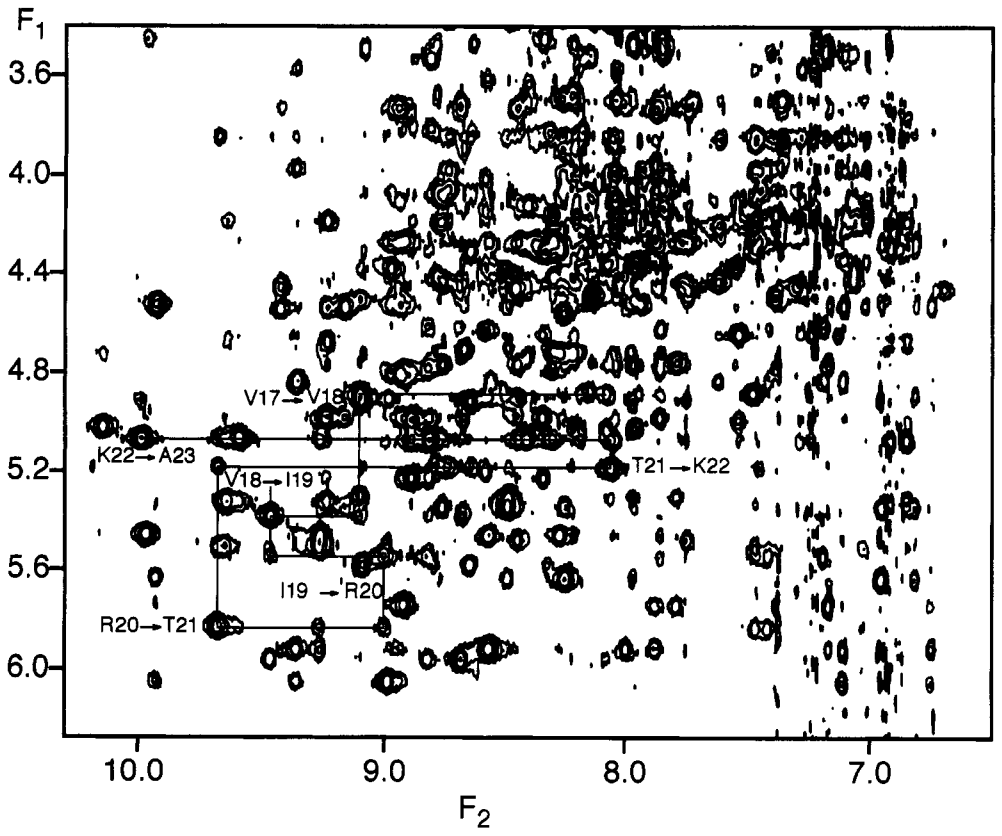

FIGURE 3: The NH to αCH region from a 125 ms mixing time NOESY spectrum of ΔTIMP-2. The labeled cross peaks arise from the sequential αCH to NH NOEs, which were used to make sequence-specific resonance assignments for the stretch of the protein from V17 to A23.

of which arise due to overlap of NH and αCH signals. In order to overcome this problem, sequential assignments for ΔTIMP-2 were made using a combination of 2D NOESY and 3D NOESY-TOCSY spectra. In the 3D NOESY-TOCSY experiment (Figure 4), protons to which NOE transfer has occurred during the NOE mixing period can be characterized not only by their own chemical shift but also by the chemical shift of the protons to which they are through-bond coupled and to which magnetization can be subsequently transferred during the TOCSY mixing time. In practice, most medium and strong sequential NOEs observed for ΔTIMP-2 (equivalent to more than two exponentially spaced contours above the noise threshold in NOESY spectra) gave rise to detectable 3D peaks of this type. The analysis strategy adopted for ΔTIMP-2 was to use the extra through-bond coupling information obtained from the 3D spectrum to resolve chemical shift ambiguities present in the 2D NOESY spectra. For example, in the NOESY spectrum of the protein we observe a strong cross peak between an αCH at 4.27 ppm and an NH at 8.30 ppm. Assuming an uncertainty of ±0.01 ppm in the chemical shifts, this NOE could arise from an interaction between any combination of five α (K51, V74, K82, S117, or E127) and five amide protons (D33, E57, V78, E83, or Y122). However, examination of the appropriate F_1 slices of the 3D NOESY-TOCSY spectrum (F_1 = 8.32 and 4.30 ppm) resulted in the identification of four 3D cross peaks, NH_{E83}–αCH_{K82}–NH_{K82}, NH_{E83}–αCH_{K82}–βCH_{K82}, NH_{E83}–αCH_{K82}–$\beta' CH_{K82}$, and αCH_{K82}–NH_{E83}–αCH_{E83} (general format: H_A–*NOE*–H_B–*TOCSY*–H_C (F_1, F_2, F_3)), which are labeled at the top of Figure 4B, and conclusively showed that the original NOESY correlation arose from an NOE between signals subsequently assigned to the αCH of K82 and the NH of E83. In a small number of cases chemical shift degeneracies could not be resolved even by using the 3D spectrum, thus producing branch points in the sequential assignment pathways. Under these circumstances it was usually possible to discriminate between the possibilities on the basis of sequence compatibility.

The approach described above greatly facilitated the identification of the sequential NH to NH, αCH to NH, and βCH to NH NOEs required to make sequence-specific resonance assignments, allowing the assignments to be determined for long stretches of neighboring amino acids, as illustrated by the sequential walks from Q9 to D16 (NH to NH) and V17 to A23 (αCH to NH) shown on the NOESY spectra in supplementary material Figure 2 and in Figure 3, respectively. Similarly, in Figure 4, pairs of F_2F_3 strips from the 3D NOESY-TOCSY spectrum, taken at F_1 chemical shifts closest to NH_i and αCH_{i-1}, illustrate a sequential αCH to NH walk from K82 to A87. The 2D and 3D spectra acquired at pH* 6.7 allowed sequence-specific resonance assignments to be made for 120 of the 127 residues in ΔTIMP-2, on the basis of identification of the following continuous stretches of neighboring residues linked by one or more sequential NOEs: S4–K48, Q49–K51, M52–P67, C72–G79, K81–G92, D93–S111, Q114–N119, R121–Y122, and C126–E127. The breaks in the sequential assignment pathways occurred because at pH* 6.7 no NH signals could be detected for C1, S2, Q49, M52, S68, S69, A70, V71, C72, K81, D93, T112, T113, H120, R121, and G125. To try to address this problem, 2D TOCSY and NOESY spectra were acquired from a ΔTIMP-2 sample at pH* 4.6 where NH exchange with water would be expected to be slower (Wüthrich, 1986). Under these conditions eight additional NH signals were detected, and many of the sequential assignment pathways identified at pH* 6.7 could be extended or linked together to form just three continuous stretches of residues linked via sequential NOEs, C3–P67, C72–Y122, and C126–E127 (Figure 5).

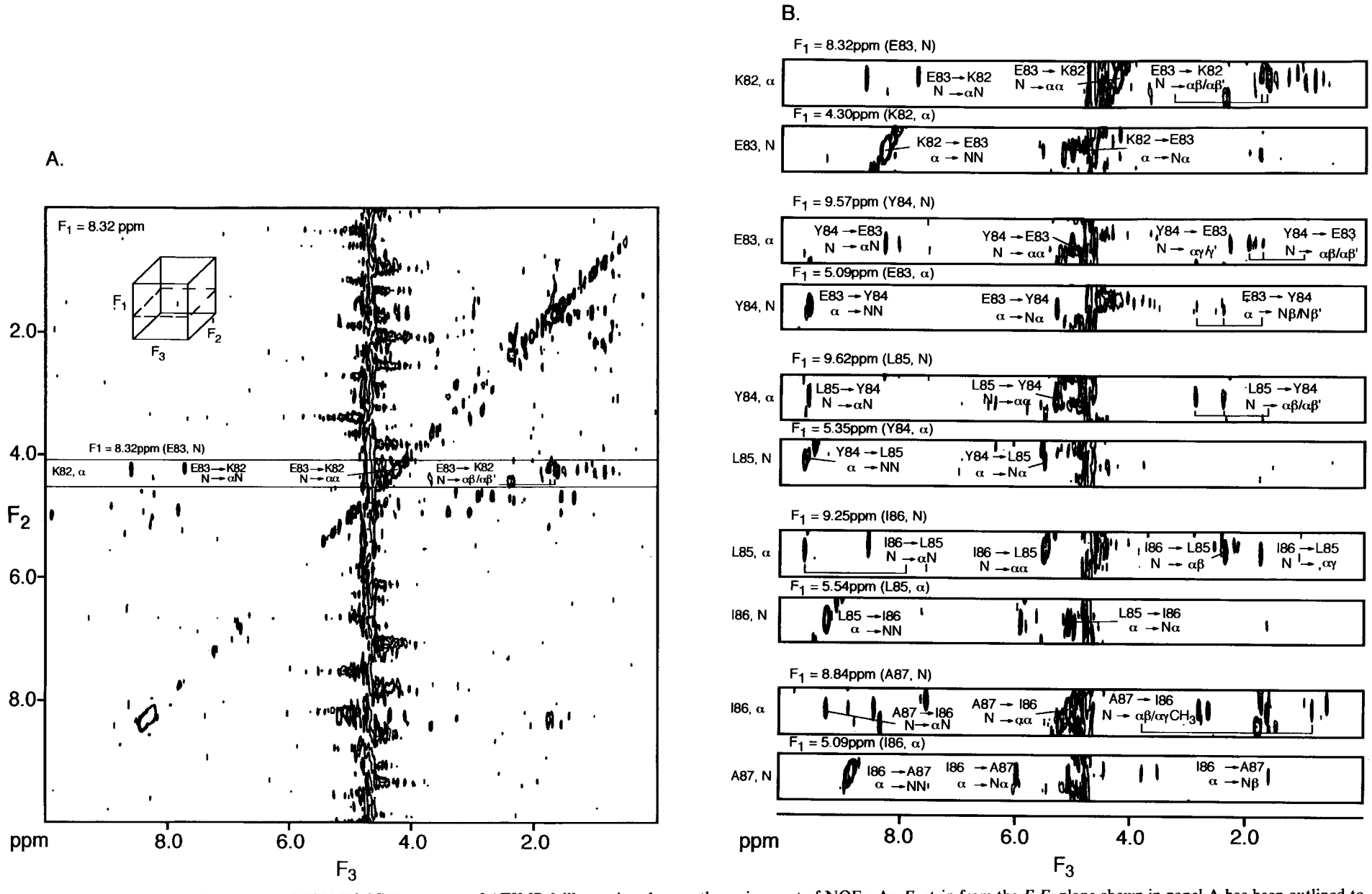

FIGURE 4: Selected regions from a 3D NOESY-TOCSY spectrum of ΔTIMP-2 illustrating the sequential walk from K82 to A87. Panel A shows an F_1 slice from the 3D experiment taken at a frequency of 8.32 ppm, which corresponds to the slice closest to the chemical shift of the NH of E83 (8.30 ppm). The diagonal in this F_2F_3 plane arises from proton magnetization that has undergone NOE transfer during the first (NOE) mixing period in the pulse sequence and so identifies pairs of hydrogen atoms, at chemical shifts F_1 (starting) and F_2 (finishing), which are close in space. The off-diagonal peaks, or cross peaks, arise from a subsequent through-bond transfer of magnetization during the second (TOCSY) mixing time in the sequence. This results in the protons to which NOE transfer occurred being characterized not only by their own chemical shifts (F_2) but also by the chemical shifts of the protons (F_3) to which they are through-bond coupled, thereby greatly facilitating the assignment of NOEs. An F_1 strip from the F_2F_3 plane shown in panel A has been outlined to illustrate the pattern of peaks arising from interresidue NOE transfer from the NH of E83 (F_1) to the αCH of K82 (F_2) followed by through-bond TOCSY transfer to the NH, βCH, and β'CH (F_3) of K82. In panel B, a series of F_1 strips analogous to the one described above are shown to illustrate a sequential αCH to NH walk from K82 to A87. For each sequential NOE, confirmation of the identity of both the NH and αCH protons involved is readily obtained from the pattern of subsequent TOCSY transfer. Thus for every residue pair, two F_1 strips are shown, corresponding to the regions from the F_2F_3 planes in which an NOE has arisen at the αCH (upper) and NH (lower). The relevant cross peaks have been labeled in the format H_A–NOE–H_B–TOCSY-H_C (F_1,F_2,F_3).

At the end of the sequential assignment procedure, complete sequence-specific ^{1}H assignments had been obtained for 118 of the 127 residues of ΔTIMP-2 and partial assignments had been obtained for a further 5 (Table 1). The only four residues for which no sequence-specific assignments could be made were S2, S68, S69, and G125. Frustratingly, the αCH–βCH$_2$ spin systems from the three serines can be clearly identified in DQF-COSY and TOCSY spectra but cannot be assigned to the specific sequence position since their NH signals are not observed. It is perhaps worth noting at this point that, in addition to being essential for assigning many NOEs, the TOCSY connectivities observed in the 3D NOESY-TOCSY experiment also aided greatly in extending the completeness of the side-chain assignments beyond those obtained during the initial identification of amino acid spin systems, principally by resolving chemical shift ambiguities present in DQF-COSY and TOCSY spectra.

Secondary Structure of ΔTIMP-2. The elements of regular secondary structure found in proteins give rise to characteristic patterns of sequential ($i, i + 1$), medium-range ($i, i < 5$), and long-range NOEs involving backbone protons (Wüthrich, 1986), the majority of which are identified as part of the sequential assignment procedure. For example, helices are characterized by relatively large NH to NH ($i, i + 1$) and weaker αCH to NH ($i, i + 3$) NOEs, whereas the extended backbone conformation found in β-sheets gives rise to large sequential αCH to NH as well as long-range NH to NH, αCH to NH, and αCH to αCH NOEs.

The sequential and medium-range backbone to backbone and backbone to side chain NOEs identified for ΔTIMP-2 are summarized in Figure 5. Residues 8–14 and 112–118 are characterized by NH to NH ($i, i + 1$), NH to NH ($i, i + 2$), αCH to NH ($i, i + 3$), and αCH to βCH ($i, i + 3$) NOEs and therefore clearly form two short helices. This conclusion is further supported by the significant upfield shifts observed for the αCH signals of residues 8–13 and 112–118 compared to their random coil values (Figure 5; Wishart et al., 1991). In contrast, many stretches of the ΔTIMP-2 sequence are characterized by strong αCH to NH sequential NOEs, large downfield shifts for the αCH signals (Figure 5), and the observation of a large number of long-range backbone to backbone NOEs (Figure 6 and supplementary material Figure 3). Analysis of this data clearly indicates that the protein contains a six-stranded β-sheet involving residues 17–34 (A), 38–53 (B), 62–65 (C), 84–91 (D), 95–98 (E), and 105–107 (F), which is shown schematically in Figure 6. Strands A, B, C, and D are arranged in a "Greek key" motif, which together with strand E is rolled over into a β-barrel type structure. The strands are all antiparallel, except for C and E, consistent with the observation of many long-range αCH to αCH NOEs (supplementary material Figure 3). Three β-bulges (short breaks in the regular structure of the β-sheet) are seen at positions V24–S25, S31–G32, and I50–K51. Strands A and B are linked via a tight β-hairpin type turn, formed from amino acids 34–37. The high β-sheet content of the protein (approximately 40%) almost certainly accounts for the excellent dispersion of the ^{1}H NMR signals.

Tertiary Structure of ΔTIMP-2. Despite the favorable dispersion of its ^{1}H NMR signals, the relatively large size of ΔTIMP-2, and consequently the large number of signals, results in very complex 2D NOESY spectra, as discussed previously. The extent of the cross-peak overlap in the region reflecting NOEs between aliphatic side chains is such that any detailed analysis is impossible. However, in some areas of the spectra, in particular those arising from NOEs involving backbone amide and aromatic ring protons, many reasonably resolved cross peaks could be identified and assigned with the help of the 3D NOESY-TOCSY data. These regions of the ΔTIMP-2 NOESY spectra were therefore analyzed in detail, and the NOE data obtained were used to determine the low-resolution structure of the protein reported here.

The low-resolution ΔTIMP-2 structures were calculated using the distance geometry program DIANA (Güntert et al., 1991; Güntert & Wüthrich, 1991) using a data set containing upper distance limits derived from 358 sequential, 52 medium-range, and 229 long-range interresidue NOEs from backbone amide and aromatic ring protons to all other proton groups in the structure and αCH to αCH cross-strand NOEs (639 NOEs in total, an average of 5.0 NOEs/residue). Upper distance limits were categorized into five different levels according to the strength of the NOE signal (see Materials and Methods). Visual inspection of ΔTIMP-2 structures obtained after preliminary structure calculations showed that the chemically determined disulfide bond assignments (Williamson et al., 1990) were consistent with the NMR data; thus, subsequent calculations included the appropriate upper and lower distance constraints for the three disulfide bonds. A total of 16 well-converged structures were obtained from 50 random starting conformations after 13 cycles of the REDAC protocol (Güntert & Wüthrich, 1991). These structures have no residual NOE, van der Waals, or disulfide bond violations exceeding 0.5 Å and were used for further analysis. The sums of violations for upper distance limits, lower distance limits, and steric constraints were 7.02 ± 1.80, 1.19 ± 0.17, and 11.0 ± 2.86 Å, respectively. Similarly, maximum violations for the converged structures were 0.38 ± 0.05, 0.12 ± 0.10, and 0.35 ± 0.08 Å, respectively. The average value of the final DIANA target function was 3.62 ± 1.35Å^2.

The backbone topology of ΔTIMP-2 is illustrated in Figure 7. As expected from a knowledge of the secondary structure, the main feature of the protein molecule is the five-stranded closed β-barrel involving strands A, B, C, D, and E. The sixth strand, F, forms an antiparallel connection with strand D in the same sense as strand E on the top rim of the barrel (Figure 7B). The backbone topology reveals several interesting features of the β-barrel fold. The β-bulges V24–S25 and I50–K51 fall within the confines of the β-barrel, and by breaking the regular hydrogen bond network of the β-sheet, they allow strands A and B to bend or coil, thereby considerably increasing the barrel radius. Also, strands A and B are considerably longer than the other elements in the β-barrel (C, D, and E) and extend beyond the barrel confines (Figures 6 and 7). Residues D30–K41 form a well-defined β-hairpin like structure; however, to date no long-range NOEs have been identified from these residues to other sites in the protein. This suggests that residues 30–41 may be linked to the barrel via a hingelike region, probably involving the β-bulge S31–G32, with the β-hairpin structure free to move through a relatively large conformational space. Superposition of the backbone atoms in the 16 converged structures of ΔTIMP-2 for the residues in the β-strands, excluding D30–K41, gives an rmsd about the average coordinates of 1.17 ± 0.20 Å, showing that the β-barrel is fairly well defined. Strands A/B and D/E are joined by tight four- or five-residue turns, whereas strands B/C and C/D are joined by loops of eight and 18 residues, respectively. These two loops cover the top and bottom of the β-barrel, although their precise conformations (particularly for the B/C loop) are poorly defined in our low-resolution structure. In addition to the β-sheet, ΔTIMP-2 contains two short helices,

Table 1: ^{1}H Resonance Assignments for ΔTIMP-2 at pH* 6.7[a]

residue	NH	α	α′	β	β′	γ	γ′	δ	δ′	ε	ε′	other
Cys-1												
Ser-2												
Cys-3	8.26	5.06		3.38	3.03							
Ser-4	8.76	4.99		4.03	3.90							
Pro-5		4.57		2.40	2.00	2.22	2.14	4.02	3.83			
Val-6	8.25	4.39		2.18		1.07	1.04					
His-7	8.50	4.89		3.22	3.22							8.04(2), 7.26(4)
Pro-8		3.54		1.92	1.78	2.08	1.97	3.85	3.26			
Gln-9	9.35	3.95		2.20	2.20	2.49	2.49					
Gln-10	7.49	4.21		2.18	2.18	2.44	2.44					6.91, 7.54(NH_2)
Ala-11	8.05	4.10		1.27								
Phe-12	8.21	4.16		3.26	3.04							7.03(2, 6), 7.08 (3, 5), 7.13(4)
Cys-13	7.94	4.55		3.37	3.14							
Asn-14	8.15	4.76		2.97	2.85							
Ala-15	7.55	4.37		1.62								
Asp-16	7.97	4.78		3.15	2.99							
Val-17	8.08	4.90		2.04		1.27	1.17					
Val-18	9.09	5.40		2.10		1.00	0.89					
Ile-19	9.46	5.55		1.82		1.60	1.13	0.82				1.08(γCH_3)
Arg-20	8.98	5.85		1.52	1.52	1.66	0.91	1.38				
Thr-21	9.67	5.21		3.85		1.02						3.38(OH)
Lys-22	8.06	5.08		1.76	1.76	1.39	1.39	1.69	1.69	2.90	2.90	
Ala-23	9.98	4.92		1.27								
Val-24	9.09	4.51		2.23		0.94	0.83					
Ser-25	7.98	4.79		4.08	4.08							
Glu-26	8.75	5.37		2.11	2.11	2.01	1.94					
Lys-27	8.48	4.71		1.96	1.86	1.57	1.40	1.73	1.73	3.05	3.05	
Glu-28	8.66	5.00		2.05	2.05	2.29	2.15					
Val-29	9.23	4.78		2.25		0.93	0.93					
Asp-30	8.44	5.09		2.84	2.67							
Ser-31	8.80	4.73		3.81	3.54							
Gly-32	8.18	4.32	3.83									
Asn-33	8.31	5.61		2.61	2.46							
Asp-34	9.07	4.73		3.49	2.77							
Ile-35	7.86	4.05		1.81		0.98	0.71	0.78				0.71(γCH_3)
Tyr-36	7.98	4.72		3.48	3.04							7.24(2, 6), 6.94(3, 5)
Gly-37	8.22	4.41	3.68									
Asn-38	8.78	5.18		3.10	2.79							7.18, 8.63(NH_2)
Pro-39		4.92		2.41	1.89	2.20	2.20	4.08	3.83			
Ile-40	8.63	4.32		1.69		1.57	1.09	0.79				0.92(γCH_3)
Lys-41	8.34	5.25		1.85	1.80	1.51	1.51	1.62	1.62	2.99	2.93	
Arg-42	8.88	4.82		1.79	1.76	1.76	1.76	3.10	3.10			
Ile-43	8.92	4.56		1.76		1.50	1.24	0.85				0.31(γCH_3)
Gln-44	9.15	4.54		2.11	2.11	1.84	1.43					
Tyr-45	9.92	5.66		2.89	2.73							6.97(2, 6), 6.83(3, 5)
Glu-46	8.25	4.77		2.17	2.11	2.42	2.42					
Ile-47	7.79	5.31		2.08		1.10	0.58	0.58				0.68(γCH_3)
Lys-48	8.47	4.67		1.77	1.64	1.37	1.33	1.71	1.71	2.99	2.99	
Gln-49	9.46	4.06		2.10	1.87	1.77						
Ile-50	8.73	3.99		1.18		1.55	1.16	0.69				0.70(γCH_3)
Lys-51	7.45	4.27		1.09	1.09	0.68	0.57	1.57	1.55	2.90	2.77	
Met-52	8.96	4.79		2.01						2.23		
Phe-53	8.81	4.62		3.51	2.62							7.20(2, 6), 7.43(3, 5), 7.50(4)
Lys-54	7.86	4.65		1.60	1.57	1.94	1.36	1.48				
Gly-55	8.57	5.22	4.37									
Pro-56		4.81		2.46	2.40	2.26	2.21	3.61	3.61			
Glu-57	8.30	4.16		2.19	2.11	2.44	2.39					
Lys-58	7.39	4.50		1.89	1.82	1.55	1.55	1.48	1.48	3.13	3.05	
Asp-59	8.15	4.48		2.47	2.45							
Ile-60	8.45	3.74		2.14		1.91	1.44	0.78				0.84(γCH_3)
Glu-61	8.93	4.29		1.98	1.82	2.15						
Phe-62	7.88	5.77		2.78	2.73							7.18(2, 6), 7.39(3, 5), 7.17(4)
Ile-63	8.90	4.85		1.61		1.66	1.06	0.78				0.95(γCH_3)
Tyr-64	9.35	6.07		3.04	3.00							7.13(2, 6), 6.75(3, 5)
Thr-65	8.97	4.92		4.27		1.61						5.50(OH)
Ala-66	8.44	4.99		1.64								
Pro-67		4.06		2.10	2.10	1.90	1.90	4.12	3.83			
Ser-68												
Ser-69												
Ala-70	8.45	4.44		1.63								
Val-71		4.74		2.63		1.16	1.10					
Cys-72		4.45		3.70	3.65							
Gly-73	8.68	4.48	3.73									
Val-74	7.75	4.28		1.81		0.82	0.77					
Ser-75	8.37	4.89		3.88	3.82							
Leu-76	7.48	4.30		0.69	0.56	0.73		−0.20	−0.24			
Asp-77	8.27	4.72		2.97	2.74							

Table 1 (Continued)

residue	NH	α	α′	β	β′	γ	γ′	δ	δ′	ϵ	ϵ′	other
Val-78	8.30	4.43		2.22		1.09	0.97					
Gly-79	8.77	4.20	4.07									
Gly-80	8.58	4.15	4.02									
Lys-81	8.37	4.44		1.93	1.93	1.52	1.52	1.86	1.77	3.14	3.14	
Lys-82	7.77	4.27		1.82	1.73	1.37	1.22	1.79	1.79	3.04	3.04	
Glu-83	8.30	5.07		2.02	1.80	2.35	2.35					
Tyr-84	9.59	5.35		2.93	2.46							6.86(2, 6), 6.93(3, 5)
Leu-85	9.64	5.52		2.39	2.39	1.79		1.14				
Ile-86	9.26	5.09		1.67		1.65	1.14	0.54				0.91(γCH$_3$)
Ala-87	8.81	5.98		1.64								
Gly-88	8.68	4.75	3.90									
Lys-89	8.22	5.03		2.05	1.96	1.75	1.75	1.91	1.91	3.21	3.21	
Ala-90	10.14	4.74		1.63								
Glu-91	8.27	4.73		2.08	1.91	2.23	2.23					
Gly-92	8.06	4.35	4.00									
Asp-93	8.63	4.55		2.76	2.76							
Gly-94	9.42	4.45	4.45									
Lys-95	8.00	5.94		2.08	1.98	1.84	1.84	1.74	1.74	3.12	3.12	
Met-96	8.56	5.49		2.24	2.17	2.59	2.47			1.89		
His-97	9.25	5.95		3.30	3.15							7.75(2), 6.97(4)
Ile-98	8.56	5.47		2.17		1.40	1.20	0.77				1.01(γCH$_3$)
Thr-99	9.96	5.08		4.74		1.43						
Leu-100	8.41	4.13		1.79	1.68	1.74		1.15	1.13			
Cys-101	8.04	4.80		2.72	3.71							
Asp-102	7.86	5.00		3.45	3.11							
Phe-103	8.34	4.67		3.31	2.89							7.24(2, 6), 7.31(3, 5), 7.23(4)
Ile-104	7.54	5.00		1.77		1.63	1.04	0.91				0.65(γCH$_3$)
Val-105	8.87	5.09		2.03		0.85	0.75					
Pro-106		4.68		2.32	2.17	2.40	2.01	4.21	3.77			
Trp-107	9.23	4.21		3.19	2.99							10.48(1), 7.30(2), 7.48(4), 7.04(5), 7.39(6), 7.73(7)
Asp-108	8.75	4.45		2.82	2.79							
Thr-109	7.62	4.43		4.43		1.33						
Leu-110	7.07	4.39		1.31	1.12	1.60		0.73	0.47			
Ser-111	8.96	4.62		4.13	4.13							
Thr-112	9.09	3.91		4.31		1.41						
Thr-113	8.14	4.01		4.15		1.32						
Gln-114	7.83	4.05		2.35	2.35	2.45	2.45					
Lys-115	8.27	3.69		1.86	1.86	1.24	0.47	1.44	1.44	2.57	2.57	
Lys-116	8.05	4.21		2.02		1.79	1.73	1.59	1.59	3.06	3.06	
Ser-117	7.89	4.27		3.94	3.71							
Leu-118	7.37	3.85		1.53	0.82	1.63		0.59	0.46			
Asp-119	7.62	4.84		2.92	2.92							
His-120		4.85		3.26	3.19							7.96(2), 7.18(4)
Arg-121	8.51	4.22		1.94	1.86	1.65	1.65	3.28	3.28			
Tyr-122	8.31	4.66		3.01	2.88							7.13(2, 6), 6.88(3, 5)
Gln-123	8.22	4.35		2.24	2.18	2.41	2.37					
Met-124	7.96	4.53		2.23	2.16	2.61	2.51			1.94		
Gly-125												
Cys-126	8.18	4.89		3.53	3.13							
Glu-127	8.15	4.26		2.19	2.03	2.34	2.34					

[a] The ^{1}H chemical shifts (ppm) are referenced to TSP (3-(trimethylsilyl)propionic acid). NH resonances seen only in NMR experiments at pH* 4.6 are shown underlined.

involving residues 8–13 and 112–118, which pack close to one another on the same face of the β-barrel (Figure 7). The helices are well defined by sequential and medium-range NOEs (Figure 5), but within the family of converged ΔTIMP-2 structures slight differences are seen regarding the precise orientation of the helices with respect to the β-barrel. This is reflected in the backbone rmsd value, which rises to 1.82 ± 0.31 Å if both the β-strands and the helices are superposed (Figure 7A). The assignment of more long-range NOEs in this region would improve the definition of the structure; however, many of these NOEs are between side-chain protons and fall in highly overlapped regions of the NOESY spectra, making their assignment impossible.

TIMP-1, -2, and -3 are expected to adopt similar tertiary structures as judged by their high level of sequence similarity (45%), their total conservation of 12 Cys residues known to form six disulfide bonds, and the finding that the N-terminal regions of both TIMP-1 and TIMP-2 can fold to give stable active structures. The major sequence difference between the N-terminal domains of the TIMPs is the presence of short insertions or deletions. TIMP-1 and TIMP-3 have deletions of seven and six residues, respectively, in the region corresponding to G32–Q44 of TIMP-2, while TIMP-1 has an insertion of five residues at the position equivalent to G55. The latter difference can be easily accommodated in the ΔTIMP-2 structure, as the insertion falls within the surface B/C loop (Figure 7B); however, the former difference requires a local modification in secondary structure if the TIMP-1 and TIMP-3 sequences are to be mapped onto the ΔTIMP-2 structure. The deletion in TIMP-1 maps to the beginning of strand B in ΔTIMP-2, while the deletion in TIMP-3 occurs at the turn between strands A and B. We propose that the short deletions in TIMP-1 and TIMP-3 are accommodated in the ΔTIMP-2 structure by a shortening of strands A and B, with the tight turn occurring earlier in the sequence. This region is clearly exposed on the surface of ΔTIMP-2 (Figure 7B) and believed to be flexible, as discussed earlier. This proposal would place the carbohydrate chain attached to N30

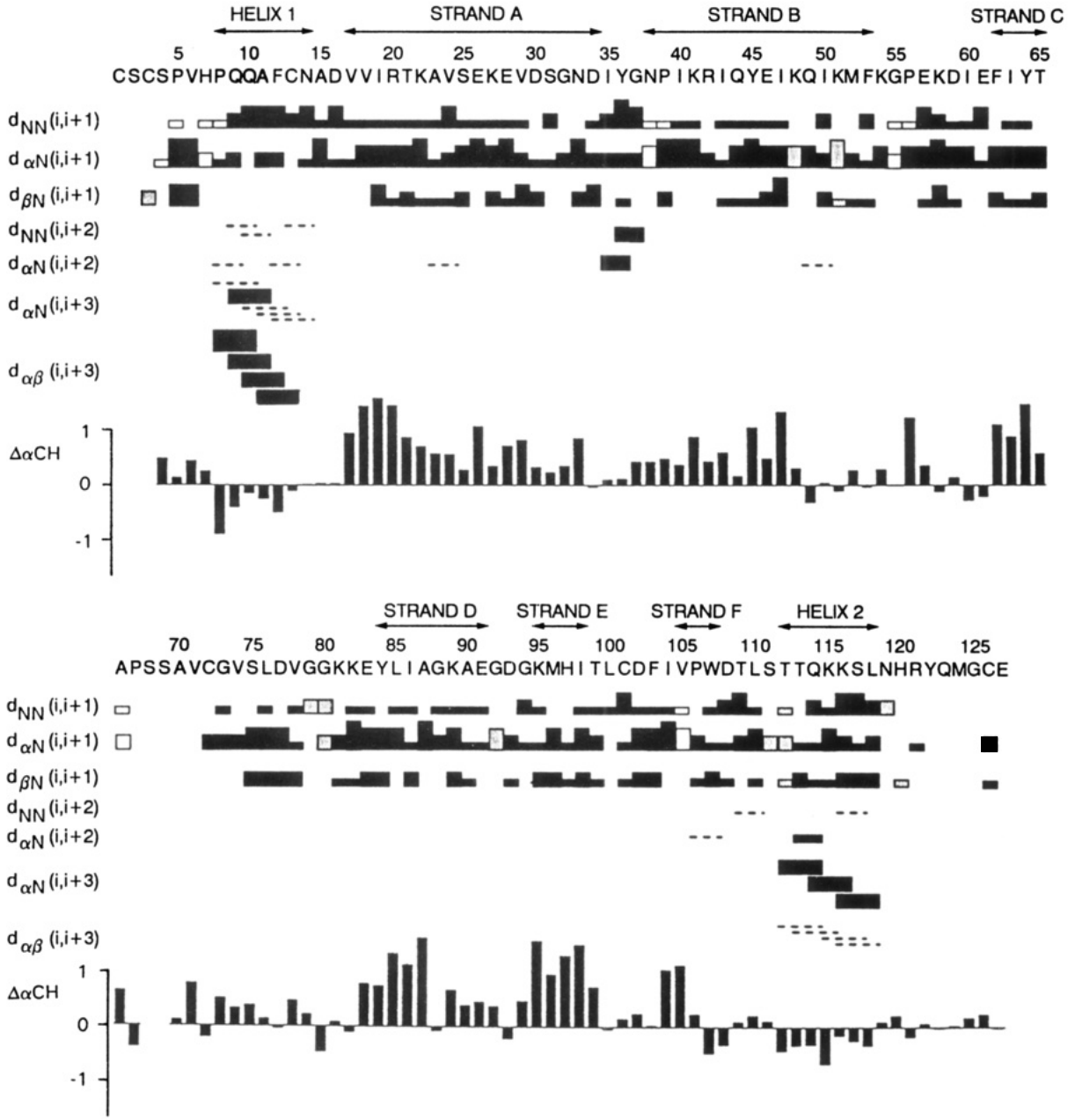

FIGURE 5: Summary of sequential ($i, i + 1$) and medium-range ($i, i < 5$) NOEs identified for ΔTIMP-2 and the differences of αCH chemical shifts from random coil values. The height of the bar used to represent an NOE is an indication of its intensity, with open bars used to represent NOEs involving proline δ protons and shaded bars for NOEs seen only at pH* 4.6. The dashed lines represent possible medium-range NOEs which could not be unambiguously assigned due to resonance overlap. The differences between the observed αCH chemical shifts for ΔTIMP-2 and random coil values (ΔαCH = αCH observed – αCH random coil) are shown beneath the NOE data, and the positions of elements of regular secondary structure are shown above the protein sequence.

in TIMP-1 at the start of the tight turn and on the surface of the protein, an environment consistent with that expected for a glycosylation site.

Structural Homology with the OB Protein Fold. The secondary and tertiary structures of ΔTIMP-2 clearly identify the protein fold as an α-β roll (Orengo & Thornton, 1993) where the β-sheet is "rolled" over to protect a hydrophobic core. Closer analysis revealed a striking similarity between ΔTIMP-2 and the so-called OB fold, a subgroup of the α-β rolls described by Murzin (1993). The common feature of this structural group is the closed five-stranded β-barrel. The proteins known to contain an OB fold motif include staphylococcal nuclease, yeast aspartyl-tRNA synthetase, the B subunit of *E. coli* heat-labile enterotoxin, the B subunit of *E. coli* verotoxin, bacteriophage gene-5 protein [all described by Murzin (1993)], and the cold-shock protein from *Bacillus subtilis* (Schindelin et al., 1993; Schnuchel et al., 1993). A comparison of the β-sheet arrangement of ΔTIMP-2 with that of the OB fold shows that all of the major structural features of the closed β-barrel are in common (Figure 8). The β-barrel of ΔTIMP-2 has the same shear number as the OB fold (the residue offset resulting from one traverse around the barrel), and the strands can be easily superimposed to give the same hydrogen bond network. The β-bulge at V24 is present in other members of the OB fold family and is thought to be important for the bending or coiling of strand A (Murzin, 1993). Strands E and F in ΔTIMP-2 can be considered as part of the same strand in the β-sheet with the continuity broken by a five-residue loop ($T_{99}LCDF_{103}$) in a position equivalent to a β-bulge present in some other OB fold members. In the TIMPs this loop contains a conserved Cys residue which forms a disulfide bond with the N-terminus. Also homologous

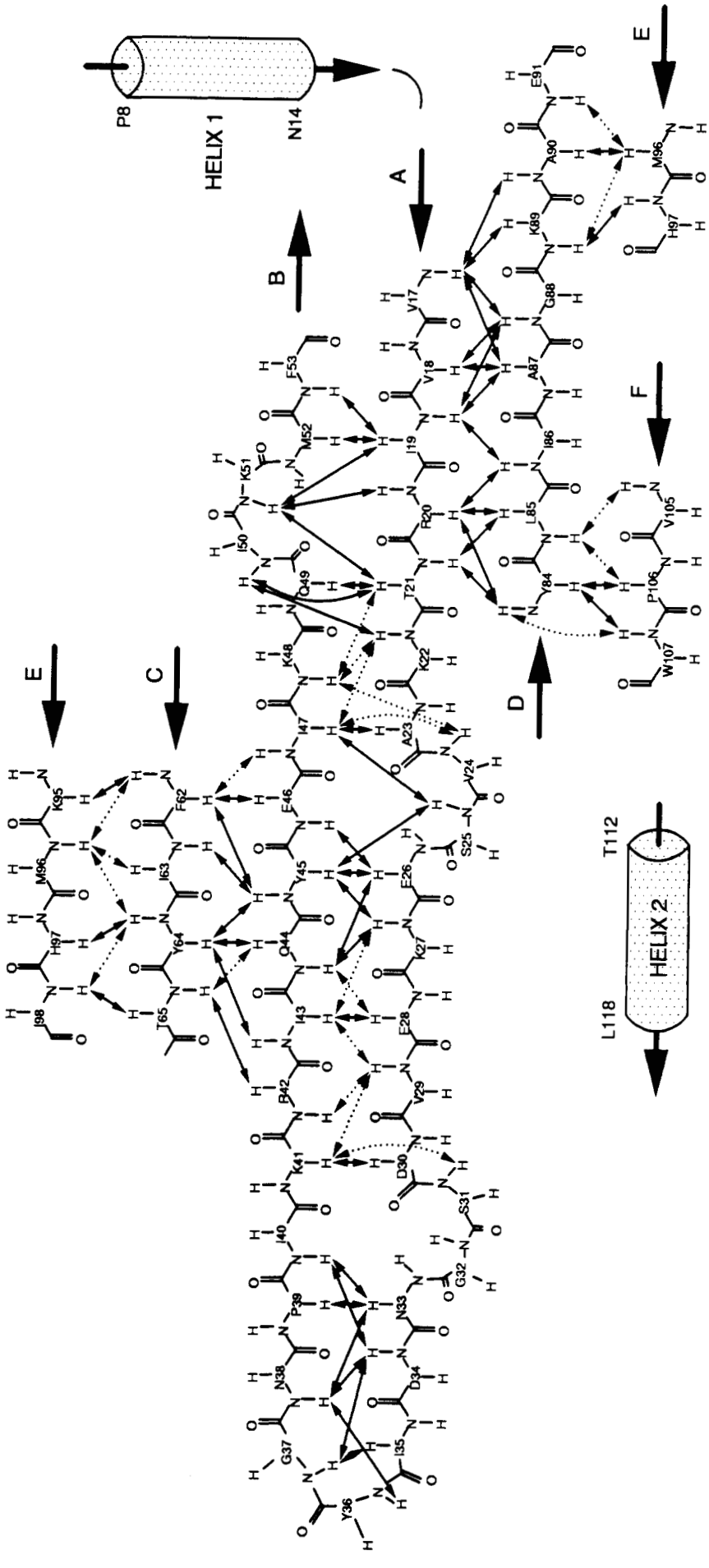

FIGURE 6: Schematic representation of the secondary structure of ΔTIMP-2. Long-range NOEs between backbone protons are denoted by double-headed arrows and were used to align the β-strands. Arrows with dotted lines represent tentatively assigned NOEs which could not be unambiguously identified due to resonance overlap and were not used in the structural calculations.

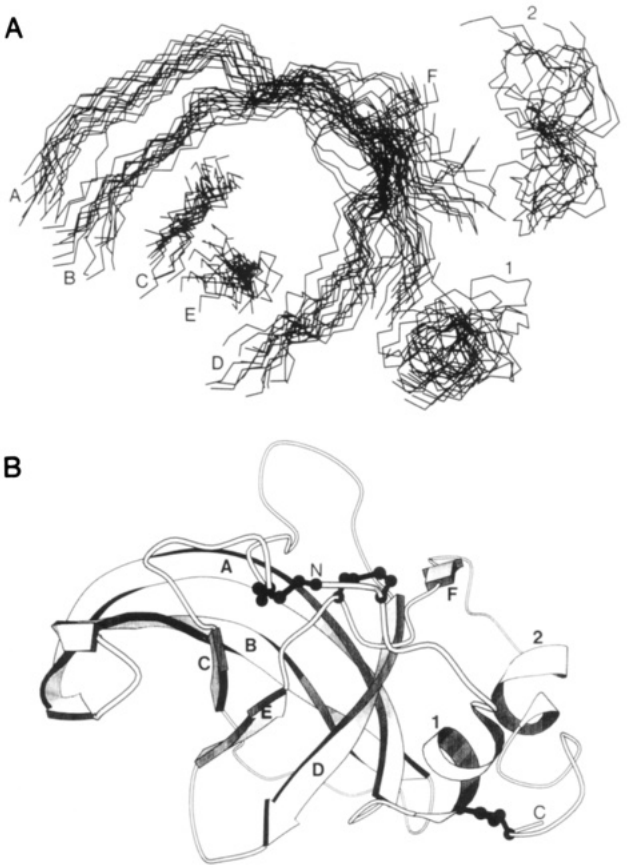

FIGURE 7: The backbone topology of ΔTIMP-2. (A) Superposition of the secondary structural elements from the 16 converged structures generated by DIANA. Backbone traces are shown with the β-strands labeled A–F and the helices labeled 1 and 2. (B) MolScript representation (Kraulis, 1991) of one of the structues shown in panel A (the one with the lowest final DIANA target function, i.e., the structure most consistent with the experimental NMR constraints). Disulfide bonds are shown in ball-and-stick representation linking together Cys residues 1–72, 3–101, and 13–126.

is the arrangement of residues which form the hydrophobic core of the barrel. In every case the core consists of three layers of residue side chains (five in each layer) which stack on top of one another perpendicular to the barrel axis. The strands forming the β-barrel contribute one residue to each layer. The residues involved in TIMP-2 are shown in Figure 8.

To compare the three-dimensional topology of the ΔTIMP-2 β-barrel to that of the OB fold, superpositions of the appropriate backbone atoms with those of staphylococcal nuclease (SN) and the B subunit of heat-labile enterotoxin (LTB) were carried out (Figure 9) and the pairwise rmsd values are summarized in Table 2. The rmsd value of the SN–LTB comparison is essentially the same as that reported by Murzin (2.30; Murzin, 1993). The β-barrel of ΔTIMP-2 is most similar in structure to that found in LTB (Figure 9), but the range of rmsd values obtained clearly shows that all three proteins share a common β-barrel topology.

All of the OB proteins so far described bind either oligosaccharides or oligonucleotides (hence the name: *o*ligosaccharide/-nucleotide *b*inding); however, there are no published reports suggesting that TIMPs might have a similar function. Nevertheless, a common ligand binding site for the OB fold has been identified (Murzin, 1993), and it is worth examining whether the analogous region of ΔTIMP-2 could be part of a functional site. The OB fold binding site does not involve homologous residues in the different family members, but common features do exist. The site is mainly composed of the loops between strands A/B and D/E which protrude outward from the barrel axis and form a shallow groove with the β-sheet at its floor. Sequence analysis of the β-sheet region in the TIMP family identifies no completely conserved residues, except for Y45 and T65, both of which are known to be located in the core of ΔTIMP-2 (Figure 8). Furthermore, the loop D/E is short in ΔTIMP-2 and is not highly conserved except for G94 (which is probably required for the tight turn), and loop A/B is variable in both length and sequence. It is interesting to note that the highly conserved residues in the β-sheet of ΔTIMP-2 tend to be clustered on the opposite side of the β-barrel on strands A an B. Thus, there is little evidence from sequence conservation in the TIMPs to suggest that they share a common ligand binding site with other previously identified members of the OB fold family. Whether there are as yet unidentified oligonucleotide/oligosaccharide ligands that bind to TIMPs or whether the TIMP and OB folds have arisen independently (i.e., by convergent evolution) remains unknown. The number of ways in which a five-stranded β-sheet can fold to enclose a hydrophobic core is limited, and so there is a reasonable probability of several independent evolution pathways leading to a β-barrel fold with a shear number of 10.

Functional Implications of the ΔTIMP-2 Structure. The solution structure of ΔTIMP-2 reported here allows for the first time some assessment of the potential functional importance of several highly conserved regions of the TIMP proteins. The conserved sequence motifs V_{18}IRA/TK_{22} in strand A and Y/F_{84}LI/L/VA/TG_{88} in strand D both form an integral part of the closed β-barrel structure. Thus alternate residues in these sequences either form part of the hydrophobic core or point out from the barrel toward the solvent or other elements of the protein. Clearly, side chains that are internalized cannot take part in binding interactions with the MMPs and are conserved for structural reasons. In addition, a number of residues in these two regions with surface side chains are involved in the packing of helices 1 and 2 onto the β-barrel core. Woessner (1991) suggested that one of these two sequence motifs may form the inhibitory site on the TIMPs responsible for interacting with the active site of the metalloproteinase, a proposal based on their sequence similarity to known MMP cleavage sites. This suggestion now seems very unlikely since the highly structured nature of both of these regions is inconsistent with the properties normally associated with a proteinase "bait" region. The importance of both of these sequences to the structural integrity of TIMPs was indicated by recent site-directed mutagenesis experiments. Deletion or modification of either of these sequences in TIMP-1 (removal of V_{18}IRAK_{22} or substitution of D_{16}LVIR_{20} to AAAAA, L_{17}V to TE, E_{81}EFLI$_{85}$ to AAAAA, and F_{83}L to TE) resulted in clones which showed very little or no protein expression (O'Shea et al., 1992; M. O'Shea and G. Murphy, unpublished work). Presumably disruption to parts of the core structure of the protein results in a molecule that is unable to fold and so fails to be secreted from the cell.

Analysis of the surface of the ΔTIMP-2 structure for patches of highly conserved residues should help to locate the position of the MMP binding site or sites. For truncated proteins, such as ΔTIMP-2, and subunits from multimeric proteins, care must be taken not to confuse ligand binding sites with regions normally involved in interaction with other protein domains. Two potentially interesting surface regions were identified on ΔTIMP-2. The first, consisting of conserved positively charged residues R20, K22, K48, and K51, forms

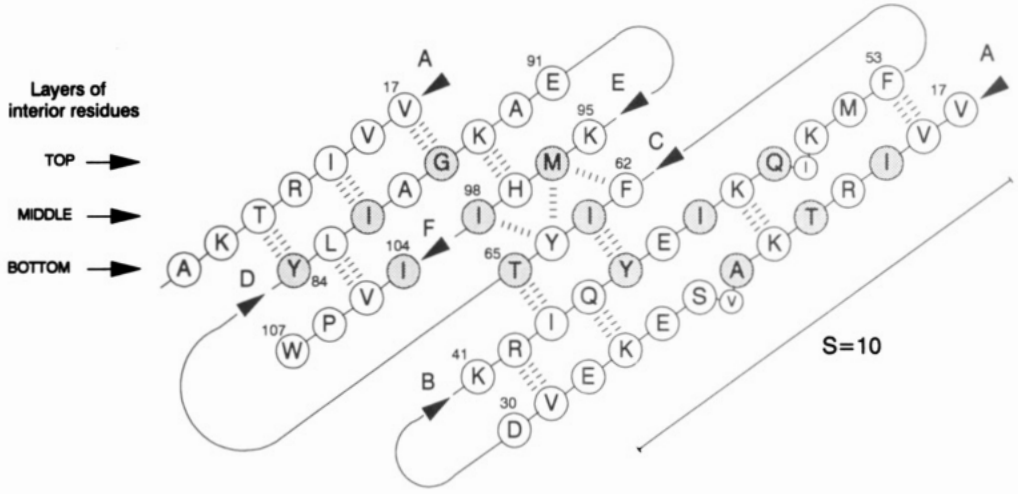

FIGURE 8: A flattened view of the β-sheet arrangement of ΔTIMP-2 drawn to show structural homology with the OB fold protein family [see Figure 4 in Murzin (1993)]. Strand A is shown twice in order to show how the β-barrel is formed. Hydrogen bonds are shown by dashed lines and are as predicted from the β-sheet alignment shown in Figure 6. The residues with side chains which contribute to the hydrophobic core of the barrel are shown as shaded circles. *S* represents the barrel shear number and is equal to the residue offset resulting from one traverse around the barrel.

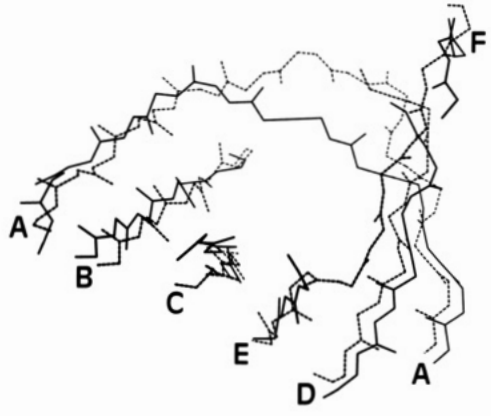

FIGURE 9: Superposition of structurally homologous β-strands from ΔTIMP-2 (solid line) and the B subunit of heat-labile enterotoxin (LTB, dashed line). Backbone atoms were matched as described in Table 2 and the text. LTB contains an additional two-residue loop in strand A which is absent from ΔTIMP-2 and so was omitted from the superposition. The figure was drawn using MolScript (Kraulis, 1991).

Table 2: Root Mean Square Deviation (rmsd) Values (Å) Calculated for the Backbone Atoms (N, αC, C, O) following Superposition of Structurally Homologous Residues within the OB Fold β-Barrel[a]

	SN	LTB
ΔTIMP-2	2.60 (29)	1.99 (34)
SN		2.27 (32)

[a] The number of residues matched in each case is shown in parentheses. The coordinates for staphylococcal nuclease (SN) were obtained from the Brookhaven database (1SNC; Loll & Lattman, 1989), the coordinates for the B subunit of heat-labile enterotoxin (LTB) were kindly provided by T. K. Sixma and W. G. J. Hol (Sixma et al., 1992), and the coordinates for ΔTIMP-2 were those of the structure with the lowest final variable target function in the DIANA calculations as used in Figure 7B.

a convex surface at the back of the β-barrel as it is shown in Figure 7B. The second site, consisting of conserved residues F103, L85, W107, Q114, L118, Y122, and H7, is rather hydrophobic in nature and forms a large, oval, gently concave face which is close to helices 1 and 2, strand F, and the loop E/F. In the case of TIMP-1, residues equivalent to R20, K22, and K51 have each been mutated individually to Ala with no substantial effect on TIMP activity (O'Shea et al., 1992), which clearly argues against these residues forming the main site of MMP interaction. The second site is close to the C-terminus of ΔTIMP-2 and as such is an obvious candidate for the site of interaction between the N- and C-terminal domains. Site-directed mutagenesis studies of TIMP-1 identified three positions (H7, Q9, and W105) at which substitution to Ala resulted in proteins with a reduced affinity for matrilysin (O'Shea et al., 1992). The equivalents of these residues lie either within or on the edge of this hydrophobic surface region. This result could still be consistent with our domain-interface proposal if modification at these sites results in a change in the orientation between the two TIMP-1 domains which in turn disrupts the proper TIMP-1–matrilysin interaction. Further site-directed mutagenesis on other conserved residues in this region may give a greater insight into its relative importance for TIMP function.

The solution structure of ΔTIMP-2 reported here provides the first molecular framework on which to begin to interpret protein sequence and site-directed mutagenesis data for the TIMP family. Equally exciting, the structure itself identifies conserved surface regions which may prove to be important for the TIMP–MMP interaction. It is now possible for the first time to take a rational approach to site-directed mutagenesis work and alter residues in known structural environments to test specific theories on the mechanism of TIMP inhibition of MMP activity.

ACKNOWLEDGMENT

The NMR spectra were recorded using facilities at the Medical Research Council Biomedical NMR Centre, NIMR, Mill Hill. The authors would like to thank Dr. C. J. Bauer for help with the 3D NOESY-TOCSY experiment, M. Cockett for ΔTIMP-2 expression, and P. Gane for help with QUANTA software at Kent University.

SUPPLEMENTARY MATERIAL AVAILABLE

DQF-COSY spectrum from the aromatic region of ΔTIMP-2 and NOESY spectra from the NH–NH and αCH–αCH regions (3 pages). Ordering information is given on any current masthead page.

REFERENCES

Apte, S. S., Mattei, M.-G., & Olsen, B. R. (1994) *Genomics 19*, 86–90.

Bax, A., & Davis, D. G. (1985) *J. Magn. Reson. 65*, 355–360.
Boone, T. C., Johnson, M. J., DeClerck, Y. A., & Langley, K. E. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 2800–2804
Borkakoti, N., Winkler, F. K., Williams, D. H., D'Arcy, A., Broadhurst, M. J., Brown, P. A., Johnson, W. H., & Murry, E. J. (1994) *Nat. Struct. Biol. 1*, 106–110.
Braunschweiler, L., & Ernst, R. R. (1983) *J. Magn. Reson. 53*, 521–528.
Brown, S. C., Weber, P. L., & Müller, L. (1988) *J. Magn. Reson. 77*, 166–169.
Carmichael, D. F., Sommer, A., Thompson, R. C., Anderson, D. C., Smith, C. G., Welgus, H. G., & Stricklin, G. P. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 2407–2411.
Carr, M. D. (1992) *Biochemistry 31*, 1998–2004.
Carr, M. D., Birdsall, B., Frenkiel, T. A., Bauer, C. J., Jimenez-Barbero, J., Polshakov, V. I., McCormick, J. E., Roberts, G. C. K., & Feeney, J. (1991) *Biochemistry 30*, 6330–6341.
Carr, M. D., Bauer, C. J., Gradwell, M. J., & Feeney, J. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 2206–2210.
Cawston, T. E., Murphy, G., Mercer, E., Galloway, W. A., Hazleman, B. L., & Reynolds, J. J. (1983) *Biochem. J. 211*, 313–318.
Davis, D. G., & Bax, A. (1985) *J. Am. Chem. Soc. 107*, 2820–2821.
Docherty, A. J. P., Lyons, A., Smith, B. J., Wright, E. M., Stephens, P. E., Harris, T. J. R., Murphy, G., & Reynolds, J. J. (1985) *Nature (London) 318*, 66–69.
Docherty, A. J. P., O'Connell, J., Crabbe, T., Angal, S., & Murphy, G. (1992) *Trends Biotechnol. 10*, 200–207.
Goldberg, G. I., Marmer, B. L., Grant, G. A., Eisen, A. Z., Wilhelm, S., & He, C. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 8207–8211.
Gooley, P. R., O'Connell, J. F., Marcy, A. I., Cuca, G. C., Salowe, S. P., Bush, B. L., Hermes, J. D., Esser, C. K., Hagmann, W. K., Springer, J. P., & Johnson, B. A. (1994) *Nature Struct. Biol. 1*, 111–118.
Güntert, P., & Wüthrich, K. (1991) *J. Biomol. NMR 1*, 447–456.
Güntert, P., Braun, W., & Wüthrich, K. (1991) *J. Mol. Biol. 217*, 517–530.
Hayakawa, T., Yamashita, K., Tanzawa, K., Uchijima, E., & Iwata, K. (1992) *FEBS Lett. 298*, 29–32.
Jeener, J., Meier, B. H., Bachmann, P., & Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546–4553.
Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946–950.
Loll, P. J., & Lattman, E. E. (1989) *Proteins: Struct., Funct., Genet. 5*, 183–201.
Lovejoy, B., Cleasby, A., Hassell, A., Longley, K., Luther, M. A., Weigl, D., McGeehan, G., McElroy, A. B., Drewry, D., Lambert, M. H., & Jordan, S. R. (1994) *Science 263*, 375–377.
Macura, S., Huong, Y., Suter, D., & Ernst, R. R. (1981) *J. Magn. Reson. 43*, 259–281.
Matrisian, L. M. (1990) *Trends Genet. 6*, 121–125.
Matrisian, L. M. (1992) *BioEssays 14*, 455–463.
Murphy, G., & Werb, Z. (1985) *Biochim. Biophys. Acta 839*, 214–218.
Murphy, G., Koklitis, P., & Carne, A. F. (1989) *Biochem. J. 261*, 1031–1034.
Murphy, G., Houbrechts, A., Cockett, M. I., Williamson, R. A., O'Shea, M., & Docherty, A. J. P. (1991) *Biochemistry 30*, 8097–8102.
Murzin, A. G. (1993) *EMBO J. 12*, 861–867.
Nguyen, Q., Willenbrock, F., Cockett, M. I., O'Shea, M., Docherty, A. J. P., & Murphy, G. (1994) *Biochemistry* (in press).
Orengo, C. A., & Thornton, J. M. (1993) *Structure 1*, 105–120.
Oschkinat, H., Griesinger, C., Kraulis, P. J., Sorensen, O. W., Ernst, R. R., Gronenborn, A., & Clore, G. M. (1988) *Nature 332*, 374–376.
O'Shea, M., Willenbrock, F. W., Williamson, R. A., Cockett, M. I., Freedman, R. B., Reynolds, J. J., Docherty, A. J. P., & Murphy, G. (1992) *Biochemistry 31*, 10146–10152.
Pavloff, N., Staskus, P. W., Kishnani, N. S., & Hawkes, S. P. (1992) *J. Biol. Chem. 267*, 17321–17326.
Press, W. H., Flannery, B. P., Teukolsky, S. A., & Vetterling, W. T. (1988) *Numerical Recipes in C, The Art of Scientific Computing*, Cambridge University Press, Cambridge.
Rance, M., Sorensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 117*, 479–485.
Schindelin, H., Marahiel, M. A., & Heinemann, U. (1993) *Nature 364*, 164–168.
Schnuchel, A., Wiltscheck, R., Czisch, M., Herrier, M., Willimsky, G., Graumann, P., Maraheil, M. A., & Holak, T. A. (1993) *Nature 364*, 169–171.
Sixma, T. K., Pronk, S. E., Kalk, K. H., van Zanten, B. A. M., Berghuis, A. M., & Hol, W. G. J. (1992) *Nature 355*, 561–564.
Stams, T., Spurlino, J. C., Smith, D. L., Wahl, R. C., Ho, T. F., Qoronfleh, M. W., Banks, T. M., & Rubin, B. (1994) *Nature Struct. Biol. 1*, 119–123.
States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286–292.
Stricklin, G. P. (1986) *Collagen Relat. Res. 6*, 219–228.
Tolley, S. P., Davies, G. J., O'Shea, M., Cockett, M. I., Docherty, A. J. P., & Murphy, G. (1993) *Proteins: Struct., Funct., Genet. 17*, 435–437.
Wagner, G., & Wüthrich, K. (1982) *J. Mol. Biol. 155*, 347–366.
Ward, R. V., Hembry, R. M., Reynolds, J. J., & Murphy, G. (1991) *Biochem. J. 278*, 179–187.
Welgus, H. G., Jeffrey, J. J., Eisen, A. Z., Roswit, W. T., & Stricklin, G. P. (1985) *Collagen Relat. Res. 5*, 167–179.
Wilhelm, S. M., Collier, I. E., Marmer, B. L., Eisen, A. Z., Grant, G. A., & Goldberg, G. I. (1989) *J. Biol. Chem. 264*, 17213–17221.
Willenbrock, F., Crabbe, T., Slocombe, P. M., Sutton, C. W., Docherty, A. J. P., Cockett, M. I., O'Shea, M., Brocklehurst, K., Phillips, I. R., & Murphy, G. (1993) *Biochemistry 32*, 4330–4337.
Williamson, R. A., Marston, F. A. O., Angal, S., Koklitis, P., Panico, M., Morris, H. R., Carne, A. F., Smith, B. J., Harris, T. J. R., & Freedman, R. B. (1990) *Biochem. J. 268*, 267–274.
Williamson, R. A., Smith, B. J., Angal, S., & Freedman, R. B. (1993) *Biochim. Biophys. Acta 1203*, 147–154.
Williamson, R. A., Bartel, H., Murphy, G., & Freedman, R. B. (1994) *Protein Eng. 7*, 1035–1040.
Wishart, D. S., Sykes, B. D., & Richards, F. M. (1991) *J. Mol. Biol. 222*, 311–333.
Woessner, J. F. (1991) *FASEB J. 5*, 2145–2154.
Wüthrich, K. (1986) *NMR of Proteins & Nucleic Acids*, Wiley, New York.
Yang, T.-T., & Hawkes, S. P. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 10676–10680.