Human plasma fibronectin (Fn) is a large glycoprotein present in blood plasma at about 0.3 mg/ml. It plays an important role in many biological processes. It consists of two identical 230–250-kDa subunits that are joined by two disulfide bonds near their carboxyl termini. Each subunit contains various binding domains composed of three types of homologous repeats. Recent work has determined the three-dimensional structures of various repeat fragments, but little is known about the three-dimensional structure of the carboxyl-terminal region. A recent NMR study of a plasmin-digested carboxyl-terminal inter-chain disulfide-linked heptapeptide dimer has proposed that the two subunits are arranged in an antiparallel fashion (An et al. 1992) Biochemistry 31, 9927–9933). We have now determined the three-dimensional structure for a substantial portion of a trypsin-digested interchain disulfide-linked 52-residue (6 kDa) fragment of the carboxy-terminal of human plasma fibronectin (which includes the above-mentioned heptapeptide dimer) using two-dimensional NMR methods and a new strategy for NMR-based protein structure determination. The NMR data requires that the two chains in the dimer be linked in a symmetric, antiparallel arrangement. The resulting monomer conformation consists of two twisted or coiled segments, Thr<sup>2</sup>-Asn<sup>7</sup> and Ile<sup>8</sup>-Phe<sup>13</sup>, connected by the Cys<sup>2</sup>-Pro<sup>5</sup> residues in extended conformations, with the two monomer chains cross-linked at residues Cys<sup>7</sup> and Cys<sup>13</sup>. The conformation of the heptapeptide dimer region differs substantially from the conformations proposed by An et al.

**References**

- An, G., et al. (1992) Biochemistry 31, 9927–9933.
- Baron, D. et al. (1990) J. Biol. Chem. 265, 10222–10229.
- Hynes, R. O. (1990) J. Biol. Chem. 265, 10222–10229.
- Kornbluth, S., et al. (1985) J. Biol. Chem. 260, 4970–4976.
- Patthy, L. (1990) J. Mol. Biol. 210, 699–708.
- Yamada, K. M. (1987) J. Cell Biol. 105, 105A.

Human plasma fibronectin (Fn)<sup>1</sup> is a large glycoprotein present in blood plasma at about 0.3 mg/ml. It plays an important role in many biological phenomena including cell adhesion and spreading, wound healing, phagocytosis, and differentiation (Hynes, 1990; Mosher, 1989; Akiyama and Yamada, 1987). The protein consists of two nearly identical subunits of 230–250 kDa each, containing binding domains for various biomolecules including fibrin, heparin, collagen, DNA, and cell surface molecules. The various binding domains of Fn are composed of homologous repeats of three different types, namely, type I, II, and III, containing 45, 60, and 90 amino acids, respectively (Petersen et al., 1983; Kornbluth et al., 1985). These three types of homologous repeats are also found in other proteins involved in blood coagulation and in proteins of the extracellular matrix (Patthy, 1990). 2D NMR investigation has shown that the seventh type I repeat of Fn (48 residues) primarily consists of two antiparallel β-sheets (Baron et al., 1990). The type II repeats of Fn resemble the Kringle structure derived from the crystal structure of the prothrombin Kringle 1 unit (Holland et al., 1987; Constantine et al., 1992). The solution structure of the tenth type III repeat (94 residues) has been determined by 2D and 3D NMR (Baron et al., 1992), and the crystal structure of a Fn type III domain from tenasin, an extracellular matrix protein has been determined by x-ray crystallography (Leahy et al., 1992); both structures consist of two antiparallel β-sheets with an immunoglobulin-like fold.

The two subunits of Fn are joined near their carboxyl termini by two disulfide bonds. Despite recent progress in obtaining detailed information concerning the primary sequence and gene structure of Fn, and the three-dimensional structures of the various repeat fragments, comparatively little is known about the three-dimensional structure of the carboxyl-terminal region. In particular, the spatial arrangement of the subunits about these disulfide bonds, i.e. whether the monomeric chains are arranged in a parallel or an antiparallel fashion in the dimer, has not been conclusively determined. A recent report, based on analysis of 2D NMR data for a 14-residue fragment of the carboxyl-terminal region, has suggested that the two subunits are arranged in an antiparallel fashion, and has proposed two alternative three-dimensional structures for aqueous and dimethyl sulfoxide environments. There are, however, problems with the analysis and resulting structures presented in that work that are discussed in more detail in the discussion section below.

In this work, we have purified the carboxyl-terminal 6-kDa Fn fragment containing two 26-residue fragments with interchain disulfide bonds, and have determined the three-dimensional structure of the disulfide linked region by 2D NMR methods. We report NMR assignments for 25 of the 26 residues in the monomer of the 6-kDa carboxyl-terminal fragment. The NOE data and a new strategy for NMR-based...
protein structure determination have been used to build the three-dimensional structure of the Thr-Pro segment containing the interchain disulfide bonds. We show that the NMR data are consistent with a set of constraints: those in the two interchain disulfide bonds linking the monomers of the Fn molecule are arranged in an antiparallel fashion. The resulting structure for the disulfide-linked region of our 52-residue fragment differs substantially from those reported by An et al. (1992) for their 14-residue fragment. We suggest that the highly truncated heptapeptide dimer may not retain its native conformation in either aqueous or dimethyl sulfoxide solution.

**EXPERIMENTAL PROCEDURES**

Sample Preparation—Fn was purified from fresh-frozen human plasma, obtained from the Blood Center of Southeastern Wisconsin, on a Sephacore 4B column and a gelatin-Sephacore 4B column, arranged in tandem (Engvall and Ruoslahti, 1977). The integrity and purity of the protein were confirmed by SDS-PAGE (Stokes, 1992), with sulfo-polyacrylamide gel electrophoresis. Purification of the carboxyl-terminal 6-kDa fragment was performed essentially as described by Garcia-Pardo et al. (1984) with some modifications. Fn (typically 60 mg) was incubated with trypsin at a 3300:1 (w/v) ratio of Fn:trypsin for 50 min. Digestion was stopped by adding phenylmethylsulfonyl fluoride to a final concentration of 0.1 mM. The 6-kDa fragment was eluted with 10 mM Tris at pH 7.4, containing 0.5 mM NaCl. The fractions were pooled, dialyzed, concentrated by partial lyophilization, and loaded onto a Sephadex G-50 column (1.5 x 95 cm, approximately 170 ml, flow rate of 8 ml/h) for final purification. The 6-kDa fragment from this step was free of higher molecular weight fragments as judged by SDS-PAGE (silver stained, 15% acrylamide gel). The typical yield was about 0.5 mg of the 6-kDa fragment per 60 mg of plasma Fn (Peterson et al., 1983). For NMR measurements, about 3 mg of the 6-kDa fragment was lyophilized to dryness and redissolved in 0.5 ml of either D2O, or H2O containing 5% D2O, giving a final peptide concentration of about 1 mM.

**1H NMR Spectroscopy**—NMR data were accumulated on a General Electric GN500 spectrometer equipped with a 1280 Nicolet computer. Phase-sensitive two-dimensional (2D) COSY and NOESY data sets were collected on 88- and 95-cm-diameter hypercomplex tubes using standard pulse sequences and phase expressions (Jeener et al., 1979, Wider et al., 1984). NOESY data were acquired with mixing times of 125 and 250 ms. Relayed COSY experiments in the absolute value mode (Wagner, 1983) were used to help identify spin systems of side chains. All 2D data sets consisted of 1024 complex points in the \( t_2 \) dimension and either 256 or 320 complex points in the \( t_1 \) dimension. All data processing was done on a Silicon Graphics IRIS computer, using the FTNMR software (Hare Inc., Woodinville, WA). Digital filtering was used prior to Fourier transformation in every case. All NOESY spectra were base-line-corrected by a fifth-order polynomial. Chemical shifts were referenced to the water signal, which is 4.79 ppm from 4,4'-dimethyl-4-silapentane-1-sulfonate at 25°C.

**Computational Methods**—Two different approaches were examined for the determination of spatial structure from NMR data: (i) the deterministic distance geometry (DG) approach (Wuthrich, 1989), followed by energy refinement; and (ii) a build-up strategy (BUILD), using a combination of 2D NOESY, COSY, and RELAY COSY experiments to generate structures consistent with covalency constraints and semiquantitative estimates of inter-proton distances, starting with random initial atomic coordinates (Nerdel et al., 1988). 2.0 Å was used as a lower limit of distance constraints, and 2.6, 3.3, and 4.0 Å were used as upper limits for qualitatively observed strong, medium, and weak cross-peak intensities, based on the estimation of distances from the experimentally observed volume integrals for a sampling of NOE cross-peaks (using a proportionality constant derived from the volume integrals of the cross-peaks relating \( dH \) and \( dH \) of Phe across a distance of 2.5 Å). Structure refinement was obtained by randomization of coordinates followed by a cycle of simulated annealing and conjugate gradient minimization of penalty functions. All atoms in the segment considered were subjected to simulated annealing. Similarly, all atoms were considered in the calculations of penalty function and gradient. Since the NMR data indicate that energy minimization was used as a constraint in the energy refinement process. Additional structures were generated by random embeds, and the energy refinement and simulated annealing algorithms were repeated to minimize the constraint violations for the new structures. Structures were defined by DSPACE, and selected qualitatively for their conformational diversity by comparison of \( \phi, \psi \) plots, was also used as starting structures for constrained energy minimization (DGREM) and restrained molecular dynamics (DGRMD) calculations using the CHARMM software package, with empirical energy potentials taken from Brooks et al. (1983). All calculations were performed on a Silicon Graphics work station.

The BUILD approach uses, in addition to NMR data, a priori information on empirical distribution functions for backbone conformations, generated from the high resolution x-ray structures in the Protein Data Bank. NMR information regarding the presence or absence of sequential \( d \) connectivities (NOE cross-peaks among near-neighbor residues) is used for statistical prediction of backbone conformations of individual amino acid residues. A three-step procedure was followed: (i) estimation of a starting set of angular coordinates (local conformations) from the NMR data (Sherman et al., 1987); (ii) determination of the spatial structure by a gradual build-up process (Sherman et al., 1988); (iii) structure refinement by energy minimization on unconstrained structures. The FISINOE approach (Sherman and Johnson, 1992) was used to estimate the \( \phi, \psi \) values with corresponding probabilities for each residue, given the \( d_{\alpha\alpha} \), \( d_{\alpha\beta} \), and \( d_{\beta\beta} \) connectivities. The most probable \( \phi, \psi \) values for each residue were used as the starting set of angular coordinates. The BUILD approach to obtain a spatial structure from the starting set of backbone conformations utilizes an optimality principle in which the fragment considered at any stage has a minimum number of residues and a maximum number of restrictions. Long-range NOE requirements and the value of conformational energy were used as steering parameters to guide the BUILD process. Two general assumptions were used: (i) the upper limit of distance constraints is 3.5 Å for sequential \( d \) connectivities, and 4 Å for long-range NOEs; and (ii) all dihedral angles must lie within sterically allowed regions in conformational space. For the sake of convenience, all NOEs other than sequential are termed long-range. The interactive graphics package INSIGHT (Bioym Technologies, San Diego, CA) was used to construct the starting structures, with backbone conformations determined by FISINOE. All energy minimizations were performed using CHARMM on a Silicon Graphics work station. The force constant for the dihedral constraints was reduced in gradual steps from 50.0 (kcal mol\(^{-1}\) rad\(^{-2}\)) to 2.5 and finally to 0.0, decreasing it by about a factor of two following each cycle of 250 steps of conjugate gradient minimization. A macroscopic dielectric constant of 10 was used for these calculations. Calculations using dielectric constants of 1, 2, 4, 10, 30, and 80 showed that conclusions based on energy considerations were unaffected for dielectric constants greater than 4. Although the same starting values were used for the dihedral angles of a particular residue in the two chains, symmetry was not used explicitly as a constraint in the energy minimization process. In general, distance constraints also were not included explicitly in the calculations, except where specifically stated.

**RESULTS AND DISCUSSION**

**NMR Assignments**—The 6-kDa carboxyl-terminal fragment from human plasma fibronectin has the following sequence for the 26-residue monomer, with two monomers joined either parallel or antiparallel to form the dimer Thr-Asn-Thr-Asn-Val-Asp-Val-Asp-Val-Cys-Pro-Val-Lys-Glu-Lys-Phe-Ase-Leu-Asp-Val-Glu-Ala-Ala-Asp-Arg-Glu-Asp-Arg-Ase.

The Thr\(_2\) to Arg\(_{25}\) segment was readily assigned by a combination of 2D NOESY, COSY, and RELAY COSY experiments in D2O and H2O. Resonances for Glu\(_{25}\) and the amide resonances for Thr\(_1\) and Asn\(_{2}\) were not readily observable. The threonine and valine side chains were identified by a relay experiment in D2O that showed the cross-peak between...
were obtained from a 1 mM solution in 95% H2O and 5% D2O at pH 7.0 to 7.5, with a 250-ms mixing time for the NOESY experiment. The spin systems for Proα and Ileβ are traced out using large dashes and solid lines, respectively. Small dashed lines show the NOEs relating the α- and β-protons of Cys7 to α- and β-protons of Pro6.

**Fig. 1.** Backbone amide and aromatic region of a 250 ms NOESY spectrum of the 6-kDa fibronectin fragment. Spectra were obtained from a 1 mM solution in 95% H2O and 5% D2O at pH 7.0 to 7.5. The dNN connectivity pathways for residues Thrα-Cysβ (solid lines) and Ileα-Metα (large dashed lines) are indicated in the bottom panel. Note that proline breaks the dNN connectivity pathway at Thrα and Ileβ. NOE contacts between aromatic and amide protons of Pheα are shown with small dashes. Small dashed lines are also used to show the NOE contacts relating the aromatic protons of Pheβ in the bottom panel, to the β-protons of Cysγ, Cysγ', and Pheγ, and to the Cα-proton of Ileε in the top panel. The αN cross-peak relating Proα and Ileβ is shown in the top panel, along with the long-range dNN connectivities for (9-11) and (10-13). The β-protons of residues 4, 5, 6, 7, 9, 10, 11, 12, and 13 have been labeled in the top panel so as to show the presence of intraresidual NOE cross-peaks.

**Fig. 2.** Combined COSY-NOESY spectra showing the aliphatic region. Spectra were obtained in D2O at pH 7.0 and 25°C, with a 250-ms mixing time for the NOESY experiment. The spin systems for Proα and Ileβ are traced out using large dashes and solid lines, respectively. Small dashed lines show the NOEs relating the α- and β-protons of Cys7 to α- and β-protons of Pro6.

**Data Interpretation**—An equilibrium situation with multiple conformations is typically encountered with short linear peptides in solution. However, the 6-kDa fragment of Fn, a dimer containing 52 residues, is roughly the size of bovine pancreatic trypsin inhibitor. The size of this fragment, the fact that it is a dimer, and the presence of two interchain disulfide bonds, are expected to lend some conformational rigidity to its structure in solution. Comparison of H-chemical shifts (Table I) of the Thrα to Argα segment with those of random coil structures (Bundi and Wuthrich, 1979) shows that several resonances within the Thrα-Metε segment were observed in the 125-ms NOESY spectrum in D2O (Fig. 3). A strong NOE was observed between the Pheε2 β-H and the Ileε α-H. More importantly, the ε- and γ-protons of Pheε2 showed definite NOEs to both β-methylene protons of Cys7 and Cys7'. Also, one of the β-methylene protons of Cys7 showed NOEs to both β-methylene protons of Cys7'. Chemical shifts of assigned resonances are shown in Table I. Relevant NOE data are summarized in Table II. Only a few minor NOE cross-peaks remain unassigned, some of which may be due to impurities in the sample.
NMR-based 3-D Structure of Fibronectin C-terminal Fragment

FIG. 3. NOESY spectrum of the 6-kDa fragment in D$_2$O. The COSY connectivity of the Phe$^{35}$ ring protons is shown in an insert at the bottom-left corner. The mixing time in the NOESY experiment was 125 ms. Other experimental conditions are as in Fig. 1. Long-range NOEs relating the $\beta$-protons of Cys$^7$ and Cys$^{11}$, and the Ile$^9$ $\alpha$-proton to the Phe$^{35}$ ring protons are shown by dashed lines. Solid lines are used to show NOEs relating the C$_{a}$- and C$_{b}$-protons of Cys$^7$, Cys$^{11}$, and Ile$^9$.

TABLE I

| Residue  | NH  | $C_H$ | $C'_{H}$ | Other  |
|----------|-----|-------|---------|--------|
| Thr$^1$  | 3.88| 4.15  |         | $\gamma CH_2$ 1.29 |
| Asn$^2$  | 4.86| 2.88  | 2.77    |        |
| Thr$^3$  | 8.30| 4.32  | 4.23    | $\gamma CH_2$ 1.17 |
| Asn$^4$  | 8.44| 4.73  | 2.86    | $\gamma NH_2$ 7.58, 6.94 |
| Val$^5$  | 9.05| 4.09  | 2.09    | $\gamma CH_2$ 0.91, 0.91 |
| Asn$^6$  | 8.49| 4.77  | 2.81    | $\gamma NH_2$ 7.61, 6.92 |
| Cys$^7$  | 8.27| 5.04  | 3.08    |        |
| Pro$^8$  | 4.53| 2.46  | 2.19    | $\gamma CH_2$ 2.08, 2.08; $\delta CH_2$ 3.84, 3.80 |
| Ile$^9$  | 8.46| 4.09  | 1.96    | $\gamma CH_2$ 1.53, 1.39; $\gamma CH_2$ 1.03; $\delta CH_2$ 0.96 |
| Gln$^{10}$ | 9.02| 4.16  | 2.19, 2.04 | $\gamma CH_2$ 2.39, 2.27 |
| Cys$^{11}$ | 8.14| 4.11  | 2.88, 2.48 | $\delta H$ 7.30; $\chi H$ 7.12; $\chi H$ 7.16 |
| Phe$^{12}$ | 7.74| 4.77  | 3.38, 2.96 | $\gamma CH_2$ 2.61, 2.61; $\delta CH_2$ 2.12 |
| Met$^{13}$ | 7.81| 4.81  | 2.11, 1.96 | $\gamma CH_2$ 1.96, 1.96; $\delta CH_2$ 3.80, 3.67 |
| Pro$^{14}$ | 8.33| 4.32  | 1.64, 1.59 | $\gamma H$ 1.69, $\delta CH_2$ 0.94, 0.88 |
| Leu$^{15}$ | 8.39| 4.61* | 2.79, 2.69 | $\gamma CH_2$ 0.94, 0.94 |
| Val$^{16}$ | 7.99| 4.11  | 2.11    | $\gamma CH_2$ 2.37, 2.37 |
| Glu$^{17}$ | 8.38| 4.32  | 2.11, 1.99 | $\gamma CH_2$ 2.37, 2.37 |
| Ala$^{18}$ | 8.24| 4.30  | 1.40    |        |
| Asp$^{19}$ | 8.32| 4.61* | 2.79, 2.70 | $\gamma CH_2$ 1.63, 1.63; $\delta CH_2$ 3.22, 3.22 |
| Arg$^{20}$ | 8.21| 4.37  | 1.90, 1.77 |        |
| Glu$^{21}$ | 8.45| 4.31  | 2.10*, 1.94 |        |
| Asp$^{22}$ | 8.39| 4.61* | 2.79*, 2.69* |        |
| Ser$^{23}$ | 8.58| 4.53  | 4.32, 4.23 |        |
| Arg$^{24}$ | 8.26| 4.39  | 1.92, 1.80 | $\gamma CH_2$ 1.63, 1.63; $\delta CH_2$ 3.22, 3.22 |

*The uncertainty (±0.02 ppm) in these chemical shift values is larger than the uncertainty (±0.01 ppm) in the rest.

connectivities (Table II, column 6) are observed only within this segment. For proline-containing peptides, the presence of multiple conformers is often indicated by the observation of two distinct sets of resonances corresponding to cis- and trans-prolines, since the rate of exchange between the species containing the two isomers is slow on the NMR time scale (Wüthrich, 1976; Larive et al., 1992). Resonances corresponding to only the trans conformers were observed for both Pro$^8$ and Pro$^{14}$ in the 6-kDa Fn fragment, confirmed by the presence of strong $\alpha_i-\delta$ NOE cross-peaks for both prolines. Based on these observations, it was assumed that a substantial population of the conformers of the 6-kDa fragment in solution had a preferred conformation for the Thr$^3$-Pro$^{14}$ segment, the structure being more flexible further away from the interchain disulfide bonds. Calculation of the three-dimensional structure, therefore, was attempted only for the Thr$^3$-Pro$^{14}$ segment. This was found to be more than adequate to demonstrate that the two interchain disulfide bonds linked the Fn monomers in an antiparallel fashion in the 6-kDa carboxyl-terminal dimer fragment.

The NMR information indicates a symmetrical structure: chemical shifts are identical for the same residue in both chains. This potentially complicates the interpretation of NOE data, since no distinction can be made between intrachain and interchain NOEs. Hence, a cross-peak relating residues $i$ and $i+1$ through space, might not necessarily indi-
Table II

Summary of NOE information for the C-terminal fragment

| Residue | d_{N,N} | d_{N,N} | d_{N,N} | d_{N,N} | Other NOEs |
|---------|---------|---------|---------|---------|------------|
| Thr^9   | +       | +       | +       | +       | d_{N,N} (3,4) |
| Asn^10  | +       | +       | +       | +       | d_{N,N} (6,6) |
| Val^11  | +       | +       | +       | +       | d_{N,N} (5,6) |
| Asn^8   | +       | +       | +       | +       | d_{N,N} (5,6) |
| Cys^7   | ++      | -       | -       | -       | d_{N,N} (7,8), d_{N,N} (7,8), d_{N,N} (7,11), d_{N,N} (7,11), d_{N,N} (7,12), d_{N,N} (7,12), d_{N,N} (7,12), d_{N,N} (7,12) |
| Pro^6   | -       | +       | +       | +       | d_{N,N} (9,11), d_{N,N} (9,12), d_{N,N} (9,12) |
| Ile^5   | +       | +       | +       | +       | d_{N,N} (10,12), d_{N,N} (10,13) |
| Glu^10  | ++      | +       | +       | +       | d_{N,N} (11,12), d_{N,N} (11,12), d_{N,N} (11,12), d_{N,N} (11,12), d_{N,N} (11,12), d_{N,N} (11,12) |
| Cys^11  | ++      | +       | +       | +       | d_{N,N} (17,18) |
| Phe^12  | ++      | ?       | ?       | ?       | |
| Met^13  | ++      | -       | -       | -       | d_{N,N} (13,14), d_{N,N} (13,14) |
| Pro^14  | -       | +       | +       | +       | |
| Leu^15  | ++      | +       | +       | +       | d_{N,N} (15,16), d_{N,N} (15,17) |
| Asp^16  | ++      | +       | +       | +       | |
| Val^17  | +       | +       | +       | +       | |
| Gin^18  | ++      | ?       | ?       | ?       | |
| Ala^19  | +       | +       | +       | +       | |
| Asp^20  | ++      | +       | +       | +       | |
| Arg^21  | ++      | +       | +       | +       | |
| Glu^22  | ++      | +       | +       | +       | |
| Asp^23  | ++      | ?       | ?       | ?       | |
| Ser^24  | ++      | ?       | ?       | ?       | |
| Arg^26  | ++      | ?       | ?       | ?       | |

| | Intra-residual NOE: d_{N,N} (i,i) + denotes NOE to the single β proton for these residues; ++ denotes NOE to both β protons. 
| | Sequential connectivities: d(i,i+1). And blank spaces, respectively, indicate the presence and absence of NOEs. – denotes the absence of a proton in the residue (e.g. NH in Pro), such that the particular NOE cannot be present. "+" indicates the absence of NOE information (e.g. due to bleaching of cross-peaks caused by water suppression using presaturation). In such cases, both possibilities (presence and absence of the corresponding d connectivity) were considered, and two regions in ϕ,ψ space were predicted by FISINOE (see Table 3). |
| | d_{N,N} (i,j) are not repeated again as d_{N,N} (i,j). While these NOEs are represented here as intrachain only (i.e. involving residues (i,j)), the data does not exclude interchain NOEs (i.e. those involving residues (i',j')); both possibilities are considered in the analysis (see text). Subscripts (δ,β) and (δ',β') have been used to represent the two Cγ and Cβ protons (column 6) differentiating them in terms of their chemical shifts only (see Table 1), and do not represent stereospecific assignments, which can be obtained from the χ1-χ2 values in Table III. |

\[ \delta_{\text{N,N}} (i,i) + \] denotes NOE to the single β proton for these residues; ++ denotes NOE to both β protons. 

Sequential connectivities: d(i,i+1). And blank spaces, respectively, indicate the presence and absence of NOEs. – denotes the absence of a proton in the residue (e.g. NH in Pro), such that the particular NOE cannot be present. "+" indicates the absence of NOE information (e.g. due to bleaching of cross-peaks caused by water suppression using presaturation). In such cases, both possibilities (presence and absence of the corresponding d connectivity) were considered, and two regions in ϕ,ψ space were predicted by FISINOE (see Table 3). 

d_{N,N} (i,j) are not repeated again as d_{N,N} (i,j). While these NOEs are represented here as intrachain only (i.e. involving residues (i,j)), the data does not exclude interchain NOEs (i.e. those involving residues (i',j')); both possibilities are considered in the analysis (see text). Subscripts (δ,β) and (δ',β') have been used to represent the two Cγ and Cβ protons (column 6) differentiating them in terms of their chemical shifts only (see Table 1), and do not represent stereospecific assignments, which can be obtained from the χ1-χ2 values in Table III.

cibrate a sequential connectivity, and could, in principle, represent a long-range NOE cross-peak relating residues i on chain 1 and (i+1)' on chain 2 of the dimer. However, a statistical analysis of short proton-proton distances in a collection of data from high resolution protein crystal structures (Billeter et al., 1982) shows that 88% of the αH-NH distances ≤ 3.0 Å represent sequential connectivities. The corresponding probabilities for NH-NH and βH-NH are 88 and 76%, respectively. The probabilities are even higher when such short interproton distance limits are imposed simultaneously: the probability that the connectivities are sequential when (d_{N,N} ≤ 3.6 Å and d_{N,N} ≤ 3.0 Å) is 99%; it is 95% when (d_{N,N} ≤ 3.6 Å and d_{N,N} ≤ 3.4 Å) and 90% when (d_{N,N} ≤ 3.0 Å and d_{N,N} ≤ 3.0 Å). In other words, protein folding patterns in nature rarely lead to short-proton-proton distances relating main chain and Cγ protons of residues that are not immediate neighbors on a polypeptide chain. This conclusion was applied to the NOE data from the 6-kDa fragment, although the uniqueness of the interchain disulfide bonds makes it an unusual case. A conservative estimate of the upper limit for the observed d_{N,N}, d_{N,N}, d_{N,N} connectivities is 3.3 Å, based on the estimation of distances from the measured volume integrals for a sampling of αH-NH, NH-NH, and βH-NH NOE cross-peaks (using a proportionality constant derived from the volume integrals of the cross-peaks relating δH and δH of Phe^22 across a distance of 2.5 Å). Hence, the simplest interpretation of the NOE data in Table II, is that all NOEs showing connectivities between main chain protons, and between main chain and Cγ protons (i.e. d_{N,N}, d_{N,N}, d_{N,N}) represent intrachain sequential connectivities. The statistical analysis noted above indicates that this should constitute a fairly accurate representation of the three-dimensional structure. This interpretation was used initially in both the deterministic and the probabilistic analyses. Because of the observed chemical shift symmetry, the set of sequential NOE connectivities were assumed to apply equally to both chains in the dimer. All d_{N,N} connectivities were assumed to be intraresidual, and hence intrachain only.

The ambiguity concerning intra- versus interchain NOEs arises solely due to the fact that the dimer structure is symmetric; otherwise residues n and n' (i.e. the same residue on the two chains of the dimer) would be distinguishable by their different chemical shifts. If only intrachain symmetry is assumed then the assumption that all "sequential" NOEs are intrachain is valid. On the other hand, if even a few of these sequential NOEs are interchain, then the identical chemical shift requirement imposes interchain symmetry in addition to intrachain symmetry. In other words, if the d_{N,N} connectivities were assumed between residues n and n+1 (and between n' and n+1'), the chemical shift symmetry requires that there must be a d_{N,N} connectivity between n' and n+1. This additional interchain symmetry would place more stringent constraints on the structure. Also, from consideration of steric hindrance alone, it appears unlikely that the interchain symmetry requirement could be satisfied in any arrangement of the dimer structure. This is supported by the results of a set of calculations using conformational analysis alone (described under "Conformational Analysis"), without any consideration of the NOE data.

For the deterministic approach, both intra- and interchain distance constraints were initially included for all NOE connectivities involving side chain protons. Later, during the process of minimization of the penalty function for distance constraints, those particular constraints that contributed to large violations were gradually removed in an attempt to minimize the distance violations.

For the probabilistic approach, the sequential NOEs were the only NOEs used. Since distance constraints were not used explicitly in this method, building the spatial structure did not require that any prior decisions be made regarding whether the other observed NOEs (involving side chain protons) were intra- or interchain NOEs. All unconstrained energy-refined structures were carefully examined to check for consistency with experimentally observed NOEs. During this examination, both intra- and interchain distances were checked to ensure that the selected structures were consistent with at least the minimum number of experimentally observed NOEs involving side chain protons in the particular dimer fragment in question.

Structure Determination Assuming All Sequential Connectivities to be Intrachain—The set of 55 approximate interproton distance constraints per chain, derived from the observed NOE data, was used both with, and without the additional symmetry constraint, to calculate three-dimensional structures of the Thr^9 to Pro^14 segment, with the specific purpose of determining whether the NMR data indicated a parallel or an antiparallel arrangement of the two chains linked by the two interchain disulfide bonds.

The Deterministic Approach—The more commonly used DG approach, followed by energy refinement, was used first.
Several DG, DGREM, and DGRMD structures of the Thr-
Met segment were obtained, all of which roughly satisfied
the NOE constraints within this dimer fragment (with small
total distance violations). The structures did not fall into any
closely related sets, and showed large variations in the main
chain conformation when compared in pairs. Addition of
symmetry as a constraint in the energy refinement also gave
numerous possible solutions (with identical monomer confor-
mations in the dimer), again with a wide variation in the main
chain conformation, as described above. This approach
showed that “constrained” structures, roughly satisfying all
NOE requirements, were possible for both the parallel and
the antiparallel dimer forms. However, all of the structures
calculated in this way contained several dihedral angles well
outside the sterically allowed regions. Therefore, comparison
of the calculated conformational energies was not an ade-
cquate criterion, either for selecting a set of preferred structures from
the many converged constrained structures, or for deciding
whether the NMR data indicated a parallel or an antiparallel
arrangement of the two chains in the dimer. We concluded that
the number of constraints available was not sufficient for
this strategy to work.

The Probabilistic Approach—We then applied the BUILD
procedure described above, using the d connectivity shown in
Table 2 (dASN, dDON, dPE) to estimate corresponding regions in
φ,ψ space for each residue, shown in Table III (column 2).
The most probable φ,ψ values corresponding to each region
(Table III, column 3) were used as the starting set of φ and ψ
angles for the 12-residue monomer segment, Thr1-Pro14. The
d connectivity data for three of these 12 residues, (Asn3, Asn6, and Phe13), are ambiguous, permitting two possible regions in φ,ψ space (see Table III, column 2). However, the two possible regions are close to each other in conformational space, and an average of the most probable φ,ψ values corre-
sponding to the two regions was used as the starting dihedrals in these cases. Extension of this method, and consideration of the intraresidue NOE data between amide- and Cp-protons was used to obtain the possible combinations of x1 and x2 angles for the side chain conformations (Sherman and John-
son, 1991; Sherman and Johnson, 1993). A four-step "build-
up" procedure was followed to construct the final three-
dimensional structure: (i) The Cys7-Cys" segment, containing
the interchain disulfide bonds was constructed first, because of
the four long-range NOE restrictions between the Cp-
protons of Cys7 and Cys" (see Table II) present in this
segment. The Ile6 and Glu10 side chains were initially trun-
cated to alanine, making the testing sequence Cys-Pro-Ala-
 Ala-Cys. (ii) Phe12 was added to this sequence, and, in two
subsequent steps, alanines at positions 9 and 10 were replaced
by Ile6 and Glu10. (iii) The Thr1-Asn4 segment was then added. As in (i) and (ii), calculations were first performed with Asn4 and Asn4 replaced by alanine. (iv) Met13 and Pro14 were finally added to complete the segment. This procedure
reduced the total number of calculations required from 4096 × 2 (for parallel and antiparallel structures) to only 90. Pro8 and Pro14 were modeled to be in the trans configuration, as indicated by the presence of strong α,1-β,1 NOE cross-peaks (Table II).

The dihedral angles estimated by the FISINOE program
were used as the starting angular coordinates. The Cys7-Cys11 segment was built using the interactive graphics package, INSIGHT. CHARMm was then used to "patch" two such segments in either a parallel or an antiparallel fashion, via the two interchain disulfide bonds. CHARMm was also used for energy minimization, with gradual reduction, and finally elimination, of force constants for dihedral constants, as described under "Experimental Procedures."

Arrangement of Monomer Chains in the Dimer—The ques-
tion of parallel versus antiparallel arrangement of the mono-
mers in the dimer fragment was answered at the first phase
of the calculations in step (i), using the Cys7-Cys11 segment
only. A summary of the results for this step of the calculations,
using both parallel and antiparallel arrangements of the mon-
omer segment Cys-Pro-Ala-Ala-Cys to form the dimer, is
given in Table IV. Of the four possible conformations corre-
sponding to different combinations of x1 values for the pair of
cystines in the monomer, none satisfied all four long-range
NOE requirements between the Cys-protons of Cys7 and Cys11
(considering both intra- and interchain connectivities) for a
parallel dimer structure. Two of the parallel structures (see
rows i and ii in Table IV) failed to satisfy all sequential NOE
requirements, as well. This is apparent in the large deviations (Table IV) in backbone conformation from that estimated
using sequential NOE data in FISINOE. The use of distance
constraints, in addition to dihedral constraints, also did not
lead to any unconstrained parallel dimer structures consistent
with all NOE data. In the antiparallel dimer structure, it was
possible to select one of the four combinations of x1 conform-
ers for the pair of cystines in the monomer (with x1 = ϕ =
−60° for both Cys7 and Cys11; see row 1 in Table IV), since

| Table III |
| Comparison of re-estimated and final conformations |
| Residue | Region | φ,ψ | x₁,x₂ | φ,ψ | x₁,x₂ |
|---------|--------|-----|-------|-----|-------|
| Thr1    | S      | −60,−55 | g   | −68,−44 | g   |
| Asn4    | R      | −65,−35 | t   | −54,−50 | t   |
| Val6    | K      | −75,−15 | t   | −65,−41 | 180 |
| Asn6    | Q      | −85,−15 | t   | −60,−40 | −60,173 |
| Cys7    | T      | −85,20  | g   | −47,−50 | −66,−68 |
| Pro8    | A      | −50,160 | t   | −73,166 |
| Ile6    | T      | −85,20  | t   | −54,−41 | −178,60 |
| Glu10   | K      | −75,−15 | t   | −66,158 | −63,−178 |
| Cys11   | R      | −65,−35 | t   | −57,−39 | −62 |
| Phe12   | Q      | −85,−15 | t   | −57,−42 | −64,70 |
| Met13   | B      | −100,135| t   | −66,158 | −63,−178 |
| Pro14   | B      | −50,145 | t   | −73,171 |

* Region denotes areas in φ,ψ space that correspond to particular combinations of sequential d connectivities and are represented by a set of most probable φ,ψ values (Sherman et al., 1987). For Asn4, Asn6, and Phe12 the NOE data indicates two regions in φ,ψ space. Both regions are shown in column 2 for each of these residues. Column 3, therefore, contains two sets of the most probable φ,ψ values (in degrees) corresponding to these two regions.

* Initial conformations. The initial x1 and x2 values are indicated as rotamers g, t, and β, representing 60, −60, 180, and 90°, respectively. The IUPAC-IUB conventions (Hoffmann-Ostenhof, 1974) were followed in naming the torsion angles. (For valine, it should be noted that these conventions are different from those used in CHARMm and INSIGHT.)

* Final conformation. Column 5 shows the rotamers describing the various side chain conformations, including the energetically indistinguishable set of conformations for the side chains of Ile6, Glu10, and Met13 that lead to the set of high energy refined unconstrained structures consistent with the NMR data. Values of φ,ψ and x₁,x₂ (in degrees) shown in columns 6 and 7 are for one of the 16 conformations (see text for angular root mean square deviations for pairs of these 16 structures).
Summary of results for the Cys-Pro-Ala-Ala-Cys sequence in step (i)

| Cys$^a$ Cys$^{11}$ | Conformational energy$^a$ | Maximum deviation$^a$ in $\psi$ | Number of long-range NOEs satisfied$^a$ |
|-----------------|----------------------|----------------------|----------------------|
| x$_1$ | x$_1$ | kcal/mol | degrees |
| 1 | $g^-$ | $g^-$ | $-232.3$ | $-48.5$ |
| 2 | $g^-$ | $g^-$ | $-290.2$ | $-45.9$ |
| 3 | $g^-$ | $g^-$ | $-218.4$ | $-42.3$ |
| 4 | $g$ | $t$ | $-214.2$ | $-44.7$ |
| 5 | $g^-$ | $g^-$ | $-235.8$ | $-50.0$ |
| 6 | $g$ | $t$ | $-215.3$ | $-48.2$ |
| 7 | $g$ | $t$ | $-195.2$ | $-33.4$ |

* Energies calculated using values of 1, 4, and 10 for the dielectric constant, $\epsilon$.

**Initial backbone conformations estimated using FISINOE are compared with conformations in the unconstrained, energy refined structures. The angular root mean square deviations (armed) for $\psi$ only are shown, since armed values for $\phi$ are similar for all eight structures (about 23°).

The maximum deviation in $\psi$ from estimated values has been included to show that some low energy conformations (see rows i and ii in B) correspond to backbone dihedrals that have deviated more than 60° from the initial estimates consistent with sequential NOEs (or more than 3σ, where the armed, $\sigma = 20°$, for FISINOE estimates of $(\phi,\psi)$ in each region). This, in turn, implies that the corresponding unconstrained structures do not satisfy all of the sequential NOEs.

Addition of distance constraints to satisfy long-range NOEs resulted in large deviations of backbone dihedrals from FISINOE estimated values, implying violation of sequential NOE requirements. When both long-range and sequential NOEs were included as distance constraints, the resulting structures contained backbone dihedrals in sterically forbidden regions.

The single structure satisfying all sequential and long-range NOE requirements.

Only this conformation satisfied all sequential and long-range NOE requirements. One interesting difference between the parallel and antiparallel structures was that the symmetry requirement was satisfied, even in the absence of any explicit symmetry constraints, in all four of the unconstrained antiparallel structures, but in none of the parallel structures. Since the probabilistic approach uses a priori information regarding empirical distribution functions for backbone conformations in addition to the NMR data, the available NOE data were sufficient for this method to distinguish between parallel and antiparallel structures.

Conformational Analysis—Since the results above have been obtained assuming all $d_{XX}$, $d_{XXN}$, $d_{NN}$, and $d_{NN}$ connectivities to be intrachain for the Cys-Pro-Ala-Ala-Cys sequence, conformational analysis was performed for this sequence to check whether other conformations (not predicted by the assumed intrachain sequential $d$ connectivity patterns) were energetically favorable for this dimer fragment containing intrachain disulfide bonds. For simplicity, only two main classes of backbone conformations were considered for each residue in creating the starting structures: (i) the twisted or $\alpha$ conformation, with $(\phi,\psi) = (-60°, -60°)$; and (ii) the extended or $\beta$ conformation, with $(\phi,\psi) = (-135°, 135°)$. Since the actual sequence has a charged residue in position 10 (Glu$^{10}$), the (60, 60) conformation (or $\alpha$), on the right side of the $(\phi,\psi)$ map, was also considered for Ala$^{10}$. The structure obtained above, by assuming all sequential $d$ connectivities to be intrachain, may be represented as the $\beta\beta\alpha\alpha\alpha$ conformation under this assumption, with $x_3$ angles for both Cys$^a$ and Cys$^{11}$ close to $g^-$ (see Table IV, row 1). In accordance with the requirement of chemical shift symmetry, identical starting conformations were used for the two monomer chains. Considering that the $x_3$ angle for each of the cystines per monomer (Cys$^a$ and Cys$^{11}$) may take three values ($g^-$, $-60°$, and $60°$), this leads to a total of 108 $\times$ 2 structures (for parallel and antiparallel forms).

When a structure with a comparable energy was obtained, the conformation of each residue was examined to see whether the sequential $d$ connectivities (previously assumed to be intrachain) were not present. Interproton distance calculations were then performed, using the coordinates of the energy-refined unconstrained structure, to check whether the corresponding interchain connectivities were present instead, in order to satisfy the NOE requirements. A summary of the results are given in Table V. Eight structures with conformational energies comparable to the $\beta\beta\alpha\alpha\alpha$ conformation (see Table V, row 1) were found. In all of these structures, the monomers were linked in the antiparallel fashion to form the dimer, and all retained conformational symmetry in the unconstrained structures (i.e. contained identical conformations for the two monomers). However, interchain $d$ connectivities (in place of missing intrachain sequential $d$ connectivities) were not found in any of these structures, so that none of them satisfied all NOE requirements. (It is important to note here that in a few of these structures, the long range NOEs relating Cys$^a$ and Cys$^{11}$ could be satisfied by intrachain, rather than interchain connectivities. Therefore, whether the structure is parallel or antiparallel, it cannot simply be assumed that NOEs relating Cys$^a$ and Cys$^{11}$ must be interchain.) Conformational energies for all parallel dimer structures were much higher than that obtained previously, assuming all sequential connectivities to be intrachain (see Table V). Interchain $d$ connectivities (in place of missing intrachain sequential $d$ connectivities) were not found in any

TABLE IV

| Structure | Cys$^a$ | Cys$^{11}$ | Conformational energy | Maximum deviation in $\psi$ | Number of long-range NOEs satisfied |
|-----------|--------|--------|-----------------|----------------------|----------------------|
| 1$^i$ | $g^-$ | $g^-$ | $-64.5$ | $-18.9$ |
| 2$^i$ | $g^-$ | $g^-$ | $-64.3$ | $-19.2$ |
| 3$^i$ | $g^-$ | $g^-$ | $-74.4$ | $-19.6$ |
| 4$^i$ | $g^-$ | $g^-$ | $-40.3$ | $-19.8$ |
| 5$^i$ | $g^-$ | $g^-$ | $-54.7$ | $-24.2$ |
| 6$^i$ | $g^-$ | $g^-$ | $-55.8$ | $-20.6$ |
| 7$^i$ | $g^-$ | $g^-$ | $-49.5$ | $-18.1$ |
| 8$^i$ | $g^-$ | $g^-$ | $-46.8$ | $-17.7$ |
| 9$^i$ | $g$ | $t$ | $-51.1$ | $-19.8$ |

* Energies calculated using values of 1, 4, and 10 for the dielectric constant, $\epsilon$.

**Initial backbone conformations estimated using FISINOE for the CPAC backbone are compared with conformations in the unconstrained, energy refined structures (where the initial values of $(\phi,\psi)$ for twisted $(\alpha)$ and extended $(\beta)$ conformations are assumed to be $(-60°, -60°)$ and $(-60°, 135°)$). The angular root mean square deviations (armed) and maximum deviation from the initial values of $(\phi,\psi)$ (columns 7 and 8) have been evaluated in order to estimate the "goodness of fit" of the unconstrained structures to the initial conformations.

The "reference" structure, i.e. the structure obtained assuming all sequential $d$ connectivities to be intrachain. Note that the nine conformations shown in this table only this conformation satisfies all NOE requirements.
of these structures; nor were all the long-range NOE requirements satisfied. Also, none of these parallel dimer forms contained conformations similar to those for the two monomers (i.e., conformations for the two monomers were not identical in the unconstrained structure), thus violating the identical chemical shift requirement. Therefore, use of the chemical shift symmetry constraint, with the help of conformational analysis, was sufficient to demonstrate that the two monomers in the carboxyl-terminal dimer fragment of Fn must be linked in an antiparallel fashion by the interchain disulfide bonds.

At this juncture, it is important to note that the observation of NOE contact between Cys-7 and Cys-11 did not, of itself, rule out a parallel structure; we had to do extensive additional conformational analysis to show that only an antiparallel structure satisfied all of the NOE constraints. The analysis by An et al. (1992) noted the existence of a ROESY cross-peak between the Cys-15 H-δ and the Cys-15 H-γ (our numbering) and stated that “a close distance between the 2 Cys residues unambiguously indicates an antiparallel” arrangement. However, the chemical shift symmetry makes it impossible to distinguish whether this ROESY cross-peak is intrachain or interchain. Thus, the ROESY data alone, without any modeling calculations, could not rule out the parallel structure in which the 2 cystines in the monomer are spatially close, but the disulfide bridge links Cys-15-Cys-15 and Cys-7-Cys-11, since ROESY cross-peaks between residues 7-7′ and 11-11′ cannot be observed due to the chemical shift symmetry. In fact, as we have shown above, it is quite possible to obtain a conformation in which Cys-7 and Cys-11 are close enough within the monomer to produce NOE contact. It is only the additional modeling that demonstrates that this monomer conformation cannot then be coupled with a second monomer to form a symmetric dimer that satisfies all NOE constraints. Thus, the analysis of An et al. (1992) assumed an antiparallel arrangement, but it did not prove it.

**Structure Consistent with NMR Data**—At least nine different antiparallel dimer structures were found with comparable conformational energies for the Cys-Pro-Ala-Ala-Cys sequence, with four different main chain conformations (ββαaaα, βαααα, ββββ, βαβα). However, as described in the above discussion on conformational analysis, only one of these (ββαaaα) in Table V, row 1 satisfied all NOE data. The structure obtained assuming all sequential d connectivities to be intrachain. Therefore, even if the Cys-Pro-Ala-Ala-Cys segment did exhibit multiple conformations in solution, it appears likely that a substantial population exists in the ββαaaα conformation. Only this structure was extended to build the three-dimensional structure of the Thr-Pro disulfide segment. Intraresidual and sequential d connectivities shown in Table II were assumed to be intrachain in all subsequent calculations. Only the antiparallel arrangement of the two chains was used, with x1 = g′ for both Cys−7 and Cys−11.

On addition of Phe-12 to the Cys-7-Cys-11 segment, only one of two possible conformations, with x1 = −60° (or g′) and x2 = 90° (or p), for the Phe-12 side chain was found to satisfy the long-range NOEs between the C6-protons of Cys′ and the ring protons of Phe′ (taking into account both intra- and interchain connectivities). Since there were no strong long-range NOE constraints relating the Cys′-Phe′ segment to the rest of the structure, all subsequent calculations in the build-up procedure used the following energy criteria to choose the set of most probable conformations at each step: for a set of x1 or x2 rotamers, all conformations with energy greater than the lowest energy conformation by 4 kcal/mol (calculated using a dielectric constant of 10), were eliminated. On replacing alanine by Ile′ in the Cys′-Pro-Ala-Ala-Cys′-Phe segment, four energetically equivalent conformations were obtained. Replacing the second alanine by Glu′ led to 16 conformations (4 x 4), eight of which were eliminated by energy criteria. Similarly, extending the sequence by the Thr′-Asn-4 segment led to 32 possible conformations for the Thr′-Phe′ segment (8 x 2 x 2), 24 of which were eliminated by energy criteria. Addition of Met′-Pro′ resulted in 32 possible conformations (8 x 4 x 1), 16 of which were selected, using energy criteria.

To be the final set of structures consistent with the NMR data. Although symmetry was not used explicitly as a constraint in the energy minimization process, the symmetry was built into the backbone conformation at the start was largely retained in the final set of energy-refined unconstrained structures. Similarly, long-range NOE requirements were also found to be satisfied in the final 16 structures, and were obtained as a by-product of energy refinement, without the use of distance constraints in the minimization process.

**Estimation of Precision**—The backbone conformations of the set of 16 final structures were very similar, and were indistinguishable by x2 statistical criteria. The 16 structures consisted of combinations of side chain conformations for Ile′ (t, t′, g, g′, g″, g‴), Glu′ (g′, t, g″, g‴) and Met′ (g′, t, g″, g‴) that are consistent with the NMR data, and are indistinguishable by energy criteria. The rotamers representing these 16 side chain conformations are shown in Table III, column 5. The angular root mean square deviation (rmsd) in backbone conformations for pairs of structures within this set of 16 is 6° for ϕ and 10° for χ (average of rmsd values for all pairs within the 16 structures). The rmsd for the estimated values of ϕ and χ in each region is about 2° (Sherman and Johnson, 1992). Table III also compares the dihedral angles estimated by the FISINOE program on the basis of the available NOE data, with the dihedral angles in the 16 representative final structures. The backbone dihedral angles are consistent with the NMR data, and are indistinguishable, using x2 criteria.

Unlike the types I, II, and III repeat units of Fn, which contain dominant structural features that are common to many proteins, the structure of the carboxyl-terminal dimer segment reported here is somewhat unusual. Fig. 4 shows the conformation of the Thr-Pro′ segment in the dimer form, with a single set of conformations for the side chains of Ile′, Glu′, and Met′ (as shown in Table III). The structure consisting of the Thr-Pro′ segment, Thr′-Asn-4, and Ile′-Phe′, connected by an extended region at Cys′-Pro′. These twisted regions may serve as recognition sites for the monomers, and thus help to bring the 2 cystines in each monomer into close proximity, specifically in the antiparallel orientation, so that the required interchain disulfide bridges are formed in the correct manner in the dimer. The two ends of the Thr-Pro′ segment are connected by interchain H-bonds involving backbone atoms (NH of Asn′ of one chain to O of Pro′ of the second chain). In addition to several inter- and intrachain H-bonds involving side chains, there are 4 pairs of intrachain H-bonds involving backbone atoms in each monomer. The donor-acceptor pairs are: (7, 3), (11, 8), (12, 9), and (13, 10). The exposed surface of the two twisted regions contains several hydrophobic side chains (Val5, Ile′, and Phe′) followed by Met′ and Pro′. However, these may be covered by the rest of the monomer (Leu′-Glu′) folding back over itself. The aromatic ring of Phe′ in each monomer lies across the two disulfide bridges, and the hydrophobic interactions involving these aromatic side chains may help stabilize the disulfide bonds. The backbone structure of the Cys′-Cys′ segment forms a loop that brings the 2 cystines in the same chain (Cys′ and Cys′) close enough to make intrachain disulfide bonds possible. Why interchain and not intra-
binding domains, such as the gelatin and cell-binding domains of the antiparallel arrangement of the interchain disulfide bond in the monomer, and that this hindrance is removed through stronger interchain hydrophobic interactions. In all 16 final structures, the long-range NOEs relating the Cp-protons of Cys7 and Cys11 were found to be interchain, while NOEs relating the Cp-protons of Cys7 and ring protons of Phe19 were found to be intrachain.

A thorough understanding of the functions of Fn requires knowledge of the precise spatial arrangements of various binding domains and their interactions in the two similar subunits of the dimeric Fn molecule. Our NMR demonstration of the antiparallel arrangement of the interchain disulfide bridge near the carboxyl termini of Fn suggests that similar binding domains, such as the gelatin and cell-binding domains in different subunits, may be arranged in a diagonal manner rather than in a mirror image. This is consistent with previous work by Skorstengaard et al. (1986), which also suggested an antiparallel arrangement for the interchain disulfide bridge of Fn, based on HPLC patterns of peptides derived from the carboxyl-terminal 6-kDa fragment. The conformation of the Val6-Cys11 region differs substantially from either of the average conformations proposed by An et al. (1992). Several of the dihedral angles in the work of An et al. (1992) fall outside the usual sterically allowed regions. Since only the average conformations were reported by An et al. (1992), we could not compare individual structures. There are significant differences between the observed NOE contacts for our work and those reported by An et al. (1992), thus the structural differences reflect real differences in solution conformations, and not simply different approaches to determining the structure. A detailed comparison suggests that truncation of the peptide chains at Val6 and Cys11 (our numbering, corresponding to their V1 and C7) permits substantially increased freedom for the disulfide-linked cyclic peptide ring due to loss of constraints from the extended peptide chain, and probably does not maintain that fragment within its native environment and conformation. Thus, we propose that the structure determined here from a larger fragment of the carboxyl-terminal region is more likely to represent the native conformation of the disulfide-linked segments of the carboxyl-terminal region of the two subunits.

Fn is an essential component of the extracellular matrix that controls cell growth, cell shape, and differentiation (Hynes, 1990; Mosher, 1988). In vitro, Fn molecules self-assemble into fibrils reminiscent of the fibrillar structures seen in the matrix (Vuontanto et al., 1980). Although the detailed structure of the Fn fibrils is not known, the current model for the assembly of the Fn fibrils, proposed by Hormann (1982), was based on a parallel interchain disulfide bridge pattern. According to this model, the fully extended form of Fn assembles into a half-staggered array to form 5-nm fibrils. Our determination of an antiparallel arrangement for the two Fn chains suggests that the fibril formation process may be far more intricate than presently perceived. It is of some interest to note that, relative to each other, the two chains extend away from the bridge region in an essentially antiparallel fashion, in sharp contrast to the "folded," parallel arrangement suggested by An et al. (1992). Thus, one might speculate that the two chains may extend away from each other in fibril formation, although more extended structural information would clearly be required to definitively determine this arrangement.

In conclusion, a three-dimensional structure has been determined for 24 residues of the dimer segment that contains the interchain disulfide bonds within the 6-kDa carboxy-terminal fragment of human plasma fibronectin. A build-up
strategy for obtaining spatial structures was described, using the FISINOE method for evaluating local conformations from sequential d connectivities. The 2D NMR data are consistent with only an antiparallel arrangement of the two monomers connected via the two interchain disulfide bridges.

Acknowledgments—We thank Dr. M. Prabhakaran for help and initial guidance in the use of the INSIGHT and CHARMM software packages, and for some initial calculations using DSPACE and CHARMM.

REFERENCES

Akiyama, S. K., and Yamada, K. M. (1987) Adv. Enzymol. Related Areas Mol. Biol. 56, 1-57
An, S. A. A., Jiménez-Barbero, J., Petersen, T. E., and Llinás, M. (1992) Biochemistry 31, 9927-9933
Bacon, M., Norman, D., Willis, A., and Campbell, D. (1990) Nature 346, 642-646
Baron, M., Main, A. L., Driscoll, P. C., Mardon, H. J., Boyd, J., and Campbell, I. D. (1992) Biochemistry 31, 2068-2073
Billette, M., Braun, W., and Wüthrich, K. (1982) J. Mol. Biol. 155, 321-346
Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983) J. Comput. Chem. 4, 187-217
Bundi, A., and Wüthrich, K. (1979) Biopolymers 18, 285-297
Constantine, K. L., Brew, S. A., Ingham, K. C., and Llinás, M. (1992) Biochem. J. 283, 247-254
Engvall, E., and Ruoslahti, E. (1977) Int. J. Cancer 20, 1-5
Garcia-Pardo, A., Pearse, R., and Frangione, B. (1984) Biochem. Biophys. Res. Commun. 120, 1015-1021
Hoffmann-Ostenhof, O. (1974) Pure Appl. Chem. 40, 291-308
Holland, S. K., Hartwig, K., and Blake, C. F. (1987) EMBO J. 6, 1875-1880
Hormann, H. (1982) Klin. Wochenschr. 60, 1285-1277
Hynes, R. O. (1990) Fibronectins, Springer-Verlag, New York
Jeener, J., Meier, B. H., Bachman, P., and Ernst, R. R. (1979) J. Chem. Phys. 71, 4546-4553
Kornblitt, A. R., Umezawa, K., Vibe-Pedersen, K., and Baralle, F. E. (1985) EMBO J. 4, 1765-1769
Larive, C. K., Guerra, L., and Rabenstein, D. L. (1992) J. Am. Chem. Soc. 114, 7331-7337
Lesby, D. J., Hendrickson, W. A., Aukhil, I., and Erickson, H. P. (1992) Science 256, 287-291
Mosher, D. F. (ed) (1969) Fibronectin, Academic Press, New York
Nordal, W., Hare, D. R., and Reid, B. R. (1988) J. Mol. Biol. 201, 717-739
Pathy, L. (1990) Semin. Thromb. Hemost. 16, 245-259
Petersen, T. E., Thogersen, H. C., Skorstengaard, K., Vibe-Pedersen, K., Sahl, P., Søttrop-Jensen, L., and Magnusson, S. (1983) Proc. Natl. Acad. Sci. U. S. A. 80, 137-141
Sherman, S. A., and Johnson, M. E. (1991) in Proteins: Structure, Dynamics and Design (Remington, K. et al., eds) pp. 62-67, Escom Sci., Leiden, The Netherlands
Sherman, S. A., and Johnson, M. E. (1992) J. Magn. Reson. 96, 457-472
Sherman, S. A., and Johnson, M. E. (1993) Prog. Biophys. Mol. Biol. 69, 285-339
Sherman, S. A., Andrianov, A. M., and Akhrem, A. A. (1987) J. Biol. Chem. 262, 589-594
Sherman, S. A., Andrianov, A. M., and Akhrem, A. A. (1988) J. Biol. Chem. 263, 785-801
Skorstengaard, K., Jensen, M. S., Sahl, P., Petersen, T. E., and Magnusson, S. (1986) Eur J. Biochem. 161, 441-453
States, D. J., Haberkorn, R. A., and Roben, D. J. (1982) J. Magn. Reson. 48, 286-292
Vuento, M., Salonen, E., Koskimies, A., and Stemman, U. H. (1980) Hoppe-Seyler's Z. Physiol. Chem. 361, 1453-1456
Wagner, G. (1983) Quart. Rev. Biophys. 16, 1-57
Wider, G., Macura, S., Anil-Kumar, Ernst, R. R., and Wüthrich, K (1984) J. Magn. Reson. 59, 237-234
Wüthrich, K. (1989) Science 243, 45-50
Wüthrich, K. (1976) NMR in Biological Research: Peptides and Proteins, North Holland, Amsterdam.