Transforming Growth Factor β1: Secondary Structure As Determined by Heteronuclear Magnetic Resonance Spectroscopy†

Sharon J. Archer,‡ Ad Bax,§ Anita B. Roberts,∥ Michael B. Sporn,∥ Yasushi Ogawa,⊥ Karl A. Piez,# James A. Weatherbee,▲ Monica L.-S. Tsang,▲ Roger Lucas,▲ Bo-Ling Zheng,▲ Jeff Wenker,▲ and Dennis A. Torchia*,‡

Bone Research Branch, National Institute of Dental Research, Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, and Laboratory of Chemoprevention, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892, Celtrix Pharmaceuticals, Inc., Santa Clara, California 95054, Department of Biochemistry and Molecular Biology, Jefferson Medical College, Philadelphia, Pennsylvania 19107, and R&D Systems Inc., Minneapolis, Minnesota 55413

Received September 21, 1992; Revised Manuscript Received November 16, 1992

ABSTRACT: Virtually complete backbone NMR signal assignments have been reported for transforming growth factor β1 (TGF-β1) [Archer et al. (1993) *Biochemistry* (preceding paper in this issue)]. Herein we report the secondary structure of the protein in solution on the basis of these assignments and proton NOE's observed in a variety of 2D and 3D heteronuclear NMR spectra. Regular elements of secondary structure derived from the NOE data consist of (a) three helices spanning residues Y58–H68, F24–G29, and N5–F8 and (b) several pairs of two-stranded antiparallel β-sheets. The longest two-stranded sheet runs from residue L83 to V106 with a type II reverse turn at G93–R94 and a chain twist at residue N103–M104. These elements of regular structure were confirmed by hydrogen exchange, chemical shift, and coupling constant data. With the exception of residues G46–S53, which exhibit relatively few and weak intraresidue NOE's, residues in the rest of the protein adopt an irregular but well-defined structure. All peptide bonds are trans except for a cis peptide bond between Glu35 and Pro36. The structural characteristics observed for TGF-β1 in solution generally agree closely with the recently derived crystal structures of TGF-β2 [Daopin et al. (1992) *Science 257*, 369–374; Schlunegger & Grütter (1992) *Nature 358*, 430–434]. Several noteworthy differences were observed that may be related to function.

The eukaryotic proteins transforming growth factor βs (TGF-β)[1] are important regulators in a number of diverse cellular processes including bone growth and wound healing (Roberts & Sporn, 1990; Sporn & Roberts, 1990; Massagué, 1990). TGF-β1 is a member of the TGF-β superfamily which includes the TGF-βs (TGF-β1, -β2, -β3, -β4, and -β5), the inhibins, the bone morphogenetic proteins (BMPs), and other related growth factors. Proteins within the TGF-β superfamily exhibit 30% sequence homology and seven conserved cysteine residues, while the TGF-βs themselves exhibit 70% sequence identity and nine conserved cysteines. In order to understand the diverse activities of TGF-β1 and of other proteins in the TGF-β superfamily at the molecular level, it will be necessary to characterize the structures of these proteins in detail.

Over the last six years, it has been shown that reliable information about the structures of proteins in solution can be determined using NMR spectroscopy once NMR signal assignments are known (Wüthrich, 1986, 1989; Clore & Gronenborn, 1987, 1989). Because of the size of TGF-β1, 25 kDa, isotopically enriched protein is required in order to obtain signal assignments. Cloning of TGF-β1 and expression in Chinese hamster ovary (CHO) cells has provided the means for preparing large quantities of isotopically enriched protein. In the preceding paper (Archer et al., 1993), we reported the sequential assignments for TGF-β1 as determined using heteronuclear NMR spectroscopy on uniformly ^{15}N-enriched TGF-β1 and three selectively ^{13}C- and ^{15}N-enriched samples of TGF-β1. Herein we report the use of these sequential assignments along with NOE, hydrogen exchange, chemical shift, and coupling constant data to determine the secondary structure of TGF-β1 in solution.

MATERIALS AND METHODS

Sample Preparation. Unlabeled natural TGF-β1 was isolated from bovine bone following standard protocols as described previously (Ogawa & Seyedin, 1991). TGF-β1 that was uniformly ^{15}N-enriched and samples of TGF-β1 isotopically enriched with ^{13}C and ^{15}N in particular amino acids were obtained by expression of recombinant TGF-β1 in Chinese hamster ovary (CHO) cells grown on isotopically enriched medium (Archer et al., 1993). All samples were purified and prepared for NMR spectroscopy as described therein, and all spectra were recorded at pH 4.2 in H_2O unless noted otherwise. The protein concentration of the samples was approximately 1 mM (dimer) unless noted otherwise. The protein samples were shown to retain full biological activity

† This work was supported by the AIDS Targeted Antiviral Program of the Office of the Director of the National Institutes of Health (to D.A.T.) and Public Health Service National Research Service Award GM13620 (to S.J.A.). Part of this work was done while K.A.P. was a Scholar-in-Residence at the Fogarty International Center, National Institutes of Health.

‡ National Institute of Dental Research.
§ National Institute of Diabetes and Digestive and Kidney Diseases.
∥ National Cancer Institute.
⊥ Celtrix Pharmaceuticals, Inc.
Jefferson Medical College.
▲ R&D Systems Inc.

[1] Abbreviations: BMP, bone morphogenetic protein; CHO, Chinese hamster ovary cells; DQF-COSY, double quantum filtered correlated spectroscopy; HMQC, heteronuclear multiple quantum correlation; HNHB, proton to amide nitrogen to β-proton correlation; HOHAHA, homonuclear Hartmann–Hahn; HSQC, heteronuclear single quantum correlation; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, NOE spectroscopy; TFA, trifluoroacetic acid; TGF-β, transforming growth factor β; 2D, two-dimensional; 3D, three-dimensional.

0006-2960/93/0432-1164$04.00/0

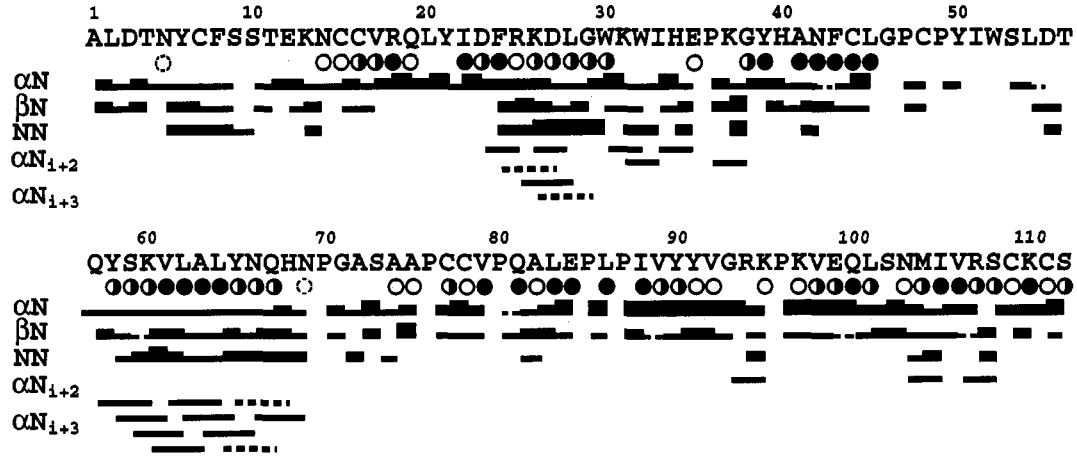

FIGURE 1: Diagram of NOE connectivities and hydrogen exchange rates for uniformly ^{15}N-labeled TGF-β1. The NOE correlations were determined from 3D NOESY–HMQC spectra at 37 °C and 45 °C and from a 3D ^{15}N–^{15}N–^{1}H HMQC–NOESY–HMQC spectrum at 37 °C. Sequential connectivities Hα_i–HN$_{i+1}$, Hβ_i–HN$_{i+1}$, and HN$_i$–HN$_{i+1}$ are designated αN, βN, and NN, respectively. The height of the bar indicates the strength of the NOE correlation (strong, medium, weak), and a dashed bar indicates that correlation is assigned tentatively due to degenerate chemical shifts. Circles indicate slow amide hydrogen exchange rates as determined using 2D ^{1}H–^{15}N HSQC spectra as described in the text: 5 min < $T \leq$ 1 h, ○; 1 < $T \leq$ 24 h, ◐; and $T >$ 24 h, ●, where T = 1/exchange rate. A broken circle indicates that the exchange rate is assigned tentatively due to degeneracies in both the amide nitrogen and proton chemical shifts.

as determined by assaying for the inhibition of proliferation of mink lung epithelial cells or mouse lymphocytes in culture (Ogawa & Seyedin, 1991; Tsang et al., 1990).

NMR Spectroscopy. NMR spectra of TGF-β1 were acquired on Bruker AMX 500 and 600 spectrometers and on a Bruker AM 500 spectrometer modified to reduce overhead time (Kay et al., 1990). 2D NMR spectra were processed using either Bruker or NMRi (New Methods Research, Inc., Syracuse, NY) software available on ASPECT 1000 and SUN data stations, respectively. The 3D NMR spectra were processed using a combination of NMRi and in-house software (Kay et al., 1989; Garrett et al., 1991; S. Grzesiek, unpublished).

2D nuclear Overhauser effect (NOE) spectra in D_2O and in H_2O were acquired with a 100-ms mixing time at 37 °C and at 45 °C. In the H_2O sample, solvent presaturation ($\gamma B_2/2\pi$ = 20–25 Hz) was used. 2D ^{13}C-edited NOESY (Bax & Weiss, 1987) and 2D ^{1}H–^{13}C HSQC–NOESY spectra of the selectively ^{13}C,^{15}N-labeled samples in D_2O were acquired at 45 °C with either a 100-ms or 200-ms mixing time. 2D ^{1}H–^{15}N HSQC (Bodenhausen & Ruben, 1980) spectra of uniformly ^{15}N-labeled TGF-β1 were acquired at 45 °C. The water resonance was suppressed with either 25-Hz solvent presaturation or ca. 15-Hz solvent presaturation and a 1-ms spin-lock pulse (Messerle et al., 1989). In all heteronuclear experiments, WALTZ-16 modulation (Shaka et al., 1983) was used to decouple ^{13}C or ^{15}N from the protons during acquisition.

3D spectra were acquired on uniformly ^{15}N-labeled TGF-β1. 3D 500-MHz and 600-MHz NOESY–HMQC (Marion et al., 1989a; Kay et al., 1989) and 3D 600-MHz HOHAHA–HSQC (Marion et al., 1989b) spectra were acquired at 37 °C and at 45 °C. A DIPSI-2 mixing sequence (Shaka et al., 1988) was used in the 3D HOHAHA–HSQC with mixing times of 39 ms at 37 °C and 58 ms and 80 ms at 45 °C. 3D NOESY–HMQC spectra were acquired at 45 °C with a 50-ms mixing time at 500 MHz and with a 90-ms mixing time at 600 MHz and at 37 °C with a 100-ms mixing time at 500 MHz. The 3D HNHB spectrum was acquired at 600 MHz as described previously (Archer et al., 1991). A 500-MHz 3D ^{15}N–^{15}N–^{1}H HMQC–NOESY–HMQC (Ikura et al., 1990; Frenkiel et al., 1990) spectrum was acquired with a 100-ms mixing time at 37 °C. The acquisition parameters for the HOHAHA–HSQC, NOESY–HMQC, and HMQC–NOESY–HMQC experiments were described previously (Archer et al., 1993). A 500-MHz 3D ROESY–HMQC (Clore et al., 1990) was acquired with a 50-ms mixing time at 37 °C, with spectral widths of 10.00, 22.9, and 16.12 ppm in F_1(^{1}H), F_2(^{15}N), and F_3(^{1}H), respectively, and with 128 complex points in t_1, 32 complex points in t_2, 1024 real points in t_3, and 16 scans per t_3 point. The ^{1}H carrier was set on water and the ^{15}N carrier at 120 ppm.

Hydrogen exchange rates of the amide protons were determined by lyophilizing the protein from H_2O, dissolving the dry protein in D_2O, and immediately acquiring the first in a series of 2D ^{1}H–^{15}N HSQC spectra (Marion et al., 1989c). HSQC spectra were acquired at 0.25, 0.50, 1, 2, 3, 5, 10, 20, 43, 78, and 117 h. Hydrogen exchange rates were determined from exponential fits of integrated peak intensities as a function of time.

RESULTS

Once sequential resonance assignments are known, elements of regular secondary structure are obtained from NOE patterns, $^{3}J_{HNH\alpha}$ coupling constants, and amide proton exchange rates (Wüthrich, 1986). Helical regions are identified by sequences of strong HN$_i$–HN$_{i+1}$ NOE correlations, slow amide hydrogen exchange rates, small $^{3}J_{HNH\alpha}$ coupling constants (<6 Hz) and Hα_i–HN$_{i+3}$ NOE correlations. Residues involved in β-strands exhibit strong Hα_i–HN$_{i+1}$ NOE correlations and large $^{3}J_{HNH\alpha}$ coupling constants (>7 Hz). β-Sheets are delineated by long-range HN$_i$–HN$_j$ and Hα_i–Hα_j NOE correlations and slow amide hydrogen exchange rates for the amide protons involved in hydrogen bonding across the β-sheet. Various types of tight turns are identified by characteristic HN$_i$–HN$_{i+1}$ and Hα_i–HN$_{i+1}$ NOE correlation patterns and coupling constants.

Short- and medium-range NOE correlations in 3D NOESY–HMQC spectra of uniformly ^{15}N-labeled TGF-β1 were assigned using the proton signal assignments for TGF-β1 (Archer et al., 1993) and were categorized as strong, medium, and weak (Figure 1). Amide proton chemical shift degeneracies were resolved by observation of ^{1}H–^{1}H correlations in the 3D ^{15}N–^{15}N–^{1}H HMQC–NOESY–HMQC spectrum in which cross-peaks are labeled with the amide nitrogen chemical

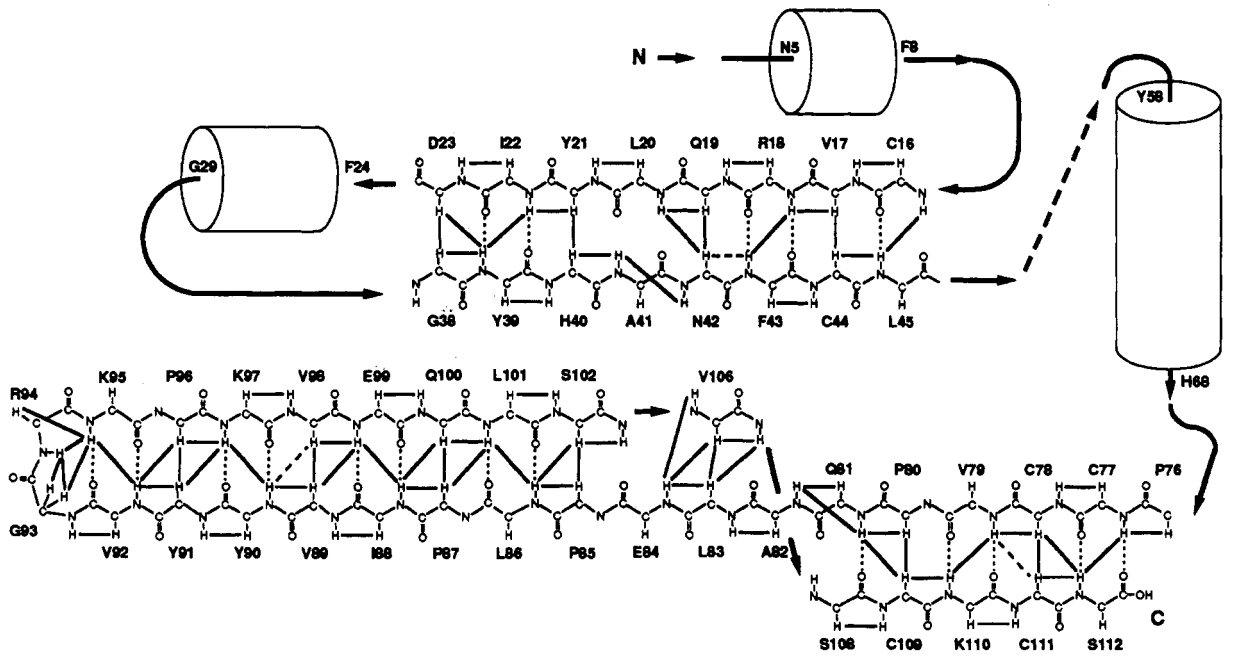

FIGURE 2: Schematic diagram of the secondary structure of TGF-β1 showing the HN–HN and Hα–Hα NOE correlations (solid lines) observed in 3D NOESY–HMQC, 3D ^{15}N–^{15}N–^{1}H HMQC–NOESY–HMQC and 2D NOESY spectra. Dashed lines represent NOE correlations that were assigned tentatively because of chemical shift degeneracies, and dotted lines represent hydrogen bonds derived from amide hydrogen exchange rates and NOE patterns.

shift in both the $f1$ and $f2$ dimensions. Long-range NOE correlations observed in 2D homonuclear NOESY spectra of TGF-β1, 2D heteronuclear NOESY spectra of specifically labeled TGF-β1, and 3D NOESY-HMQC spectra of uniformly ^{15}N-labeled TGF-β1 were used to define β-sheet structures (Figure 2). Hydrogen exchange, coupling constant, and chemical shift data were used to confirm elements of secondary structure (Figures 1, 4, and 5). The secondary structure of TGF-β1 derived from NOE and hydrogen exchange data is described in detail below.

α-Helices. There are three helical segments in TGF-β1. The longest helical segment in TGF-β1 stretches from Y58 to H68 and exhibits all the classical NMR features of a well-defined α-helix (Figures 1 and 3). In the 3D NOESY–HMQC spectrum, this region exhibits characteristic medium HN_i–HN_{i+1} and weak $H\alpha_i$–HN_{i+1} NOE correlations as well as $H\alpha_i$–HN_{i+3} NOE's (Figures 1 and 3). In addition, the amide protons exhibit slow hydrogen exchange rates.

The second helical segment stretches from F24 to G29 as identified from a series of strong HN_i–HN_{i+1} and weak $H\alpha_i$–HN_{i+1} NOE correlations in the 3D NOESY–HMQC spectrum (Figure 1). In the case of the second helix, unambiguous identification of several of the $H\alpha_i$–HN_{i+3} NOE correlations is difficult because the Hα proton chemical shifts cluster around 4.0 ppm. The overlap of the Hα signals often causes the $H\alpha_i$–HN_{i+3} NOE's to be degenerate with intraresidue and/or sequential Hα–HN NOE's. The amide protons in this helix have slow hydrogen exchange rates (Figure 1), indicating the presence of hydrogen bonds expected for a helix.

Residues N5–F8 are assigned to a one-turn helix on the basis of their sequential medium intensity HN_i–HN_{i+1} and weak $H\alpha_i$–HN_{i+1} NOE correlations in the 3D NOESY–HMQC spectrum (Figure 1). Because the helix is very short and presumably not highly stable, amide hydrogen exchange rates are not slow.

β-Sheets. There are several two-stranded antiparallel β-sheets in TGF-β1, but there is no evidence of multistrand β-sheet structure. Most of the C-terminal region of TGF-β1 (from C77 to S112) is in an extended β-strand conformation as indicated by strong $H\alpha_i$–HN_{i+1} NOE correlations in the 3D NOESY–HMQC spectrum (Figure 1). Indeed, residues C77–S112 form a nearly continuous two-stranded antiparallel sheet having a twist at N103–M104 and a short break at residues Q81–A82/R107–S108. The sheet contains a single reverse turn at residues V92–K95 which form a type II β-turn as indicated by (1) a strong HN_i–HN_{i+1} NOE between R94 and K95, (2) a strong $H\alpha_i$–HN_{i+1} NOE between G93 and R94, and (3) an $H\alpha_i$–HN_{i+2} NOE between G93 and K95 observed in the 3D NOESY–HMQC spectrum (Figure 1). A strong intraresidue $H\alpha_i$–HN_i NOE for R94 distinguishes this type II turn from a type II′ turn (Baldisseri et al., 1991). The antiparallel β-sheet formed by strands P85–V92 and K95–S102 is clearly delineated by numerous long-range HN_i–HN_j NOE's in 3D ^{15}N–^{1}H NOESY–HMQC and ^{15}N–^{15}N–^{1}H HMQC–NOESY–HMQC spectra and by $H\alpha_i$–$H\alpha_j$ NOE's in 2D homonuclear and heteronuclear NOESY spectra (Figure 3). Slow hydrogen exchange rates are seen for the amide protons involved in hydrogen bonding across the β-sheet (Figure 1). Residues L83–E84 are in an extended conformation as indicated by medium to strong $H\alpha_i$–HN_{i+1} NOE's in 3D NOESY–HMQC spectra. Residues I105–V106 lie antiparallel to residues L83–E84 as indicated by long-range $H\alpha_i$–$H\alpha_j$ NOE's in 2D NOESY spectra and HN_i–HN_j NOE's in 3D NOESY–HMQC. Residues N103 and M104 bend out of the β-sheet as determined by medium and weak $H\alpha_i$–HN_{i+1} NOE's for N103 and M104, medium and strong HN_i–HN_{i+1} NOE's for N103 and M104, respectively, and the absence of NOE correlations across the sheet. Molecular modeling based upon NOE correlations shows that one strand crosses over the other near residue N103. After a brief break in the two-stranded sheet at residues Q81–A82 and R107–S108, the structure continues and terminates with a short antiparallel β-sheet formed by strands C77–P80 and C109–S112. This structure is identified by long-range $H\alpha_i$–$H\alpha_j$ NOE's in 2D NOESY spectra and HN_i–HN_j NOE's in 3D ^{1}H–^{15}N NOESY–HMQC and ^{15}N–^{15}N–^{1}H HMQC–NOESY–HMQC spectra (Figure 3). The amide protons involved in hydrogen bonding across the β-sheet exhibit slow hydrogen

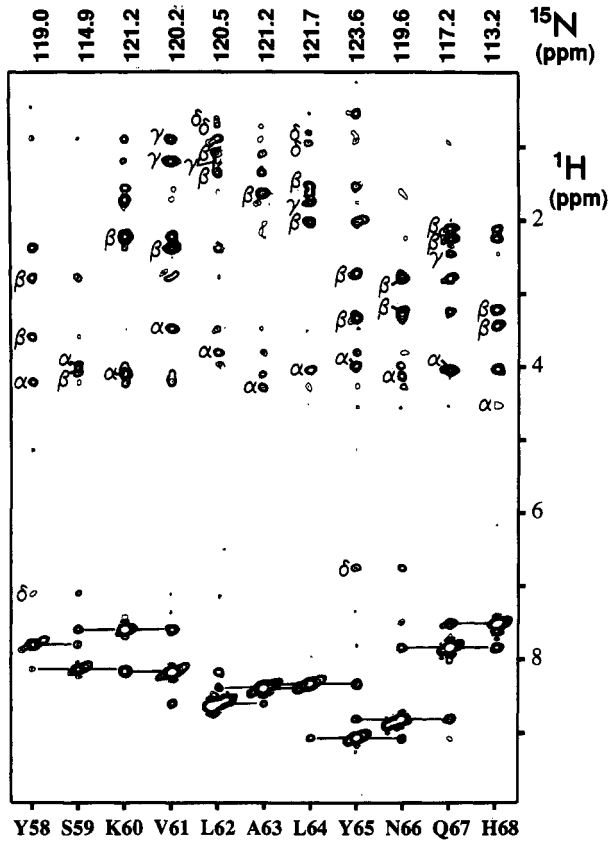

FIGURE 3: NOE correlations to the amide protons of residues Y58–H68. Strips were extracted from ^{15}N planes of a 600-MHz 3D NOESY–HMQC spectrum of uniformly ^{15}N-labeled TGF-β1. For clarity, NOE's to amide protons with chemical shifts close to the amides of interest are not shown. The spectrum was acquired at 45 °C with a 90-ms mixing time and processed with 60°-shifted sinebell-squared digital filtering in t_1 and t_3 and 60°-shifted sinebell digital filtering in t_2. ^{15}N chemical shifts of the amide peaks are listed at the top of the strips. Lines indicate NOE correlations that were also observed in the 3D ^{15}N–^{15}N–^{1}H HMQC–NOESY–HMQC spectrum.

exchange rates (Figure 1). Although we cannot always distinguish between intra- and intermonomer NOE's at the present time, we were able to determine that this long two-stranded β-sheet is an intramolecular β-sheet because we were able to trace sequential NOE's through the β-hairpin and throughout the β-strands.

In the N-terminal region of the protein, a distorted two-stranded antiparallel sheet is formed by strands V17–D23 and G38–C44. Residues V17–D23 exhibit medium to strong $H\alpha_i$–HN_{i+1} NOE correlations in the 3D NOESY–HMQC spectrum (Figure 1), while in the other strand, only residues G38–Y39 and C44 exhibit medium to strong $H\alpha_i$–HN_{i+1} NOE correlations (Figure 1). Long-range HN_i–HN_j NOE's in 3D NOESY–HMQC and ^{15}N–^{15}N–^{1}H HMQC–NOESY–HMQC spectra and $H\alpha_i$–$H\alpha_j$ NOE's in 2D NOESY spectra between the ends of the strands indicate that the strands are antiparallel. Slow amide hydrogen exchange rates for the amide protons located between the two strands (except for the amide of L20) indicate that there is hydrogen bonding between the strands consistent with the presence of β-sheet structure. A strong HN_i–HN_{i+1} NOE between A41 and N42 in the 3D NOESY–HMQC spectrum and the fast amide exchange rate for L20 indicate that the center of the β-sheet is distorted. At the present time, we cannot determine whether this short β-sheet is within a monomer or between monomers on the basis of NMR data alone because of the difficulty of distinguishing between intramonomer and intermonomer NOE's.

Other Secondary Structural Features. There are several turns in TGF-β1 in addition to the type II reverse turn for residues V92 to K95. At the end of the second helix, there is a series of turns from W30 to G38. Residues K31–I33, H34–E35, and K37–G38 exhibit strong HN_i–HN_{i+1} NOE correlations and W30, K31, and I33 show $H\alpha_i$–HN_{i+2} NOE's in addition in $H\alpha_i$–HN_{i+1} NOE's in the 3D NOESY–HMQC spectrum (Figure 1).

Flexible Region. Residues in the rest of the protein adopt an irregular but well-defined structure with the exception of residues G46–S53 which exhibit relatively few and weak intraresidue NOE's. Due to the lack of interresidue NOE's, we were unable to assign the amide and aliphatic proton signals of residues G46 and I51 and the amide of S53. In addition to very weak or absent ^{1}H–^{15}N correlations for G46, I51, and S53, the ^{1}H–^{15}N correlations for Y50, W52, and L54 were relatively broad and weak in the HSQC spectrum. In the ^{1}H–^{13}C HSQC NOESY of TGF-β1 labeled with [$^{13}C\gamma$]Pro, the NOE cross-peaks to the Hγ protons of P49 were also weak.

cis- and trans-Proline. NOE correlations between proline and its preceding residue, X, can be used to distinguish between *cis*- and *trans*-X-Pro peptide bonds. The presence of strong NOE correlations between the Hδ protons of Pro and the Hα proton of its preceding residue indicate the presence of a trans peptide bond while NOE's between the Hα proton of Pro and the Hα proton of its preceding residue indicate the presence of a cis peptide bond (Wüthrich, 1986). The presence of an NOE correlation between the Hα proton of P36 and the Hα proton of E35 in the ^{13}C-edited NOESY spectrum of TGF-β1 labeled with [2-^{13}C]Pro and the absence of NOE correlations between the Hδ protons of P36 and the Hα proton of E35 in the homonuclear 2D NOESY spectrum indicated that P36 was in the cis conformation.

The chemical shifts of the proline Cγ also indicate that the E35–P36 peptide bond is cis and the eight other X-Pro peptide bonds in TGF-β1 are trans. In ^{13}C-edited spectra of TGF-β1 labeled with [4-^{13}C]Pro, the chemical shifts of the Cγ carbons are clustered at 28 ppm ± 1.3 ppm, except for P36 which has a Cγ carbon shift of 25 ppm. This upfield shift for P36 is characteristic of *cis*-prolines (Howarth & Lilley, 1978; Torchia et al., 1989; Stanczyk et al., 1989) and supported the assignment of a cis E35–P36 peptide bond. The Cγ chemical shifts of the other eight prolines suggested that they were in the trans conformation. This conclusion was confirmed by identification of $H\delta_i$–$H\alpha_{i-1}$ NOE correlations in the 2D homonuclear NOESY spectrum for five prolines and by the absence of $H\alpha_i$–$H\alpha_{i-1}$ NOE correlations in the ^{13}C-edited NOESY spectrum of TGF-β1 labeled with [2-^{13}C]Pro for all prolines except P36.

Coupling Constants. Estimates of $^{3}J_{HNH\alpha}$ coupling constants were also used to probe the secondary structure of TGF-β1. In a 3D HOHAHA–HSQC spectrum obtained with a short mixing time, the major cross-peaks correlate HN and Hα signals and arise from magnetization transfer between HN and Hα protons via $^{3}J_{HNH\alpha}$ coupling. Further magnetization transfer between the Hα and Hβ protons, via $^{3}J_{H\alpha H\beta}$ coupling, sometimes results in minor HN–Hβ cross-peak intensity. For mixing times much less than $(^{3}J_{HNH\alpha})^{-1}$, the intensities of these cross-peaks are approximately (neglecting differences in T_2 of HN and aliphatic protons) related to $^{3}J_{HNH\alpha}$ by the following equation:

$$[I(H\alpha) + I(H\beta)]/I(HN) = \tan^2(\pi Jt) \qquad (1)$$

where I(Hα), I(Hβ), and I(HN) are the signal intensities of

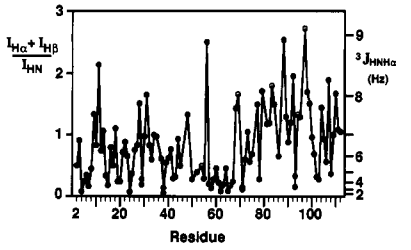

FIGURE 4: Ratio of (Hα + Hβ):HN peak intensities as a qualitative determination of $^3J_{HNH\alpha}$ coupling constants. The righthand scale shows the J couplings derived from eq 1. The peak intensities were measured in a 3D HOHAHA–HSQC acquired at 37 °C with a 38.7-ms mixing time. Open circles indicate that the ratio of (Hα + Hβ): HN peak intensities had to be estimated because of amide proton chemical shift degeneracies.

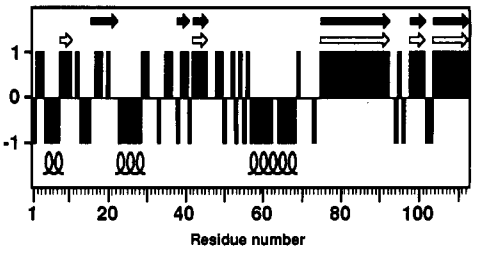

FIGURE 5: Comparison of the secondary structure of TGF-β1 determined from NOE data with the secondary structure predicted from Hα chemical shifts using the method of Wishart et al. (1992). Residues were assigned values of +1, 0, or −1 as described in the text. Solid arrows represent β-sheet regions as determined from NOE and hydrogen exchange data; open arrows represent β-strands predicted from Hα chemical shifts, and coils represent helical regions determined using either method.

the HN–Hα, HN–Hβ, and HN (diagonal) cross-peaks, respectively, $J = 0.93 \times {}^3J_{HNH\alpha}$ (the scale factor, 0.93, takes into account, in an average sense, off-resonance effects on the DIPSI-2 mixing sequence; Rucker & Shaka, 1989), and t is the isotropic mixing time. According to eq 1, the intensity ratio calculated assuming t = 38.7 ms increases from 0.05 to 2.62 as $^3J_{HNH\alpha}$ increases from 2 Hz to 9 Hz. We note that eq 1 is not valid for a typical Gly residue because the α-protons have significantly smaller T_2 values than the amide proton. A plot of the experimental intensity ratios observed in the 3D HOHAHA–HSQC spectrum of TGF-β1 recorded with a 38.7-ms mixing time (Figure 4) falls within the predicted range. This is a gratifying result considering that eq 1 contains no adjustable parameters. The right scale of Figure 4 shows the J couplings derived from the intensity ratio measurements and eq 1. We emphasize that the 3D HOHAHA–HSQC used for this coupling constant analysis had been acquired for signal assignments such that no additional experiments were necessary to obtain this coupling constant information.

Comparison of Figures 2 and 4 shows that almost all residues in the three helical regions exhibited relatively weak $^3J_{HNH\alpha}$ couplings as was anticipated. Exceptions were found near the ends of the helices; for example, L28 and H68 exhibited strong $^3J_{HNH\alpha}$ couplings. Considering that the ends of the helices are not precisely defined at the current level of structural analysis, the coupling constant data are in agreement with the three helical regions defined by NOE and hydrogen exchange data.

Most of the residues involved in the C-terminal two-stranded β-sheet exhibited strong $^3J_{HNH\alpha}$ couplings, while the residues involved in the G93–R94 turn, in the twist near N103, and in the cross-over near R107 showed weaker couplings as was expected. Residues in the N-terminal β-sheet exhibited weaker $^3J_{HNH\alpha}$ couplings than in the C-terminal β-sheet. Residues R18, Y21–D23, H40, and C44 exhibited medium to strong couplings, while other residues, such as Q19, exhibited weak $^3J_{HNH\alpha}$ couplings. This disparity in coupling constants could be attributed to a disruption of the β-sheet in the N-terminal region of TGF-β1 that was noted in the NOE patterns.

Chemical Shifts of Hα Protons. Recently, it has been shown that quantitative analysis of the chemical shift of Hα protons can be used to predict secondary structure of proteins (Wishart et al., 1992). The chemical shift of an Hα proton experiences a downfield or upfield shift when the residue is in a β-strand or helical conformation, respectively. The method of Wishart et al. assigns the value +1, 0, or −1 to a residue depending on whether the chemical shift of the Hα proton is downfield of, within, or upfield of a chemical shift range for its amino acid type. A high density of sequential +1's or −1's predicts the presence of β-strand or helix, respectively.

By applying this analysis to TGF-β1 (Figure 5), three regions of TGF-β1 were found that had a high density of upfield-shifted Hα protons and were predicted to be helical. These regions correlated well with the three helices delineated from the NOE and hydrogen exchange data. On average, the ends of the helices determined by the chemical shift and NOE data differed by less than one residue. Several regions had high densities of downfield-shifted Hα protons suggesting that these regions were β-strand. All of these regions, except F8–S10, correlated well with β-sheet regions delineated by NOE and hydrogen exchange data. In the 3D NOESY–HMQC F8–S10 exhibited relatively weak NOE's compared with other regions in the protein. Although the chemical shifts of the Hα protons of S9 and S10 were 4.65 and 4.66 ppm, respectively, at 45 °C, the weak water presaturation (at 4.56 ppm) cannot account for the weak HN_i–$H\alpha_{i-1}$ NOE's observed for both S9 and S10. In addition, the presence of a weak HN_i–HN_{i+1} NOE between S9 and S10 indicated that this region was not a β-strand. NOE and hydrogen exchange data also indicated the presence of β-strands from G38 to Y39 and from C16 to D23 (excluding L20) that were not predicted as β-strand following the Wishart prediction rules strictly. The Hα proton chemical shifts of Y39–H40 were downfield, but due to the presence of upfield-shifted Hα protons for G38 and A41, this region was not classified as β-strand. Likewise, strict adherence to the rules excluded residues between C16 and I22 from β-strand classification, although three of the seven residues were shifted downfield and the other residues had neutral Hα proton chemical shifts. Accounting for these minor discrepancies, the secondary structure predicted by chemical shift analysis is in close agreement with the structure delineated by NOE and hydrogen exchange data.

Disulfide Linkages. There are nine cysteines in each TGF-β1 monomer which form nine disulfides in the fully oxidized TGF-β1 dimer. The pairs of cystine side chains that were close in 3D space were delineated from NOE correlations between Cys Hβ protons in a 2D ^1H–^{13}C HSQC–NOESY of TGF-β1 in which the Cβ carbons of Cys were specifically labeled with ^{13}C (Figure 6). Two disulfide linkages were clearly suggested by the NMR data: C7 to C16 and C15 to C78. The Cβ proton of C7 showed NOE correlations to the Hα proton and one Hβ proton of C16 while the Cβ proton of C16 showed an NOE correlation with the Hα proton of C7 in the 2D ^1H–^{13}C HSQC–NOESY of TGF-β1 labeled with ^{13}C at the Cβ carbon. In addition, NOE correlations between the Hβ protons of C7 and the HN of V17 observed in the 3D NOESY–HMQC of uniformly ^{15}N-enriched TGF-β1 further supported the presence of the C7–C16 disulfide bond. In the 2D ^1H–^{13}C HSQC–NOESY, the Cβ proton of C15 showed NOE correlations to the Hα proton and both Hβ protons of

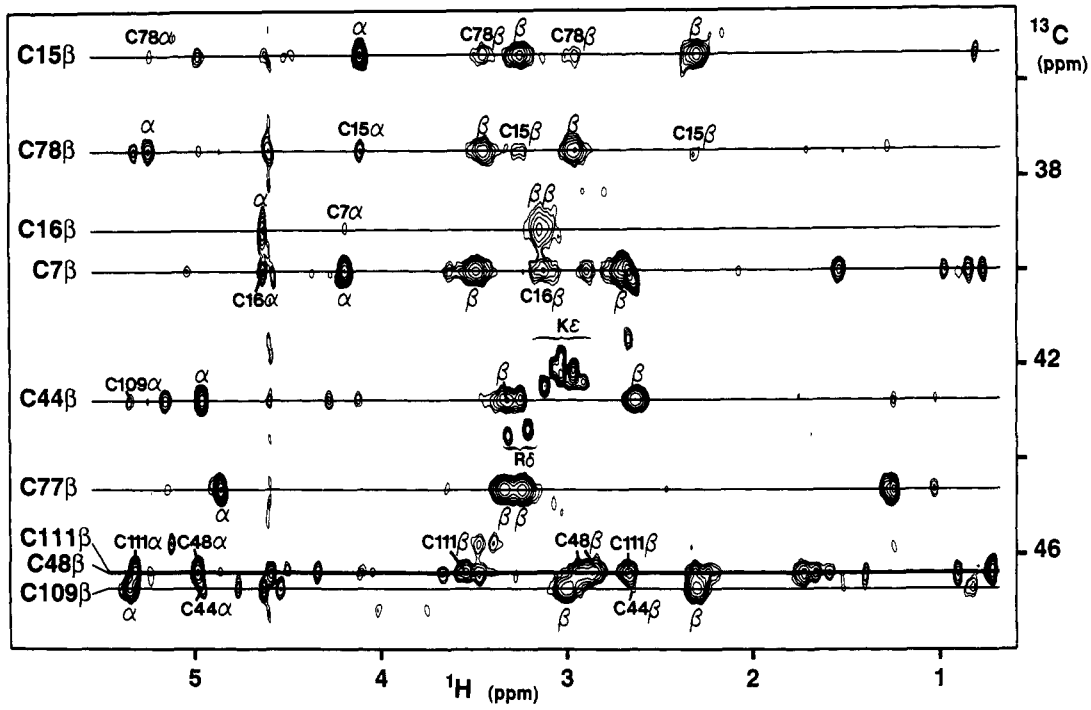

FIGURE 6: Cysteine $^{13}C\beta$ region from ^{13}C–^{1}H HSQC–NOESY of selectively labeled TGF-β1, (scheme 1) in which the Cβ carbons were ^{13}C-enriched. Intraresidue NOE cross-peaks are labeled with Greek symbols only while interresidue NOE cross-peaks between cysteines are labeled fully. The 500-MHz ^{13}C–^{1}H HSQC–NOESY spectrum was acquired on 0.56 mM protein with a 200-ms mixing time, 512 complex points in t_2, 184 complex points in t_1, and 512 scans per t_2 point. The spectrum was processed with Lorentzian to Gaussian digital filtering and zero-filled once in t_2 and twice in t_1.

C78 while the Cβ proton of C78 exhibited NOE correlations to both Hβ protons of C15, indicating the presence of a disulfide bond between C15 and C78. Because of the proximity of C15 and C16, we cannot rule out the possibility that the disulfide linkages are in fact C7–C15 and C16–C78. In addition, since at the present stage we cannot rigorously distinguish intra- from interdomain NOE's, we cannot determine whether these disulfide linkages are within a monomer or between monomers.

Minor Conformation. A major and a minor set of amide proton signals are observed for numerous residues in the 2D ^{1}H–^{15}N HSQC and the 3D NOESY–HMQC spectra of TGF-β1 at pH 4.2. The major peaks, which were ca. 4–5 times stronger than the minor peaks, were used to determine the secondary structure of TGF-β1 as described above. The minor signals were most readily identified and assigned for residues S73–A75 and C77–V79. The NOE correlations to the minor amide protons in these and other residues mimic the NOE patterns for the major component, indicating that major and the minor components have similar structures. Two sets of peaks were present in spectra of both ^{15}N-labeled and nonlabeled samples of TGF-β1, which indicated that the two components were present in both recombinant and naturally occurring TGF-β1. Raising the pH of the solution to 4.8 caused changes in the chemical shifts of some of the major and minor component signals, but the ratio of the intensities of the components was not measurably different than that found at pH 4.2.

DISCUSSION

After completion of the signal assignments of TGF-β1 (Archer et al., 1993) and the secondary structure determination reported herein, two independently determined crystal structures of TGF-β2 were reported (Daopin et al., 1992; Schlunegger & Grütter, 1992). The sequences of TGF-β1 and TGF-β2 are 71% identical, and the two proteins have similar biological activities (Roberts & Sporn, 1990). In certain cell types, however, the activities of these two proteins differ significantly. For example, the growth of endothelial cells in monolayer is inhibited by TGF-β1 nearly 100 times more strongly than by TGF-β2 (Cheifetz et al., 1990), while the formation of mesoderm in *Xenopus laevis* is induced by TGF-β2 but not TGF-β1 (Rosa et al., 1988). Hence, comparison of the structure of TGF-β1 with that of TGF-β2 is of great interest as it may provide a basis for understanding the similarities and differences in the biological activity of these two proteins.

The three helical domains, Y58–H68, F24–G29, and N5–N8 reported for TGF-β2 in the crystalline state by Daopin et al. (1992) are present in solution, and the hydrogen bonded amides inferred from the crystal structure are entirely consistent with the slow hydrogen exchange observed for amide protons of the two longer helices in solution. In the crystal structure the amide protons of C7 and F8 in the short helix form hydrogen bonds; however, in solution, the hydrogen exchange rates of C7 and F8 are large, indicating that the helix is less stable than the two longer helices.

There is a close homology between the structure of the long C-terminal β-sheet observed for TGF-β1 in solution with the C-terminal β-sheet observed in the crystal structure of TGF-β2. The interstrand hydrogen bonds identified in solution from NOE patterns and exchange rates (Figure 3) correlate almost perfectly with the regular β-strand hydrogen bonding patterns observed in the crystal structure. Loss of the β-sheet NOE pattern between N103 and R107 for TGF-β1 in solution is consistent with the loss of ideal β-sheet structure for this region in the crystal structure of TGF-β2.

The well-defined regions in the N-terminal β-sheet (Figure 3) between strands C16–R18 and F43–L45 and between strands I22–D23 and G38–Y39 are seen in both TGF-β1 and TGF-β2. The break in the middle of the sheet in TGF-β1, characterized by the short HN–HN distance between A41 and N42 in solution, correlates with a disruption of the β-sheet

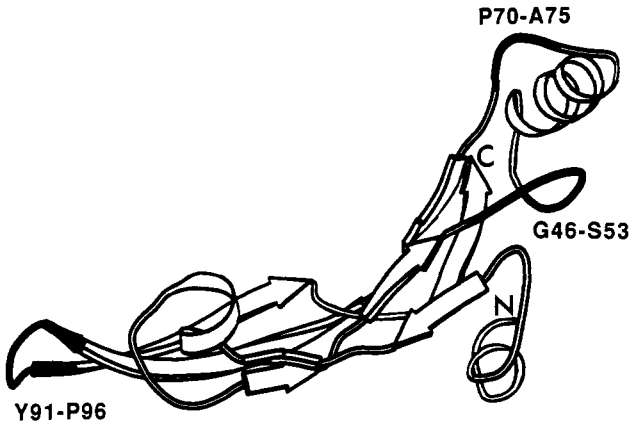

FIGURE 7: Comparison of the solution structure of TGF-β1 with the crystal structure of TGF-β2. The topology is based on the crystallographic coordinates for TGF-β2 (Daopin et al., 1992) and drawn using MOLSCRIPT software (Kraulis, 1991). Shaded regions indicate that the conformation and/or mobility of TGF-β1 in solution differs from TGF-β2 in the crystalline state.

observed in the crystal structure of TGF-β2. A single cis peptide bond, between E35 and P36, is also observed in the crystal structure of TGF-β2.

The disulfide bonds C7–C16, C15–C78, and C44–C109 observed in the crystal structure of TGF-β2 are in agreement with NOE correlations observed for TGF-β1 (Figure 6). Although a C48–C111 disulfide bond is seen in the crystal structure, its presence cannot be confirmed by the NOE data in Figure 6 because the C48 and C111 $^{13}C\beta$ chemical shifts are degenerate.

A systematic comparison of the solution structure of TGF-β1 with the crystal structure of TGF-β2 was also made in the following manner. Using the heavy atom coordinates obtained from the X-ray structure (S. Daopin and D. Davies, private communication), the set of all proton–proton distances in TGF-β2, less than 5 Å, was calculated using CHARMM. These distances were then used to predict relative NOE intensities in the solution spectra of TGF-β1. Approximately 400 short- and medium-range (involving residues i and $i+j$, where $j \leq 4$) NOE's predicted by the crystal structure were observed in the NOE spectra of TGF-β1. With the exception of three short sections of the protein sequence discussed in detail below, no significant differences between the observed and predicted NOE's were found. In addition, 100 long-range (involving residues i and $i+j$, where $j > 4$) intramonomer and intermonomer NOE's predicted by the X-ray structure were identified in NOE spectra. Ten of the long-range NOEs that were correctly predicted were between monomer units. Taken together, these results indicate that the two TGF-β isoforms have similar monomer structures and dimer interfaces.

Given the high sequence homology of the β1 and β2 isoforms, it is hardly surprising that they exhibit the close structural homology just discussed. However, as noted above, there are three regions of the amino acid sequence, residues G46–S53, P70–P76, and Y90–V98, where the structure of TGF-β1 may differ significantly from that of TGF-β2 (Figure 7). We discuss these structural differences in detail, as they may be in part responsible for the different activities of the two proteins.

In the crystal structure of TGF-β2, residues Y91–P96 are poorly defined in the electron density map indicating disorder (Daopin et al., 1992). In contrast, in solution residues Y91–P96 all exhibit strong interresidue NOE's and the amides of V92 and K95 have small exchange rates which show that G93–R94 form a well-defined type II reverse turn. Preliminary analysis of measurements of $^{15}N\{-^{1}H\}$ NOE's suggest that the reverse turn may undergo small amplitude internal motions on the subnanosecond time scale in solution. Small amplitude motions on this time scale would spoil the electron density map but have little effect on proton NOE's.

In solution, residues 46–53 in TGF-β1 exhibit relatively few interresidue NOE correlations, and several amide proton signals of residues in this region are broadened by chemical exchange. Evidently the internal motions that cause line broadening in solution are absent in the crystal structure of TGF-β2, because residues 46–53 are well-ordered as indicated by small B-factors (S. Daopin and D. Davies, personal communication). There are four amino acid substitutions from β1 to β2 (L45A, P47A, I51L, and L54S) in this region of the sequence; they could cause a significant difference in the local structure of the two isoforms.

Residues P70–P76 adopt well-defined, but different, structures for TGF-β1 in solution and TGF-β2 in the crystalline state. The NOE patterns of P70–A75 in TGF-β1 do not agree with the NOE patterns predicted by the crystal structure of TGF-β2. There are two amino acid substitutions in this region from β1 to β2: G71E and A75S. The G71E substitution may be the source of the conformational difference in this region of the sequence because the (ϕ,ψ) angles of P70 obtained in solution (from the NOE data) and in the crystalline state are ca. (60,100) and (60,–33), respectively.

Finally, we note that the comparison of the structures of the two TGF-β isoforms was limited for residues 9–13 because of sequence differences, S9R, S10N, T11V, E12Q, and K13D, and few observed long-range NOE's. Short-range backbone NOE's predicted by the crystal structure were observed; for example, a weak HN_i–HN_{i+1} NOE for residues 9–10 and a medium HN_i–HN_{i+1} NOE for residues 13–14. However, for a predicted interproton distance, the NOE cross-peak intensities in this region of the protein were generally weaker than those observed elsewhere.

In summary, the available NMR data show that the unusual monomer fold and dimer interface observed for TGF-β2 in the crystalline state (Daopin et al., 1992; Schlunegger & Grütter, 1992) are maintained by TGF-β1 in solution. Furthermore, with the exception of the three regions of the sequence discussed above, the two TGF-β isoforms have highly homologous, if not identical, secondary structures. It is perhaps significant that the regions of the sequence where structural and/or flexibility differences are evident are at, or near, the surface of the protein, and could therefore interact with TGF-β receptors. In addition, Qian et al. (1992) have shown that residues 40–82 have an important role in specifying the activity of a particular TGF-β isoform. It is noteworthy that this region of the sequence contains two sequences where the NMR structure of TGF-β1 differs from the crystal structure of TGF-β2. Although it is possible that the differences in β1/β2 structure that we have described could result from the different state (crystal vs solution) of the two proteins, we think that this is unlikely because the regions of the sequence that exhibit structural differences are not involved in significant crystalline contacts (S. Daopin and D. Davies, personal communication). In any case, this question will be resolved by further studies of TGF-β2 in solution and/or TGF-β2 in the crystalline state.

ACKNOWLEDGMENT

We gratefully acknowledge Drs. D. Garrett, S. Grzesiek, M. Ikura, L. Kay, and D. Marion and F. Delaglio for providing computer software. We thank Drs. S. Daopin and D. Davies for making the coordinates of their structure available to us

prior to publication and for helpful discussions regarding their results. We also thank R. Tschudin for expert technical support.

REFERENCES

Archer, S. J., Ikura, M., Torchia, D. A., & Bax, A. (1991) *J. Magn. Reson. 95*, 636–641.

Archer, S. J., Bax, A., Roberts, A. B., Sporn, M. B., Ogawa, Y., Piez, K., Weatherbee, J., Tsang, M., Lucas, R., Zheng, B., Wenker, J., & Torchia, D. A. (1993) *Biochemistry* (preceding paper in this issue).

Baldisseri, D. M., Torchia, D. A., Poole, L. B., & Gerlt, J. A. (1991) *Biochemistry 30*, 3628–3633.

Bax, A., & Weiss, M. A. (1987) *J. Magn. Reson. 71*, 571.

Bodenhausen, G., & Ruben, D. J. (1980) *Chem. Phys. Lett. 69*, 185–189.

Cheifetz, S. C., Hernandez, H., Laiho, M., ten Dijke, P., Iwata, K. K., & Massague, J. (1990) *J. Biol. Chem. 265*, 20533–20538.

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

Clore, G. M., & Gronenborn, A. M. (1989) *Crit. Rev. Biochem. Mol. Biol. 24*, 479–564.

Clore, G. M., Bax, A., Wingfield, P. T., & Gronenborn, A. M. (1990) *Biochemistry 29*, 5671–5676.

Daopin, S., Piez, K. A., Ogawa, Y., & Davies, D. R. (1992) *Science 257*, 369–374.

Frenkiel, T., Bauer, C., Carr, M. D., Birdsall, B., & Feeney, J. (1990) *J. Magn. Reson. 90*, 420–425.

Garrett, D. S., Powers, R., Gronenborn, A. M., & Clore, G. M. (1991) *J. Magn. Reson. 95*, 214–220.

Howarth, O. W., & Lilley, D. M. J. (1978) *Prog. NMR Spectrosc. 12*, 1–40.

Ikura, M., Bax, A., Clore, G. M., & Gronenborn, A. M. (1990) *J. Am. Chem. Soc. 112*, 9020–9022.

Kay, L. E., Marion, D., & Bax, A. (1989) *J. Magn. Reson. 84*, 72–84.

Kay, L. E., Ikura, M., Tschudin, R., & Bax, A. (1990) *J. Magn. Reson. 89*, 496–514.

Kraulis, P. (1991) *J. Appl. Crystallogr. 24*, 946–950.

Marion, D., Kay, L. E., Sparks, S. W., Torchia, D. A., & Bax, A. (1989a) *J. Am. Chem. Soc. 111*, 1515–1517.

Marion, D., Driscoll, P. C., Kay, L. E., Wingfield, P. T., Bax, A., Gronenborn, A. M., & Clore, G. M. (1989b) *Biochemistry 28*, 6150–6156.

Marion, D., Ikura, M., Tschudin, R., & Bax, A. (1989c) *J. Magn. Reson. 85*, 393–399.

Massagué, J. (1990) *Annu. Rev. Cell Biol. 6*, 597–641.

Messerle, B. A., Wider, G., Otting, G., Weber, C., & Wüthrich, K. (1989) *J. Magn. Reson. 85*, 608–613.

Ogawa, Y., & Seyedin, S. M. (1991) *Methods Enzymol. 198*, 317–327.

Qian, S. W., Burmester, J. K., Merwin, J. R., Madri, J. A., Sporn, M. B., & Roberts, A. B. (1992) *Proc. Natl. Acad. Sci. 89*, 6290–6294.

Roberts, A. B., & Sporn, M. B. (1990) *Handbook of Experimental Pharmacology. Peptide Growth Factors and Their Receptors I* (Sporn, M. B., & Roberts, A. B., Eds.) pp 419–472, Springer-Verlag, New York.

Rosa, F., Roberts, A. B., Danielpour, D., Dart, L. L., Sporn, M. B., & Dawid, I. B. (1988) *Science 239*, 783–786.

Rucker, S. P., & Shaka, A. J. (1989) *Mol. Phys. 68*, 509–517.

Schlunegger, M. P., & Grütter, M. G. *Nature 358*, 430–434.

Shaka, A. J., Keeler, J., & Freeman, R. (1983) *J. Magn. Reson. 53*, 313–340.

Shaka, A. J., Lee, C. J., & Pines, A. (1988) *J. Magn. Reson. 77*, 274–293.

Sporn, M. B., & Roberts, A. B. (1990) *Cell Regul. 1*, 875–882.

Stanczyk, S. M., Bolton, P. H., DellÀcqua, M., & Gerlt, J. A. (1989) *J. Am. Chem. Soc. 111*, 8317–8318.

Torchia, D. A., Sparks, S. W., Young, P. E., & Bax, A. (1989) *J. Am. Chem. Soc. 111*, 8315–8317.

Tsang, M. L.-S., Weatherbee, J. A., Dietz, M., Kitamura, T., & Lucas, R. C. (1990) *Lymphokine Res. 9*, 670.

Wishart, D. S., Sykes, B. D., & Richards, F. M. (1992) *Biochemistry 31*, 1647–1651.

Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley, New York.

Wüthrich, K. (1989) *Acc. Chem. Res. 22*, 36–44.