Proton and Nitrogen Sequential Assignments and Secondary Structure Determination of the Human FK506 and Rapamycin Binding Protein†

Michael K. Rosen, Stephen W. Michnick, Martin Karplus,* and Stuart L. Schreiber*
Department of Chemistry, Harvard University, 12 Oxford Street, Cambridge, Massachusetts 02138
Received December 10, 1990; Revised Manuscript Received January 31, 1991

ABSTRACT: Sequential ^1H and ^{15}N assignments of human FKBP, a cytosolic binding protein for the immunosuppressive agents FK506 and rapamycin, are reported. A combination of homonuclear and relayed heteronuclear experiments has enabled assignment of 98 of 99 backbone amide NHs, 119 of 120 $C^{\alpha}Hs$, 97 of 99 non-proline amide ^{15}Ns, and 375 of 412 side-chain resonances of this 107-residue protein. Long-range NOEs are used to demonstrate that FKBP has a novel folding topology consisting of a five-stranded antiparallel β sheet with +3, +1, −3, +1 loop connectivity.

The cyclosporin A (CsA, Figure 1) binding protein cyclophilin (Handschumacher et al., 1984) and the FK506 (Siekierka et al., 1989; Harding et al., 1989) and rapamycin (Fretz et al., 1990) (Figure 1) binding protein FKBP[1] are the best studied members of an emerging family of cytoplasmic, soluble binding proteins termed immunophilins (immunosuppressant binding proteins). Both proteins catalyze interconversion of the cis and trans rotamers of peptidyl–prolyl amide bonds of peptide and protein substrates. These rotamase enzymes bind with high affinity to and are inhibited by the immunosuppressive ligands CsA, FK506, and rapamycin. In addition to their therapeutic potential, these ligands are proving to be illuminating probe reagents for the events that comprise T-cell activation (Schreiber, 1991). Whereas CsA and FK506 potently inhibit a T-cell receptor (TCR) mediated single-transduction pathway, rapamycin appears to act at a later stage by interfering with a lymphokine receptor mediated signaling pathway (Dumont et al., 1990a; Bierer et al., 1990a). In addition to their ability to inhibit distinct T-cell activation events, the structurally related immunosuppressants FK506 and rapamycin, but not CsA, inhibit each other's actions in a variety of functional assays (Dumont et al., 1990b; Bierer et al., 1990a). The inability of CsA to inhibit the actions of either FK506 or rapamycin suggests that the actions of CsA are mediated by its binding to an immunophilin different from that associated with either FK506 or rapamycin. The mutual inhibition of FK506 and rapamycin requires concentrations of the drug acting as the antagonizing agent that are 10–100 times greater than the effective drug concentration of either FK506 or rapamycin (IC_{50} = 0.5 nM). This can be accounted for by the existence of separate immunophilins with different affinities for the two agents or by an excess of a common immunophilin that is responsible for mediating the actions of both FK506 and rapamycin. FKBP is present in T lymphocytes at a 10-fold higher concentration than the IC_{50} of these drugs, has been shown to be the predominant binding protein for both FK506 and rapamycin (Fretz et al., 1991), and binds to FK506 (dissociation constant K_d = 0.4 nM) and rapamycin (K_d = 0.2 nM) with high affinity (Bierer et al., 1990a). An attractive explanation for the concentration dependence observed in the mutual inhibition studies is that the abundant, uncomplexed FKBP may serve to sequester the antagonizing agent. The concentration of the antagonizing agent would rise to levels that are sufficient to displace the drug from its biological receptor only after the excess FKBP binding sites are saturated (Bierer et al., 1990a; Bierer et al., 1990b). In addition to implicating immunophilin–drug complexes as inhibitors of two distinct signal-transduction pathways, the concentration dependency of drug action and the abundance of FKBP also illustrate that the inhibition of the rotamase activity of FKBP is an insufficient requirement for mediating the actions of FK506 and rapamycin. This latter point is confirmed by the different biological properties of FK506, rapamycin, and a nonnatural FKBP ligand, 506BD (Bierer et al., 1990b), despite their potent rotamase inhibitory properties.

The actions of the inhibitory immunophilin–drug complexes appear to be distal to early membrane events and proximal to nuclear processes. Although these intermediary events are in general poorly understood, the existence of immunophilin–drug complexes that inhibit two distinct pathways provides hope for determining general membrane-to-nucleus signaling mechanisms. We have suggested that the structure of the immunophilin–drug complex determines the specificity for the factors associated with the different signaling pathways leading to T-cell activation. The high abundance and highly conserved nature of these ubiquitous enzymes suggest that they may have a second function, possibly in facilitating the folding of newly synthesized proteins (Fischer et al., 1989; Schreiber, 1991).

For these reasons, determination of the three-dimensional structures of immunophilins, free and complexed with their

† This work was supported by the National Institute of General Medical Sciences (GM-38627, awarded to S.L.S.; GM-30804, awarded to M.K.). A National Science Foundation Predoctoral Fellowship to M.K.R. is gratefully acknowledged. NMR spectra were obtained through the auspices of the Harvard University Department of Chemistry Instrumentation Center, which was supported by NIH Grant 1-S10-RR04870 and NSF Grant CHE 88-14019.

[1] Abbreviations: NMR, nuclear magnetic resonance; COSY, homonuclear two-dimensional correlated spectroscopy; COSY-type experiments, experiments manifesting cross-peaks due to scalar coupling; DQF, double-quantum filtered; DQ, double-quantum (two-quantum) spectroscopy; FID, free induction decay; GARP, heteronuclear decoupling using globally optimized alternating-phase rectangular pulses; HMQC, heteronuclear multiple-quantum coherence spectroscopy; NOE, nuclear Overhauser effect; NOESY, two-dimensional NOE spectroscopy; ppm, parts per million; SQC, heteronuclear single-quantum coherence spectroscopy; TOCSY, two-dimensional total correlation coherence spectroscopy; TPPI, time-proportional phase incrementation; TSP, 3-(trimethylsilyl)propionate-d_4; three-spin, the $C^{\alpha}H$-$C^{\beta}H_2$ spin subsystem of serine, cysteine, aspartic acid, asparagine, and the aromatic amino acid residues; five-spin, the $C^{\alpha}H$-$C^{\beta}H_2$-$C^{\gamma}H_2$ spin subsystem of glutamic acid, glutamine, and methionine residues.

0006-2960/91/0430-4774$02.50/0

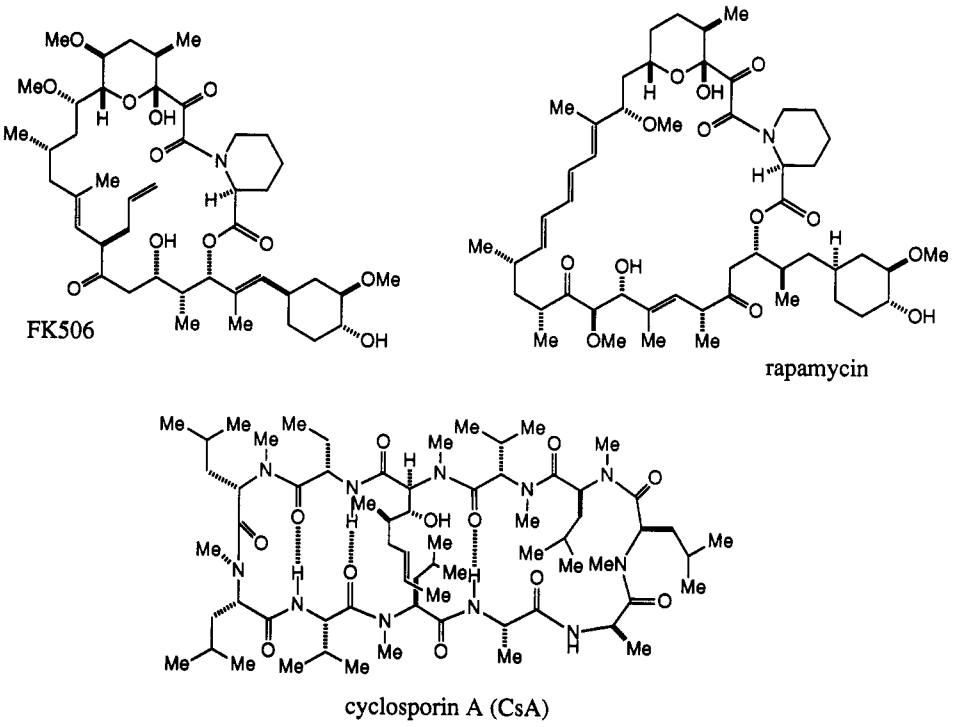

FIGURE 1: Structures of the immunophilin ligands cyclosporin A, FK506, and rapamycin.

immunosuppressive ligands (Rosen et al., 1990), is an important aspect of research into cytoplasmic signaling mechanisms. As a first step in this direction, we have undertaken the structure determination of human FKBP by NMR. The prerequisite NMR sequential assignment and secondary structure determination of FKBP are detailed in this report. NMR has proven to be an established method for the determination of secondary structure in proteins up to approximately 18 kDa (Driscoll et al., 1990; Dyson et al., 1989; Ikura et al., 1990; LeMaster & Richards, 1988; McIntosh et al., 1990; Torchia et al., 1989; Wang et al., 1990). However, with 107 residues FKBP is one of the larger proteins to have been sequentially assigned prior to determination of its own solid-state structure or of that of a highly homologous protein (Chazin & Wright, 1988; Dyson et al., 1989) by X-ray crystallography. The NMR study has been aided by the fact that FKBP is stable over a pH range of 4.5–7.5 and a temperature range of 298–314 K and is soluble at concentrations as high as 5 mM.

A wide variety of standard homonuclear (^{1}H–^{1}H) techniques and more novel heteronuclear (^{15}N–^{1}H) experiments have been used in the sequential assignment. Heteronuclear experiments such as HMQC (and SQC) (Mueller, 1979; Redfield, 1983; Bax et al., 1983; Glushka & Cowburn, 1987), HMQC–COSY (Clore et al., 1988), and HMQC–NOESY (Gronenborn et al., 1989) applied to uniformly (>95%) ^{15}N-labeled protein offer several advantages over standard homonuclear spectra. First, large, uniform $^1J_{^{15}N\text{-}NH}$ couplings and freedom of resonance overlap with the irradiated water signal afford HMQC and SQC inherently greater sensitivity than standard COSY spectra in determining the amide fingerprint of a protein. In addition, the large ^{15}N spectral dispersion allows unambiguous assignment in cases of NH overlap. Finally, characteristic ^{15}N chemical shifts and cross-peak shapes in these spectra provide information not available in standard proton spectra.

Spin system identification was performed according to methods described by Weber et al. (1987), Chazin et al. (1988), and Gronenborn et al. (1989). Briefly, protons and nitrogens belonging to individual spin systems were identified through correlations to amide protons and ^{15}N atoms in SQC, COSY, HMQC–COSY, DQ, relay COSY, double-relay COSY, and TOCSY spectra recorded in H_2O. This preliminary data base was then augmented by linking these amide-based spin subsystems to side-chain-based spin subsystems generated from correlations to characteristic side-chain protons (e.g., methyl protons of aliphatic residues and low-field protons of basic residues) found in spectra recorded in D_2O. The sequential assignment and secondary structure determination of FKBP were achieved through a combination of strategies (Wagner & Wüthrich, 1982; Weber et al., 1987; Chazin & Wright 1988; Gronenborn et al., 1989; MacIntosh et al., 1990; Wang et al., 1990). Neighboring spin systems were first connected through strong $C^\alpha H_i$–NH_{i+1}, NH_i–NH_{i+1}, and $C^\beta H_i$–NH_{i+1} NOEs by use of both ^{1}H–^{1}H and ^{1}H–^{15}N spectra. The sequences of connected residues were mapped onto the known protein primary sequence through identification of unique peptide sequences. Gaps in the assigned sequence were then filled in, and spin system identification was completed. Finally, distinctive patterns of medium- and long-range NOEs were used to identify elements of secondary structure such as β sheets, helices, and turns.

The use of both ^{1}H–^{1}H and ^{1}H–^{15}N spectra has allowed us to achieve more complete and unambiguous assignments of FKBP, a relatively large protein, than would have been possible with either method alone. In addition, resolution of amide proton degeneracies allows definitive identification of many more long-range NOEs to amide protons, which will aid in the three-dimensional structure determination of FKBP.

By using the methodology described above, we have been able to obtain virtually complete ^{1}H and ^{15}N sequential assignments. The paper describes in some detail the methodology used in the assignments, the assignment results, and the analysis of the secondary structure, which is found to correspond to an unusual fold for an antiparallel β sheet.

Experimental Procedures

Purification of FKBP and Sample Preparation. Recombinant human FKBP was purified from *Escherichia coli* strain

XA90 containing the overexpression plasmid FKBP-pHN1+ as previously reported (Standaert et al., 1990). The recombinant protein is identical with native protein in all respects, indicating the absence of posttranslational modifications of the native protein. Uniformly (>95%) ^{15}N-labeled FKBP was produced by growing the bacteria on M9 minimal media containing $^{15}NH_4Cl$ (Isotech, Inc.) as the sole nitrogen source (Muchmore et al., 1989).

All solutions used for NMR were 1–5 mM in FKBP/25 mM acetate-d_4 (Cambridge Isotope Labs) buffer at pH 5.0 (uncorrected for isotope effects). H_2O samples were prepared by exchanging the protein five times with 90% H_2O/10% D_2O buffer and concentrating the final solution by ultrafiltration. D_2O samples were prepared by lyophilizing the protein from H_2O and dissolving the residue in 99.96% D_2O buffer to desired concentration.

NMR Spectroscopy. NMR spectra were recorded on either Bruker AM-500 or Varian VXR500 spectrometers equipped with either inverse proton-broad-band or proton-selective probes. For the ^{15}N experiments, a separate BFX-5 amplifier unit was added to the Bruker instrument to generate soft pulses for ^{15}N GARP (Shaka et al., 1985) decoupling. Quadrature detection in t_1 was achieved through time-proportional phase incrementation (Piantini et al., 1982; Shaka & Freeman, 1983) on the Bruker instrument and by the method of States–Haberkorn–Ruben (States et al., 1982) on the Varian instrument. In spectra recorded in H_2O solution, water suppression was achieved by selective on-resonance irradiation of the water signal during a relaxation delay of 1.8–2.0 s and in NOESY experiments during the mixing period as well. An HMQC spectrum was also acquired using a $1-\tau-1$ water suppression sequence (Sklenar & Bax, 1987). Most experiments were performed at 300 K. Typically, spectra in H_2O and D_2O solutions were recorded by use of sweep widths of 6500 and 4237 Hz, respectively, although those in freshly prepared D_2O solutions required a sweep width of 5618 Hz. Between 400 and 800 t_1 increments of 2K complex data points were acquired in all experiments.

COSY and DQF-COSY spectra (Aue et al., 1976; Rance et al., 1983, 1984, 1985) were recorded in H_2O and D_2O at 298, 300, and 314 K according to standard pulse sequences and phase cycling; the different temperatures were particularly useful for identifying protons near the water resonance. Relay COSY (Eich et al., 1982; Wagner, 1983; Bax & Drobny, 1985) spectra were acquired in H_2O solution with τ = 25, 40, and 60 ms and in D_2O solution with τ = 30 ms. A double-relay COSY spectrum (Bax & Drobny, 1985) was recorded in H_2O solution with τ_1 = 28 ms and τ_2 = 35 ms. DQ spectra (Braunschweiler et al., 1983; Wagner & Zuiderweg, 1983; Rance et al., 1984, 1985, 1986; Otting & Wüthrich, 1986) were acquired with excitation periods of 32 ms for the H_2O sample and 30 ms for the D_2O sample. TOCSY spectra (Braunschweiler & Ernst, 1983; Bax & Davis, 1985; Davis & Bax, 1985) and clean TOCSY spectra (Griesenger et al., 1988) were recorded on H_2O and D_2O samples by use of an MLEV-17 sequence for spin lock with a 10.4-kHz spin-lock field. The r factors in the clean TOCSY experiments were either 2.6 or 0.6. Mixing times were 75 and 100 ms in H_2O and 75 ms in D_2O. NOESY spectra (Macura & Ernst, 1980; Dobson et al., 1982; Olejniczak et al., 1985) were recorded on H_2O and D_2O samples at 298, 300, 308, and 314 K with mixing times of 50, 100, 150, and 200 ms.

HMQC, SQC (Mueller, 1979; Redfield, 1983; Bax et al. 1983; Glushka & Cowburn, 1987), HMQC–COSY (Clore et al., 1988), and HMQC–NOESY (Gronenborn et al., 1989) experiments were recorded at 298 K by use of standard phase cycling. Fixed delays in the HMQC (and relayed HMQC) and SQC experiments were 4.4 and 2.2 ms, respectively. The HMQC–NOESY mixing time was 200 ms. In all cases, the ^1H and ^{15}N sweep widths were 8000 and 3500 Hz, respectively. ^{15}N decoupling during the detection period was achieved in all cases with a GARP sequence (Shaka et al., 1985) by use of an ^{15}N 90° pulse width of 250 ms.

All data were transferred to a Silicon Graphics Iris 4D/85GT workstation for processing and analysis with the FELIX software package (D. Hare, unpublished). Signal processing of individual FIDs consisted of apodization with the use of skewed sine-bell functions with phase shifts of 0 or 5° for resolution enhancement or 20, 35, or 50° for sensitivity enhancement. For NOESY and TOCSY data, a fourth-order polynomial base-line correction was applied in ω_2. Data were typically zero-filled in t_1 to yield a 1K × 1K real matrix after Fourier transformation and reduction.

RESULTS

Spin System Identification. In this section we describe the identification of the spin systems and their sequential assignment. As in other such studies, these two aspects are interwoven; that is, the logical stages of spin system identification and sequential assignment do not always follow in rigid order. This is particularly true of three-spin and five-spin systems, where unambiguous assignment of the spin systems is difficult at the spin system assignment stage. The apparent three-spin and five-spin systems may be more complicated spin systems whose other resonances are not seen because of spectral overlap or a lack of correlation cross-peaks. In these cases, final confirmation of the spin system assignment is dependent on unambiguous sequential assignment.

Of the 99 possible unique backbone amide protons (107 residues minus 7 Pro and the N-terminal Gly), 95 are visible in an SQC spectrum taken in H_2O (Figure 2). Of these, 92 could be connected to $C^\alpha H$ protons in COSY and HMQC-COSY spectra taken in H_2O. A DQ spectrum showed four additional $C^\alpha Hs$ that were under the saturated H_2O resonance at 300 K (one of these, Phe 48, has NH and ^{15}N chemical shifts degenerate with those of His 94 and therefore was not identified initially in the SQC spectrum). These protons were also later observed in a COSY spectrum recorded at 314 K. This gave a total of 96 ^{15}N-NH-$C^\alpha H$ spin systems. The amide protons belonging to Tyr 82 and Gly 86 were unambiguously determined only during the sequential assignment procedure. The amide proton of Ala 84 was not determined. Examination of relay COSY, double-relay COSY, and TOCSY spectra completed the amide-based spin subsystem search. DQF-COSY, relay COSY, and TOCSY spectra recorded in D_2O provided side-chain-based spin subsystem assignments. Classification of the spin systems into particular residue types and assignment of the various side-chain resonances are detailed below. Residue specific assignments are listed in Table I.

The 12 methyl resonances of the five alanines and seven threonines in FKBP were initially identified by their characteristic strong cross-peaks to $C^\alpha H/C^\beta H$, respectively, in DQFC spectra. Six of these protons were easily assigned to threonine residues on the basis of strong $C^\gamma H_3$–$C^\alpha H$ peaks in relay COSY spectra. The last Thr side chain (Thr 14) was more difficult to identify due to near-degeneracy of its $C^\alpha H$ and $C^\beta H$ resonances. However, tentative assignment was made at this stage based on distortion of the $C^\gamma H_3$–$C^\alpha H$ cross-peak in relay COSY spectra, suggesting overlap of two peaks ($C^\gamma H_3$–$C^\alpha H$ and $C^\gamma H_3$–$C^\beta H$), and on the presence of a peak on the pseu-

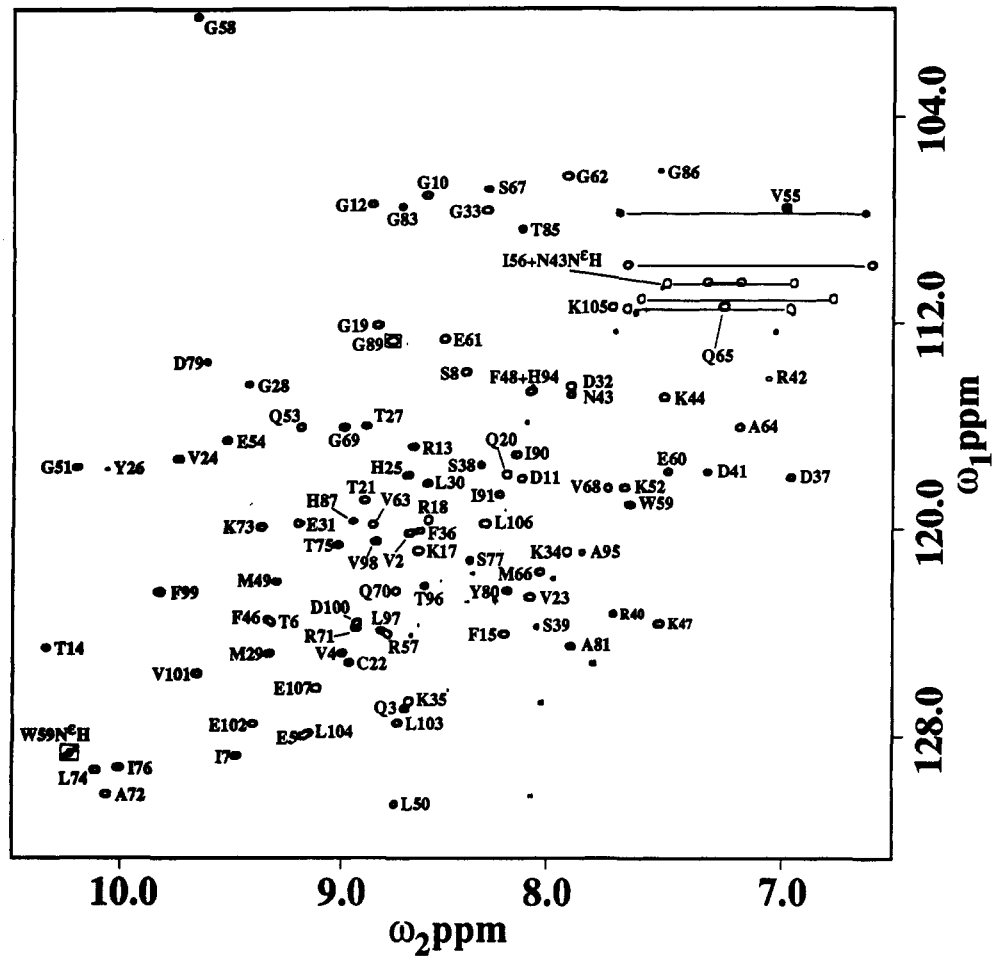

FIGURE 2: SQC spectrum of uniformly ^{15}N-labeled FKBP in H_2O buffer showing amide proton to amide nitrogen correlation. Side-chain amide protons are joined by lines. Boxed cross-peaks are shown at a lower contour level than that used for the remainder of the spectrum.

dodiagonal of a DQ spectrum ($\omega_1 = 2\omega_2$) at the ω_2 chemical shift of the $C^{\alpha}H$ resonance. This assignment was subsequently confirmed through sequential assignment. The five remaining methyl resonances were assigned to alanines by default. All alanine and threonine side chains except that of Ala 84 (discussed later) were easily linked to their respective amide protons through strong NH–CH_3 peaks in relay COSY, double-relay COSY, and TOCSY spectra recorded in H_2O.

Glycine spin systems were identified by characteristic NH–$C^{\alpha}H$ peaks in COSY spectra and by high-field ^{15}N chemical shifts in heteronuclear spectra recorded in H_2O (Figure 2). All 12 expected glycine amides could be seen in SQC spectra. Eleven of these showed at least one of the expected two NH–$C^{\alpha}H$ peaks in the fingerprint region of the COSY spectrum. The 12th, Gly 86, had only a weak, broad peak in the SQC and HMQC spectra and showed no COSY cross-peaks. Its identity was not confirmed until sequential assignment. In all glycine spin systems other than Gly 86, both NH–$C^{\alpha}H$ peaks were seen in TOCSY spectra recorded in H_2O. Gly 58 also showed an additional NH–$C^{\alpha}H$ cross-peak in TOCSY spectra recorded with long mixing times; this anomaly is discussed later. The identities of all glycine residues except Gly 86 and Gly 58 were confirmed through observation of characteristic $C^{\alpha}H$–$C^{\alpha}H$ cross-peaks in DQFC spectra recorded in D_2O and of remote cross-peaks at ω_2 = NH and $\omega_1 = C^{\alpha 1}H + C^{\alpha 2}H$ in DQ spectra recorded in H_2O (Rance et al., 1983).

The nine valine residues were identified similarly to alanine and threonine residues. Initially, 15 of the expected 16 pairs of characteristic valine and leucine methyl cross-peaks ($C^{\gamma}H_3$–$C^{\beta}H$ for Val and $C^{\delta}H_3$–$C^{\gamma}H$ for Leu) were tentatively identified in the high-field region of a DQFC spectrum recorded in D_2O. Eight of these pairs showed strong relay peaks to the $C^{\alpha}H$ region in a relay COSY spectrum and were assigned as valine protons. Seven of these side chains were then linked in a straightforward manner to amide protons showing correct $C^{\alpha}H$, $C^{\beta}H$, and $C^{\gamma}H_3$ correlations in TOCSY spectra recorded in H_2O. The eighth pair of methyl protons (those of Val 68) could not be linked to an amide proton by use of COSY-type spectra due to attenuation of coherence transfer by a small $^3J_{N\alpha}$ coupling. Similarly, the ninth pair of valine methyl protons (those of Val 55) could not be linked to any amide-based spin system due to small $^3J_{\alpha\beta}$. These last two spin systems were completed through strong NOESY cross-peaks from NH and $C^{\alpha}H$ to all side-chain protons after sequential assignment had identified the backbone protons as belonging to valines.

The seven leucines proved more difficult to assign than the valines because correlations from leucine amide protons did not extend beyond the $C^{\beta}Hs$ in relay COSY and TOCSY spectra in most cases. Also, correlations were not generally observed from the α protons to the methyl protons, even in D_2O TOCSY spectra recorded with long mixing times. Hence, for these residues we also used NOESY data to link the amide-based side-chain-based spin subsystems. For Leu 74 and Leu 106, correlations were observed between the NH and all side-chain protons in H_2O TOCSY spectra. These were the only two leucines definitively assigned before sequential assignment. The five other leucine residues, Leu 30, Leu 50, Leu 97, Leu 103, and Leu 104, showed correlations from the

Table I: ^{1}H and ^{15}N Chemical Shifts[a] of Recombinant Human FKBP

residue	amide ^{15}N	amide ^{1}H	$C^{\alpha}H$	$C^{\beta}H$	other
Gly 1			3.42, 3.39		
Val 2	119.9	8.68	5.28	1.97	$C^{\gamma}H$ 1.06, 0.970
Gln 3	126.8	8.72	4.82	2.08, 2.24	$C^{\gamma}H$ 2.44; $N^{\epsilon 1}H$ 6.97, 7.71; $N^{\epsilon 1}$ 111.6
Val 4	124.7	8.99	4.67	2.17	$C^{\gamma}H$ 0.95, 0.78
Glu 5	128.0	9.17	4.95	2.09, 2.24	
Thr 6	123.4	9.31	4.04	4.23	$C^{\gamma}H$ 1.25
Ile 7	128.6	9.47	4.21	1.42	$C^{\gamma 2}H$ 1.00; $C^{\gamma 1}H$ 1.70 1.18; $C^{\delta}H$ 0.86
Ser 8	113.6	8.44	5.13	4.03, 3.94	
Pro 9			4.47	2.43, 2.23	$C^{\gamma}H$ 2.17, 2.07; $C^{\delta}H$ 3.86, 3.94
Gly 10	106.7	8.61	4.29, 3.66		
Asp 11	117.8	8.18	4.43	3.10, 2.85	
Gly 12	107.0	8.86	3.63, 4.13		
Arg 13	116.5	8.69	4.85	1.68, 1.87	$C^{\gamma}H$ 1.56, 1.68; $C^{\delta}H$ 3.29, 3.28; $C^{\zeta}H$ 7.27; N^{ϵ} 83.9
Thr 14	124.4	10.34	4.35	4.35	$C^{\gamma}H$ 1.23
Phe 15	123.8	8.26	5.39	3.21, 2.74	$C^{\delta}H$ 7.18; $C^{\zeta}H$ 7.18; $C^{\xi}H$ 7.18
Pro 16			4.49		$C^{\gamma}H$ 1.93; $C^{\delta}H$ 4.45, 4.17
Lys 17	120.6	8.65	4.72	1.67	$C^{\gamma}H$ 1.56, 1.13; $C^{\delta}H$ 1.85, 1.86; $C^{\zeta}H$ 3.14
Arg 18	119.4	8.61	3.76	1.67, 1.67	$C^{\gamma}H$ 1.64, 1.83; $C^{\delta}H$ 3.35, 3.35; $C^{\zeta}H$ 7.46; N^{ϵ} 83.5
Gly 19	111.8	8.83	4.53, 3.79		
Gln 20	117.6	8.26	4.65	2.43, 2.01	$C^{\gamma}H$ 2.43; $N^{\epsilon 2}H$ 6.59, 7.69; $N^{\epsilon 2}$ 109.9
Thr 21	118.6	8.89	4.60	4.03	$C^{\gamma}H$ 0.95
Cys 22	124.9	8.97	4.41	1.93, 1.82	
Val 23	122.4	8.14	4.74	2.13	$C^{\gamma}H$ 0.85, 0.73
Val 24	116.9	9.75	6.00	2.51	$C^{\gamma}H$ 1.40, 1.04
His 25	117.7	8.70	6.00	3.00, 2.84	$C^{\delta 2}H$ 8.39; $C^{\epsilon 1}H$ 7.27
Tyr 26	117.8	10.07	6.60	3.03, 3.13	$C^{\gamma}H$ 7.13; $C^{\epsilon}H$ 6.72
Thr 27	115.7	8.89	4.68	4.12	$C^{\gamma}H$ 1.35
Gly 28	114.1	9.43	4.19, 2.31		
Met 29	124.6	9.33	5.39	2.07, 1.97	$C^{\gamma}H$ 2.49, 2.18
Leu 30	118.0	8.61	4.91	2.34, 2.34	$C^{\gamma}H$ 2.04; $C^{\delta}H$ 0.94, 1.16
Glu 31	119.6	9.19	3.98	2.20, 2.24	$C^{\gamma}H$ 2.40, 2.42
Asp 32	114.1	7.96	4.57	3.15, 2.73	
Gly 33	107.3	8.34	4.42, 3.79		
Lys 34	120.7	7.96	4.25	2.03, 1.92	$C^{\gamma}H$ 1.55, 1.43; $C^{\delta}H$ 1.82 1.86; $C^{\zeta}H$ 3.15
Lys 35	126.5	8.68	4.65	1.83, 1.83	$C^{\gamma}H$ 1.42, 1.56; $C^{\delta}H$ 1.86 1.78; $C^{\zeta}H$ 3.08, 2.97
Phe 36	119.9	8.65	5.27	3.47, 2.78	$C^{\delta}H$ 7.10; $C^{\zeta}H$ 7.39; $C^{\xi}H$ 7.15
Asp 37	117.8	6.96	4.98	2.31, 3.50	
Ser 38	117.3	8.36	4.78	4.15, 3.73	
Ser 39	123.6	8.11	4.29	3.91, 5.15	
Arg 40	123.1	7.78	3.73	1.50, 1.62	$C^{\gamma}H$ 1.54, 1.62; $C^{\delta}H$ 2.96, 2.87; $C^{\zeta}H$ 7.11; N^{ϵ} 83.1
Asp 41	117.5	7.35	4.44	2.90, 2.74	
Arg 42	113.8	7.08	4.58	2.11, 2.14	$C^{\gamma}H$ 1.89, 1.81; $C^{\delta}H$ 2.75, 2.75; $C^{\zeta}H$ 7.72; N^{ϵ} 82.4
Asn 43	114.5	7.96	4.56	3.33, 2.83	$N^{\delta 2}H$ 6.96, 7.53; $N^{\delta 2}$ 110.6
Lys 44	114.6	7.53	5.06	1.95, 1.82	$C^{\gamma}H$ 1.61; $C^{\delta}H$ 1.88, 1.89; $C^{\zeta}H$ 3.20
Pro 45			4.08	1.54, 1.54	$C^{\delta}H$ 3.77, 3.72
Phe 46	123.3	9.33	5.08	3.34, 3.73	$C^{\delta}H$ 7.49; $C^{\zeta}H$ 7.24
Lys 47	123.5	7.55	5.72	1.59, 1.38	$C^{\gamma}H$ 1.17, 1.53; $C^{\delta}H$ 1.64, 1.64; $C^{\zeta}H$ 2.95, 2.76
Phe 48	114.3	8.14	4.85	3.03, 2.94	$C^{\delta}H$ 7.09; $C^{\zeta}H$ 7.16; $C^{\xi}H$ 7.03
Met 49	121.8	9.29	4.84	1.93, 1.97	$C^{\gamma}H$ 2.25
Leu 50	130.6	8.77	4.36	2.08, 1.49	$C^{\gamma}H$ 1.75; $C^{\delta}H$ 1.19, 0.86
Gly 51	117.3	10.19	4.47, 3.97		
Lys 52	118.1	7.72	4.59	2.23, 1.87	$C^{\gamma}H$ 1.83, 1.82; $C^{\delta}H$ 1.68, 1.45; $C^{\zeta}H$ 3.15, 3.15
Gln 53	115.8	9.18	4.10	2.32, 2.37	$C^{\gamma}H$ 2.40; $N^{\epsilon 2}H$ 6.77, 7.63; $N^{\epsilon 2}$ 111.1
Glu 54	116.3	9.52	4.36	2.30, 2.22	$C^{\gamma}H$ 2.41
Val 55	107.4	6.98	4.46	1.59	$C^{\gamma}H$ 0.26, 0.21
Ile 56	110.6	7.52	3.99	2.31	$C^{\gamma 2}H$ 0.36; $C^{\gamma 1}H$ 0.96, 1.27; $C^{\delta}H$ 0.87
Arg 57	123.9	8.78	4.18		
Gly 58[b]	99.6	9.65	4.55, 4.09, 3.92		
Trp 59	118.8	7.70	4.38	2.93, 2.86	$C^{\delta 1}H$ 6.50; $N^{\epsilon 1}H$ 10.24; $N^{\epsilon 1}$ 128.4; $C^{\epsilon 3}H$ 6.75; $C^{\zeta}H$ 5.85; $C^{\zeta 2}H$ 6.32; $C^{\zeta 3}H$ 6.79
Glu 60	117.5	7.53	4.07	2.19, 2.30	$C^{\gamma}H$ 2.48
Glu 61	112.3	8.53	4.31	2.03, 2.27	$C^{\gamma}H$ 2.58, 2.39
Gly 62	105.9	7.99	3.86, 3.77		
Val 63	119.6	8.86	4.07	2.40	$C^{\gamma}H$ 1.22, 1.13
Ala 64	115.8	7.20	4.18	1.71	
Gln 65	111.0	7.28	4.52	2.27, 2.53	$C^{\gamma}H$ 2.68; $C^{\epsilon 2}H$ 7.19, 7.33; $N^{\epsilon 2}$ 110.5
Met 66	121.4	8.10	4.94	2.10, 2.27	$C^{\gamma}H$ 2.34, 2.89
Ser 67	106.4	8.34	4.97	3.90, 3.48	
Val 68	118.2	7.79	3.24	1.94	$C^{\gamma}H$ 1.00, 0.97
Gly 69	115.8	8.99	4.57, 3.97		
Gln 70	122.2	8.76	4.18	1.50, 1.93	$C^{\gamma}H$ 2.32; $N^{\epsilon 2}H$ 6.62, 7.73; $N^{\epsilon 2}$ 107.8
Arg 71	123.6	8.93	5.82	2.00, 1.69	$C^{\gamma}H$ 1.71, 1.54; $C^{\delta}H$ 3.12, 3.06; $C^{\zeta}H$ 9.15; N^{ϵ} 84.3
Ala 72	130.2	10.07	5.26	1.42	
Lys 73	119.7	9.37	5.40	1.91, 1.74	$C^{\gamma}H$ 1.28; $C^{\delta}H$ 1.42; $C^{\zeta}H$ 2.92, 3.12
Leu 74	129.2	10.12	5.63	1.93, 1.45	$C^{\gamma}H$ 1.78; $C^{\delta}H$ 0.82, 0.82
Thr 75	120.4	9.02	5.21	4.15	$C^{\gamma}H$ 1.21
Ile 76	129.1	10.01	4.72	1.95	$C^{\gamma 2}H$ 1.26; $C^{\gamma 1}H$ 1.66, 1.36; $C^{\delta}H$ 1.08
Ser 77	120.0	8.42	4.86	4.06, 4.51	

Table I (Continued)

residue	amide ^{15}N	amide ^{1}H	$C^{\alpha}H$	$C^{\beta}H$	other
Pro 78					$C^{\delta}H$ 4.07, 3.92
Asp 79	113.1	9.61	4.55	2.93, 2.84	
Tyr 80	122.2	8.26	4.55	3.08, 2.93	$C^{\delta}H$ 6.82; $C^{\epsilon}H$ 6.95
Ala 81	124.3	7.96	4.56	1.55	
Tyr 82		9.24	4.72	3.28, 2.77	$C^{\delta}H$ 7.21; $C^{\epsilon}H$ 6.70
Gly 83	107.1	8.73	4.07, 3.75		
Ala 84[b]			3.58, 3.66	1.56	
Thr 85	108.0	8.19	4.23	4.38	$C^{\gamma}H$ 1.37
Gly 86	105.6	7.56	3.83, 3.66		
His 87	119.4	8.93	5.13	2.82, 2.25	
Pro 88			4.32	1.98, 2.35	$C^{\gamma}H$ 1.58, 1.58; $C^{\delta}H$ 3.51, 3.66
Gly 89	112.8	8.77	4.37, 3.75		
Ile 90	116.9	8.19	4.41	1.58	$C^{\gamma 2}H$ 1.00; $C^{\gamma 1}H$ 1.12 1.44; $C^{\delta}H$ 0.86
Ile 91	118.3	8.27	4.60	1.42	$C^{\gamma 2}H$ 0.99; $C^{\gamma 1}H$ 0.50, 0.88; $C^{\delta}H$ 0.01
Pro 92			4.82	2.51, 1.99	$C^{\gamma}H$ 2.07; $C^{\delta}H$ 3.88, 3.46
Pro 93			3.84		$C^{\delta}H$ 3.60, 3.89
His 94	114.3	8.14	4.00	3.40, 3.40	
Ala 95	120.7	7.91	4.62	1.37	
Thr 96	119.4	8.61	4.75	4.10	$C^{\gamma}H$ 1.17
Leu 97	123.7	8.82	5.24	1.89, 1.66	$C^{\gamma}H$ 1.96; $C^{\delta}H$ 1.33, 0.95
Val 98	120.3	8.85	5.39	1.87	$C^{\gamma}H$ 1.03, 0.95
Phe 99	122.2	9.83	5.92	2.89, 2.82	$C^{\delta}H$ 7.22; $C^{\epsilon}H$ 7.36; $C^{\zeta}H$ 7.75
Asp 100	123.5	8.93	5.30	2.94, 2.60	
Val 101	125.4	9.66	5.18	1.90	$C^{\gamma}H$ 0.84, 0.81
Glu 102	127.6	9.41	5.54	1.75, 2.17	$C^{\gamma}H$ 2.13, 1.98
Leu 103	127.4	8.75	4.81	2.40, 1.18	$C^{\gamma}H$ 1.44; $C^{\delta}H$ 0.91, 0.77
Leu 104	127.8	9.15	4.33	1.64, 1.58	$C^{\gamma}H$ 1.83; $C^{\delta}H$ 0.93, 0.80
Lys 105	111.1	7.78	4.47	1.84, 1.88	$C^{\gamma}H$ 1.33, 1.30; $C^{\delta}H$ 1.72, 1.73; $C^{\epsilon}H$ 3.05, 3.02
Leu 106	119.5	8.35	5.36	1.67, 1.31	$C^{\gamma}H$ 1.48; $C^{\delta}H$ 0.98, 0.77
Glu 107	126.0	9.11	4.45	1.79, 2.13	$C^{\gamma}H$ 2.09, 2.02

[a] Chemical shifts are reported for FKBP at pH 5.0 and 300 K in 25 mM sodium acetate-d_3 buffer. Proton chemical shifts are referenced to internal TSP and are accurate to ±0.02 ppm. Nitrogen chemical shifts are referenced to external liquid NH_3 and are accurate to ±0.2 ppm. [b] These residues undergo slow exchange.

NH to $C^{\alpha}H$ and both $C^{\beta}Hs$ only. These amide-based spin systems were linked to previously identified methyl-based spin systems through NOEs from the amide proton to at least two of the three protons $C^{\delta 1}H_3$, $C^{\delta 2}H_3$, and $C^{\gamma}H$. In all five cases, side-chain assignment was completed after sequential assignment had established the residues as leucines.

Prior to sequential assignment, only one of the five isoleucines was completely assigned. This residue, Ile 91, showed relayed connectivities from the amide proton to all side-chain protons in H_2O TOCSY spectra. The $C^{\beta}H$ and $C^{\gamma 2}H_3$ resonances of three other isoleucines, Ile 7, Ile 76, and Ile 90, were linked to NH–$C^{\alpha}H$ resonances on the basis of $C^{\alpha}H$–$C^{\gamma 2}H_3$ peaks in relay COSY spectra, and NH–$C^{\beta}H$/$C^{\gamma 2}H_3$ correlations in H_2O relay COSY, double-relay COSY, and TOCSY spectra. The $C^{\beta}H$ and $C^{\gamma 2}H_3$ resonances of the last isoleucine were identified on the basis of strong NOESY cross-peaks to the NH and $C^{\alpha}H$ protons only after sequential assignment had clearly established the residue as Ile 56. The $C^{\delta}H_3$/$C^{\gamma 1}H_2$ protons of these last four isoleucines were initially found in DQ spectra, where in all cases remote peaks at $\omega_2 = C^{\delta}H_3$ and $\omega_1 = C^{\gamma 1,1}H + C^{\gamma 1,2}H$ were observed in addition to direct peaks between the $C^{\gamma 1}H$ and $C^{\delta}H$ protons. After sequential assignment, these methyl-based spin subsystems were linked to the amide-based spin subsystems by use of NOEs from the NH and $C^{\alpha}H$ protons.

Arginine and lysine spin system assignments were determined as follows. First, amide-based spin subsystem assignments were made by use of H_2O TOCSY and H_2O relay and double-relay COSY spectra. Of particular importance here was the detection of cross-peaks between backbone NH and $C^{\delta}H$ (for arginine) or NH and $C^{\epsilon}H$ (for lysine), in the range of 2.50–3.40 ppm. Additional assignments and confirmation of the amide-based subassignment depended on establishing correlations to unique $C^{\beta}H$ and $C^{\gamma}H$ chemical shifts from the $C^{\delta}H$ of arginine and on correlations to unique $C^{\beta}H$, $C^{\gamma}H$, and $C^{\delta}H$ chemical shifts from the $C^{\epsilon}H$ of lysine. Many of the $N^{\epsilon}H$ and $C^{\delta}H$ chemical shifts of the lysine and $C^{\delta}H$ and $C^{\gamma}H$ chemical shifts of the arginines were degenerate; however, it was possible to dissect these in D_2O DQ spectra on the basis of two direct cross-peaks from $C^{\epsilon}H$ chemical shifts to $C^{\delta}H$ chemical shifts for lysine and on the basis of the same pattern from $C^{\delta}H$ to $C^{\gamma}H$ chemical shifts of arginine (Chazin et al., 1987). Arginine chemical shifts were distinguished from those of lysine by the correlation of $N^{\epsilon}H$ to $C^{\beta}H$, $C^{\gamma}H$, and $C^{\delta}H$ chemical shifts in H_2O COSY, TOCSY, relay COSY, double-relay COSY, and DQ spectra. Finally, arginine N^{ϵ} chemical shifts were determined by the identification of characteristic $N^{\epsilon}H$–$^{15}N^{\epsilon}$ cross-peaks in HMQC and SQC spectra (Muchmore et al., 1989). Furthermore, $N^{\epsilon}H$–$C^{\delta}H$ correlations were confirmed from HMQC-NOESY spectra.

Proline spin systems were only identified during sequential assignment. Chemical shifts of six of the seven proline $C^{\alpha}Hs$ were determined through $C^{\alpha}H_i$–NH_{i+1} NOEs (Figure 3B,C). Chemical shifts for all seven proline $C^{\delta}H$ pairs were assigned on the basis of $C^{\alpha}H_i$–$C^{\delta}H_{i+1}$ NOEs (Figure 7B). Some proline $C^{\beta}Hs$ and $C^{\gamma}Hs$ were determined late in the assignment procedure on the basis of correlations from $C^{\alpha}H$ and $C^{\delta}Hs$ to remaining unassigned protons in TOCSY, COSY, and DQ spectra.

The 25 three-spin residues (Asp, Asn, Ser, Phe, Tyr, Cys, Trp, and His) were assigned through application of H_2O COSY, TOCSY, relay COSY, and double-relay COSY spectra. Complete NH-based assignments could be made for all 25 three-spin residues on the basis of NH–$C^{\alpha}H$ and NH–$C^{\beta}H$ connectivities. $C^{\alpha}H$–$C^{\beta}H$ connectivities were confirmed with D_2O DQF-COSY, relay COSY, and DQ data. The relay COSY and DQ experiments were particularly useful in determining connections to both $C^{\beta}Hs$ in cases where only one $C^{\alpha}H$–$C^{\beta}H$ cross-peak was observed due to weak coupling between the $C^{\alpha}H$ and the second $C^{\beta}H$. It should be stressed

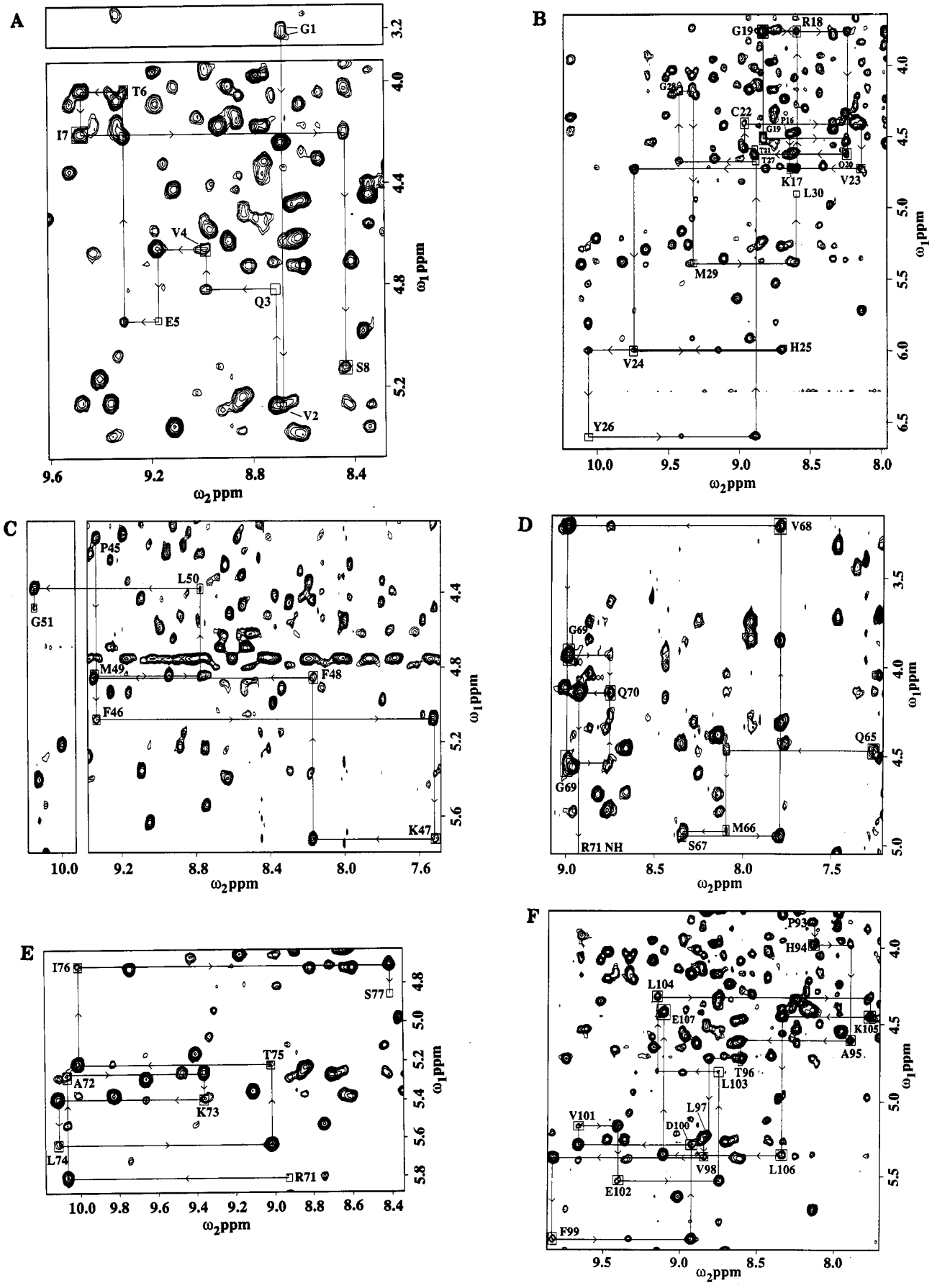

FIGURE 3: Fingerprint regions (NH–C$^{\alpha}$H) of 500-MHz ^{1}H NOESY spectra of FKBP. Sequential C$^{\alpha}$H$_i$–NH$_{i+1}$ connectivity data are shown for (A) Gly 1–Ser 8; (B) Pro 16–Leu 30; (C) Pro 45–Gly 51; (D) Gln 65–Gln 70; (E) Arg 71–Ser 77; and (F) Pro 93–Glu 107. Intraresidue C$^{\alpha}$H$_i$–NH$_i$ cross-peaks are boxed (in some cases these peaks are weak and can only be seen in a contour level lower than that shown). Arrows indicate the direction of the sequential connectivity from N to C terminus. Panels A–C, E, and F are taken from spectra recorded on a 4 mM solution of FKBP in H_2O buffer at pH 5.0 with τ_m = 200 ms. Probe temperature was 300 K in panels A, B, E, and F and was 314 K in C. Panel D is taken from a spectrum recorded at 298 K on a 1.5 mM sample of FKBP freshly dissolved in D_2O buffer at pH 5.0 (uncorrected) with τ_m = 200 ms.

that final confirmation of the three-spin systems required sequential assignment.

Identification of NH, $C^{\alpha}H$, and $C^{\beta}H$ for all Asn, His, Phe, Trp, five of the six Asp, and two of the three Tyr residues was completed as described above. However, Cys 22, Asp 32, and Tyr 82 proved more difficult. These specific cases are described below. The spin system of Cys 22 was originally assigned to either a five-spin or an aliphatic residue because of the unusually high-field chemical shifts of the $C^{\beta}Hs$. However the spin system was assigned to the Cys 22 position during sequential assignment on the basis of an unambiguous $C^{\alpha}H_i$–NH_{i+1} NOE to Val 23 and an NH_i–$C^{\alpha}H_{i-1}$ NOE to Thr 21. The Asp 32 spin system was difficult to resolve due to near degeneracy of both its NH and $C^{\alpha}H$ chemical shifts with those of Ala 81 and Asp 34. However, this spin system was established as unique during the sequential assignment stage and confirmed by use of HMQC, HMQC-COSY, and HMQC-NOESY spectra. The spin system for Tyr 82 was particularly difficult to assign, as its amide proton was not seen in any of the COSY-type or heteronuclear experiments recorded in H_2O. Its $C^{\alpha}H$ and both $C^{\beta}Hs$ were tentatively established through strong NOEs from the tyrosine ring $C^{\delta}H$. The $C^{\alpha}H$–$C^{\beta}H$ connections were then confirmed through correlations seen in DQF-COSY, DQ, relay COSY, and TOCSY spectra recorded in D_2O. The amide proton of Tyr 82 was not identified until sequential assignment, as will be discussed later.

Two of the serine spin systems, Ser 8 and Ser 38, were easily distinguished from the other three-spin systems by their low-field $C^{\beta}H$ chemical shifts. One other serine, Ser 77, was more difficult to identify, as its $C^{\alpha}H$ was coincident with the irradiated water signal at 300 K. However, at 308 K this signal was easily observed, and all protons in the spin system were assigned unambiguously. The Ser 39 and Ser 67 were not assigned until the sequential assignment stage, as both showed no NH–$C^{\beta}H$ and only weak $C^{\alpha}H$–$C^{\beta}H$ connections in DQF-COSY, DQ, relay COSY, and TOCSY spectra. During sequential assignment, NH–$C^{\beta}H$ connectivities were observed for these two serines in H_2O NOESY spectra. It was then possible to correlate these assignments with the weak connectivities observed in the COSY-type spectra.

Aromatic side-chain chemical shift assignments were completed by establishing characteristic correlations for the Tyr, Phe, and Trp ring protons in the aromatic region of D_2O DQF-COSY, relay COSY, TOCSY, and DQ spectra. Once the aromatic ring protons were assigned, connections to an amide-based three-spin system were achieved with D_2O NOESY spectra. Here, connectivities to $C^{\alpha}H$ and $C^{\beta}H$ chemical shifts were established.

Tyrosine $C^{\delta}H$ and $C^{\epsilon}H$ chemical shifts were identified by strong correlation cross-peaks in DQF-COSY and DQ spectra. Phe $C^{\delta}H$, $C^{\epsilon}H$, and $C^{\zeta}H$ protons were primarily assigned in the DQ spectra on the basis of direct connectivities between $C^{\delta}H$ and $C^{\epsilon}H$ and between $C^{\epsilon}H$ and $C^{\zeta}H$ and remote connectivities at $\omega_2 = C^{\epsilon}H$ and $\omega_1 = C^{\delta}H + C^{\zeta}H$. All Tyr and Phe ring connectivities were established with the above procedure except for Tyr 80, where the usual low- to high-field relationship of the $C^{\delta}H$ and $C^{\epsilon}H$ chemical shifts was reversed. As a result, NOESY cross-peaks between $C^{\delta}H$ and $C^{\beta}H$ chemical shifts, normally observed from the lower field of the two tyrosine ring proton chemical shifts, were observed to arise from the higher of the two.

Tryptophan indole protons are usually assigned on the basis of $N^{\epsilon 1}H$–$C^{\delta}H$ cross-peaks in H_2O COSY spectra and $N^{\epsilon 1}H$–$C^{\zeta 2}H$ cross-peaks in H_2O NOESY spectra. However, the $N^{\epsilon 1}H$–$C^{\delta 1}H$ cross-peak for the single Trp in FKBP was not observed in any homonuclear spectra. However, a cross-peak observed in HMQC, SQC, and HMQC-NOESY spectra was tentatively assigned as $^{15}N^{\epsilon 1}$–$N^{\epsilon 1}H$ on the basis of low-field ^1H and ^{15}N chemical shifts. This cross-peak was very broad in both dimensions (^1H 45 Hz; ^{15}N 40Hz) and weak (Figure 2), explaining the absence of cross-peaks arising from this chemical shift in ^1H–^1H correlation spectra. Assignment of the ring protons was made only after sequential assignment through NOEs to both $C^{\beta}Hs$ and $C^{\alpha}H$ from an unassigned singlet resonance at 6.50 ppm ($C^{\delta}H$) and an unassigned aromatic resonance at 6.77 ppm ($C^{\zeta 3}H$). The remaining ring protons were assigned through connections to $C^{\zeta 3}H$ in DQ and TOCSY spectra.

All three sets of histidine $C^{\delta 2}H$ and $C^{\epsilon 1}H$ cross-peaks, including one $C^{\delta 2}H$ that may be in slow exchange (see below), were observed in H_2O COSY and TOCSY spectra, but it was not possible to connect the $C^{\delta 2}H$ to the proximal $C^{\beta}Hs$ in H_2O or D_2O NOESY spectra. However, the $C^{\delta 2}H$ and $C^{\epsilon 1}H$ of His 25 were assigned during the determination of secondary structure from cross-strand interresidue NOEs between its $C^{\delta 2}H$ and the $C^{\alpha}Hs$ of Lys 44 and Pro 45 and to a $C^{\beta}H$ (3.73 ppm) of Phe 46.

Six of the 15 five-spin residues (Met 29, Glu 31, Glu 61, Met 66, Glu 102, and Glu 107) were completely assigned on the basis of connectivities to amide protons from side-chain protons in relay COSY, double-relay COSY, and TOCSY spectra recorded in H_2O. Of particular note was the use of relay and double relay spectra. Here, five-spin systems were distinguished from other aliphatic side-chain spin systems by the fact that the $C^{\gamma}H$ chemical shifts are often at slightly lower field than the $C^{\beta}H$ chemical shifts. Therefore, H_2O relay COSY spectra provided NH to high-field $C^{\beta}H$ (1.75–2.2 ppm) connectivities and double-relay COSY spectra provided NH–$C^{\gamma}H$ connectivities (about 1.9–2.5 ppm). These NH-based assignments were confirmed with D_2O DQF-COSY, relay COSY, TOCSY, and DQ spectra. In each case, the $C^{\beta}H$ resonances were distinguished from $C^{\gamma}H$ resonances on the basis of direct and remote cross-peaks to $C^{\alpha}H$ in DQ spectra.

For seven additional five-spin residues (Gln 3, Met 49, Gln 53, Glu 54, Glu 60, Gln 65, and Gln 70), only three of the four expected side-chain protons could be found from correlations to amide protons. In each case, direct and remote connectivities to $C^{\alpha}H$ in the DQ spectra allowed identification of the 2 $C^{\beta}Hs$, establishing the missing proton as a $C^{\gamma}H$.

One five-spin residue, Gln 20, shows only two side-chain protons in all spectra recorded in both H_2O and D_2O. Direct and remote peaks in DQ spectra have established that these resonances are the $C^{\beta}H$ protons. However, TOCSY, relay COSY, and NOESY spectra in D_2O show that the peak from the lower field resonance to $C^{\alpha}H$ is exceptionally large. In addition, this proton (but not the proton at higher field) shows strong NOEs to two side-chain terminal NH_2 protons (identified in HMQC and SQC spectra). This would only be expected of the $C^{\gamma}H$ resonances. Thus, for this spin system we have assigned a single $C^{\gamma}H$ resonance as degenerate with the low-field $C^{\beta}H$.

Assignment of the final five-spin residue, Glu 5, was complicated by near degeneracy of the amide proton with those of two other five-spin residues, Glu 31 and Gln 53, and of the α proton with that of another five-spin residue, Met 66. Because of these degeneracies, only the $C^{\beta}Hs$ have been definitively identified for this spin system.

Side-chain terminal amide protons of Asn 43 and the five glutamine residues were identified on the basis of their

characteristic cross-peak shapes (due to residual passive ^1H–^1H geminal coupling in ω_1; Clore & Gronenborn, 1989) in HMQC and SQC spectra recorded in H_2O. These were connected to $C^\beta H$ and $C^\gamma Hs$ through cross-peaks in NOESY and HMQC-NOESY spectra.

Terminal methyl groups of the three methionine residues, which are usually determined through NOEs to $C^\gamma H$, could not be definitively identified due to severe overlap in the aliphatic regions of the NOESY spectra.

Sequential Assignment. Peptide segments with extended conformation were easily identified on the basis of strong $C^\alpha H_i$–NH_{i+1} (2.2-Å) and medium intraresidue $C^\alpha H_i$–NH_i (2.9-Å) NOE cross-peaks in both the NOESY and HMQC-NOESY spectra. Sequential assignment of helical residues was based on medium NH_i–NH_{i+1} (2.8-Å) and strong intraresidue $C^\alpha H_i$–NH_i (2.2-Å) NOEs. Sequential assignment of loop regions, of which there were several in FKBP, was based on combinations of $C^\alpha H_i$–NH_{i+1} and NH_i–NH_{i+1} NOEs. In all cases, amide proton degeneracy was resolved by use of HMQC-type spectra. Coincidence of α protons with the irradiated H_2O signal was overcome either by recording spectra on protein freshly dissolved in D_2O, where some amide protons are unexchanged, or by recording spectra at different temperatures.

The $C^\alpha H_i$–NH_{i+1} connectivities used to assign segment Gly 1–Ser 8 are shown in Figure 3A. The NOEs between Gln 3 and Val 4 and between Glu 5 and Thr 6 are attenuated due to partial saturation of $C^\alpha H$ of Gln 3 and $C^\alpha H$ of Glu 5, respectively, by water irradiation and are not visible in Figure 3A. These peaks appear with much greater intensity in spectra recorded at 298 and 308 K, respectively. In addition, the connections were confirmed by either $C^\beta H_i$–NH_{i+1} or weak NH_i–NH_{i+1} NOEs in NOESY spectra recorded with τ_m of 150 and 200 ms.

Pro 9 was assigned through NOEs from $C^\alpha H$ of Ser 8 to both $C^\delta Hs$ of Pro 9 (Figure 7B) and from NH of Gly 10 to $C^\alpha H$ of Pro 9. Assignment of residues Gly 10 to Phe 15 was aided by the fact that the Gly 10, Gly 12, and Thr 14 spin systems had been tentatively assigned at the spin system identification stage. Furthermore, a continuous set of both NH_i–NH_{i+1} and $C^\alpha H_i$–NH_{i+1} NOEs was observed through this sequence. Additional connectivities confirmed the sequential assignment, including two NH_i–NH_{i+2} NOEs between Asp 11 and Arg 13 and between Gly 12 and Thr 14, and a NH_i–NH_{i+3} between Asp 11 and Thr 14 (Figure 6). Phe 15 was connected to Pro 16 through NOEs from the Phe $C^\alpha H$ to both Pro 16 $C^\delta Hs$ (Figure 7B).

The NOE data for peptide segment Pro 16–Leu 30 are shown in Figure 3B. Two portions of this sequence were complicated by $C^\alpha H$ overlap. Degeneracy of the $C^\alpha H$ of Val 24 and the $C^\alpha H$ of His 25 was resolved by observation of NOEs from the NH of His 25 to all of the side-chain protons of Val 24 and also from the NH of Tyr 26 to the side-chain of His 25 in both NOESY and HMQC-NOESY spectra. Degeneracy of the $C^\alpha H$ of Gly 19 and the $C^\alpha H$ of Arg 18 was overcome by observation of NH_i–NH_{i+1} NOEs between Arg 18 and Gly 19 and between Gly 19 and Gln 20 and also of a weak nonsequential NH_i–NH_{i+2} NOE between Arg 18 and Gln 20. These additional NOEs suggest that this short segment may form a short loop or turn at the end of a more extended β strand (see below).

The sequence Glu 31–Lys 44 contains discontinuous segments of NH_i–NH_{i+1}, $C^\alpha H_i$–NH_{i+1}, and $C^\beta H_i$–NH_{i+1} connectivities (Figure 5). Leu 30 was initially connected to Glu 31 via a $C^\beta H_i$–NH_{i+1} connectivity; a $C^\alpha H_i$–NH_{i+1} cross-peak was only observed in a NOESY spectrum recorded at 314 K, as the Leu 30 $C^\alpha H$ was saturated by solvent irradiation in spectra recorded at lower temperatures. Connectivities from Glu 31 to Lys 34 were complicated by the degeneracy of the NH chemical shifts of Asp 32 and Lys 34. This difficulty was resolved in two ways. First, in a NOESY spectrum taken at 298 K, the NH chemical shifts were slightly different, allowing tentative assignment in this region. Second, in an HMQC-NOESY spectrum, the ^{15}N chemical shifts of Asp 32 and Lys 34 differ by more than 6 ppm, allowing unambiguous differentiation of NOEs to the two amide protons (Figure 5). Connectivities from Lys 34 to Ser 39 could be followed directly. However, connectivities from Ser 39 to Pro 45 were not simple. A weak $C^\alpha H$–NH cross-peak between Ser 39 and Arg 40 was observed in the NOESY spectrum taken at 308 K. However, unambiguous assignment was only achieved with the HMQC-NOESY spectrum where NH_i–NH_{i+1} cross-peaks were observed between Ser 39 and Arg 40, as well as between Arg 40 and Asp 41. Connectivities from Asp 41 to Pro 45 were also made by use of the HMQC-NOESY spectrum, as degeneracy of the NH chemical shift of Asn 43 with those of three other residues complicated assignment of Arg 42–Lys 44 in the NOESY spectra.

The assignment of residues Pro 45–Gly 51, shown in Figure 3C, was problematic due to near degeneracy of the $C^\alpha H$ of Phe 48 with the $C^\alpha H$ of Met 49, and coincidence of these resonances with the residual H_2O signal at 300 K. In a NOESY spectrum recorded at 314 K, however, both resonances are seen clearly and their connection can be established without difficulty. The assignments of Phe 48 and Met 49 were subsequently confirmed through observation of characteristic long-range cross-strand NOEs to residues in the sequence Thr 21–Val 23, known to be part of the central β sheet of the protein (see below and Figure 7).

Sequential assignment from Gly 51 to Ala 64 showed the longest continuous stretch of NH_i–NH_{i+1} connectivities in FKBP (Figure 4). Assignment of residues Gly 51–Val 55 was made on the basis of NH_i–NH_{i+1} and $C^\alpha H_i$–NH_{i+1} sequential NOEs and numerous NH_i–NH_{i+2} and $C^\alpha H_i$–NH_{i+3} NOEs as well. The only ambiguous connections were those to Val 55, as the amide-proton chemical shift of this residue is nearly degenerate with two other protons. However, correlations to Val 55 were made unambiguously with the H_2O HMQC-NOESY spectrum. Assignment of residues Ile 56–Trp 59 was complicated by difficulties in assigning the chemical shifts for these residues during spin system identification. $C^\alpha H_i$–NH_{i+1} cross-peaks for all of the residues, except that of Trp 59, were weak in the COSY spectra. Also, TOCSY and relayed COSY spectra showed no NH to side chain chemical shift correlations for Ile 56 and Arg 57, and Gly 58 showed an extra intraresidue $C^\alpha H_i$–NH_i cross-peak in a 100-ms TOCSY spectra, perhaps due to slow exchange. However, strong or medium-strength NOESY $C^\alpha H_i$–NH_{i+1} and confirming NH_i–NH_{i+1} cross-peaks from Val 55–Gly 58 were observed. In addition, $C^\beta H_i$–NH_{i+2} NOESY correlations were observed from Val 55 to Arg 57 and from the $C^\delta H$ of Ile 56 to the NHs of Arg 57 and Gly 58. Assignment of the sequence Glu 61–Ala 64 was routine because spin system assignments of Gly, Val, and Ala residues were virtually complete, the Gly-Val-Ala sequence was unique, and there was no chemical shift overlap for NH and $C^\alpha H$ protons for these residues.

Assignment of residues Gln 65–Gln 70 was complicated by the proximity of the $C^\alpha H$ of Met 66 and the $C^\alpha H$ of Ser 67 to the residual H_2O line. However, both the NH of Met 66 and the NH of Ser 67 can be seen in spectra of protein freshly

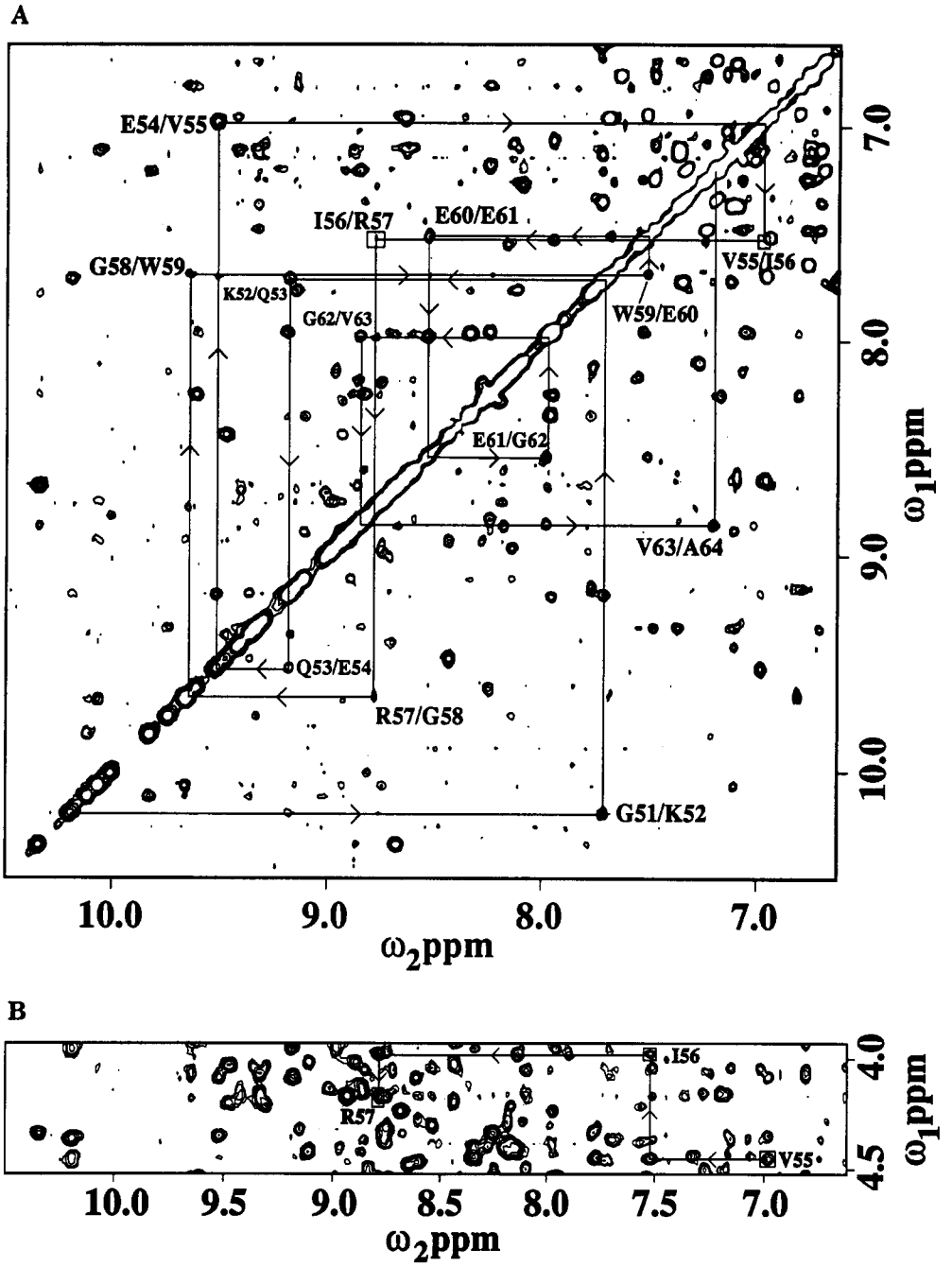

FIGURE 4: NH–NH region of ^{1}H NOESY spectrum presented in Figure 3A. Sequential NH_i–NH_{i+1} connectivities from Gly 51 to Ala 64 are shown. Panel B shows $C^\alpha H_i$–NH_{i+1} connectivities from Val 55 to Arg 57 in the fingerprint region of the same spectrum. Arrows indicate the direction of the sequential connectivity from N to C terminus.

dissolved in D_2O buffer, allowing clear assignment of the entire sequence. The relevant portion of a NOESY spectrum of such a sample is shown in Figure 3D.

The NOEs used to assign the sequence Arg 71–Ser 77 are shown in Figure 3E. The series of strong $C^\alpha H_i$–NH_{i+1} connections was one of the easiest to follow in FKBP, as the NH and $C^\alpha H$ resonances of all residues fall in the least crowded regions of the spectrum. In addition, all connections were confirmed through $C^\beta H_i$–NH_{i+1} NOEs. The sequence is completed by NOEs from the $C^\alpha H$ of Ser 77 to both $C^\delta Hs$ of Pro 78 in a NOESY spectrum recorded in D_2O (see Figure 7B).

Initial assignment of Asp 79 to Ala 81 was made on the basis of NH_i–NH_{i+1} connectivities. However, it was not clear which of the two three-spin systems (Asp 79 or Tyr 80) was Tyr 80 because the low-field Tyr ring proton (assumed to be $C^\delta H$) showed only a single weak NOE to a chemical shift common to both residues. Strong NOE cross-peaks were observed from the high-field Tyr ring proton to all protons in the Tyr 80 spin system. This evidence, combined with observation of NOESY cross-peaks from the Ala 81 NH to the Tyr 80 $C^\beta Hs$ led to the conclusion that the order of the assignments was correct and that the aberrant pattern of NOESY cross-peaks from the Tyr aromatic protons mentioned earlier was caused by the reversal of the expected low- to high-field relationship of the $C^\delta H$ and $C^\epsilon Hs$ (Bundi & Wüthrich, 1979).

Assignment of the sequence Ala 81–Pro 88 was particularly troublesome for several reasons and therefore was not completed until late in the assignment process. As stated in the spin system identification section, an amide proton could not be definitively identified for Tyr 82 on the basis of any of the COSY-type or heteronuclear experiments. The $C^\alpha H$ and both $C^\beta Hs$ of this residue were unambiguously determined, however,

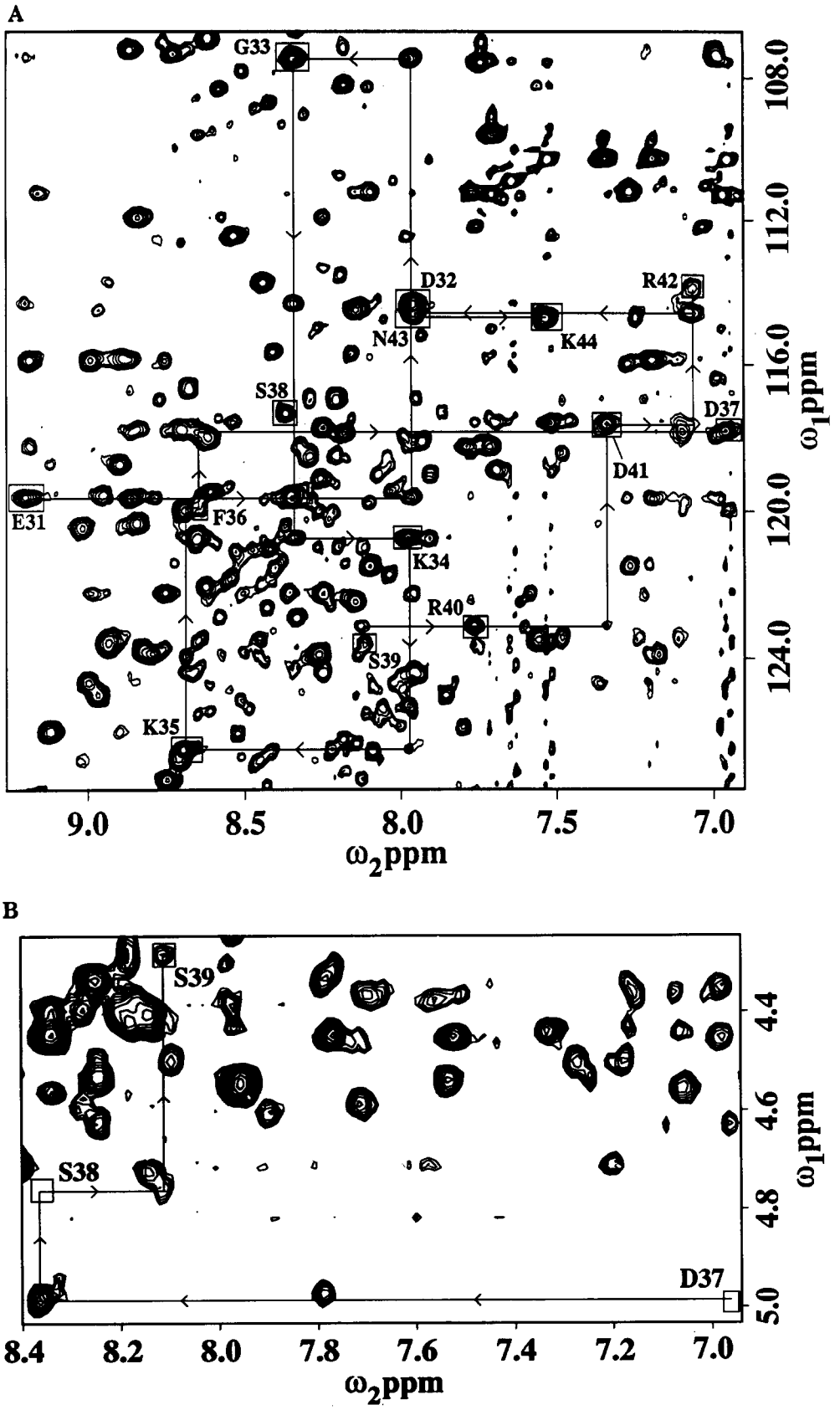

FIGURE 5: ^1H–^{15}N HMQC-NOESY spectrum (τ_m = 200 ms) of FKBP in H_2O buffer. NH–NH sequential correlations of residues Glu 31–Lys 44 are indicated. Intraresidue cross-peaks are boxed. Panel B shows $C^\alpha H_i$–NH_{i+1} connections from Asp 37 to Ser 39 observed in a NOESY spectrum of FKBP recorded at 300 K in H_2O buffer. Note that the ω_2 scales are different in the two panels.

through strong NOEs from the aromatic $C^\delta H$ and through evidence found in DQ, DQF-COSY, and TOCSY spectra performed on D_2O solutions. An amide proton at 9.24 ppm was established as that of Tyr 82 on the basis of extremely weak but reproducible NOEs from this resonance to $C^\alpha H$, both $C^\beta Hs$, and the $C^\delta H$ of the putative Tyr spin system and to $C^\alpha H$ and $C^\beta H$ of Ala 81. The connection to Gly 83 was established through medium $C^\alpha H_i$–NH_{i+1} and weak NH_i–NH_{i+1} NOEs to an unassigned glycine residue.

Assignment of the sequence Ala 84–Pro 88 was also difficult

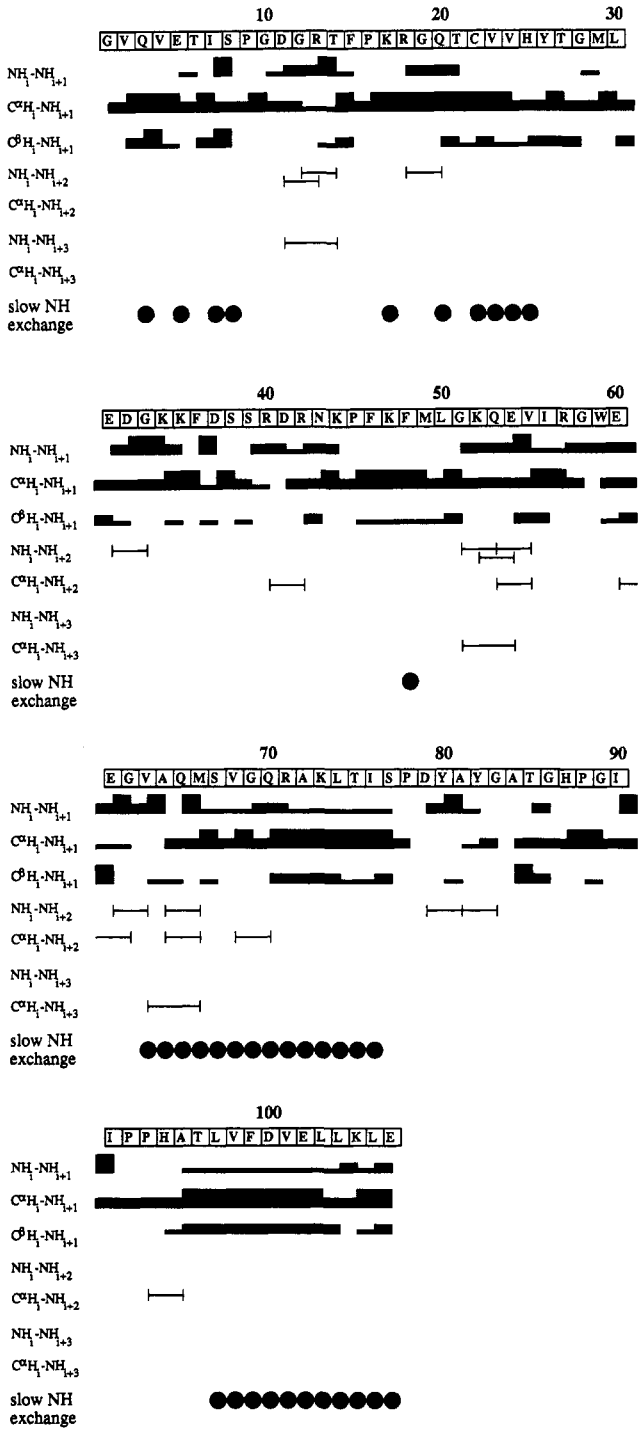

FIGURE 6: Summary of sequential NOESY correlations for FKBP. Intensities of sequential NOESY cross-peaks are indicated by the thickness of bars beneath the primary sequence and are classified as strong (triple), medium (double), and weak (single). Slowly exchanging amide protons, defined as those present in COSY spectra of FKBP 24 h after the protein is dissolved in D_2O buffer, are indicated by dots.

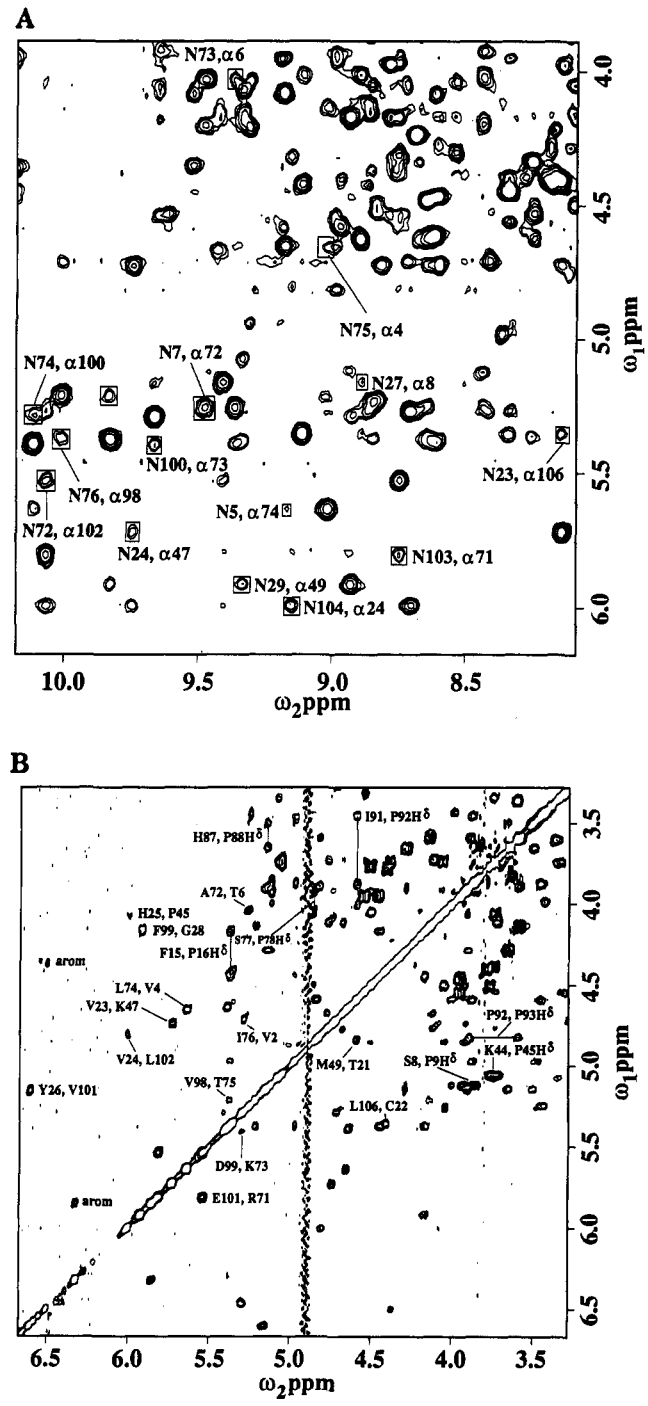

FIGURE 7: NOESY correlations indicative of antiparallel β sheet: (A) ^{1}H NOESY spectrum (τ_m = 200 ms) of FKBP in H_2O solution showing interstrand C$^\alpha$H–NH cross-peaks; (B) ^{1}H NOESY spectrum (τ_m = 100 ms) of FKBP in D_2O solution showing interstrand C$^\alpha$H–C$^\alpha$H correlations as well as sequential C$^\alpha$H$_i$–C$^\delta$H$_{i+1}$ proline correlations.

because two of the spin systems involved, Ala 84 and Gly 86, had not been determined unequivocally at the spin system identification stage. Sequential assignment through this region was based on several pieces of evidence. First, at this stage only one threonine and one alanine spin system remained unassigned. These, therefore, had to be Ala 84 and Thr 85. The Ala-Thr connection was established through weak C$^\alpha$H$_i$–NH$_{i+1}$ and medium C$^\beta$H$_i$–NH$_{i+1}$ NOEs between these residues. There were, however, no intraresidue NH–C$^\alpha$H or NH–C$^\beta$H cross-peaks observed in any spectra, other than the H_2O NOESY, for Ala 84. The absence of NH–C$^\alpha$H cross-peaks in the COSY and TOCSY spectra could be explained by the fact that the NH appeared to be undergoing slow to medium rate exchange between two distinct chemical shifts; evidence of this was the presence of a weak cross-peak between the NH chemical shift of the putative Ala and another NH chemical shift that did not correlate with any other spin system, observed only in TOCSY spectra recorded with a 100-ms mixing time. This exchange may also explain the lack of NOEs between Ala 84 and Gly 83. It must be noted that, in addition to showing NOESY cross-peaks to C$^\alpha$H and C$^\beta$H of Ala 84, the slow exchange chemical shifts (8.39 and 8.61 ppm)

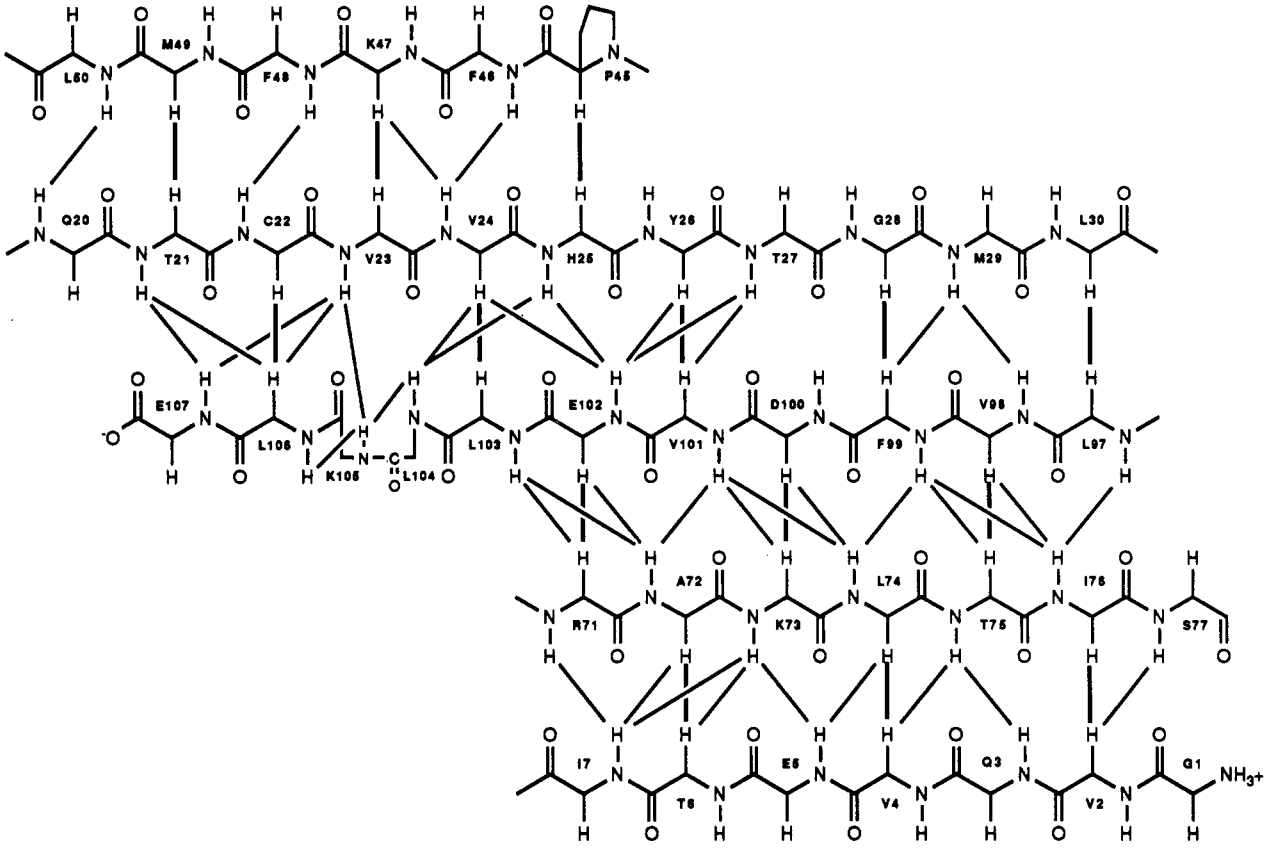

FIGURE 8: Antiparallel β sheet in FKBP deduced from interstrand NOEs indicated in Figure 7. Interstrand, NH–NH, NH-C$^{\alpha}$H, and C$^{\alpha}$H-C$^{\alpha}$H NOEs are indicated by lines.

show TOCSY cross-peaks to an unassigned His (either His 87 or His 94) side-chain C$^{\delta 2}$H proton at 7.30 ppm. It is therefore possible that these chemical shifts correspond to the C$^{\delta 2}$H of either His 87 or His 94 rather than the NH of Ala 84. Thus, the NH of Ala 84 is not definitively assigned.

Prior to sequential assignment, the only evidence for Gly 86 was a weak, broad peak with high-field ^{15}N chemical shift in HMQC and SQC spectra recorded in H_2O. No C$^{\alpha}$H protons could be found for this spin system in any of the COSY-type spectra. However, the Gly 86 NH showed NH$_i$–NH$_{i+1}$, C$^{\alpha}$H$_i$–NH$_{i+1}$, and C$^{\beta}$H$_i$–NH$_{i+1}$ NOEs to Thr 85 in NOESY spectra recorded in H_2O and was thus assigned to Gly 86. The C$^{\alpha}$H protons of this residue were assigned by observation of two additional NOESY cross-peaks from the amide proton to two unassigned C$^{\alpha}$Hs that showed NOEs to the NH of a three-spin residue. This residue was assigned as His 87 on the basis of this evidence and on strong C$^{\alpha}$H$_i$–C$^{\delta}$H$_{i+1}$ NOEs to Pro 88 in D_2O NOESY spectra (Figure 7B). All connections in the sequence Ala 81–Pro 88 were finally confirmed from the HMQC, HMQC-COSY, and HMQC-NOESY spectra, with the exception of those from the NH of Tyr 82, which was not seen in the heteronuclear experiments.

The longest sequence of extended strand in FKBP, Pro 93–Glu 107, appears in Figure 3F. The pattern of strong C$^{\alpha}$H$_i$–NH$_{i+1}$ and weak NH$_i$–NH$_{i+1}$ sequential NOEs is interrupted only at residue Leu 104, which shows medium NH$_i$–NH$_{i+1}$ and C$^{\alpha}$H$_i$–NH$_{i+1}$ NOEs to Lys 105. These data, along with a relatively strong intraresidue C$^{\alpha}$H$_i$–NH$_i$ cross-peak at Leu 104, suggest that there is probably a β bulge at this residue.

Secondary Structure Analysis (β Sheet). The main secondary structural element of FKBP is a five-stranded β sheet comprised of peptide segments Gly 1–Ile 7, Gln 20–Leu 30, Pro 45–Leu 50, Arg 71–Ser 77, and Leu 97–Glu 107. The

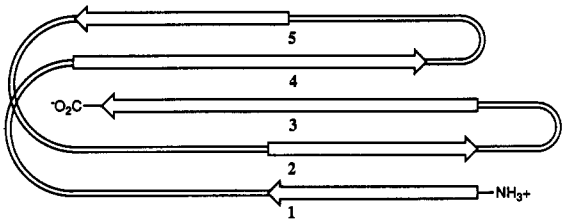

FIGURE 9: Topology of the FKBP fold.

expected sequential NOE connectivities for these residues (i.e., strong C$^{\alpha}$H$_i$–NH$_{i+1}$ and weak NH$_i$–NH$_{i+1}$) can be seen in Figure 6. The topology of the central sheet and the NOEs used to define it are shown in Figure 8. All of the strands are oriented in an antiparallel fashion as indicated by the pattern of strong interstrand C$^{\alpha}$H–C$^{\alpha}$H (Figure 7) and weaker NH–NH and NH–C$^{\alpha}$H NOEs (Wüthrich, 1986; Weber et al., 1987; Chazin & Wright, 1988); parallel β sheets would have similar patterns of interstrand NH–NH and interstrand NH–C$^{\alpha}$H NOEs but no strong C$^{\alpha}$H–C$^{\alpha}$H NOEs. The only break in these trans-strand interactions occurs at residue Leu 104, which is part of the β bulge discussed above. The connectivity pattern of the strands indicates that the sheet has a novel topology characterized by +3, +1, −3, +1 loop connections (Richardson, 1977) (Figure 9). One necessary feature of this strand topology is the crossing of the loops connecting the first and fourth and the fifth and second strands (i.e., the sequences Ser 8–Gly 19 and Gly 51–Gln 70). While such topological crossings have been observed for parallel sheets (e.g., glyceraldehyde phosphate dehydrogenase; Moras et al., 1975), there are no examples for antiparallel sheets. The absence of strand crossings has been discussed extensively (Richardson, 1977; Ptitsyn & Finkelstein, 1980; Chothia, 1984) and formulated as a rule of protein folding (Chirgadze,

1987). A thermodynamic rationale that has been applied to topological, but not to physical, crossings is based on the difficulty of packing a loop between a solvent-exposed loop and the hydrophobic core of the protein. The supposed difficulty arises from the requirement for satisfying the hydrogen-bond propensities of the peptide backbone amide and carbonyl groups. It has been observed (Baker & Hubbard, 1984) that essentially all such groups in proteins participate in hydrogen bonds. Sequences of a protein in which all residues have secondary structural hydrogen bonds (i.e., two β sheets, two helices, a helix and a sheet) can and do cross in many of the well-characterized patterns of super secondary structure (Richardson, 1977). If two loops are involved, the outer one can be solvent-exposed to satisfy its hydrogen-bond propensity but the inner one would require interactions with itself or with other parts of the protein. The two "loops" that cross (the 1-to-4 and 5-to-2 loops; Figure 9) in FKBP both contain secondary structural elements (see below), and the loop between 5 and 2 is partially helical (i.e., between residues 51 and 64). At this stage in the refinement it is not possible to determine exactly where loops contact each other and what types of contacts are involved.

Another argument against loop crossings is based on the desirability of simple folding pathways such as that for going from a β hairpin to a Greek key motif (Richardson, 1977; Ptitsyn & Finkelstein, 1980; Chothia, 1984). All of the β strands are connected by long interstrand regions (i.e., 12 residues between strands 1 and 4, 14 residues between strands 4 and 5, 20 residues between strands 5 and 2, and 19 residues between strands 2 and 3). Therefore, the loops may follow a circuitous pathway from one strand to the next, and a simple folding model may not be applicable to the proposed secondary structure of FKBP.

One additional striking feature of the β sheet is its amphiphilicity. The uniform chirality of natural amino acids dictates that side chains of residues in a β strand alternate direction; that is, they are oriented above and below the plane of a β sheet, which gives the structure two distinct faces (Creighton, 1984; Richardson & Richardson, 1989). One side of the FKBP sheet is hydrophobic, being composed entirely of aliphatic and aromatic side chains. The other side is hydrophilic, containing largely acidic, basic, and polar side chains. The pattern is most notable in strands 1–3. In fact, the β bulge at residue Leu 104 in strand 3 serves to maintain the correct register of the hydrophobic–hydrophilic alternation even through the Leu 103–Leu 104 dipeptide.

Secondary Structure Analysis (Loops). The peptide sequence Ser 8–Gly 19 forms the 12-residue +3 loop connecting the first and fourth strands of the β sheet. The strong NH_i–NH_{i+1} connectivity between Ile 7 and Ser 8 indicates that the peptide backbone changes direction sharply at Ile 7, curving away from the β sheet. A dipeptide conformation placing successive amide protons in close proximity is typical of residues at the ends of a β-strand. Such a conformation is often seen at the beginning of +3 loops between strands in Greek key β-barrel-type proteins (Richardson & Richardson, 1989). After Ser 8 the peptide continues in extended conformation, as evidenced by sequential $C^\alpha H_i$–NH_{i+1} connections to Asp 11, where the backbone again changes direction. The pattern of medium and strong NH_i–NH_{i+1} NOEs, weak $C^\alpha H_i$–NH_{i+1} NOEs, and weak medium-range NH_i–NH_{i+2} and NH_i–NH_{i+3} NOEs in the sequence Asp 11–Thr 14 indicates a type I β turn or possibly a single turn of 3_{10} helix in this region. Support for a turn-like conformation comes from the very low field chemical shift of the amide proton of Thr 14 (10.34 ppm) and the medium exchange rate of this proton with D_2O solvent (it can be seen clearly in a freshly prepared D_2O sample). Both facts indicate a hydrogen bond between Asp 11 and Thr 14, which would be expected in a β turn. The conformation of the turn does not, however, appear to be strictly classical, as the expected $C^\alpha H_i$–NH_{i+2} NOE between Gly 12 and Thr 14 is not observed. Medium NH_i–NH_{i+1} NOEs between Arg 18 and Gly 19 and between Gly 19 and Gln 20 and a weak NH_i–NH_{i+2} NOE between Arg 18 and Gln 20 also suggest a second β turn between Lys 17 and Gln 20. This turn, in contrast to the first, appears to be type II, on the basis of a characteristic strong $C^\alpha H_i$–NH_{i+1} NOE between Arg 18 and Gly 19. As before, however, the conformation does not appear to be strictly classical, as a $C^\alpha H_i$–NH_{i+2} NOE expected between Arg 18 and Gln 20 is not observed (Wüthrich, 1986).

The 14-residue loop from Glu 31 to Lys 44 shows only a single identifiable element of secondary structure. A turn is indicated from Leu 30 to Gly 33 by a strong NH_i–NH_{i+1} NOE between Asp 32 and Gly 33 and by a weak NH_i–NH_{i+2} NOE between Glu 31 and Gly 33. It is unclear, however, which type of turn is present, as the connection between the second and third residues, which is usually used to distinguish between type I and type II turns (Wüthrich, 1986) is made by both NH_i–NH_{i+1} and $C^\alpha H_i$–NH_{i+1} NOEs. Although the remainder of the loop contains several strong NH_i–NH_{i+1} NOEs, the lack of medium-range connections prevents a more detailed analysis at this point.

In the 20-residue loop from Gly 51 to Gln 70, the sequence consisting of residues Gly 51–Val 55 forms a single helical turn. This is followed by extended structure between Ile 56 and Arg 57 and then another helical turn between Glu 60 and Val 63. Evidence for this structure is based, first, on continuous NH_i–NH_{i+1} and NH_i–NH_{i+2} NOEs from Gly 51–Val 55 and Glu 60–Val 63 and, second, on small $^3J_{N\alpha}$ coupling constants through these sequences, as inferred from weak, nonsaturated NH–C^αH cross-peaks in H_2O COSY spectra (Pardi et al., 1984). Residues in the sequence Ile 56–Gly 58 also have small $^3J_{N\alpha}$ coupling constants; however, all of the sequential connectivities in this sequence are based on strong $C^\alpha H_i$–NH_{i+1} and weak NH_i–NH_{i+1} NOEs. In addition, the absence of NH_i–NH_{i+2} and $C^\alpha H_{i+2}$ NOEs suggests a more extended structure, possibly with poly(proline) conformation. The helical turn between Glu 60 and Val 63 is followed by a classical type I β turn between Val 63 and Met 66. Evidence for this structure is provided by the characteristic NH_i–NH_{i+1} NOEs from Ala 64 to Met 66, NH_i–NH_{i+2} and $C^\alpha H_i$–NH_{i+2} NOEs between Ala 64 and Met 66, and a $C^\alpha H_i$–NH_{i+3} NOE between Val 63 and Met 66. The sequence Ser 67–Gln 70 may form either a half-turn or a type II β turn as suggested by a medium-strength NH_i–NH_{i+1} NOE between Gly 69 and Gln 70, a strong $C^\alpha H_i$–NH_{i+1} NOE between Val 68 and Gly 69, and a weak $C^\alpha H_i$–NH_{i+2} NOE between Val 68 and Gln 70. In addition, a half-turn or type II β turn structure for this sequence is suggested by a small $^3J_{N\alpha}$ for Val 68 and a large $^3J_{N\alpha}$ for Gly 69.

The sequence including residues Asp 79–Gly 89 is the only sequence in FKBP where no periodic or semiperiodic structure was deduced from the available data. As discussed above, sequential assignment through this region required care, as three of the spin systems, Tyr 82, Ala 84, and Gly 86, could not be assigned initially; NH–C^αH cross-peaks were not observed for these residues in COSY and TOCSY spectra. Intraresidue NH–C^αH cross-peaks were observed in NOESY spectra, but they were weak. The chemical shifts of both the

NH and $C^{\alpha}H$ of Thr 85 and the $C_{\alpha}H$ of Ala 84 were also very temperature sensitive. Three possible explanations for very weak NH–$C^{\alpha}H$ cross-peaks are small $^3J_{N\alpha}$, saturation of NHs during solvent irradiation, or slow to medium chemical exchange (Robertson & Jardetsky, 1982; Wemmer, 1986). The first possibility seems unlikely due to the fact that even in regions of the protein where $^3J_{N\alpha}$ is known to be small, such as the helical turns mentioned above, NH–$C^{\alpha}H$ cross-peaks could be observed in COSY-type spectra; furthermore, even if the $^3J_{N\alpha}$ were small, we would have expected to observe some evidence of NH–$C^{\alpha}H$ cross-peaks in TOCSY spectra. The second possibility also seems unlikely since NH–^{15}N cross-peaks for these residues were not enhanced in an HMQC spectrum taken by use of a $1-\tau-1$ solvent suppression sequence instead of presaturation. The third possibility seems most likely for several reasons. The NH of Ala 84 or the $C^{\delta 2}H$ of either His 87 or His 94 shows slow chemical exchange as deduced from cross-peaks in an H_2O TOCSY and H_2O and D_2O NOESY spectra. Also, the NH-^{15}N cross-peak for Gly 86 in HMQC and SQC spectra was weak and broad relative to other cross-peaks. Finally, the temperature sensitivity of the Ala 84 $C^{\alpha}H$ and the Thr 85 NH and $C^{\alpha}H$ chemical shifts supports a chemical exchange argument. The cause of this chemical exchange cannot be determined at this point but may become clear when the structure of FKBP is deduced. However, long-range NOEs were observed between protons in this sequence to protons in the sequence Ile 56–Trp 59, where slow exchange was also observed, specifically in the Gly 58 $C^{\alpha}H$ and Trp $N^{\epsilon}H$ chemical shifts. As discussed below, these two loops show extensive contacts with other aromatic side chains. It is possible that some time-resolved exchange is occurring in these regions. For example, reorientation of aromatic side chains or of the polypeptide loops themselves could cause chemical shift perturbations of the various protons (Jardetzky & Roberts, 1981; Karplus & Dobson, 1986, Williams, 1989).

Discussion

In assessing a new folding topology, especially one that violates an established trend in known protein structures, it is important to consider the possibility that errors in the analysis led to an incorrect result. In a sequential assignment there are several types of possible errors. First, there may be errors in determination of connectivities between the various spin systems due to overlap in the spectra. We have used both heteronuclear and homonuclear spectroscopy along with variations in experimental conditions and techniques to ensure a high level of redundancy in the data. In this way we have been able to resolve all problems resulting from degeneracies of the proton chemical shifts. Second, it is possible that long-range NOE connections may be mistaken for sequential connections, leading, by analogy to X-ray crystallography, to an incorrect chain trace. In β sheets, sequential connections are supported by an extensive network of strong $C^{\alpha}H_i$–$C^{\alpha}H_j$ NOEs, which appear in a relatively uncrowded region of the spectrum. In the most severe case, an incorrect chain trace would lead to large regions where the sequence of connected spin systems had been incorrectly mapped onto the known primary sequence of the protein. Unlike X-ray crystallography, where it is often difficult at low to medium resolution to identify residue types unambiguously prior to chain tracing, the wide range of experiments employed in the spin system identification allowed us to identify 82% of the spin systems according to type prior to sequential assignment. Thus, when we began the sequential assignment procedure there were no sequences longer than five residues that were not punctuated by a spin system whose type was known. The strands of the β sheet contain a particularly high proportion of the Gly, Ala, Thr, and Val residues, which are among the simplest to determine in the spin system identification. For these reasons, the chance of an incorrect chain trace appears highly unlikely. Furthermore, it should be noted that there is evidence, cited above, that the crossing loop regions have their hydrogen-bond propensity satisfied, at least in part, by β-turn and helical stretches.

Preliminary restrained dynamics simulations using the program X-PLOR (Brünger et al., 1986; Brünger, 1990) with a limited set of long-range interresidue NOE-derived distance restraints have allowed us to infer characteristics of the three-dimensional global fold of FKBP. This can be described as follows. The central β sheet has a right-handed twist, and one side of it forms a hydrophilic, solvent-exposed face of the protein. The connecting loops, including many secondary structural elements (β turns and α-helical regions), form the remainder of the structure. They cross over the hydrophobic surface of the β sheet in contact with the sheet and with each other. There is an aromatic-rich core region comprised of hydrophobic side chains from the β sheet and strand-connecting loops. A second surface, involving the FK506 binding site, is defined by the strand-connecting loops, which are packed against the sheet and against each other. While the orientations of the loops are not yet determined, the global fold for FKBP is consistent with a twisted orientation of the β sheet that has a full complement of interstrand hydrogen bonds and relatively buried aliphatic and aromatic side chains. Further simulations with additional NOE-derived distance restraints are presently underway, and a complete description of the three-dimensional solution structure of FKBP will be reported when these studies are completed.

Added in Proof

Following review of this paper, a report appeared that noted the identity of FKBP to endogenous inhibitor 2 of protein kinase C (PKC-I2) (Goebl, 1991).

Acknowledgments

We thank Dr. Dennis Hare for generously providing a copy of his software package FELIX, Dr. Jonathan Lee of the Harvard Medical School NMR facility for collection of some of the H_2O spectra, and Tom Wandless, Soroosh Shambayati, Bob Standaert, and Mark Albers for useful discussions and assistance in preparation of the manuscript.

References

Aue, W. P., Batholdi, E., & Ernst, R. R. (1976) *J. Chem. Phys. 64*, 2229–2246.

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

Bax, A., & Davis, D. G. (1985) *J. Magn. Reson. 65*, 355–360.

Bax, A., & Drobny, G. (1985) *J. Magn. Reson. 61*, 306–320.

Bax, A., Griffey, R. H. & Hawkins, B. L. (1983) *J. Am. Chem. Soc. 105*, 7188–7190.

Bierer, B. E., Mattila, P. S., Standaert, R. F., Herzenberg, L. A., Burakoff, S. J., Crabtree, G. R., & Schreiber, S. L. (1990a) *Proc. Natl. Acad. Sci. U.S.A. 87*, 9231–9235.

Bierer, B. E., Somers, P. K., Wandless, T. J., Burakoff, S. J., & Schreiber, S. L. (1990b) *Science 250*, 556–559.

Billeter, M., Braun, W., & Wüthrich, K. (1982) *J. Mol. Biol. 155*, 321–346.

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

Braunschweiler, L., Bodenhausen, G., & Ernst, R. R. (1983) *Mol. Phys. 48*, 535–560.

Brünger, A. T. (1990) X-PLOR Manual, Version 2.1, Yale University, New Haven, CT.
Brünger, A. T., Clore, G. M., Gronenborn, A. M., & Karplus, M. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 3801–3805.
Bundi, A., & Wüthrich, K. (1979) *Biopolymers 18*, 285–297.
Chazin, W. J., & Wright, P. E. (1988) *J. Mol. Biol. 202*, 623–636.
Chazin, W. J., Rance, M., & Wright, P. E. (1987) *FEBS Lett. 222*, 109–114.
Chazin, W. J., Rance, M., & Wright, P. E. (1988) *J. Mol. Biol. 202*, 603–622.
Chirgadze, Y. N. (1987) *Acta Crystallogr. A43*, 405–427.
Chothia, C. (1984) *Annu. Rev. Biochem. 53*, 537–572.
Clore, G. M., Bax, A., Wingfield, P. T., & Gronenborn, A. M. (1988) *FEBS Lett. 238*, 17–21.
Creighton, T. E. (1984) in *Proteins: Structure and Molecular Principles*, W. H. Freeman & Co., New York.
Davis, D. G., & Bax, A. (1985) *J. Am. Chem. Soc. 107*, 2820–2821.
Dobson, C. M., Olejniczak, E. T., Poulson, F. M., & Ratcliffe, R. G. (1982) *J. Magn. Reson. 48*, 97–110.
Driscoll, P. C., Gronenborn, A. M., Wingfield, P. T., & Clore, G. M. (1990) *Biochemistry 29*, 4668–4682.
Dumont, F. J., Staruch, V. J., Koprak, S. L., Melino, M. R., & Sigal, N. H. (1990a) *J. Immunol. 144*, 251–258.
Dumont, F. J., Melino, M. R., Staruch, V. J., Koprak, S. L., Fischer, P. A., & Sigal, N. H. (1990b) *J. Immunol. 144*, 1418–1424.
Dyson, H. J., Holmgren, A., & Wright, P. E. (1989) *Biochemistry 28*, 7074–7087.
Eich, G., Bodenhausen, G., & Ernst, R. R. (1982) *J. Am. Chem. Soc. 104*, 3731–3732.
Fischer, G., Wittmann-Liebold, B., Lang, K., Kiefhaber, T., & Schmid, F. X. (1989) *Nature 337*, 476–478.
Fretz, H., Albers, M. W., Galat, A., Standaert, R. F., Lane, W. S., Burakoff, S. J., Bierer, B. E., & Schreiber, S. L. (1991) *J. Am. Chem. Soc. 113*, 1409–1411.
Glushka, A., & Cowburn, D. (1987) *J. Am. Chem. Soc. 109*, 7879–7881.
Goebl, M. G. (1991) *Cell 64*, 1051–1052.
Griesinger, C., Otting, G., Wüthrich, K., & Ernst, R. R. (1988) *J. Am. Chem. Soc. 110*, 7870–7872.
Gronenborn, A. M., Bax, A., Wingfield, P. T., & Clore, G. M. (1989) *FEBS Lett. 243*, 93–98.
Handschumacher, R. E., Harding, M. W., Rice, J., & Speicher, D. W. (1984) *Science 226*, 544–547.
Harding, M. W., Galat, A., Uehling, D. E., & Schreiber, S. L., (1989) *Nature 341*, 758–760.
Ikura, M., Kay, L. E., & Bax, A. (1990) *Biochemistry 29*, 4659–4667.
Jardetzky, O., & Roberts, G. C. K. (1981) *NMR in Molecular Biology*, Academic Press, London, New York.
Karplus, M., & Dobson, C. M. (1986) *Methods Enzymol. 131*, 362–389.
LeMaster, D. M., & Richards, F. M. (1988) *Biochemistry 27*, 142–150.
Macura, S., & Ernst, R. R. (1980) *Mol. Phys. 41*, 95–117.
McIntosh, L. P., Wand, A. J., Lowry, D. F., Redfield, A. G., & Dahlquist, F. W. (1990) *Biochemistry 29*, 6341–6362.
Moras, D., Olsen, K. W., Sabesan, M. N., Buehner, M., Ford, G. C., & Rossmann, M. G. (1975) *J. Biol. Chem. 250*, 9137–9162.
Muchmore, D. C., McIntosh, L. P., Russell, C. B., Anderson, D. E., & Dahlquist, F. W. (1989) *Methods Enzymol. 177*, 44–73.
Olejniczak, E. T., Hoch, J. C., Dobson, C. M., & Poulsen, C. M. (1985) *J. Magn. Reson. 64*, 199–206.
Otting, G., & Wüthrich, K. (1986) *J. Magn. Reson. 64*, 359–363.
Pardi, A., Billeter, M., & Wüthrich, K. (1984) *J. Mol. Biol. 180*, 741–751.
Piantini, U., Sørensen, O. W., & Ernst, R. R. (1982) *J. Am. Chem. Soc. 104*, 6800–6801.
Ptitsyn, O. B., & Finkelstein, A. V. (1980) *Quart. Rev. Biophys. 13(3)*, 339–386.
Rance, M., & Wright, P. E. (1986) *J. Magn. Reson. 66*, 372–378.
Rance, M., Sørensen, O. W., Bodenhausen, O. W., Wagner, O. W., Ernst, R. R., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 117*, 475–498.
Rance, M., Sørensen, O. W., Leupin, W., Kogler, H., Wüthrich, K., & Ernst, R. R. (1985) *J. Magn. Reson. 61*, 67–80.
Rance, M., Wagner, G., Sørensen, O. W., Wüthrich, K., & Ernst, R. R. (1984) *J. Magn. Reson. 59*, 250–261.
Redfield, A. G. (1983) *Chem. Phys. Lett. 96*, 537–540.
Richardson, J. S. (1977) *Nature 268*, 495–500.
Richardson, J. S., & Richardson, D. C., (1989) in *Prediction of Protein Structure and the Principles of Protein Conformation* (Fasman, G. D., Ed.) Plenum Press, New York.
Rosen, M. K., Standaert, R. F., Galat, A., Nakatsuka, M., & Schreiber, S. L. (1990) *Science 248*, 863–866.
Schreiber, S. L. (1991) *Science 251*, 283–287.
Shaka, A. J., & Freeman, R. (1983) *J. Magn. Reson. 51*, 169–173.
Shaka, A. J., Barker, P. B., & Freeman, R. (1985) *J. Magn. Reson. 64*, 547–552.
Siekierka, J. J., Hung, S. H. Y., Poe, M., Lin, C. S., & Sigal, N. H. (1989) *Nature 341*, 755–757.
Sklenar, V., & Bax, A. (1987) *J. Magn. Reson. 71*, 379–383.
Standaert, R. F., Galat, A., Verdine, G. L., & Schreiber, S. L. (1990) *Nature 346*, 671–674.
States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286–292.
Torchia, D. A., Sparks, S. W., & Bax, A. (1989) *Biochemistry 28*, 5509–5524.
Wagner, G. (1983) *J. Magn. Reson. 55*, 151–156.
Wagner, G., & Wüthrich, K. (1982) *J. Mol. Biol. 155*, 347–366.
Wagner, G., & Zuiderweg, E. R. P. (1983) *Biochem. Biophys. Res. Comm. 113*, 854–860.
Wang, J., Hinck, A. P., Loh, S. N., & Markley, J. L. (1990) *Biochemistry 29*, 102–113.
Weber, P. L., Drobny, G., & Reid, B. R. (1985) *Biochemistry 24*, 4549–4552.
Weber, P. L., Brown, S. C., & Mueller, L. (1987) *Biochemistry 26*, 7282–7290.
Wemmer, D. E., Kumar, N. V., Metrione, R. M., Lazdunski, M., Drobny, G., & Kallenbach, N. R. (1986) *Biochemistry 25*, 6842–6849.
Williams, R. J. P. (1989) *Eur. J. Biochem. 183*, 479–497.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.