Solution Conformation of the Type I Collagen α-1 Chain N-Telopeptide Studied by ^{1}H NMR Spectroscopy†

Albin Otter,‡ George Kotovych,‡ and Paul G. Scott*,§

Department of Chemistry, University of Alberta, Edmonton, Alberta, Canada T6G 2G2, and Department of Oral Biology, University of Alberta, Edmonton, Alberta, Canada T6G 2N8

Received December 14, 1988; Revised Manuscript Received June 9, 1989

ABSTRACT: The solution conformation of the type I collagen α-1 chain N-telopeptide has been studied by CD and ^{1}H NMR spectroscopy at 600 MHz in CD_3OH/H_2O (60/40 v/v) and H_2O solutions. The 19 amino acids form the N-terminal end of the α-1 polypeptide chain. By the combined application of several two-dimensional, phase-sensitive NMR techniques (COSY, RELAY, ROESY), a complete assignment of all proton resonances was achieved, and the conformation of the backbone could be established on the basis of the coupling constant and NOE data. In CD_3OH/H_2O solutions the spectroscopic evidence clearly indicates that two sections of the molecule (pE^1–Y^6 and T^{11}–M^{19}) are extended and that the D^7–S^{10} segment forms a β-turn, stabilized by a hydrogen bond between NH(S^{10}) and CO(D^7). The data suggest that the turn is of the type I kind (minor) and that it coexists with an extended structure (major conformer). Interactions between the two extended parts of the peptide were not observed, thus excluding the existence of a β-sheet. In H_2O solution the conformation is significantly different, with no β-turn, but a completely extended structure is observed.

Proteolytic removal of the nonhelical peptides ("telopeptides") at N- and C-terminal ends of the type I collagen triple helix impairs the ability of collagen monomers to assemble into fibrils in vitro (Rubin et al., 1963; Drake et al., 1966). It is these fibrils that are the functional form of the protein in fibrous connective tissues such as dermis, tendon, ligaments, bone, and dentine. Bovine type I collagen carries a 16-residue N-telopeptide and a 24-residue C-telopeptide on each of the two α-1 chains. The N-telopeptide of the α-2 chain consists of nine residues and the C-telopeptide of six residues (Glanville & Kuhn, 1979). From the results of selective digestion with proteinases it has been deduced that the N-telopeptide of the α-1 chain directs linear growth of type I collagen fibrils (Leibovich & Weiss, 1970) while the C-telopeptide promotes both linear and lateral growth (Helseth & Veis, 1981a; Capaldi & Chapman, 1982). To function in this way, the telopeptides of one molecule must interact in the growing fibril with the triple-helical domain of one or more adjacent collagen monomers. Helseth et al. (1979), in a theoretical study, pointed out a specific set of polarized interactions involving the N-telopeptide of the α-1 chain of type I collagen that could account for its apparent role in driving the linear growth of collagen fibrils in which all collagen monomers are parallel rather than antiparallel or randomly oriented. Protein–protein interactions are, almost by definition, conformation dependent, and hence there has been much interest in defining the three-dimensional structures of the telopeptides. Several attempts have been made to predict secondary structures from the amino acid sequences (Helseth et al., 1979; Helseth & Veis, 1981b; Scott, 1986; Dion & Meyers, 1987; Jones & Miller, 1987), but the results even from the application of the same algorithm have often varied between laboratories and the physical evidence available has, until recently, been insufficiently detailed to support or refute the various models proposed.

In this laboratory we have begun to study the conformations of the collagen telopeptides in solution using modern techniques of high-field NMR spectroscopy. Previously we studied the structure of a pentapeptide fragment of the type I collagen C-telopeptide (Otter et al., 1987) and more recently of the entire C-telopeptide (Otter et al., 1988), for which a role in fibrillogenesis was proposed on the basis of the results obtained. The present paper describes an NMR investigation of the N-telopeptide of type I collagen.

Materials and Methods

Synthetic N-Telopeptide. The peptide ($pGlu^1$-Leu^2-Ser^3-Tyr^4-Gly^5-Tyr^6-Asp^7-Glu^8-Lys^9-Ser^{10}-Thr^{11}-Gly^{12}-Ile^{13}-Ser^{14}-Val^{15}-Pro^{16}-Gly^{17}-Pro^{18}-Met^{19}) was synthesized by Peninsula Laboratories Inc., Belmont, CA. Purity was assessed by thin-layer chromatography on cellulose and silica gel and by electrophoresis at pH 1.9 and 6.4. In each case only a single spot was seen. Theoretical and actual amino acid compositions were in close agreement (data not shown).

Circular Dichroism Spectroscopy. Samples were dissolved in water and diluted with water or with methanol and water to a final concentration of 0.2 mg/mL. Spectra were recorded at ambient temperature in a cell of path length 0.0994 cm in a Jasco J-20A spectropolarimeter that had been previously calibrated with *d*-10-camphorsulfonic acid (DeTar, 1969). The mean residue weight was calculated from the amino acid composition. Other details are given in the legend to Figure 1.

Proton NMR Spectroscopy. Six milligrams of the purified solid material was dissolved in water and the pH adjusted to 7.0. After freeze-drying, the sample was redissolved in 0.5 mL of a CD_3OH/H_2O (60/40 v/v) mixture. Half of the water was D_2O to provide an internal deuterium lock signal for the NMR measurements. Prior to sealing, the sample was degassed by bubbling argon through it for approximately 15 min. The concentration of the sample was 5 mM. To check for possible aggregation, a 1.0 mM sample was also prepared. In the same way a sample was prepared by dissolving 6 mg of

† The Alberta Heritage Foundation for Medical Research, the Natural Sciences and Engineering Research Council of Canada, the Canadian Medical Research Council, and the University of Alberta are acknowledged for generous financial support.

* Author to whom correspondence should be addressed.
‡ Department of Chemistry.
§ Department of Oral Biology.

0006-2960/89/0428-8003$01.50/0

Table I: Summary of Experimental Parameters Used in the Two-Dimensional NMR Experiments at 600 MHz[a]

parameter	COSYPH	RELAY	ROESYPH
sweep width in F_2 (Hz)	6410	6410	6410
sweep width in F_1 (Hz)	3205	3205	3205
matrix size ($F_1 \times F_2$)			
before zero-filling	512 × 4K	256 × 4K	512 × 4K
after zero-filling	1K × 8K	1K × 8K	1K × 8K
evolution time			
initial value (μs)	1	1	1
increment (μs)	78	156	78
no. of scans (dummy scans)	16 (2)	16 (2)	32 (2)
acquisition time (s)	0.32	0.32	0.32
relaxation delay (s)	1.8	1.8	2.4
other delays (ms)		80[b]	125[c]
window functions for 2D FT (F_1/F_2)	S/S	S/S	S/S
shifts of window function in fractions of π	4/8	4/8	3/4

[a] All spectra were recorded at 295 K, pH 7.0, concentration 5.0 mM in 60% CD_3OH/40% H_2O. S stands for sine bell window function. [b] Coherence transfer time. [c] Spin locking time at an average field strength of 5.6 kHz.

the peptide in H_2O/D_2O (80/20 v/v). The final concentration was 5 mM, and the pH was 7.2. We refer to the two samples as the methanol sample and the water sample, respectively.

Most NMR experiments were performed on a Bruker AM600 NMR spectrometer at 295 K. Some experiments, mainly the study of the peptide in water, were carried out on a Bruker AM400 NMR spectrometer. In both cases, data collection and data processing were controlled by an Aspect 3000 computer equipped with an array processor and using 1987 DISNMR software. The chemical shifts were determined relative to the undeuterated fraction of the methyl group of CD_3OH (Merck Sharp and Dohme, MD-67, 99.2% D) at 3.30 ppm with respect to TMS. For the aqueous sample an external reference of 5% TSP[1] in H_2O, measured separately under the same experimental conditions, was used.

One-dimensional 600-MHz spectra were recorded with a sweep width of 6400 Hz and 32K (zero-filled to 64K) data points. A total of 256 scans were accumulated to produce an excellent signal-to-noise ratio that allowed extensive resolution enhancements without compromising the detectability of the relatively low intensity exchangeable protons. Typically, Gaussian multiplications with line-broadening factors of up to −5 Hz were used. The two strong solvent signals (H_2O and −OH of CD_3OH at 4.72 and the undeuterated CD_3OH fraction at 3.30 ppm) were both presaturated during a 2.4-s relaxation delay. To account for the big difference in the intensity of the two solvent peaks, the frequency of the decoupler was changed every 0.4 s and the water signal was presaturated twice as long as the methyl group. The effect of presaturation on the exchangeable protons was carefully monitored, and no wipeout due to saturation transfer was observed. In the case of the aqueous sample, the presaturation procedure was, of course, simplified to the suppression of the water signal. The experimental conditions were basically identical at 600 and 400 MHz except for a shorter sweep width in the latter case and, accordingly, a reduced number of data points.

The parameters used in the various two-dimensional NMR techniques (Benn & Günther, 1983) at 600 MHz are summarized in Table I. The COSY (Aue et al., 1976) and ROESY (Bothner-By et al., 1984; Bax & Davis, 1985) experiments were carried out in the phase-sensitive mode by using time-proportional phase increments (Marion & Wüthrich, 1983), whereas the RELAY spectrum (Eich et al., 1982; Bolton & Bodenhausen, 1982) was recorded in the absolute value mode. The solvent suppression was achieved in the same way as described previously for the one-dimensional spectra. In addition, the water signal was continuously irradiated at a reduced power level during the evolution time (Wider et al., 1983). Phasing of the phase-sensitive experiments was achieved in the following way: the second FID was multiplied with the weighting function as outlined in Table I, Fourier transformed, and then phased. The determined phase constants were then applied to the entire set of FIDs in the t_2 domain. No phase correction was applied along t_1 in the COSY experiiment, whereas the ROESY required a basic 90° phase correction in this domain. The resulting 2D spectra were then displayed, and slight phase adjustments were usually necessary to obtain the best possible data representation. Due to the appearance of considerable t_1 noise and ridges (Mehlkopf et al., 1984), caution must be exercised when information is extracted from the spectra. As long as only connectivities through J couplings were of interest, but not the actual values of the coupling constants, we found it helpful to process the data with unshifted sine bell functions in both dimensions and then to calculate the magnitude spectrum of the COSY data matrix. Then, two spectra parallel to t_2 and outside the absorption range of the telopeptide (at approximately 9.5 and 0.3 ppm) were summed. The resulting one-dimensional spectrum contained the base line and ridges only, but did not have any signals from the molecule under investigation. Two traces were used to average differences in the intensities of the ridges on the low-field and the high-field end of the spectrum. Finally, the above-mentioned one-dimensional spectrum was subtracted from every spectrum of the 2D matrix parallel to t_2 by means of a subtraction routine. The resulting two-dimensional spectrum was almost free of t_1 noise and ridges, and connectivities could be found much more easily. In the so-called "fingerprint" region (Wagner & Wüthrich, 1982), where the phase-sensitive data representation must be retained, there were basically no obscuring ridges, and thus the coupling constants could be determined with good accuracy. The phase-sensitivity was also retained in the ROESY experiment as spurious resonances due to magnetization transfer between scalar-coupled spins (Marion, 1985; Neuhaus & Keeler, 1986) are easily identified by their phase being the same as the diagonal signals. In contrast, real ROE cross-peaks have opposite signs with respect to the diagonal peaks. Under the selected experimental conditions (Kessler et al., 1987) the two-dimensional spectrum was almost free of any such peaks. However, some base-line distortions around intense diagonal peaks such as residual solvent signals and methyl groups were present, and the spectrum was therefore subjected to a base-line correction routine (Griesinger & Ernst, 1987). The Bruker ABS base-line correction routine, using a polynomial fitting of fifth degree to the base line, eliminated this problem almost completely in both dimensions.

The temperature coefficients of the amide protons were determined by recording the COSY experiment at two different temperatures, namely, 295 and 300 K.

RESULTS

The circular dichroism spectrum of the synthetic N-telopeptide in water (Figure 1) shows no discernible features between 240 and 200 nm, suggesting that it is present under

[1] Abbreviations: 2D NMR, two-dimensional nuclear magnetic resonance; COSY, two-dimensional correlated spectroscopy; ROESY, two-dimensional rotating-frame Overhauser enhancement spectroscopy; RELAY, two-dimensional relayed coherence transfer spectroscopy; CD, circular dichroism spectroscopy; FID, free induction decay; TSP, 3-(trimethylsilyl)-1-propanesulfonate.

Table II: ^{1}H Chemical Shifts,[a] Coupling Constants,[b] and NH Temperature Dependence[c] of Type I α-1 Chain N-Telopeptide in CD_3OH/H_2O[d]

	NH	H_α	H_β[e]	H_γ	H_δ	others	J_{NH-H_α}	ΔT
pE^1	8.03	4.30	2.08, 1.92	2.36			7.8	9.1
L^2	8.35	4.37	1.59, 1.48	1.57	0.89, 0.85		7.8	8.3
S^3	8.17	4.40	3.85, 3.77				7.8	8.3
Y^4	8.20	4.50	3.04, 2.87			H_2/H_6, 7.06 H_3/H_5, 6.74	7.0	7.7
G^5	8.31	3.90, 3.71					5.7/5.7	6.7
Y^6	7.95	4.50	3.01, 2.88			H_2/H_6, 7.03 H_3/H_5, 6.74	7.0	6.0
D^7	8.26	4.58	2.71, 2.63				7.0	6.4
E^8	8.90	4.14	2.05, 2.00	2.27			4.5	4.1
K^9	8.36	4.26	1.86	1.45	1.68	H_ϵ, 2.97	7.8	4.1
S^{10}	8.06	4.40	3.91, 3.72				7.8	2.0
T^{11}	7.90	4.27	4.24	1.22			7.8	4.7
G^{12}	8.26	3.97, 3.87					5.5/5.5	5.7
I^{13}	7.85	4.22	1.82	1.45, 1.12	0.84	CH_3^γ, 0.87	8.6	5.4
S^{14}	8.21	4.50	3.76				7.8	8.1
V^{15}	8.06	4.46	2.08	0.97, 0.92			7.8	8.6
P^{16}		4.41	2.20, 1.98	1.97	3.81, 3.67			
G^{17}[f]	8.10	4.20, 3.92					5.5/5.5	8.6
p^{18}		4.42	2.23, 1.98	2.00	3.62, 3.57			
M^{19}	7.76	4.27	2.07, 1.89	2.51	2.08		7.8	7.3

[a] Chemical shifts in parts per million relative to the methyl group of CD_3OH (3.30 ppm downfield from TMS). [b] Coupling constants in hertz. [c] The effect of temperature change on the NH protons is expressed in ppb/K. [d] All spectra were measured at 600 MHz at 295 K, pH 7.0, at a concentration of 5.0 mM in 60% CD_3OH/40% H_2O. [e] The pairs of protons x ($x = \alpha, \beta, \gamma, \ldots$) could not be assigned stereospecifically. The larger chemical shift was assigned to atom x, the smaller shift to atom x'. [f] The G^{17}–P^{18} peptide bond exists as cis/trans isomers in the ratio 22:78. The chemical shifts for the cis isomer are (only the resonances that can be assigned unambiguously) as follows: G^{17}, NH 8.06, $H_{\alpha,\alpha'}$ 4.08, 3.69; P^{18}, H_α 4.54, $H_{\beta,\beta'}$ 2.34, 2.14, $H_{\delta,\delta'}$ 3.51, 3.46; M^{19}, NH 7.77.

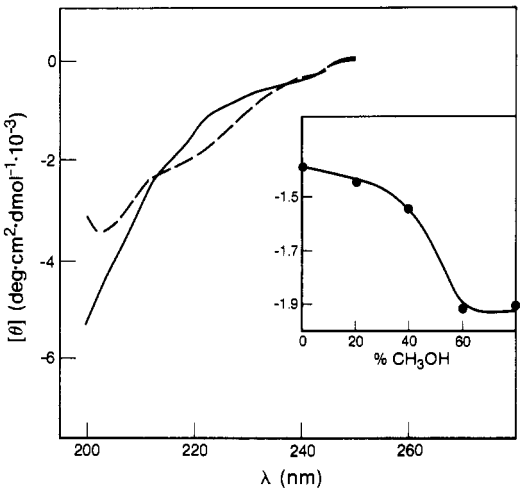

FIGURE 1: Circular dichroism spectra of the synthetic C-telopeptide at 25 °C in water (—) and in 80% (v/v) methanol (--). Each spectrum shown is the average of four. The inset shows the effect of increasing concentrations of methanol on the mean residue ellipticity of 220 nm.

these conditions in a predominantly random conformation. Addition of methanol caused a discontinuous transition toward a spectrum indicative of some periodic structure, possibly β-sheet and/or β-turn, together with random coil. Attempts to fit theoretical spectra calculated by using parameters derived for proteins (Chang et al., 1978) or peptides (Greenfield & Fasman, 1969) or to apply the computerized method of Provencher and Glockner (1981) to these results were unsuccessful, so that a quantitative interpretation cannot be given. On the basis of these results we selected a mixed solvent system (60% methanol, 40% water) for the NMR experiments.

The NMR spectral analysis of the N-telopeptide, whose one-dimensional ^{1}H spectrum is shown in Figure 2, in the CD_3OH/H_2O solvent mixture is based on the results of various two-dimensional experiments suitable for structure elucidations of biomolecules (Wüthrich, 1986). First, a phase-sensitive COSY spectrum (Marion & Wüthrich, 1983) was recorded

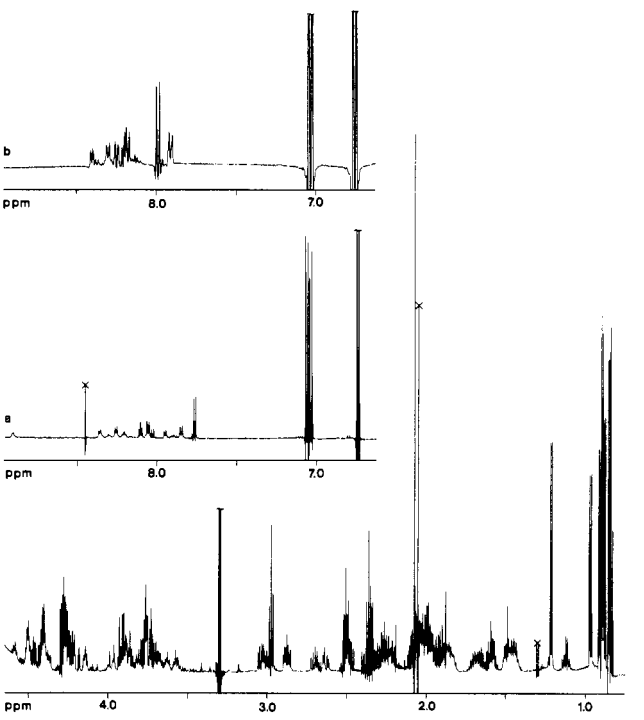

FIGURE 2: One-dimensional 600-MHz ^{1}H spectrum of type I collagen α-1 chain N-telopeptide recorded in CD_3OH/H_2O (60/40 v/v) at 5.0 mM, pH 7.0, temperature 295 K. The insets depict the exchangeable and aromatic proton region in (a) the same solvent as described above and (b) H_2O at 400 MHz, pH 7.2, concentration and temperature as above. Impurities are denoted with an ×.

and analyzed with respect to NH–H_α coupling constants and NH–H_α connectivities. The contour plot of the fingerprint region (Wider et al., 1984) is shown in Figure 3. The digital resolution along the F_2 axis is approximately 1.5 Hz, allowing an accuracy of about 0.8 Hz in the determination of NH–H_α coupling constants, which are summarized in Table II. Further valuable information for the assignments of amino acids was provided by the relayed COSY experiment (RELAY) (Eich et al., 1982; Bolton & Bodenhausen, 1982) by

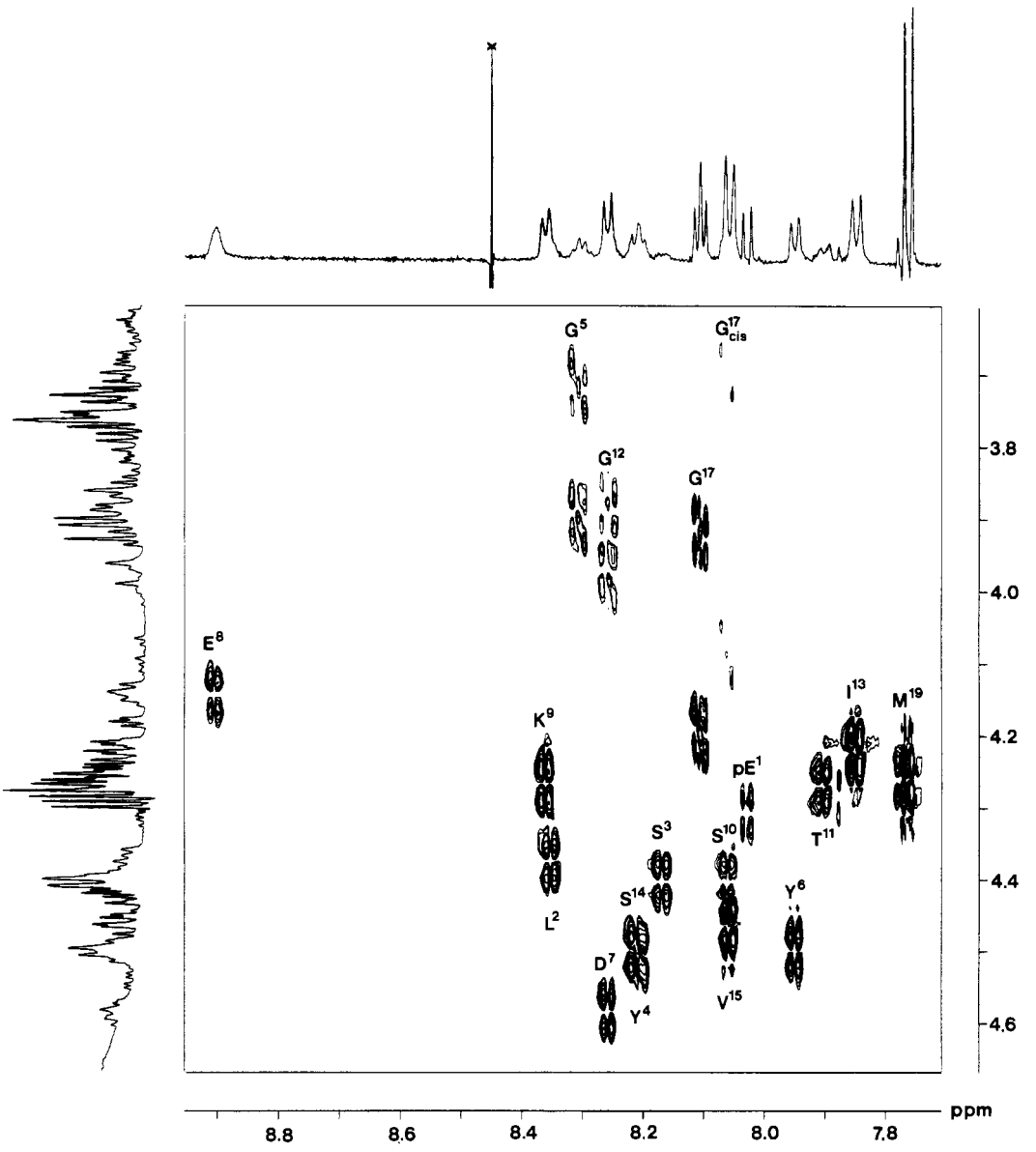

FIGURE 3: 600-MHz two-dimensional contour plot of a phase-sensitive COSY experiment of type I collagen α-1 chain N-telopeptide recorded under the same experimental conditions as described in Figure 2. The fingerprint region with the corresponding one-dimensional spectra is shown only. The digital resolution along F_2 is 1.5 Hz. The spectrum is unsymmetrized, and positive and negative contours are plotted without distinction.

means of NH–H_β cross-peaks and, to a somewhat smaller extent, through H_α–H_γ interactions. Finally, the sequential assignment could be deduced unambiguously from the results of phase-sensitive rotating frame NOE experiments (ROESY) (Bothner-By et al., 1984; Bax & Davis, 1985), which provide distance information based on the nuclear Overhauser effect (Sanders & Mersh, 1982). Several experiments with different spin-locking times τ_m showed that the NOEs reach maximum intensities around τ_m = 125 ms. The most relevant section of the ROESY at 125 ms is shown in Figure 4. It is noteworthy that under the selected experimental conditions, outlined under Materials and Methods, the two-dimensional spectrum provided a large number of cross-peaks of remarkably strong intensity, and the commonly described artifacts due to COSY and Hartmann–Hahn-type transfers (Marion, 1985; Neuhaus & Keeler, 1986) were not observed in any section of the spectra important for the conformational analysis. On the basis of the observation of $H_\alpha(i)$–NH(i+1) cross-peaks, the sequential assignment could be achieved easily for the NH-bearing amino acids and for the two proline moieties where the complementary $H_\alpha(i)$–$H^{Pro}_{\delta\delta'}(i+1)$ effects were observed. By comparison to the previously investigated C-telopeptide (Otter et al., 1988), the number and the strength of the NOEs observed here were much higher.

The temperature dependence of the amide protons was high for most parts of the molecule (between ca. 5 and ca. 9 ppb/K), indicating that there were no hydrogen bonds present. There was, however, one clearly outstanding value of only 2.0 ppb/K, recorded for NH(S^{10}). From this value and from the reduced temperature dependence (lower than 5 ppb/K) of the amino acids adjacent to S^{10}, we conclude that this part of the peptide must be involved in some tight turn (see next section). The results of the temperature study, together with the assignment of all proton resonances of the N-telopeptide are summarized in Table II. Since the one-dimensional spectra recorded at 5.0 (80 scans) and 1.0 mM (2000 scans) are completely congruent, we rule out the possibility of aggregation with confidence.

The sample prepared in H_2O was analyzed by use of the same basic approach as outlined above for the methanol sample. All data obtained indicate that the overall behavior of the peptide is considerably different in this solvent, thus

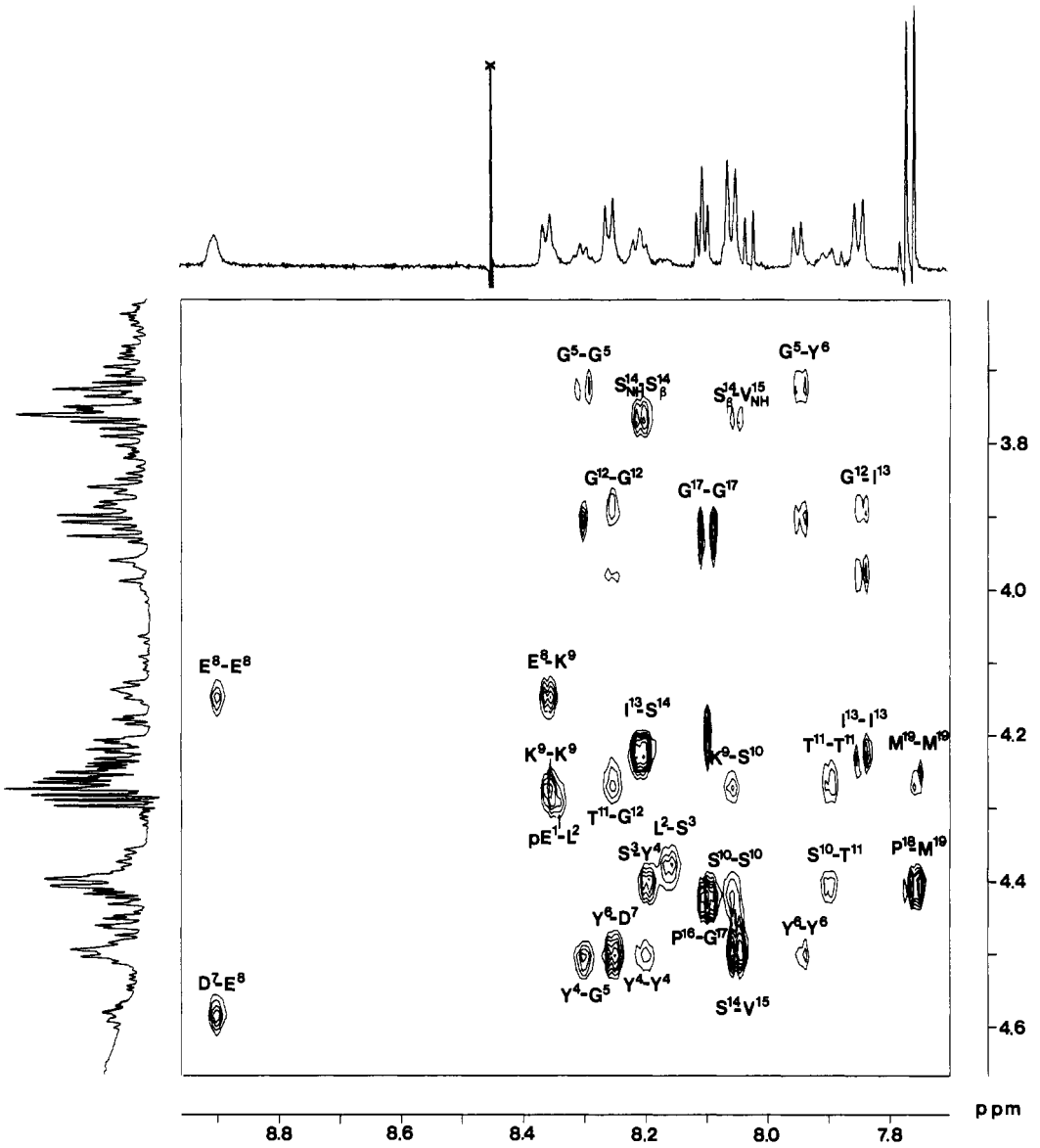

FIGURE 4: NH–H_α region of a 600-MHz ROESY experiment of type I collagen α-1 chain N-telopeptide recorded under the experimental conditions outlined in Figure 2. The spin-locking time was 125 ms at a field strength γB_2 of 5.6 kHz. The contour levels are equidistant, separated by 500 vertical display units, and the cross-peaks are of opposite sign with respect to the diagonal. The spectrum is unsymmetrized but base-line corrected in both dimensions. Unless otherwise indicated, the first number refers to H_α and the second to NH.

confirming the result of the previously described CD study. Four reasons lead to the conclusion that the N-telopeptide adopts a completely extended, if not partially random, conformation in water. (i) The NH temperature coefficients are higher and much more uniformly distributed. (ii) None of the amide protons is shifted to low field, indicating the absence of hydrogen bonds. (iii) All NH–H_α coupling constants are $\geq$6.5 Hz. (iv) No nuclear Overhauser effect is observed between any of the amide protons. Therefore, we limited our efforts to a partial assignment of the NH and H_α resonances, which are listed in Table III.

Discussion

The NMR study provides conformational information in the form of nuclear Overhauser enhancements, NH–H_α coupling constants, and the temperature dependence of the amide protons. We consider first the torsion angle ϕ, describing the rotation about the N–C_α bond. The analysis of coupling constants by means of the well-known Karplus equation (Karplus, 1963) provides an estimate of the angle θ that can then be used to calculate ϕ. In general, up to four different dihedral angles may correlate with a given coupling constant. In peptides and proteins, however, the range of ϕ is greatly reduced (Richardson, 1981), leaving only two possible ranges for ϕ, namely, around −80° and around −160° for a 7-Hz coupling constant. Due to conformational averaging, a numerically precise interpretation of the coupling constants is impossible. Nevertheless, the observation of medium to strong NH(i)–$H_\beta(i)$ NOEs renders ϕ angles around −80° more likely since only in this case are the NH and β-protons of the same amino acid close enough to entail such effects (Leach et al., 1977).

ψ angles cannot be calculated from ^{1}H coupling constants (Bystrov, 1976), but they can be estimated from the $H_\alpha(i)$–NH(i+1) and $H_\alpha(i)$–$H^{Pro}_{\delta\delta'}(i+1)$ NOEs. In general, we classified the observed nuclear Overhauser enhancements into three categories (weak, medium, and strong) on the basis of a careful visual comparison of the cross-peak volumes. As can be seen in Figure 5, we observed a large number of such effects of different strengths throughout the peptide. Assuming the peptide bond to be trans (ω = 180°), the sequential NOE is maximum for the C_α–$H_\alpha(i)$ bond in the trans position to the carbonyl group (i+1), which is equivalent to ψ = +120°. The observation of strong NOEs therefore requires ψ = +120° ±

Table III: Chemical Shifts[a] of the NH and H_α Protons, Coupling Constants,[b] and NH Temperature Dependence[c] of Type I α-1 Chain N-Telopeptide in Aqueous Solution[d]

	NH	H_α	J_{NH-H_α}	ΔT
pE^1	8.28	4.30	7.9	13.0
L^2	8.36	4.34	7.4	10.8
S^3	8.19	4.46	7.9	8.1
Y^4	8.22	4.53	7.9	11.6
G^5	8.32	3.82	6.0/6.0	9.3
Y^6	7.96	4.50	7.0	9.1
D^7	8.29	4.55	7.4	8.7
E^8	8.45	4.15	6.5	8.0
K^9	8.35	4.29	6.8	8.7
S^{10}	8.26	4.40	8.4	10.1
T^{11}	8.15	4.33	7.9	9.8
G^{12}	8.35	3.94	6.1/6.1	8.1
I^{13}	7.95	4.21	7.5	10.4
S^{14}	8.41	4.48	7.4	11.1
V^{15}	8.22	4.44	7.9	13.0
P^{16}		4.42		
G^{17}	8.24	4.21, 3.92	5.6/5.6	12.6
P^{18}		4.43		
M^{19}	8.03	4.25	7.5	12.1

[a] Chemical shifts relative to TSP measured in a separate sample under the same experimental conditions. [b] Coupling constants in hertz. [c] The effect of temperature change is expressed in ppb/K. [d] All spectra were measured at 400 MHz at 295 and 300 K, respectively. The concentration was 5 mM, pH 7.2, in 80% H_2O/20% D_2O.

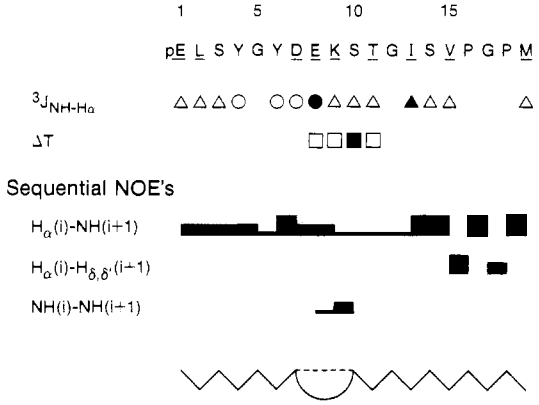

FIGURE 5: Amino acid sequence of type I α-1 chain N-telopeptide with unique residues underlined. For the ^3J coupling constants the symbols are used in the following way: solid circles, values <5.0 Hz; open circles, 7.0 ± 0.4 Hz; open triangles, 7.8 ± 0.4 Hz; solid triangle, 8.6 ± 0.4 Hz. A solid square indicates low-temperature sensitivity (≤2.0 ppb/K), whereas open squares are used for the range above 2.0 but below 5.0 ppb/K. The sequential NOEs are indicated by lines, and the strength of the effect is separated into three groups, weak, medium, and strong, according to the height of the bar. The secondary structure is schematically described at the bottom. A zig-zag line stands for an extended chain, and the β-turn is shown by a half-circle with the dashed line standing for the NH(S^{10})—C=O(D^7) hydrogen bond.

30°, whereas in the case of medium, and, in particular, weak effects, larger deviations must be considered with minimal NOEs (maximum distance) at $\psi = -60°$ [C_α–$H_\alpha(i)$ bond cis to C=O(i+1)].

Finally, the peptide bond angle ω has to be considered. The presence of proline rings, which can undergo cis/trans isomerization quite easily (Deslauriers & Smith, 1978; Otter & Kotovych, 1987), makes an all trans ($\omega = 180°$) assumption questionable. Our data indicate that the peptide bond G^{17}–P^{18} exists in a 22:78 cis/trans equilibrium. An unambiguous proof for this phenomenon is, besides the doubling of lines for the protons closest to the isomerization site, the observation of an $H_\alpha(G^{17})$–$H_\alpha(P^{18})$ NOE in addition to the trans isomer related $H_\alpha(i)$–$H^{Pro}_{\delta\delta'}(i+1)$ effect. As the difference in chemical shifts between the two isomeric forms declines to values beyond detection within short distances, only a few resonances of the cis isomer can be assigned without ambiguity. They are listed in footnote f of Table II. None of the other peptide bonds show any indication of cis/trans isomerization, and therefore $\omega = 180°$ is a fully acceptable assumption throughout the rest of the peptide.

The evaluation of the conformation of the side chains in terms of side-chain rotamer populations (Jardetzky & Roberts, 1981) requires clearly resolved H_α–H_β coupling constants. Because of overlapping resonance lines, these were not measurable for most of the amino acids. Only the two tyrosine moieties and D^7 provide a complete, reliable set of coupling constants. Analysis using literature values for the three classical rotamers (Feeney, 1976) determined that none of these revealed a predominance of a single rotamer, and thus fixed side-chain conformations can be excluded for the above-mentioned amino acids. It appears unlikely that for the other amino acids a different result would be obtained. Nevertheless, the observation of NOEs along the side chains suggests that many of them (in particular the sterically cumbersome ones) have considerably reduced mobility but not necessarily fixed conformations.

The final consideration concerning the conformation of the N-telopeptide is the evaluation of possible hydrogen bonds. By measuring the COSY spectrum at two different temperatures (295 and 300 K), we obtained the changes in chemical shift as a function of temperature, summarized in Tables II and III. In the water sample all NH resonances show a strong dependence on the temperature, indicating that no hydrogen bonds are present, and the conformation of the molecule can best be described as fully extended, probably with some degree of randomness. The methanol sample, on the other hand, exhibits a completely different behavior. NH(S^{10}) is only very slightly temperature dependent, and the measured value of 2.0 ppb/K is indicative of considerable shielding from solvent, most likely because of the existence of a hydrogen bond (Hruby, 1974). The amino acids adjacent to S^{10} also show some reduced interaction with the solvent. Toward both ends of the molecule the temperature coefficients steadily increase. The different spectroscopic behavior of the peptide in the two solvents, together with the low-temperature sensitivity of NH(S^{10}) in CD_3OH, establishes this amide proton as being part of a hydrogen bond. To find the corresponding carbonyl group, we consider first the overall change of NH chemical shifts when changing the solvent from H_2O to CD_3OH/H_2O. An average high-field shift of approximately −0.1 ppm is observed for amide groups of nonterminal amino acids. Almost the same value (−0.16 ppm) can be obtained when the NH shifts obtained from the methanol sample are compared to data of a recently published statistical analysis of a large number of peptides and proteins (Gross & Kalbitzer, 1988). One amide resonance, however, clearly stands out by displaying a completely reversed behavior. NH(E^8) experiences a low-field shift of +0.45 ppm (+0.70 ppm in comparison to the aforementioned literature data) upon change of the solvent. This large effect clearly identifies the D^7 as the other amino acid involved in the hydrogen bond (Wüthrich, 1986, p 262). Consequently, the amino acids D^7, E^8, K^9, and S^{10} form a β-turn. In contrast, the conformation is extended for the first six as well as the last nine residues of the telopeptide with no measurable interactions between the two extended sections as shown by the absence of long-range NOEs. Therefore, the existence of a true β-sheet can be excluded. The conformation of the four-residue β-turn is determined by the four torsion angles ϕ_8, ψ_8, ϕ_9, and ψ_9 as the residues E^8 and K^9 form the

actual turn, whereas for residues D^7 and S^{10} only the orientation of the carbonyl group and the NH group, respectively, is defined (Sibanda & Thornton, 1985). Depending on the orientation of the central C=O/NH unit, two major types of turns (type I with, in the ideal form, dihedral angles $\phi_8 = -60°$, $\psi_8 = -30°$, $\phi_9 = -90°$, and $\psi_9 = 0°$ and type II with $\phi_8 = -60°$, $\psi_8 = +120°$, $\phi_9 = +90°$, $\psi_9 = 0°$) have to be considered, along with the corresponding mirror images, type I′ and type II′ (Richardson, 1981). On the basis of NOE patterns and NH–H_α coupling constants, the latter two versions can be excluded (Wüthrich, 1986). The decision, whether the peptide adopts a type I or type II β-turn, has to be based on NOE arguments and their dependent on the angles ψ_8 and ϕ_9 (ϕ_8 and ψ_9 are identical in both types of turns). The NOE pattern in the amide proton region leads to the conclusion that the β-turn is type I, together with a major percentage of extended conformation. The NH(K^9)–NH(S^{10}) NOE is typical for the type I and type II β-turn and therefore not conclusive. On the other hand, the NH(E^8)–NH(K^9) interaction is very different in the two turns under question with corresponding interproton distances of 2.6 (type I) and 4.5 Å (type II), respectively (Wüthrich, 1986). The observation of an NOE between these two amide protons clearly indicates the presence of the type I β-turn, since no measurable NOE can be expected between protons separated by more than 4 Å (Brown et al., 1981). It should be mentioned that the NOEs observed between exchangeable protons are generally of lower intensity than effects between nonexchangeable nuclei irrespective of the interproton distance involved. This is mainly due to exchange with the bulk solvent. Therefore, it was necessary to examine cross sections taken through the NH signals (not shown) in addition to the examination of the two-dimensional contour plot.

The NOE between $H_\alpha(E^8)$ and NH(K^9) shows that not all of the molecules adopt the type I β-turn conformation. In an extended conformation, with ψ_8 around +120°, the two protons are very close in space (2.2 Å; Wüthrich, 1986), whereas in the type I β-turn ($\psi_8 \approx -30°$) the distance is 3.5 Å. In the first case, the nuclear Overhauser enhancement would be strong, but in the latter case only a weak effect can be expected (approximately 15 times smaller due to the r^{-6} dependence). The observation of an intermediate size effect (Figure 5) is presumably the superposition of a strong contribution from the majority of the molecules in the extended conformation and a very weak contribution from the ones in the β-turn structure (minor conformer). It should be emphasized that in a type II β-turn the $H_\alpha(E^8)$–NH(K^9) distance would be very similar to the one found in an extended conformer. However, considering the tendency of small peptides to adopt extended or random conformations to a large degree (Dyson et al., 1988a,b; Wright et al., 1988) and in view of the small effects observed in the CD spectrum, we conclude that the type I β-turn and the extended conformation coexist. On the basis of the proton NMR data alone it is not possible to assess the ratio of the two conformers.

The β-turn formed by residues D^7–S^{10} is the dominant structural feature of the N-telopeptide in methanol–water. This turn is stabilized by a hydrogen bond and perhaps also by a charge–charge interaction (salt bridge) between the side chains of E^8 and K^9. The structure-promoting effect of methanol could be explained by the reduced dielectric constant. The presence of the salt bridge is indirectly supported by empirical conformational energy calculations and by the observation that the CD spectrum in organic solvents changes very little from that shown in Figure 1 over the temperature range 5–63 °C (P. G. Scott, unpublished data), suggesting a rather high degree of stability for the β-turn. By contrast, the segments pE^1–Y^6 and T^{11}–M^{19} are expected to be flexible but with a preference for extended over folded conformations.

While there is no measurable interaction (hydrogen bonding) between the extended sequences preceding and following the β-turn, structures consistent with our data bear some resemblance to the "hairpin" model predicted by Helseth et al. (1979) on the basis of the Chou–Fasman criteria (Chou & Fasman, 1978). Jones and Miller (1987) described an alternative model, based on the combined application of several predictive algorithms, incorporating a loop around residue K^9, flanked by β-turns. Our data clearly preclude such a conformation in solution.

As first pointed out by Helseth et al. (1979), intermolecular interactions within the growing collagen fibril are expected to stabilize a hairpin conformation. It could further be suggested that the dielectric constant experienced by the telopeptides in this environment would be lower than in bulk solution. Both effects would favor β-turn formation. Folded structures for the telopeptides, with an axial rise per residue of considerably less than the 2.86 Å of the triple helix, best fit the X-ray and neutron diffraction data on collagen fibrils (Hulmes et al., 1977, 1980).

The intact collagen monomer carries two 16-residue α-1 chain N-telopeptides and a 9-residue α-2 chain N-telopeptide attached to the triple helix. It seems unlikely, given the terminal position, short length, and generally hydrophilic character of the telopeptides, that they would interact measurably or that their conformation would be influenced by the triple helix. This is in contrast to the situation in the intact N-propeptides, where the conformation would be constrained by the flanking triple-helical domains (Dombrowski & Prockop, 1988).

In summary, our data suggest that the type I collagen α-1 chain N-telopeptide possesses a flexible, extended conformation in aqueous solution but that it has the potential to adopt a type I β-turn under certain conditions. To this extent our data support the original proposal of Helseth et al. (1979) of a structural basis for the observed role of the N-telopeptide in collagen fibrillogenesis.

ACKNOWLEDGMENTS

Bruker-Spectrospin in Switzerland and especially Dr. Hans-Peter Kellerhals are gratefully acknowledged for providing access to the AM 600 system. We thank Dr. Gerhard Wider (Spectrospin, Fällanden), Dr. Hans Vogel (University of Calgary), and Elena Edwards for technical assistance.

REFERENCES

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

Bax, A., & Davis, D. G. (1985) *J. Magn. Reson. 63*, 207–213.

Benn, R., & Günther, H. (1983) *Angew. Chem., Int. Ed. Engl. 22*, 350–380.

Bolton, P. H., & Bodenhausen, G. (1982) *Chem. Phys. Lett. 89*, 139–144.

Bothner-By, A. A., Stephens, R. L., Lee, J., Warren, C. D., & Jeanloz, R. W. (1984) *J. Am. Chem. Soc. 106*, 811–813.

Braun, W., Bösch, C., Brown, L. R., Go, N., & Wüthrich, K. (1981) *Biochim. Biophys. Acta 667*, 377–396.

Bystrov, V. F. (1976) *Prog. NMR Spectrosc. 10*, 41–81.

Capaldi, M. J., & Chapman, J. A. (1982) *Biopolymers 21*, 2291–2313.

Chang, C. T., Wu, C.-S. C., & Yang, J. T. (1978) *Anal. Biochem. 91*, 13–31.

Chou, P. Y., & Fasman, G. D. (1978) *Adv. Enzymol. Relat. Areas Mol. Biol. 47*, 45–148.
Deslauriers, R., & Smith, I. C. P. (1976) in *Topics in Carbon-13 NMR Spectroscopy* (Levy, G. C., Ed.) pp 2–66, Wiley-Interscience, New York.
DeTar, Del. F. (1969) *Anal. Chem. 41*, 1406–1408.
Dombrowski, K., & Prockop, D. J. (1988) *J. Biol. Chem. 263*, 16545–16552.
Drake, M. P., Davison, P. F., Bump, S., & Schmitt, F. O. (1966) *Biochemistry 5*, 301–310.
Dyson, H. J., Rance, M., Houghten, R. A., Lerner, R. A., & Wright, P. E. (1988a) *J. Mol. Biol. 201*, 161–200.
Dyson, H. J., Rance, M., Houghten, R. A., Wright, P. E., & Lerner, R. A. (1988b) *J. Mol. Biol. 201*, 201–217.
Eich, G., Bodenhausen, G., & Ernst, R. R. (1982) *J. Am. Chem. Soc. 104*, 3731–3732.
Feeney, J. (1976) *J. Magn. Reson. 21*, 473–478.
Glanville, R. W., & Kihn, K. (1979) in *Fibrous Proteins: Scientific, Industrial and Medical Aspects* (Parry, D. A. D., & Cramer, L. K., Eds.) Vol. 1, pp 133–150, Academic Press, London.
Greenfield, N. J., & Fasman, G. D. (1969) *Biochemistry 8*, 4108–4116.
Griesinger, Ch., & Ernst, R. R. (1987) *J. Magn. Reson. 75*, 261–271.
Gross, K.-H., & Kalbitzer, H. R. (1988) *J. Magn. Reson. 76*, 87–99.
Helseth, D. L., Jr., & Veis, A. (1981a) *J. Biol. Chem. 256*, 7118–7128.
Helseth, D. L., Jr., & Veis, A. (1981b) in *Chemistry and Biology of Mineralized Connective Tissues* (Veis, A., Ed.) pp 85–92, Elsevier/North-Holland, New York.
Helseth, D. L., Jr., Lechner, J. H., & Veis, A. (1979) *Biopolymers 18*, 3005–3014.
Hruby, V. J. (1974) *Chem. Biochem. Amino Acids, Pept., Proteins 3*, 1–188.
Hulmes, D. J. S., Miller, A., White, S. W., & Doyle, B. B. (1977) *J. Mol. Biol. 110*, 643–666.
Hulmes, D. J. S., Miller, A., White, S. W., Timmins, P. A., & Berthet-Colominas, C. (1980) *Int. J. Biol. Macromol. 2*, 338–346.
Jardetzky, O., & Roberts, G. C. K. (1981) *NMR in Molecular Biology*, pp 143–186, Academic Press, New York.
Jones, E. Y., & Miller, A. (1987) *Biopolymers 26*, 463–480.
Karplus, M. J. (1963) *J. Am. Chem. Soc. 85*, 2870–2871.
Kessler, H., Griesinger, Ch., Kerssebaum, R., Wagner, R., & Ernst, R. R. (1987) *J. Am. Chem. Soc. 109*, 607–609.
Leach, S. J., Nemethy, G., & Scheraga, H. A. (1977) *Biochem. Biophys. Res. Commun. 75*, 207–215.
Leibovich, S. J., & Weiss, J. B. (1970) *Biochim. Biophys. Acta 214*, 445–454.
Marion, D. (1985) *FEBS Lett. 192*, 99–103.
Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967–974.
Mehlkopf, A. F., Korbee, D., Tiggelmann, T. A., & Freeman, R. (1984) *J. Magn. Reson. 58*, 315–323.
Neuhaus, D., & Keeler, J. (1986) *J. Magn. Reson. 68*, 568–574.
Otter, A., & Kotovych, G. (1987) *J. Magn. Reson. 74*, 293–307.
Otter, A., Scott, P. G., & Kotovych, G. (1987) *J. Am. Chem. Soc. 109*, 6995–7001.
Otter, A., Scott, P. G., & Kotovych, G. (1988) *Biochemistry 27*, 3560–3567.
Provencher, S. W., & Glockner, J. (1981) *Biochemistry 20*, 33–37.
Richardson, J. (1981) *Adv. Protein Chem. 34*, 167–339.
Rubin, A. L., Pfahl, D., Speakman, P. T., Davison, P. F., & Schmitt, F. O. (1963) *Science (Washington, D.C.) 139*, 37–39.
Sanders, J. K. M., & Marsh, J. D. (1982) *Prog. NMR Spectrosc. 15*, 353–371.
Scott, P. G. (1986) *Biochemistry 25*, 974–980.
Sibanda, B. L., & Thornton, J. M. (1985) *Nature 316*, 170–174.
Wagner, G., & Wüthrich, K. (1982) *J. Mol. Biol. 155*, 347–366.
Wider, G., Hosur, R. V., & Wüthrich, K. (1983) *J. Magn. Reson. 52*, 130–135.
Wider, G., Macura, S., Kumar, A., & Ernst, R. R. (1984) *J. Magn. Reson. 56*, 207–234.
Wright, P. E., Dyson, H. J., & Lerner, R. A. (1988) *Biochemistry 27*, 7167–7175.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley-Interscience, New York.