Solution Conformation of Thymosin β_4: A Nuclear Magnetic Resonance and Simulated Annealing Study†

Jutta Zarbock,‡ Hartmut Oschkinat,‡ Ewald Hannappel,§ Hubert Kalbacher,§ Wolfgang Voelter,§ and Tadeusz A. Holak*,‡

Max-Planck Institut für Biochemie, D-8033 Martinsried bei München, FRG, and Abteilung für Physikalische Biochemie des Physiologisch-Chemischen Institutes der Universität Tübingen, Hoppe-Seyler-Strasse 4, 7400 Tübingen, FRG

Received March 6, 1990; Revised Manuscript Received April 30, 1990

ABSTRACT: The conformation of the polypeptide thymosin β_4 in solutions of 60% (v/v) trifluoroethanol-d_3 and 50% (v/v) hexafluoroisopropyl-d_2 alcohol in water is investigated by nuclear magnetic resonance (NMR) spectroscopy. Under these conditions thymosin β_4 adopts an ordered structure. By use of a combination of two-dimensional NMR techniques, the ^{1}H NMR spectrum of thymosin β_4 is assigned. A set of 180 approximate interproton distance constraints is derived from nuclear Overhauser enhancement (NOE) measurements. These, together with 33 ϕ constraints obtained for $J_{NH\alpha}$ coupling data and the 23 ψ dihedral angles identified on the basis of the pattern of short-range NOEs, form the basis of a three-dimensional structure determination by dynamical simulated annealing. The calculations are carried out starting from three initial structures, an α-helix, an extended β-strand, and a mixed α/β structure. Ten independent structures are computed from each starting structure by using different random number seeds for the assignments of the initial velocities. All 30 calculated structures satisfy the experimental constraints, display very small deviations from idealized covalent geometry, and possess good nonbonded contacts. Analysis of the 30 converged structures indicates that there are two helical regions extending from residues 4–16 and from residues 30–40, which are well defined both in terms of atomic root mean square differences and backbone torsion angles. For the two helical regions individually the average backbone rms difference between all pairs of structures is ~2 Å. The two helices exhibit typical amino acid preferences for specific locations at the ends of helices. Pro-4 is found at the (N-end + 1) of the first helix, Gly-41 serves as a terminator of the second helix. The first helix exhibits asymmetrical charge distributions in the first and last helical turn. A loop between residues 24–29 is also evident in the structure. The relative orientation of the helices toward each other, however, could not be determined due to the paucity of NOEs for residues 17 and 18.

Thymosin β_4 is polypeptide of 43 amino acid residues first isolated from calf thymus (Low et al., 1981). The primary structure was confirmed by two different synthetic approaches, an automated solid-phase procedure (Wang et al., 1981), and a synthesis in solution (Kapurniotu et al., 1988). The peptide is widely distributed in numerous tissues (Horecker & Morgan, 1981; Goodall et al., 1983; Hannappel, 1986) and in recent years different β-thymosins were isolated from many species ranging from sea urchin (Schmid et al., 1989) to mammals (Hannappel, 1986). Thymosin β_4 has been reported to induce terminal deoxynucleotidyl transferase activity in vivo and in vitro (Low et al., 1981), inhibit the migration of guinea pig peritoneal macrophages (Thurman et al., 1981), stimulate the hypothalamic secretion of luteinizing hormone releasing factor (Rebar et al., 1981), or induce phenotypic changes in the Molt-4 leukemic cell line (Kokkinopoulos et al., 1985).

Although several biological activities of thymosin β_4 have been reported in the literature (Voelter et al., 1987), its physiological role is still unknown. According to its wide distribution in different types of cells and large quantities (in the range of 100 μg/g of fresh thymus tissue), the expression "thymosin" is misleading and it is unlikely that thymosin β_4 acts as a hormone. It has to be assumed that this polypeptide is responsible for more general functions in the cell.

Recently, a tetrapeptide—Ac-Ser-Asp-Lys-Pro-OH—was isolated from fetal calf bone marrow, exhibiting a high inhibitory activity on the proliferation of hematopoietic pluripotent stem cells (Lenfant et al., 1988). This tetrapeptide is identical with the amino terminal part of thymosin β_4 and it is possible that the polypeptide is a precursor of this inhibitor.

In the present study we present the determination by ^{1}H NMR[1] spectroscopy of the solution conformation of thymosin β_4 under conditions where it adopts an ordered structure (50% v/v HFP-d_2 or 60% TFE-d_3). By use of a variety of two-dimensional NMR techniques (Ernst et al., 1987), the ^{1}H NMR spectrum is assigned in a sequential manner and a set of 180 approximate interproton distance constraints is derived from nuclear Overhauser effect (NOE) measurements. These form the basis of a three-dimensional structure determination by dynamical simulated annealing (Nilges et al., 1988).

Experimental Procedures

Isolation of Thymosin β_4. Thymus tissue (800 g) was pulverized in liquid nitrogen and dispersed into boiling water. After the solution was cooled to room temperature, pyridine and formic acid were added to final concentrations of 0.2 and 1 M, respectively. The resultant homogenate (3800 mL) was

† This work was supported by the Max-Planck Gesellschaft and Grant 0318909A from the Bundesministerium für Forschung and Technologie.

‡ Max-Planck Institut für Biochemie.

§ Universität Tübingen.

[1] Abbreviations: CD, circular dichroism; HFP, 1,1,1,3,3,3-hexafluoropropan-2-ol; TFE, trifluoroethanol; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, two-dimensional NOE spectroscopy; HOHAHA, homonuclear Hartmann–Hahn spectroscopy; SD, standard deviation; rms, root mean square.

0006-2960/90/0429-7814$02.50/0

centrifuged at 3000g for 30 min at room temperature, and the supernatant was adsorbed on 40 g of LiChroprep RP-18 (40–63 μL, Merck) in a glass funnel and washed intensively with water. The desired polypeptide was eluated by addition of 200 mL of 30% 1-propanol containing 2.5% of ampholines (2.5 mL, pH 3.5–9, and 2.5 mL, pH 4–6.5). To recover the ampholines completely from the RP-18 column, 200 mL of 30% 1-propanol was passed through. After evaporation to a volume of 100 mL and addition of 2.5 mL of thiodiglycol, the eluates were mixed with 4 g of Ultrodex (LKB) and the gel was cast into a glass plate (11 × 24.5 cm) and electrofocused for 16 h over the long direction at 8 °C and 8 W (1.5 kV, 20 mA). With a grid, the focused gel was cut into 30 sections. Each of these sections was transferred to a disposable plastic column (1.5 × 80 mm), which contained 1 g of LiChroprep RP-18 (treated with methanol and water) between two frits. The ampholines were eluated with 9 mL of water, while the peptide was still adsorbed on the material. The pH value of each fraction at 4 °C was determined. Thymosin β_4 was usually found at pH 4.6. The peptide was eluated with 9 mL of 30% 1-propanol in water. Aliquots of the eluates were used for protein determination and HPLC analysis. One fraction contained thymosin β_4 in high purity (>98%) and a yield of 42 mg after lyophilization. For further purification semipreparative HPLC was applied under the following conditions: 10 mg of crude thymosin β_4 was dissolved in 1 mL of water and purified on a Nucleosil 7 C_{18} column (10 × 250 mm) with a linear gradient from 0.05% trifluoroacetic acid to 54% acetonitrile in 0.05% trifluoroacetic acid within 30 min. Small amounts of methionine sulfoxide thymosin β_4 (t_R = 18.5 min) could be easily separated from the main product (t_R = 20.3 min).

Sample Preparation. The samples for NMR contained ~6 mM thymosin β_4 in 2.5 mM sodium phosphate buffer, pH 5.0, 60% TFE-d_3, and 40% D_2O or 40% H_2O. The only difference for the HFP-d_2 samples is the concentration of the alcohol (50%). The deuterated, fluorinated alcohols were purchased from Cambridge Isotope Laboratories, Woburn, MA 01801. The CD spectra were run at the concentration of 100 μg/mL of the peptide at 298 K.

NMR Spectroscopy. All NMR measurements were carried out at 500 MHz on a Bruker AM-500 spectrometer. All two-dimensional spectra were recorded in the pure-phase absorption mode by using the time proportional incrementation methods (Redfield & Kuntz, 1975; Bodenhausen et al., 1980; Marion & Wüthrich, 1983). The following spectra were recorded in both D_2O and H_2O: NOESY (Jeener et al., 1979; Macura et al., 1981) with mixing times of 100 and 200 ms and HOHAHA (Braunschweiler et al., 1983; Davis & Bax, 1985) with a MLEV17 mixing sequence (Bax & Davis, 1985) of 23-, 43-, and 60-ms duration. All spectra were measured at 25 °C. For the NOESY spectra in H_2O, the water resonance was suppressed by the use of a semiselective excitation pulse in which the last 90° pulse in the sequence was replaced by the jump–return sequence ($90°_x$–τ–$90°_{-x}$) with the carrier placed at the position of the solvent (Plateau & Gueron, 1982) and value of 80 μs for τ. For the HOHAHA spectra in H_2O, the water resonance was suppressed either by presaturation of water resonance or by adding the pulse sequence $90°_x$–H–Δ–$90°_x$–τ–$90°_{-x}$ at the end of the MLEV17$_y$ sequence (Bax et al., 1987), where H is a homospoil pulse (4 ms), Δ a recovery delay (4 ms), and $90°_x$–τ–$90°_{-x}$ the semiselective jump–return sequence with a value of 80 μs for τ. For the majority of the spectra, 700 t_1 increments were collected, each with 4K data points, over a spectral width of 6 kHz in both dimensions. The

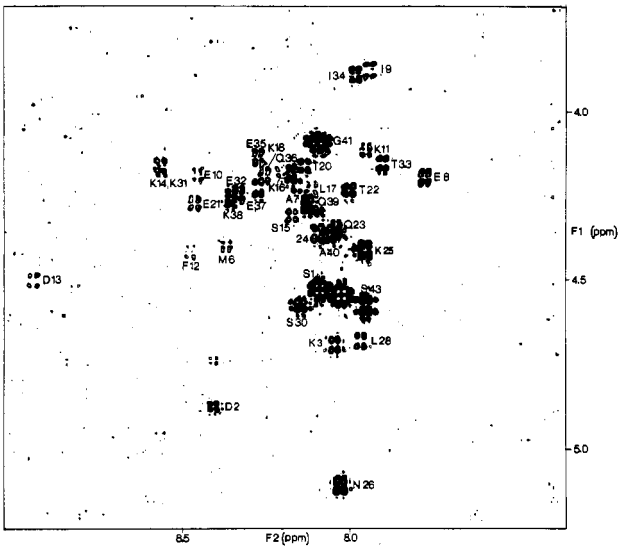

FIGURE 1: Double-quantum filtered phase-sensitive COSY spectrum (TFE-d_3 solution) of the NH ($F2$ axis)–C$^\alpha$H ($F1$ axis) region.

coupling constants were measured directly from the antiphase splitting in DQF–COSY spectrum in H_2O after zero filling the spectrum in both dimension to the final digital resolution of 0.34 Hz.

Calculations. All minimization and simulated annealing calculations were carried out as described by Holak et al. (1989a,b) on a CONVEX C1-XP2 computer using the program XPLOR (Brünger, 1988; Brünger et al., 1986). Structure display was carried out by using a modified version of the function network of FRODO (Jones, 1978) interfaced with XPLOR on an Evans & Sutherland PS390 color graphics system.

Results

Secondary Structure. The circular dichroism and ^{1}H NMR spectra of thymosin β_4 in water show no obvious regular structure. Addition of TFE (or HFP) to the aqueous peptide solution caused a change in conformation, as evidenced by changes in the CD spectrum, which were first detectable at about 30% (v/v) of the alcohol and complete by about 60%. Although the resulting spectrum was not clearly interpretable, we inferred that the presence of the alcohol had stabilized a particular secondary conformation, and therefore, the NMR investigation was carried out in the mixed solvent system: 50–60% alcohol/50–40% water. Sequence-specific resonance assignment was carried out in a sequential manner (Wüthrich et al., 1982; Billeter et al., 1982). HOHAHA spectra (Davis & Bax, 1985), recorded at several mixing times ranging from 20 to 60 ms to detect successively direct, single, and relayed through-bond (scalar) connectivities, were used to identify amino acid spin systems. NOESY spectra were used to observe through-space (<5 Å) connectivities. For the purposes of sequential assignment the most useful NOEs involve the NH, C$^\alpha$H, and C$^\beta$H protons, as well as the C$^\delta$H protons of proline, and are of the type NH(i)–NH($i\pm1$), C$^\alpha$H(i)–NH(i+1,3,4), and C$^\alpha$H(i)–C$^\beta$H(i+3). Examples of NOESY and COSY spectra are shown in Figures 1 and 2, and a complete list of assignments in HFP-d_2 is given in Table I. Another table containing chemical shifts of thymosin β_4 in TFE-d_3 together with an example of the NH–NH region of the NOESY spectrum taken in HFP-d_2 has been included in the supplementary material.

A summary of the short-range ($|i-j| < 5$) NOEs involving the NH, C$^\alpha$H, and C$^\beta$H protons is shown in Figure 3 and

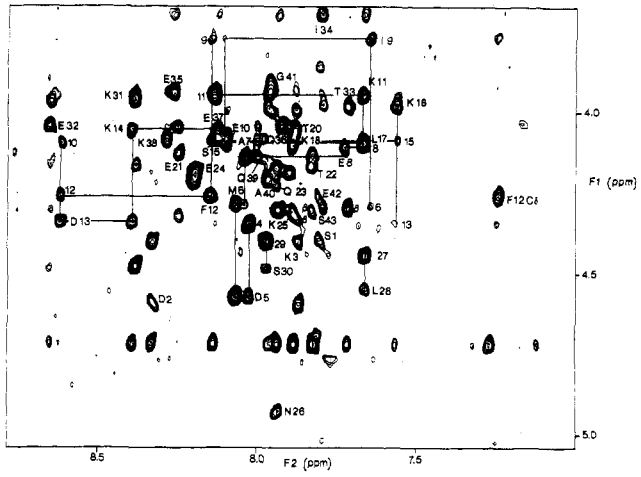

FIGURE 2: NOESY spectrum (HFP-d_2 solution) of the NH ($F2$ axis)–C$^{\alpha}$H ($F1$ axis) region of thymosin in 50% HFP-d_2 at 25 °C with a mixing time of 100 ms. The crosspeaks arise from through-space (<5 Å) connectivities and some examples of C$^{\alpha}$H(i)–NH(i+1) and C$^{\alpha}$H(i)–NH(i+3) connectivities are shown. Intraresidue crosspeaks are labeled with one-letter abbreviations and interresidue connectivities are indicated by simple numbers.

provides the basis for a qualitative interpretation of the secondary structure (Wüthrich et al., 1984, Wüthrich, 1986). The stretches of NH(i)–NH(i+1) NOEs together with the presence of C$^{\alpha}$H(i)–NH(i+3,4) and C$^{\alpha}$H(i)–C$^{\beta}$H(i+3) NOEs suggest that there are two helices extending from residues 5–16 and 30–40. Studies of other peptides in fluorinated alcohols suggest that they stabilize the helix, but not indiscriminately: regions that have a intrinsic high tendency to adopt the helix become more helical, but regions that are helix-disfavoring are not induced to become helical (Nelson & Kallenbach, 1986, 1989; Bruch et al., 1989).

Interproton Distance Constraints and Structure Computation. A set of 180 approximate interproton distance constraints were derived from the NOESY spectra recorded at 298 K with a mixing time of 100 ms and 33 ϕ angles from the COSY spectrum. The input data for calculations were derived from the spectra of thymosin in HFP-d_2. For partially overlapped peaks, a few distance constraints were estimated from the comparison with the NOESY spectrum of thymosin β_4 in TFE-d_3. Despite differences in the chemical shifts for the two solutions, the NOE patterns in the NOESY spectra in both solutions were nearly identical. The distances were classified into four distance ranges, 2.1–2.6, 2.5–3.1, 3.0–3.7, and 3.6–5.0 Å, corresponding to strong, medium, weak, and very weak NOEs, respectively (Williamson et al., 1985; Holak et al., 1989a,b). The upper limits for distances involving either methyl or methylene protons with equivalent chemical shifts were corrected with pseudoatom terms as described by

Table I: Chemical Shifts (ppm) of Thymosin β_4 in HFP-d_2/Water[a] (50%/50% v/v)

residue	NH	C$^{\alpha}$H	C$^{\beta}$H	others
S 1	7.80	4.41	4.02/3.96	
D 2	8.33	4.61	2.74/2.60	
K 3	7.89	4.42	2.37/2.30	C$^{\gamma}$H 1.83
P 4		4.35	2.20	C$^{\gamma}$H 1.86
D 5	8.03	4.58	2.78/2.69	
M 6	8.07	4.30	2.12/2.04	C$^{\gamma}$H 2.86/2.56
A 7	8.03	4.16	1.45	
E 8	7.73	4.13	2.11/1.96	C$^{\gamma}$H 2.34
I 9	7.65	3.78	1.65	C$^{\gamma}$H$_3$ 0.95, C$^{\gamma}$H$_2$ 1.17, C$^{\delta}$H$_3$ 0.82
E 10	8.10	4.11	1.90/2.10	C$^{\gamma}$H 2.36
K 11	7.67	3.96	1.85	C$^{\gamma}$H 1.65
F 12	8.15	4.27	3.25	C$^{\delta}$H 7.24
D 13	8.62	4.35	2.82/2.74	
K 14	8.39	4.17	1.90	C$^{\gamma}$H 1.39, C$^{\delta}$H 1.50
S 15	8.14	4.10	3.97	
K 16	7.57	4.00	1.80	C$^{\gamma}$H 1.37, C$^{\delta}$H 1.60
L 17	7.68	4.11	1.64	C$^{\gamma}$H 1.35, C$^{\delta}$H$_3$ 0.90
K 18	7.90	4.12		
K 19	7.88	4.64		
T 20	7.88	4.07	4.34	C$^{\gamma}$H$_3$ 1.24
E 21	8.25	4.14	2.15	C$^{\gamma}$H 2.40
T 22	7.83	4.17	4.31	C$^{\gamma}$H$_3$ 1.24
Q 23	7.94	4.23	2.15	C$^{\gamma}$H 2.40, NH$_2$ 7.30/6.52
E 24	8.20	4.20	2.07/2.01	C$^{\gamma}$H 2.30
K 25	7.90	4.31	1.83/1.75	C$^{\gamma}$H 1.47
N 26	7.94	4.94	2.82/2.70	NH$_2$ 7.32/6.37
P 27		4.45		C$^{\delta}$H 3.71
L 28	7.67	4.56	1.69/1.64	C$^{\gamma}$H 1.37, C$^{\delta}$H$_3$ 0.89/0.84
P 29		4.41	2.18/1.90	
S 30	7.98	4.50		
K 31	8.39	3.97	1.86/1.80	C$^{\gamma}$H 1.41
E 32	8.65	4.06	2.04/1.96	C$^{\gamma}$H 2.43/2.28
T 33	7.92	4.00	4.31	C$^{\gamma}$H$_3$ 1.24
I 34	7.80	3.71	1.91	C$^{\gamma}$H$_3$ 0.87, C$^{\gamma}$H$_2$, 1.60/1.10, C$^{\delta}$H$_3$ 0.83
E 35	8.26	3.96	2.08/2.02	C$^{\gamma}$H 2.45/2.28
Q 36	7.93	4.05	2.22/2.18	C$^{\gamma}$H 2.48/2.18, NH$_2$ 7.30/6.52
E 37	8.13	4.08	2.14	C$^{\gamma}$H 2.39/2.34
K 38	8.29	4.11	1.90	C$^{\gamma}$H 1.50
Q 39	8.00	4.16	2.15/2.11	C$^{\gamma}$H 2.48/2.39, NH$_2$ 7.02/6.46
A 40	7.96	4.22	1.46	
G 41	7.96	3.96		
E 42	7.81	4.30		
S 43	7.80	4.28		

[a] The chemical shifts are referenced to DSS through the calibration to the water signal in the spectra taken at 298 K, pH 5.0.

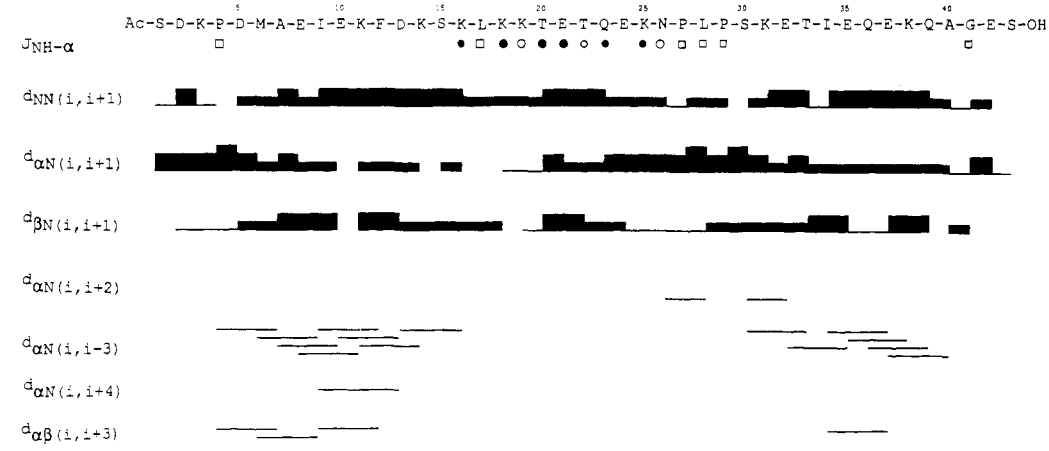

FIGURE 3: Summary of the short range ($|i - j| < 5$) connectivities involving the NH, C$^{\alpha}$H, and C$^{\beta}$H protons. Only the intensities of sequential NOEs are reflected in the thickness of the lines. The C$^{\alpha}$H(i)–C$^{\delta}$H(i+1) (Pro) NOE is shown along the same line as the C$^{\alpha}$H(i)–NH(i+1) connectivities. Ranges of coupling constants: filled circle, $^3J_{NH-C^{\alpha}H} > 11$ Hz; circle, $11 > {}^3J_{NH-C^{\alpha}H} > 10$ Hz; square, no data available; without sign, $^3J_{NH-C^{\alpha}H} < 9$ Hz.

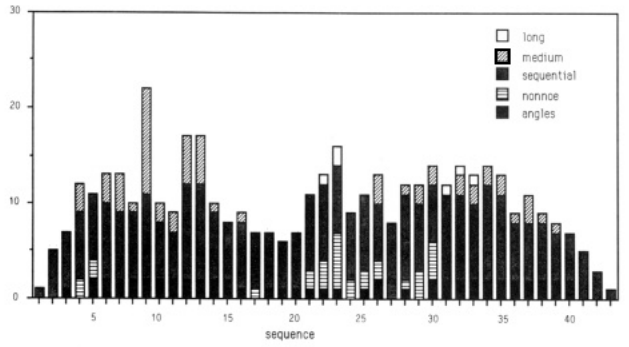

FIGURE 4: Plot of the number of NOE distance and angle constraints per residue versus the amino acid sequence of thymosin β_4. All constraints appear twice, once for each interacting residue. No intraresidue constraints were used.

Wüthrich et al. (1983). The methylene protons and methyl groups with two separate signals were allowed to flip at the prochiral centers during the simulated annealing stage of the calculations (Holak et al., 1989a,b). Although no stereospecific assignments of the side-chain conformations were obtained in the floating chirality calculations, use of the method makes the pseudoatom corrections unnecessary. Thus the information content of the input data is not reduced through introducing comparatively large pseudoatom correction terms.

The interproton distance constraints were supplemented by backbone ϕ torsion angle constraints derived from $^3J_{NH\alpha}$ coupling constant data in a manner analogous to that described by Pardi et al. (1984). Three different ranges are summarized in Figure 3. Fifteen non-NOE distance constraints were present in the calculations (Holak et al., 1989a,b). These constraints served to exclude any close contacts (< ca. 2 Å), which would contradict the lack of an observable NOE in the NOESY spectra. They were introduced for residues between 20 and 29 after a few initial calculations. The lower limit for these constraints was set to 3.0 Å, while the upper limit was not set. After including 15 such constraints, no abnormally short contacts were observed in the structures. The non-NOEs were also tried for additional constraints on residues 20–29, which have few medium-range NOEs. No noticeable reduction in the spread of the structures in this fragment was achieved with the introduction of these constraints, however. At this point, no more non-NOEs were introduced because of the variability of the fragment emerging in the calculated structures. The absence of NOEs may result from mechanisms other than nonproximity in space, for example, if the variability were due to increased local mobility of the fragment. The number of NOE distance and angle constraints per residue is given in Figure 4. The complete list of all distance constraints used in the computations is included in the supplementary material.

The approach used for the determination of the three-dimensional structure of thymosin involved the application of simulated annealing (Nilges et al., 1988). The basis of simulated annealing involves raising the temperature of the system followed by slow cooling in order to overcome local minima and locate the global minimum region of the target function (Nilges et al., 1988).

The total target function for which the global minimum is searched is made up of the following terms:

$$F_{tot} = F_{covalent} + F_{repel} + F_{NOE} + F_{\phi}$$

$F_{covalent}$ maintains correct bond lengths, angles, planes, and chirality. The force constants for the bond angle and improper torsions are set to uniform high value of 500 kcal/mol to ensure near-perfect stereochemistry of single amino acids throughout the calculation. F_{NOE} is a square-well potential for the experimental distance constraints (Nilges et al., 1988). Finally, F_{ϕ} is a square-well dihedral potential term (Clore et al., 1986) used to restrict the ϕ backbone torsion angles.

Three starting structures were used for the calculations: an α-helix, an extended β-strand, and a mixed $\alpha\beta$ structure in which residues 5–16 and 30–41 were α-helices, as predicted from the qualitative interpretation of the NOE data, and residues 1–4 and 17–29 β-strands. All the side chains were placed in an extended conformation (χ_i = 180°).

The annealing schedule was carried out in two steps. Step 1 comprised 50 cycles of 75-fs dynamics each. The initial velocities were chosen from a Maxwellian distribution at 1000 K and after each cycle the velocities were rescaled to this temperature. The value of the force constant for the experimental distance constraints, k_{NOE}, was doubled at the beginning of each cycle from an initial value of 0.2 kcal·mol^{-1}·Å^{-2} to a maximum value of 40 kcal·mol^{-1}·Å^{-2}. The maximum value of k_{NOE} is lower than that used in our previous calculations (Holak et al., 1989a,b). With a higher than 40 kcal·mol^{-1}·Å^{-2} value of k_{NOE}, the β-strand did not converge into an α-helix. The larger value of the NOE force constant makes the NOE constraints too rigid to allow for substantial movements of atoms in the peptide. In addition, to make rearrangement easily possible, the force constant for the repulsion term was multiplied by a factor of $400^{1/50}$ prior to each cycle from an initial value of 0.01 kcal·mol^{-1}·Å^{-2} to a final value of 4 kcal·mol^{-1}·Å^{-4}. Step 2 consisted of 1.5-ps dynamics at 300 K with values of the force constants k_{NOE} and k_{rep} equal to their final values at the end of step 1. The velocities were rescaled to 300 K every 150 fs during this period. This was followed by 200 cycles of restrained Powell minimization.

The simulated annealing schedule was repeated on average once for the structures obtained from the α-helix and mixed $\alpha\beta$ structure and three times for the structures starting off from an extended β-strand in order convergence with respect to the experimental constraints (e.g., the structure obtained after the first pass through the annealing schedule was used as the input structure for the second pass and so on). Ten independent calculations were carried out from each initial structure by using different random number seeds for the assignment of the initial velocities. The 10 structures calculated from an α-helix are referred to collectively as ⟨SAα⟩, those from an extended β-strand as ⟨SAβ⟩, and those from the mixed $\alpha\beta$ structure as ⟨SA$\alpha\beta$⟩. All 30 final structures are referred to collectively as ⟨SA⟩.

The Converged Structures. The structural statistics of the initial and final structures are given in Table II. All the ⟨SA⟩ structures satisfy the experimental constraints, have good nonbonded contacts as evidenced by low values of F_{repel} and negative Lennard-Jones van der Waals energies, and display very small deviations from idealized covalent geometry.

Superpositions of the final ⟨SA⟩ structures are shown in Figure 5, and plots of atomic rms distributions and ϕ,ψ rms differences are given in Figures 6, 7, and 8, respectively. Atomic rms differences between the various structures are given in Table III.

Discussion

The main structural observation to emerge from the analysis of the calculated structures is that there are two helical regions extending from 4–16 and residues 30–40. The two α-helices are very well defined in terms of both atomic rms differences (Figures 6 and 7) and backbone torsion angles (Figure 8). The average backbone atomic rms difference between all pairs of structures for the two helical regions individually is <2.0 Å

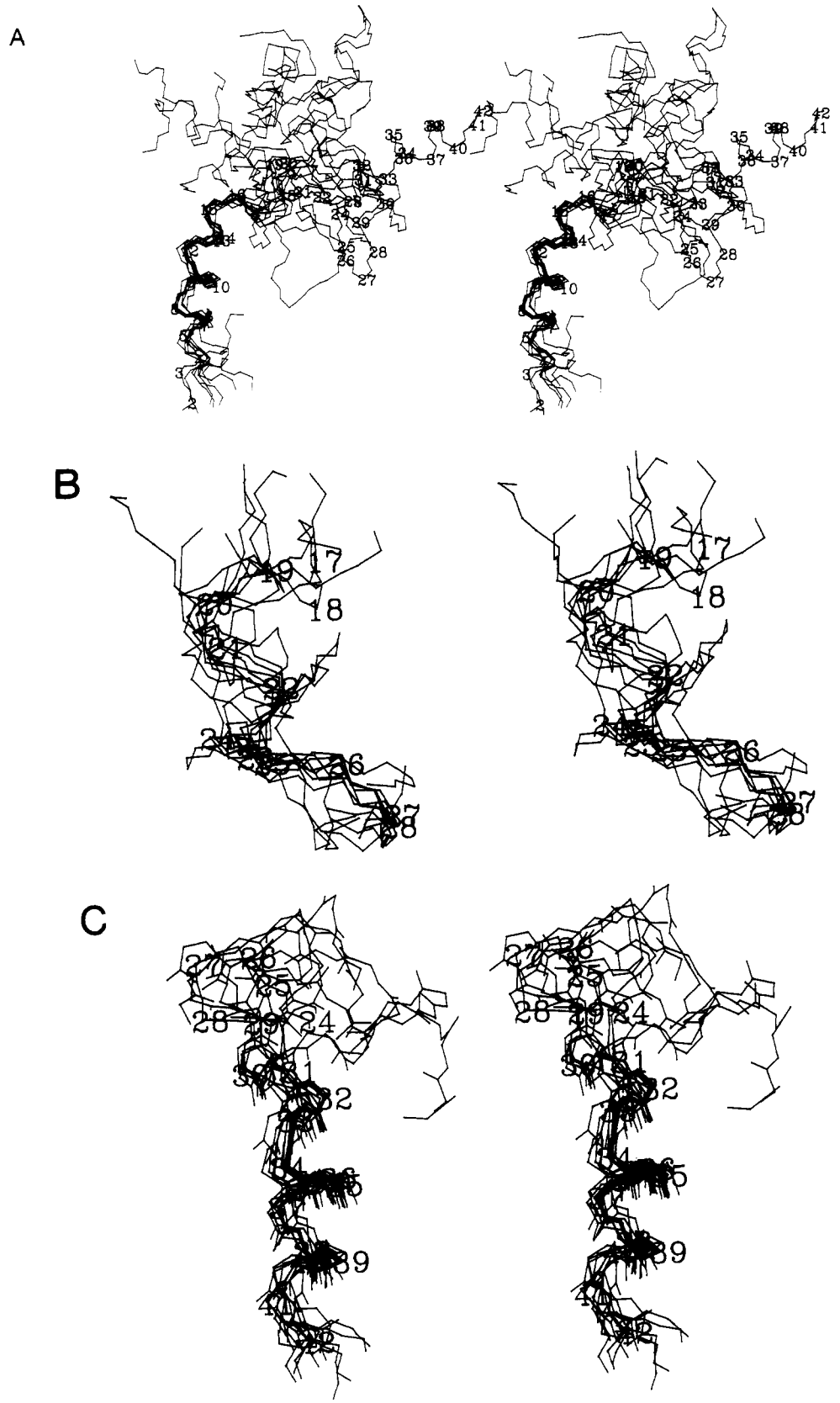

FIGURE 5: Stereoviews of the final structures. Four SAα, three SAβ, and three SA$\alpha\beta$ structures are shown, collectively. (A) Superposition of backbone atoms of residues 2–42 best fitted to residue 5–15 of helix I. (B) Superposition to backbone atoms of residues 16–28 fitted to the same region. (C) Superposition of backbone atoms of residue 24–42 best fitted to residue 30–41 of helix II.

and the average rms difference in the ϕ,ψ angles is <30°. The overall average values of 61 ± 20° and −41 ± 30° for the ϕ and ψ angles, respectively, in this region are very close to those of an idealized α-helix. Confidence in the conformation of these regions in the ⟨SA⟩ structures is also afforded by the convergence from an extended β-strand as an initial structure that is accompanied by backbone atomic rms shifts of 10.0 ± 0.2 Å for helix 1 and 9.3 ± 0.3 Å for helix 2.

Helix boundaries are usually difficult to determine precisely. They can be defined in terms of ϕ,ψ angles (IUPAC-IUB,

Table II: Structural Statistics[a]

	final structures		
	⟨SAα⟩	⟨SAβ⟩	⟨SA$\alpha\beta$⟩
rms deviations from exptl constraints (Å)[b]			
all (180)	0.123 ± 0.01	0.143 ± 0.02	0.117 ± 0.01
non-NOE distances (15)	0.02 ± 0.02	0.035 ± 0.03	0.04 ± 0.04
F_{NOE} (kcal·mol^{-1})[c]	82 ± 15	111 ± 33	75 ± 11
F_ϕ (kcal·mol^{-1})[d]	0.6 ± 0.4	2.7 ± 1.9	0.7 ± 0.5
F_{repel} (kcal·mol^{-1})[e]	2.6 ± 1.3	10.0 ± 6.9	3.7 ± 0.9
F_{L-J} (kcal·mol^{-1})[f]	−156 ± 9	−125 ± 12	−152 ± 13
deviations from idealized geometry[g]			
bonds (Å) (693)	0.011 ± 0.002	0.014 ± 0.002	0.012 ± 0.001
angles (deg) (1274)	2.10 ± 0.10	2.34 ± 0.16	2.16 ± 0.05
impropers (deg) (116)	0.337 ± 0.098	0.684 ± 0.36	0.365 ± 0.071

[a] ⟨SAα⟩, ⟨SAβ⟩, and ⟨SA$\alpha\beta$⟩ are the three final categories of converged structures derived from an α-helix, an extended β-strand, and a mixed $\alpha\beta$ structure, respectively. There are 10 individual structures in each final category (see text). [b] The rms deviations from the experimental constraints are calculated with respect to the upper and lower limits of the distance constraints. The number of distances in each category is given in parentheses next to the category name. [c] The values of the square-well NOE potential F_{NOE} are calculated with a force constant of 50 kcal·mol^{-1}·Å^{-2}. [d] The values of F_ϕ are calculated with a force constant of 200 kcal·mol^{-1}·rad^{-2}. F_ϕ is a square-well dihedral potential that is used to restrict the ϕ backbone torsion angles of all residues. [e] The values of the van der Waals repulsion term F_{repel} are calculated with a force constant of 4 kcal·mol^{-1}·Å^{-4} with the hard sphere van der Waals radii set to 0.8 times the standard values used in the CHARMM empirical energy function (Brooks et al., 1983). [f] E_{L-J} is the Lennard-Jones van der Waals energy calculated by using the CHARMM empirical energy function (Brooks et al., 1983). [g] The number of bond, angle, and improper terms is given in parentheses. The improper terms serve to maintain planarity and appropriate chirality.

Table III: Atomic Rms Differences[a]

	atomic rms differences (Å)					
	residues 5–15		residues 17–29		residues 30–40	
	A	B	A	B	A	B
⟨SAα⟩ vs ⟨SAα⟩$_{(5-15)}$	0.7 ± 0.4	0.8 ± 0.3				
⟨SAβ⟩ vs ⟨SAβ⟩$_{(5-15)}$	0.9 ± 0.5	2.3 ± 0.8				
⟨SA$\alpha\beta$⟩ vs ⟨SA$\alpha\beta$⟩$_{(5-15)}$	0.7 ± 0.5	1.6 ± 0.5				
⟨SAα⟩ vs ⟨SAα⟩$_{(17-29)}$			2.2 ± 0.9	3.1 ± 0.7		
⟨SAβ⟩ vs ⟨SAβ⟩$_{(17-29)}$			3.4 ± 1.6	5.2 ± 1.7		
⟨SA$\alpha\beta$⟩ vs ⟨SA$\alpha\beta$⟩$_{(17-29)}$			3.4 ± 2.1	4.6 ± 2.2		
⟨SAα⟩ vs ⟨SAα⟩$_{(30-40)}$					1.0 ± 0.6	2.1 ± 0.8
⟨SAβ⟩ vs ⟨SAβ⟩$_{(30-40)}$					1.0 ± 0.8	2.5 ± 1.1
⟨SA$\alpha\beta$⟩ vs ⟨SA$\alpha\beta$⟩$_{(30-40)}$					1.3 ± 0.6	2.2 ± 0.7

[a] The notation of the structures is the same as in Table II. A, backbone atoms; B, all atoms. The regions to which the coordinates are best fitted are given in parentheses.

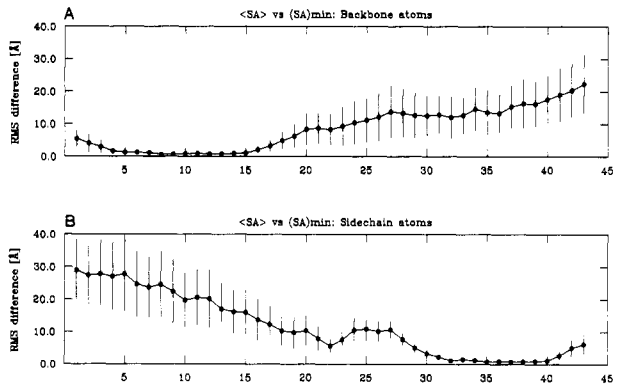

FIGURE 6: Atomic rms distributions of the 30 ⟨SA⟩ structures for all residues best fitted to residues 5–16 (A) and 30–41 (B). The backbone atoms comprise the C, C$^\alpha$, N, and O atoms. The filled-in circles represent the average values and the bars the standard deviations in these values.

1970), hydrogen bonding (Kabsch et al., 1983), or α-carbon positions (Levitt & Greer, 1977). We use here a method recently suggested by Richardson and Richardson (1988), which uses α-carbons positions with the visual examination of structures on computer graphics. In such a description Lys-3 is considered to be the N-cap, the interface residue. Pro-4 is then the first residue inside the first α-helix. The helix ends with Lys-16 as a C-cap. Although only one of the ϕ,ψ angles of Pro-4 has the correct conformational angle, Pro-4 can be considered to be the first residue inside the α-helix because in the calculated structures its carbonyl participates in a weak hydrogen bond. The helix exhibits an asymmetrical preference for negatively charged side chains in the first turn and positively charged ones in the last turn. Such a preference was observed for many α-helices (Ptitsyn, 1969; Shoemaker et al., 1987). Also, the high position-specific preference for proline at position N1 (the first residue after the N-cap) was recently found in a survey of 215 α-helices (Richardson & Richardson, 1988).

The second helix starts at residues Ser-30, which is treated as the N-cap. Although ϕ,ψ angles are typical of the α-helix, no hydrogen bond is indicated for this residue. The helix ends with the C-cap at Gly-41 (Figure 5C). This helix again exhibits strong residue preferences that have been previously observed in α-helices, e.g., accentuated known residue preferences of Gly at the C-cap (Schellmann, 1980) and Ser at the N-Cap (Kendrew et al., 1961).

The exact orientation of the two helices with respect to each other could not be defined. This is clearly evident in the atomic rms distribution plots (Figure 6) and by the large angular rms differences in the ϕ,ψ angles of residues 21–24 (Figure 8). This is not very surprising as all the interresidue NOEs are of a short-range nature, and only few NOEs involving residues more than four apart in the sequence could be detected.

The fluorinated alcohol/water solvent mixture leads to substantially broader line widths than water alone. The origin of this effect is not known (Bruch et al., 1989). Because of

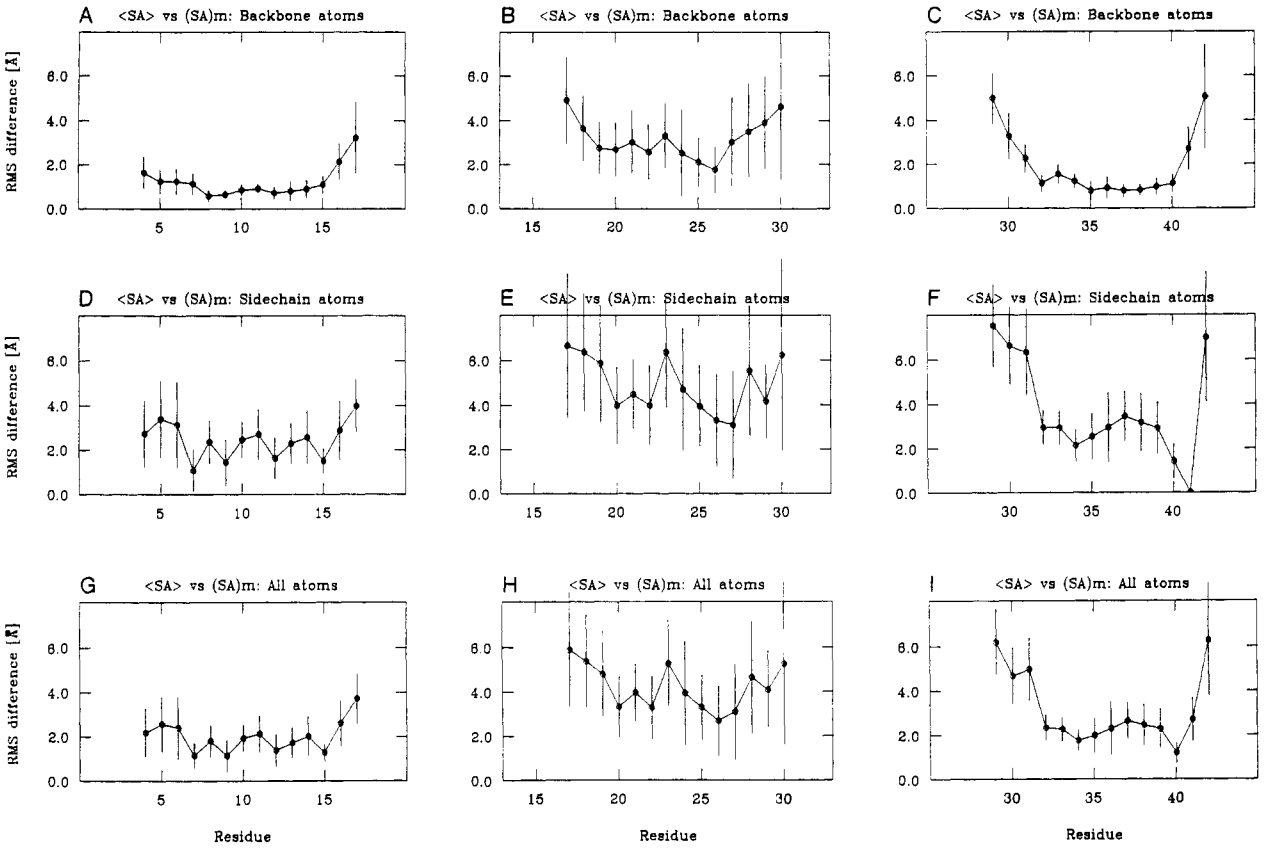

FIGURE 7: Atomic rms distributions for backbone atoms (A, B, C), side-chain atoms (D, E, F), and all atoms (G, H, I) of the 30 ⟨SA⟩ structures best fitted to three different regions (see Table III).

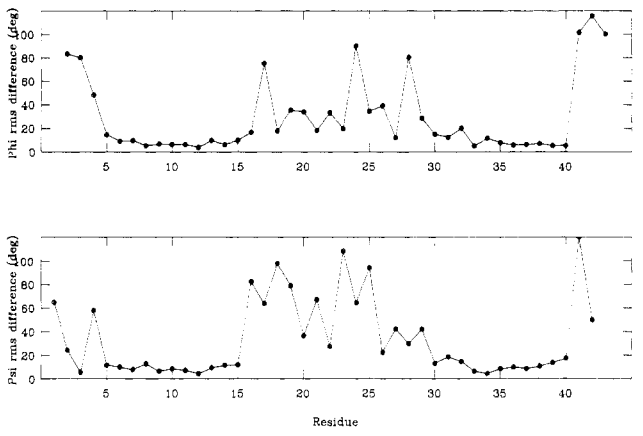

FIGURE 8: Average angular ϕ and ψ rms differences between all pairs of ⟨SA⟩ structures.

the broad line widths, we could not determine coupling constants smaller than 8 Hz (Figure 3). The lack of the coupling constant data and NH exchange rates, together with the fact that only few $d_{\alpha N}$ (i, i+2,4) connectivities were observed, precludes discussion of relative stabilities of the α-helices. It is important to remember that in the case of fast interconversion between more than one conformation, the measured NMR parameters are the population-weighted average of the parameters associated with the individual states. The implications of the presence of fast conformational equilibria and other motional averaging effects for the determination of structures of proteins by NMR were discussed by Jardetzky (1980), Kessler et al. (1988), and Pepermans et al. (1988). Taking into account the considerations presented by these authors, we conclude that we were able to generate a set of structures that satisfy all the NOE constraints. Also, no conflicting NOEs (the NOE distances that are mutually exclusive, considering a single rigid structure) were detected. Analysis of the 30 converged structures indicates that thymosin β_4 in fluorinated alcohol/water mixtures adopts a conformation with two helical regions extending from residue 4 to 16 and from residue 30 to 40, which are well defined both in terms of atomic root mean square differences and backbone torsion angles.

Supplementary Material Available

One figure showing the NH–NH region of the NOESY spectrum for thymosin β_4 in HFP-d_2, one table containing chemical shift information for the ^{1}H NMR spectrum of thymosin β_4 in TFE-d_3, and one table giving the complete list of NOE interproton distance constraints used in the computation of the three-dimensional structure of thymosin (6 pages). Ordering information is given on any current masthead page.

Registry No. Thymosin β_4, 77642-24-1.

References

Bax, A., & Davis, D. G. (1985) *J. Magn. Reson. 65*, 355–366.
Bax, A., Sklenar, V., Clore, G. M., & Gronenborn, A. M. (1987) *J. Am. Chem. Soc. 109*, 6511–6513.
Billeter, M., Braun, W., & Wüthrich, K. (1982) *J. Mol. Biol. 153*, 321–345.
Bodenhausen, G., Vold, R. L., & Vold, R. R. (1980) *J. Magn. Reson. 37*, 93–106.
Braunschweiler, L., & Ernst, R. R. (1983) *J. Magn. Reson. 53*, 521–558.
Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., & Karplus, M. (1983) *J. Comput. Chem. 4*, 187–217.

Bruch, M. D., Mc Knight, C. J., & Gierasch, L. M. (1989) *Biochemistry 28*, 8554–8561.
Brünger, A. T. (1988) *J. Mol. Biol. 203*, 803–816.
Brünger, A. T., Clore, G. M., Gronenborn, A. M., & Karplus, M. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 3801–3805.
Clore, G. M., Nilges, M., Sukumaran, D. K., Brünger, A. T., Karplus, M., & Gronenborn, A. M. (1986) *EMBO J. 5*, 2729–2735.
Davis, D. G., & Bax, A. (1985) *J. Am. Chem. Soc. 107*, 2821–2822.
Ernst, R. R., Bodenhausen, G., & Wokaun, A. (1987) *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*, Clarendon Press, Oxford.
Goodall, G. J., Morgan, J. I., & Horecker, B. L. (1983) *Arch. Biochem. Biophys. 221*, 598–601.
Hannappel, E. (1986) *Anal. Biochem. 156*, 390–396.
Hannappel, E. (1987) in *Thymusfaktoren, Thymuspraeparate* (Hager, E. D., Ed.) pp 64–73, Gustav Fischer Verlag, Stuttgart, New York.
Holak, T. A., Nilges, M., & Oschkinat, H. (1989a) *FEBS Lett. 242*, 218–242.
Holak, T. A., Gondol, D., Otlewski, J., & Wilusz, T. (1989b) *J. Mol. Biol. 210*, 635–648.
Horecker, B. L., & Morgan, J. I. (1981) *Lymphokines 9*, 15–35.
IUPAC–IUB Commission on Biochemical Nomenclature (1970) *J. Biol. Chem. 245*, 6489 ff.
Jardetzky, O. (1980) *Biochim. Biophys. Acta 621*, 227–232.
Jeener, J., Meier, B. H., Bachman, P., & Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546–4553.
Jones, T. A. (1978) *J. Appl. Crystallogr. 11*, 268–272.
Kabsch, W., & Sander, C. (1983) *Biopolymers 22*, 2577 ff.
Kapurniotu, A., Link, P., & Voelter, W. (1989) in *Peptides 1988* (Jung, G., & Bayer, E., Eds.) pp 97–99, Walter de Gruyter, Berlin, New York.
Kendrew, J. C., Watson, H. C., Strandberg, B. E., & Dickerson, R. E. (1961) *Nature 190*, 666 ff.
Kessler, H., Griesinger, C., Lautz, J., Müller, A., van Gunsteren, W. F., & Berendsen, H. J. C. (1988) *J. Am. Chem. Soc. 110*, 3393–3396.
Kokkinopoulos, D., Perez, S., & Papamichail, M. (1985) *Blut 50*, 341–348.
Lenfant, M., Wdzieczak-Bakala, J., Guittet, E., Prome, J.-C., Sotty, D., & Frindel, E. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 779–782.
Levitt, M., & Greer, J. (1977) *J. Mol. Biol. 114*, 181 ff.
Low, T. K. L., Hu, S.-K., & Goldstein, A. L. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 1162–1166.
Macura, S., Huang, Y., Suter, D., & Ernst, R. R. (1981) *J. Magn. Reson. 43*, 259–281.
Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967–974.
Nelson, J. W., & Kallenbach, N. R. (1986) *Proteins: Struct., Funct., Genet. 1*, 211–217.
Nelson, J. W., & Kallenbach, N. R. (1989) *Biochemistry 28*, 5256–5261.
Nilges, M., Gronenborn, A. M., Brünger, A. T., & Clore, G. M. (1988) *Protein Eng. 2*, 27–38.
Pardi, A., Billeter, M., & Wüthrich, K. (1984) *J. Mol. Biol. 180*, 741–751.
Pepermans, H., Tourwe, D., van Binst, G., Boelens, R., Scheek, R. M., van Gunsteren, W. F., & Kaptein, R. (1988) *Biopolymers 27*, 323–338.
Plateau, P., & Gueron, M. (1982) *J. Am. Chem. Soc. 104*, 7310–7311.
Ptitsyn, O. B. (1969) *J. Mol. Biol. 42*, 501.
Rebar, R. W., Miyake, A., Low, T. K. L., & Goldstein, A. L. (1981) *Science 214*, 669–671.
Redfield, A. G., & Kuntz, S. D. (1975) *J. Magn. Reson. 19*, 250–254.
Richardson, J. S., & Richardson, D. C. (1988) *Science 240*, 1648–1652.
Schellman, C. (1980) in *Protein Folding* (Jänicke, R., Ed.) p 53ff, Elsevier/North-Holland, New York.
Schmid, B. (1989) Ph.D. Thesis, Abteilung für Physikalische Biochemie, University of Tübingen.
Shoemaker, K. R., Kim, P. S., York, E. J., & Baldwin, R. L. (1987) *Nature 326*, 563 ff.
Thurman, G. B., Low, T. K. L., Rossio, J. L., & Goldstein, A. L. (1981) in *Lymphokines and Thymic Hormons: Their Potential Utilization in Cancer Therapeutics* (Goldstein, A. L., & Chirigos, M. A., Eds.) pp 145–157, Raven Press, New York.
Voelter, W., Echner, E., Kalbacher, H., Dinh, T. Q., Kapurniotu, A., Jahan, M., Gruebler, G., Link, P., Mihelic, M., & Stock, W. (1987) in *Peptides 1986* (Theodoropoulos, D., Ed.) pp 581–584, Walter de Gruyter, Berlin, New York.
Wang, S. S., Wang, B. S. H., Chang, J. K., Low, T. K. L., & Goldstein, A. L. (1981) *Int. J. Pept. Protein Res. 18*, 413–415.
Williamson, M. P., Havel, T. F., & Wüthrich, K. (1985) *J. Mol. Biol. 182*, 295–315.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, J. Wiley, New York.
Wüthrich, K., Wider, G., Wagner, G., & Braun, W. (1982) *J. Mol. Biol. 155*, 311–319.
Wüthrich, K., Billeter, M., & Braun, W. (1983) *J. Mol. Biol. 169*, 949–961.
Wüthrich, K., Billeter, M., & Braun, W. (1984) *J. Mol. Biol. 180*, 715–740.