J. Mol. Biol. (1982) **159**, 323–351

Assignment of the ^{1}H Nuclear Magnetic Resonance Spectrum of the Trypsin Inhibitor E from *Dendroaspis polylepis polylepis*

Two-dimensional Nuclear Magnetic Resonance at 500 MHz

ALEXANDER S. ARSENIEV[1]†, GERHARD WIDER[1]
FRANS J. JOUBERT[2] AND KURT WÜTHRICH[1]

[1]*Institut für Molekularbiologie und Biophysik*
Eidgenössische Technische Hochschule
ETH-Hönggerberg, CH-8093 Zürich, Switzerland

[2]*National Chemical Research Laboratory, Pretoria, South Africa*

(Received 18 January 1982, and in revised form 1 April 1982)

The assignment of the ^{1}H nuclear magnetic resonance spectrum of the trypsin inhibitor E from the venom of *Dendroaspis polylepis polylepis* is described and documented. A sample of 18 mg of the protein was investigated with two-dimensional nuclear magnetic resonance experiments at 500 MHz. The assignments are based entirely on the amino acid sequence and the nuclear magnetic resonance data. Individual assignments were obtained for the backbone and C^β protons of all 59 residues of inhibitor E, with the exceptions of the N-terminal amino group, Pro8, Pro13 and the amide proton of Gly37. The amino acid side-chain resonance assignments are complete, with the exception of Pro8, Pro13, $C^\gamma H_2$ of Glu49, all the lysyl and arginyl residues and the three histidyl residues 1, 34 and 53, for which the imidazole ring proton lines have not been assigned individually. The chemical shifts for the assigned resonances are listed for an aqueous solution at 50°C and pH 3·2.

1. Introduction

This paper is part of a comparative nuclear magnetic resonance investigation of three protease inhibitors with molecular weights of 6500 to 7000, i.e. the basic pancreatic trypsin inhibitor from bovine organs (Tschesche, 1974) and the trypsin inhibitors E and K from the venom of the snake *Dendroaspis polylepis polylepis* (Strydom, 1973; Joubert & Strydom, 1978). As a basis for the determination of the three-dimensional structures in solution (Wüthrich *et al.*, 1982), the ^{1}H n.m.r.‡ spectra of these three proteins have been assigned, and this paper describes these data for the inhibitor E. The individual assignments for BPTI were published

† On leave from the Shemyakin Institute of Bioorganic Chemistry, U.S.S.R. Academy of Sciences, Moscow, U.S.S.R.

‡ Abbreviations used: n.m.r., nuclear magnetic resonance; BPTI, basic pancreatic trypsin inhibitor; p.p.m., parts per million; 2D n.m.r., two-dimensional n.m.r.; COSY, 2D correlated spectroscopy; SECSY, 2D spin echo correlated spectroscopy; NOESY, 2D nuclear Overhauser enhancement spectroscopy; NOE, nuclear Overhauser enhancement.

0022-2836/82/220323-29 $03.00/0

recently (Wagner & Wüthrich, 1982), and corresponding data for the inhibitor K will be presented in a forthcoming article. Work on the elucidation of the solution conformations of the three inhibitors is in progress in our laboratory.

In the three inhibitors selected for this project, one has on the one hand readily apparent sequence homologies, i.e. after proper alignment of the six homologous cysteinyl residues, 20 out of a total of 57 to 59 residues are invariant among the three species (Joubert & Strydom, 1978). On the other hand, one finds numerous non-conservative amino acid substitutions in the variant positions, e.g. Trp for Ala, or Pro for Ile and for Thr in positions that are located in regular β-sheet structure in BPTI. Preliminary circular dichroism and ^{1}H n.m.r. (unpublished work) spectral studies indicated that overall similar conformation types prevail for the three proteins. It appeared interesting, therefore, to obtain the data needed for a comparison of the solution conformations of these three species, since this should provide detailed insight into the influence of localized primary structure variations on the spatial structures. With regard to correlations between structure and function, it is an interesting starting point that two of the three homologous proteins are potent protease inhibitors, whereas inhibitor K has practically no protease inhibitory action (Joubert & Strydom, 1978). Finally, work with these three proteins was ideal for further elaboration of the recently proposed novel techniques for spatial protein structure determination in solution (Wüthrich *et al.*, 1982; Billeter *et al.*, 1982; Wagner & Wüthrich, 1982; Wider *et al.*, 1982). While this new n.m.r. approach resulted largely from work with BPTI (Dubs *et al.*, 1979; Nagayama & Wüthrich, 1981; Wagner *et al.*, 1981), the analysis of the inhibitors E and K, for which no data on the spatial structure are available, provided important tests for a variety of different aspects of the new procedures. With regard to future developments of protein ^{1}H n.m.r. studies, structural comparisons of the three inhibitors based on nearly complete individual resonance assignments should also provide new information on correlations between structure and n.m.r. parameters.

This paper describes and documents the assignment of the ^{1}H n.m.r. spectrum of the inhibitor E with two-dimensional n.m.r. experiments. In contrast to the previously described corresponding results for BPTI (Wagner & Wüthrich, 1982), where numerous resonance assignments had been obtained with conventional, one-dimensional n.m.r. techniques (Wüthrich & Wagner, 1979), all the assignments for the inhibitor E resulted from analysis of a small number of two-dimensional ^{1}H n.m.r. spectra recorded with a single sample of about 16 mg of the protein.

2. Materials and Methods

Trypsin inhibitor E was isolated from the venom of *D. polylepis polylepis* and purified as described (Strydom, 1976). In all, 18 mg of the toxin were available for the n.m.r. experiments. The total amount of the toxin was repeatedly dissolved and lyophilized to obtain approx. 0·007 M-solutions in 0·35 ml of the following solvents: (1) ^2H_2O, p^{2}H 2·7. This sample was measured immediately after dissolving the protein in ^2H_2O, so that the slowly exchanging interior amide protons could be observed (Wüthrich, 1976). (2) ^2H_2O, p^{2}H 3·2. In this sample, the labile protons of the protein were replaced with ^{2}H by heating the solution to 55°C for 1 h, and the residual water protons were minimized by repeated lyophilization from ^2H_2O. (3) 90% H_2O + 10% ^2H_2O, pH 3·2.

The three 2-dimensional n.m.r. experiments used to obtain the resonance assignments have been described in detail. They are 2-dimensional correlated spectroscopy (COSY), which was recorded with the pulse sequence (Aue *et al.*, 1976; Nagayama *et al.*, 1980; Bax & Freeman, 1981):

$$[90^\circ - t_1 - 90^\circ - t_2]_n,$$

where t_1 and t_2 are the evolution period and the observation period, respectively. To obtain a 2-dimensional n.m.r. spectrum, the measurement is repeated for a set of equidistant t_1 values. To improve the signal-to-noise ratio, n transients were accumulated for each value of t_1. At the end of each recording, the system was allowed to reach equilibrium during a fixed delay of 1·2 s. All the COSY spectra for this study were obtained from 512 measurements, with t_1 values from 0 to 47 ms. Two-dimensional spin echo correlated spectroscopy (SECSY) used the pulse sequence (Nagayama *et al.*, 1979,1980):

$$\left[90^\circ - \frac{t_1}{2} - 90^\circ - \frac{t_1}{2} - t_2\right]_n.$$

The SECSY spectra were obtained from 512 measurements, with t_1 values from 0 to 197 ms. Two-dimensional nuclear Overhauser enhancement spectroscopy (NOESY) used a sequence of 3 pulses (Jeener *et al.*, 1979; Anil Kumar *et al.*, 1980*a*):

$$[90^\circ - t_1 - 90^\circ - \tau_m - 90^\circ - t_2]_n,$$

where τ_m is the so-called mixing time. The NOESY spectra used here were obtained from 512 measurements, with t_1 values from 0 to 47 ms. Mixing times of 50, 100 and 200 ms were used. To suppress contributions from coherent magnetization transfer to the cross peak intensities, the 50 ms and 100 ms mixing times were stochastically modulated with a modulation amplitude corresponding to 10% and 5% of τ_m, respectively (Macura *et al.*, 1981).

The 2-dimensional n.m.r. spectra were recorded at 500 MHz on a Bruker WM 500 spectrometer. Quadrature detection was used, with the carrier frequency at the low field end of the spectrum. To eliminate experimental artifacts, groups of 16 recordings with different phases were added for each value of t_1 (Nagayama *et al.*, 1979,1980). For measurements in H_2O, the solvent resonance was suppressed by selective, continuous irradiation at all times except during data acquisition (Anil Kumar *et al.*, 1980*b*). To end up with a 1024×1024 point frequency domain data matrix for COSY and NOESY, which corresponds to the digital resolution given in the Figure legends, the time domain matrix was expanded to 2048 points in t_1 and 4096 points in t_2 by "zero-filling". For the SECSY spectra, the time domain matrix was expanded to 2048×4096 data points. Prior to Fourier transformation, the time domain data matrix was multiplied in the t_1 direction with a phase shifted sine bell, $\sin(\pi(t+t_0)/t_s)$, and in the t_2 directions with a phase shifted sine-squared bell, $\sin^2(\pi(t+t_0)/t_s)$. The length of the window functions, t_s, was adjusted for the bells to reach zero at the last experimental data point in the t_1 or t_2 direction, respectively. The phase shifts, t_0/t_s, were 1/16 and 1/64 in the t_1 and t_2 directions, respectively. All the spectra are shown in the absolute value representation.

3. Results and Discussion

The complete primary structure of the inhibitor E from *D. polylepis polylepis* is known (Joubert & Strydom, 1978), but no spectroscopic data have been published. This section therefore starts with a qualitative survey of the n.m.r. spectral properties, followed by a description of the experiments used to obtain individual ^{1}H n.m.r. assignments.

(a) *^1H n.m.r. spectra of the inhibitor E*

The ^1H n.m.r. spectra used for the resonance assignments were recorded from solutions in H_2O or ^2H_2O at pH 3·2 and at temperatures between 25 and 50°C. At this slightly acidic pH value, one is at or near the minimum rates for the exchange of the polypeptide amide protons with the solvent (Englander *et al.*, 1972; Richarz *et al.*, 1979). The denaturation temperature at pH 3·2 was found to be 79°C from the temperature dependence of the circular dichroism spectrum. Thus, the globular conformation of the protein was preserved under the conditions used for the n.m.r. experiments. The temperature dependence of the aromatic resonances further showed that at 50°C and pH 3·2 all the phenylalanine and tyrosine rings in the inhibitor E are, with the single exception of Tyr35, sufficiently mobile to give symmetric A_2M_2X and A_2X_2 spin systems (Wüthrich, 1976). At 25°C, asymmetric spin systems were observed for Phe21, Tyr23, Tyr35 and Phe45, which thus have only limited rotational mobility at this temperature (Campbell *et al.*, 1975,1976; Wüthrich & Wagner, 1975; Wagner *et al.*, 1976).

A survey of the n.m.r. spectral properties of the inhibitor E is presented in Figures 1 to 6 (the resonance assignments indicated in these Figures will be discussed later). COSY and SECSY spectra recorded in ^2H_2O after complete exchange of all labile hydrogen atoms (Figs 1 to 4) were used to identify the spin systems of the amino acid side-chains. As discussed in detail elsewhere (Aue *et al.*, 1976; Bax & Freeman, 1981; Nagayama & Wüthrich, 1981; Wider *et al.*, 1982), both COSY and SECSY manifest connectivities between J-coupled protons. While fundamentally the two experiments thus provide the same information, practical experience has shown that more extensive and more reliable information on the amino acid spin systems in proteins is obtained from combined use of the two techniques. This is mainly a consequence of the different spectral resolution that arises from the combined effects of the different geometric arrangement of the J-coupling fine structure components in COSY and SECSY cross peaks (Nagayama *et al.*, 1980) and the higher digital resolution that is generally applied for SECSY (Nagayama *et al.*, 1979). Figures 1 and 2 show the complete COSY spectrum, and Figures 3 and 4 show the complete SECSY spectrum of the inhibitor E in ^2H_2O. It is seen that both spectra contain a large number of resolved cross peaks, which connect individual groups of J-coupled protons. At the digital resolution selected for these experiments, the cross peaks in COSY have largely unstructured, square shapes, whereas the J coupling fine structure is readily apparent in most of the SECSY cross peaks.

A COSY spectrum was also obtained from a freshly prepared solution of the inhibitor E in ^2H_2O, where only the solvent-accessible labile protons of the polypeptide were replaced by deuterium before the spectrum was recorded. Compared to the spectrum in Figures 1 and 2, the only additional cross peaks in a COSY spectrum obtained from such a sample are those linking unexchanged amide protons with the C^α protons of the same residues. The spectral region that contains these cross peaks is shown in Figure 5. Since each amino acid residue can give rise to only one NH–C^αH cross peak (except for Gly, which gives generally 2 peaks with identical NH chemical shifts, and Pro and the N-terminal residue, which give

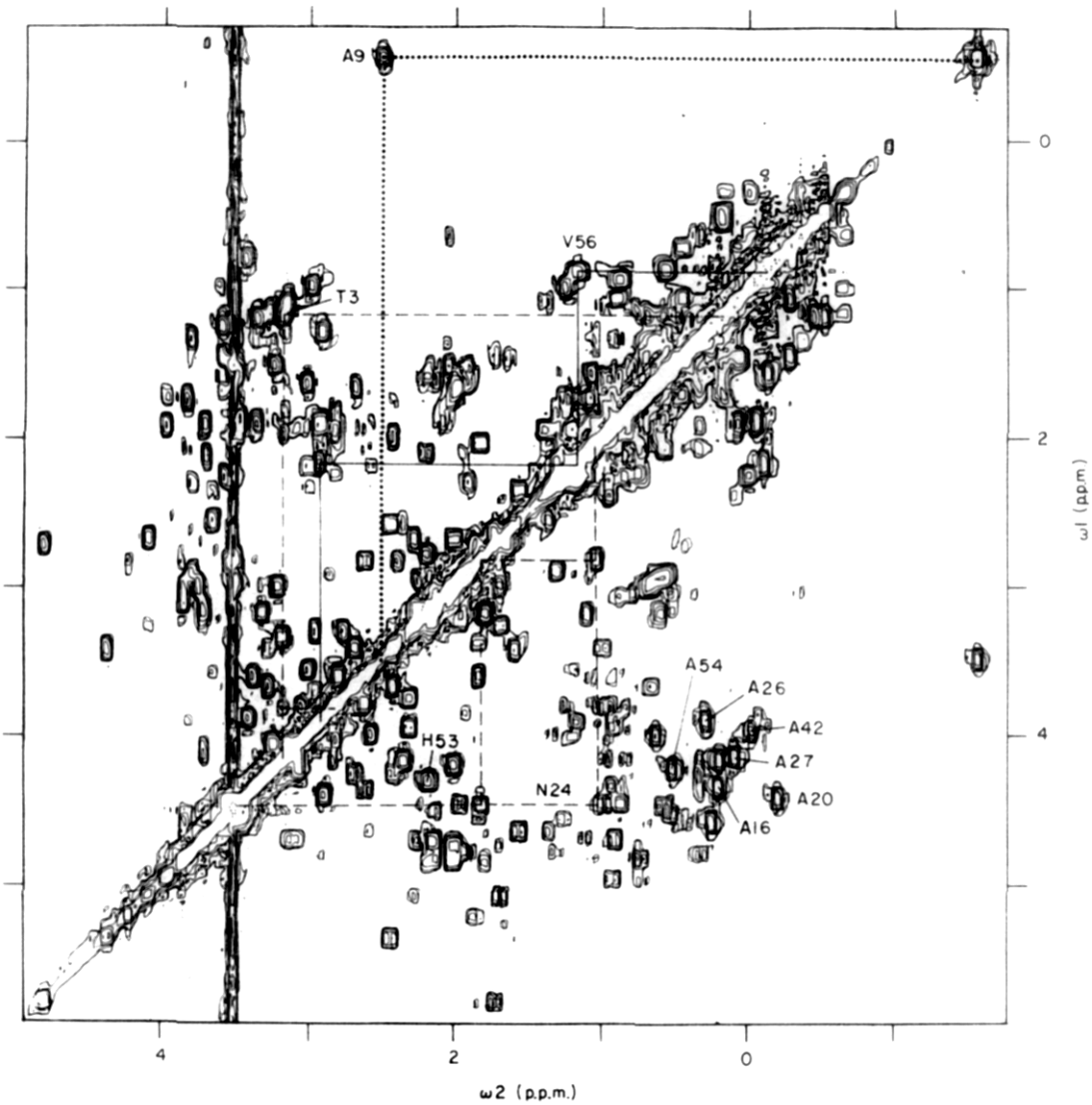

FIG. 1. Contour plot of the spectral region for −0·8 to 6 p.p.m. of a 500 MHz ^{1}H COSY spectrum of a 0·007 M solution of trypsin inhibitor E from *D. polylepis polylepis* in ^2H_2O, p^{2}H 3·2 at 50°C. The digital resolution is 5·3 Hz/point. The spectrum was recorded in approx. 22 h. The strong vertical noise band at 4·5 p.p.m. is at the chemical shift of the residual water protons. Proton–proton J connectivities are indicated for the following amino acid side-chains: Thr3 (– – –), Ala9 (· · ·), Ala16, Ala20, Asn24 (– – –), Ala26, Ala27, Ala42, His53, Ala54, Val56 (———). In order not to overcrowd the Figure, only the $C^\alpha H$–$C^\beta H_n$ cross peaks are identified for the A_2X and A_3X spin systems in the lower right triangle.

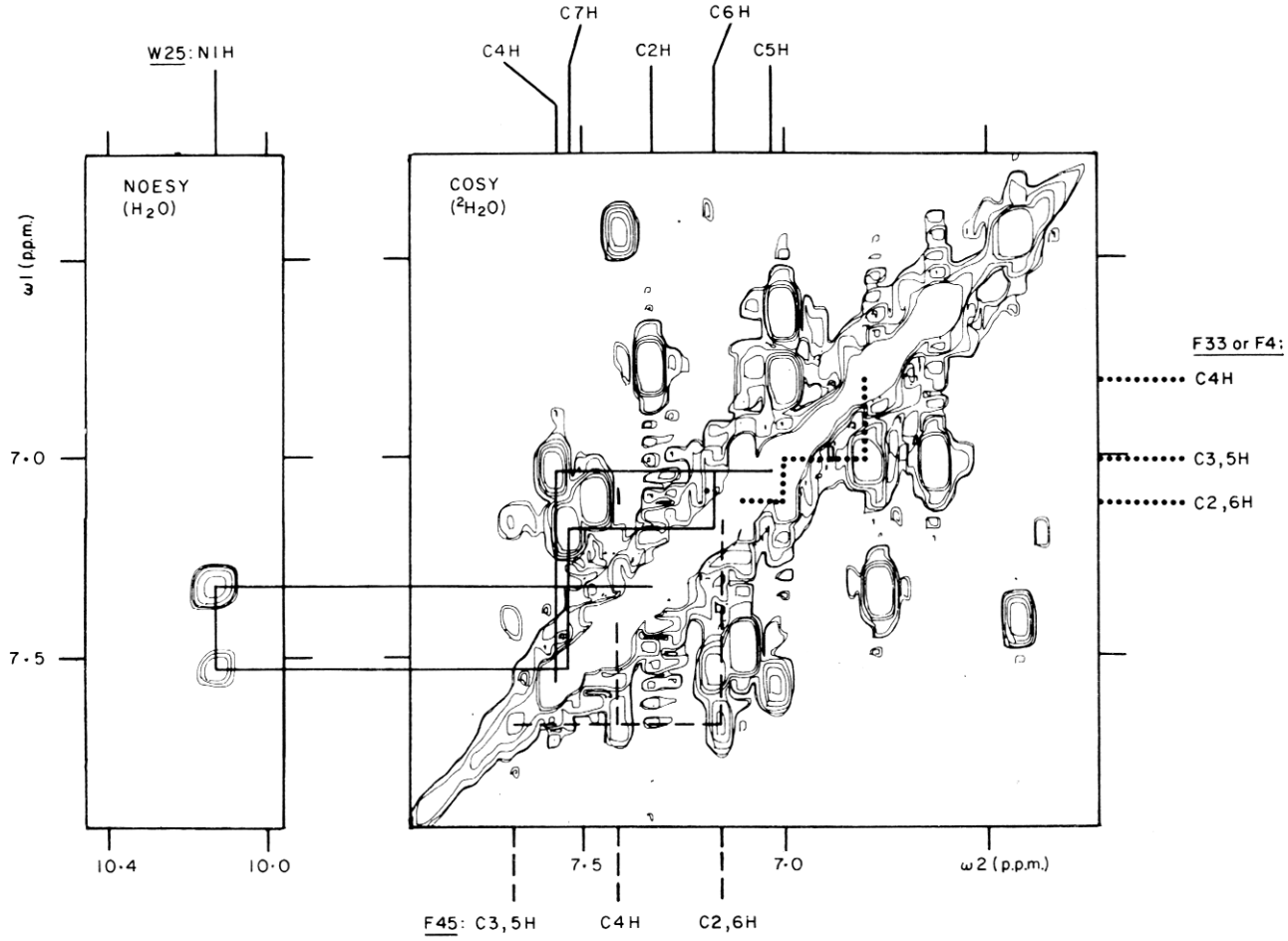

FIG. 2. Contour plot of the region from 6·2 to 7·9 p.p.m. of the same COSY spectrum as in Fig. 1. *J* connectivities are indicated for the following aromatic spin systems: Phe33 or Phe4 (· · ·), Phe45 (– – –) and Trp25 (———). To complete the spin system of Trp25 with N1H and C2H, we used the NOE connectivities observed in a NOESY spectrum recorded in H_2O and with identical parameters as Fig. 6, except that the temperature was 50°C and the mixing time was 100 ms. The spectral region (ω1 = 6·2 to 7·9 p.p.m. ω2 = 9·9 to 10·5 p.p.m.) of this NOESY spectrum is shown on the left of the Figure.

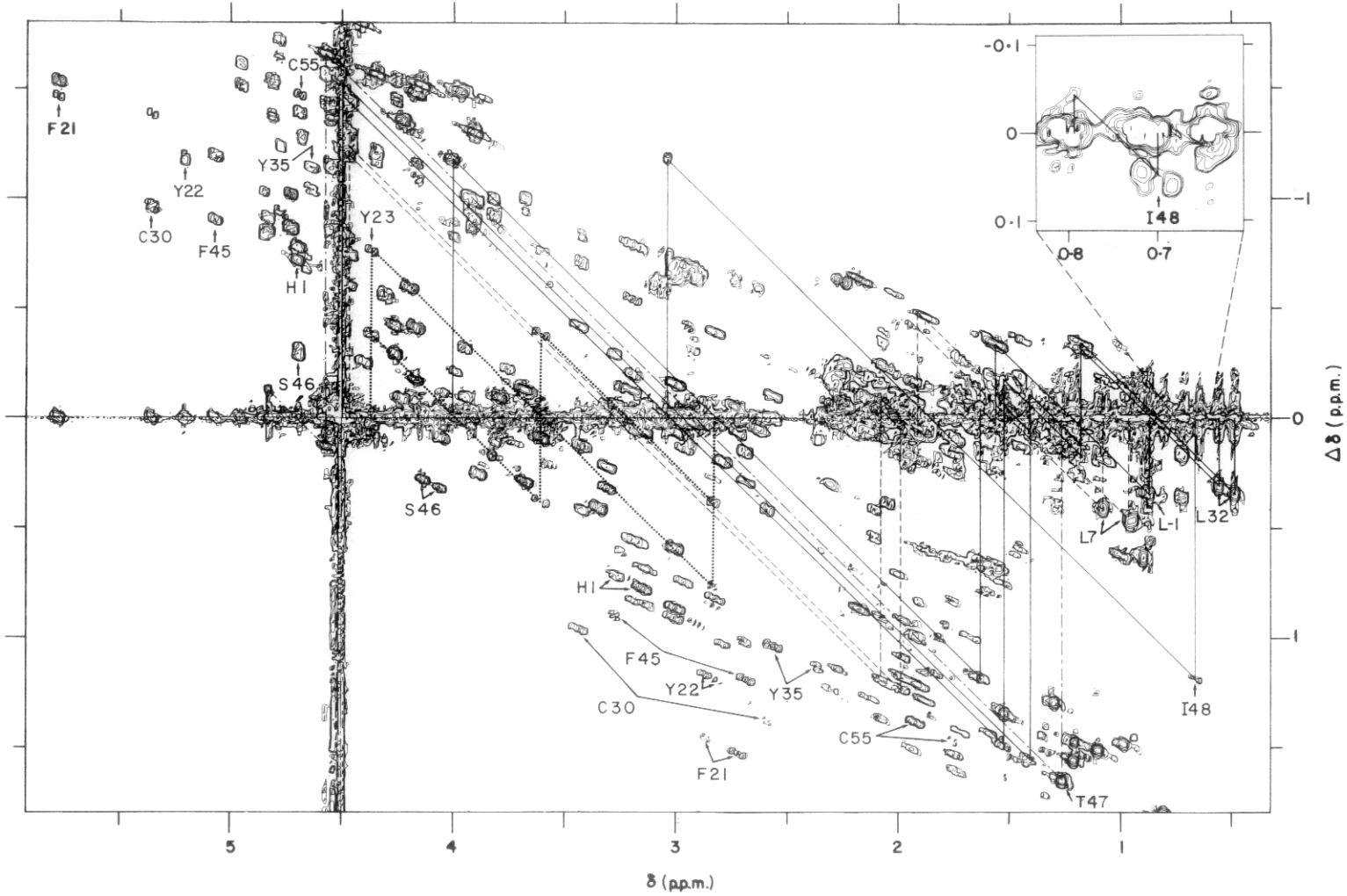

FIG. 3. Contour plot of the region from 0·3 to 5·9 p.p.m. of a 500 MHz ^{1}H SECSY spectrum of a 0·007 M-solution of trypsin inhibitor E in ^2H_2O, p^{2}H 3·2 at 50°C. The digital resolution is 1·3 Hz/point along $\Delta\delta$ and 2·5 Hz/point along δ. The spectrum was recorded in approx. 21 h. Proton–proton J connectivities are shown for the following amino acid side-chains: Leu−1 (———), Leu7 (– – –), Tyr23 (· · ·), Leu32 (———), Thr47 (– · – · –), Ile48 (———). Part of the Ile48 spin system is shown on an expanded scale in the inset in the upper right corner, where only selected higher levels have been plotted (see the text for further explanations of the assignments of this residue). In order not to overcrowd the Figure, the $C^\alpha H$–$C^\beta H_2$ cross peaks of the AMX spin systems of His1, Phe21, Tyr22, Cys30, Tyr35, Phe45, Ser46 and Cys55 have been identified, but these spin systems, which further include the $C^\beta H$–$C^\beta H$ cross peaks, are not connected by lines.

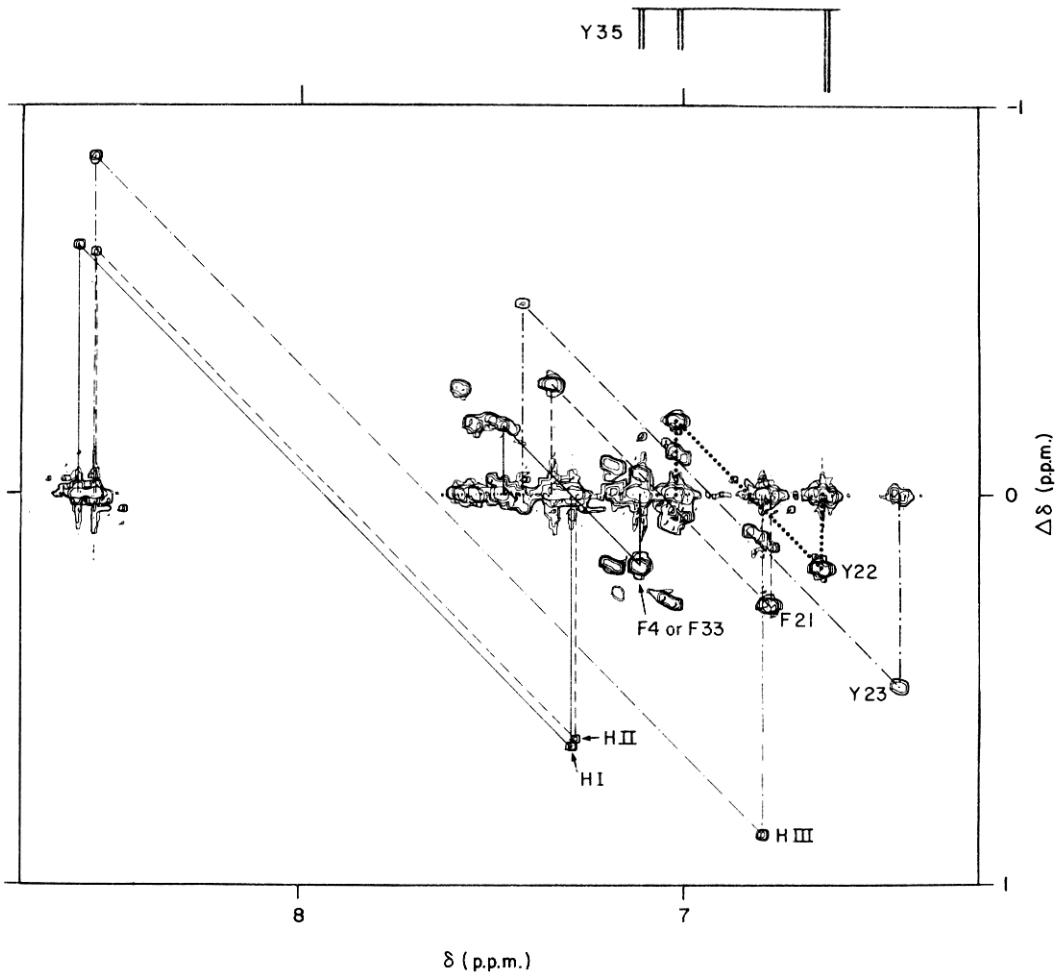

FIG. 4. Contour plot of the region from 6·3 to 8·7 p.p.m. of the same SECSY spectrum as in Fig. 3. The J connectivities are indicated for the following aromatic spin systems: Phe21 (– – –), Tyr22 (···), Tyr23 (– · – · –). Phe4 or Phe33 (———: see the text and Table 1) and the 3 histidyl residues, which have not been assigned individually and are therefore marked HI (———), HII (– – –) and HIII (– · – · –). At the top of the Figure, the spin system of Tyr35, which was identified in the 1-dimensional spectra (see the text), is indicated.

none), we conclude that at 25°C and p^{2}H 2·7 (Fig. 5), the inhibitor E contains 24 slowly exchanging amide protons. In a COSY spectrum recorded in slightly acidic H_2O, where all the polypeptide backbone amide protons are preserved, the number of NH–C$^\alpha$H cross peaks coincided closely with that expected from the amino acid sequence of the inhibitor E (see Figs 11 and 17). This "fingerprint" of the protein structure thus showed that nearly all the residues could be resolved in the ^{1}H n.m.r. spectra.

NOESY spectra of the inhibitor E recorded in H_2O (Fig. 6) or ^2H_2O contain numerous cross peaks in all the different regions. For example, it is readily seen in

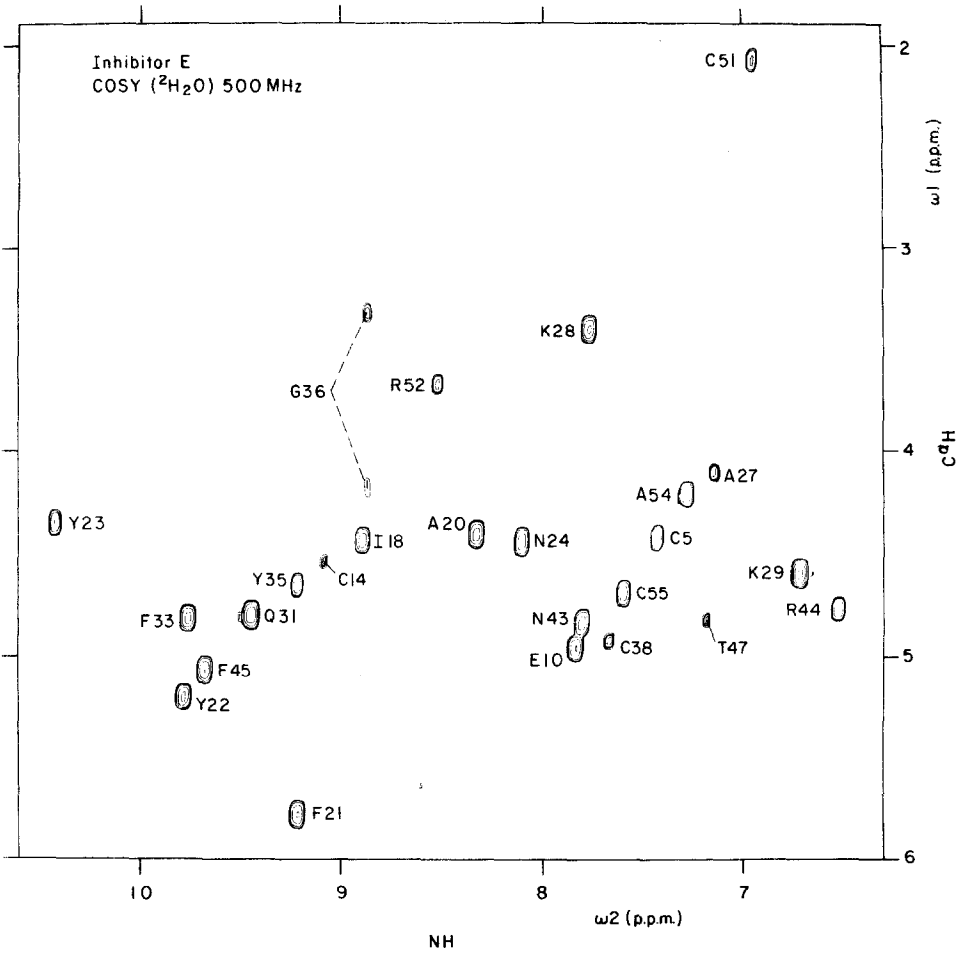

FIG. 5. Spectral region ($\omega 1 = 1{\cdot}9$ to $6{\cdot}0$ p.p.m., $\omega 2 = 6{\cdot}3$ to $10{\cdot}7$ p.p.m.) of a 500 MHz ^{1}H COSY spectrum of a 0·007 M solution of the trypsin inhibitor E in ^2H_2O, $p^2H = 2{\cdot}7$ at 25°C; digital resolution 5·3 Hz/point. The spectrum was recorded from 20 to 44 h after the solution had been prepared. The resonance of the residual solvent protons was suppressed by weak, selective irradiation so that no cross peaks were bleached out. The assignments for the individual cross peaks are indicated by the 1-letter symbol for the amino acid residue and the position in the amino acid sequence.

Figure 6 that there are nuclear Overhauser enhancement connectivities linking the high field methyl resonance at −0·56 p.p.m. with aromatic protons near 7 p.p.m., and aromatic protons and backbone amide protons between 6·5 and 10·6 p.p.m., with a large number of aliphatic side-chain protons and C^{α} protons between 0 and 6 p.p.m. Since the through-space proton–proton connectivities determined by NOESY have a pivotal role for both the individual resonance assignments and the determination of the conformation, a good NOESY spectrum, such as that in Figure 6, is an essential spectral property for a protein to be amenable for structure determination by ^{1}H n.m.r.

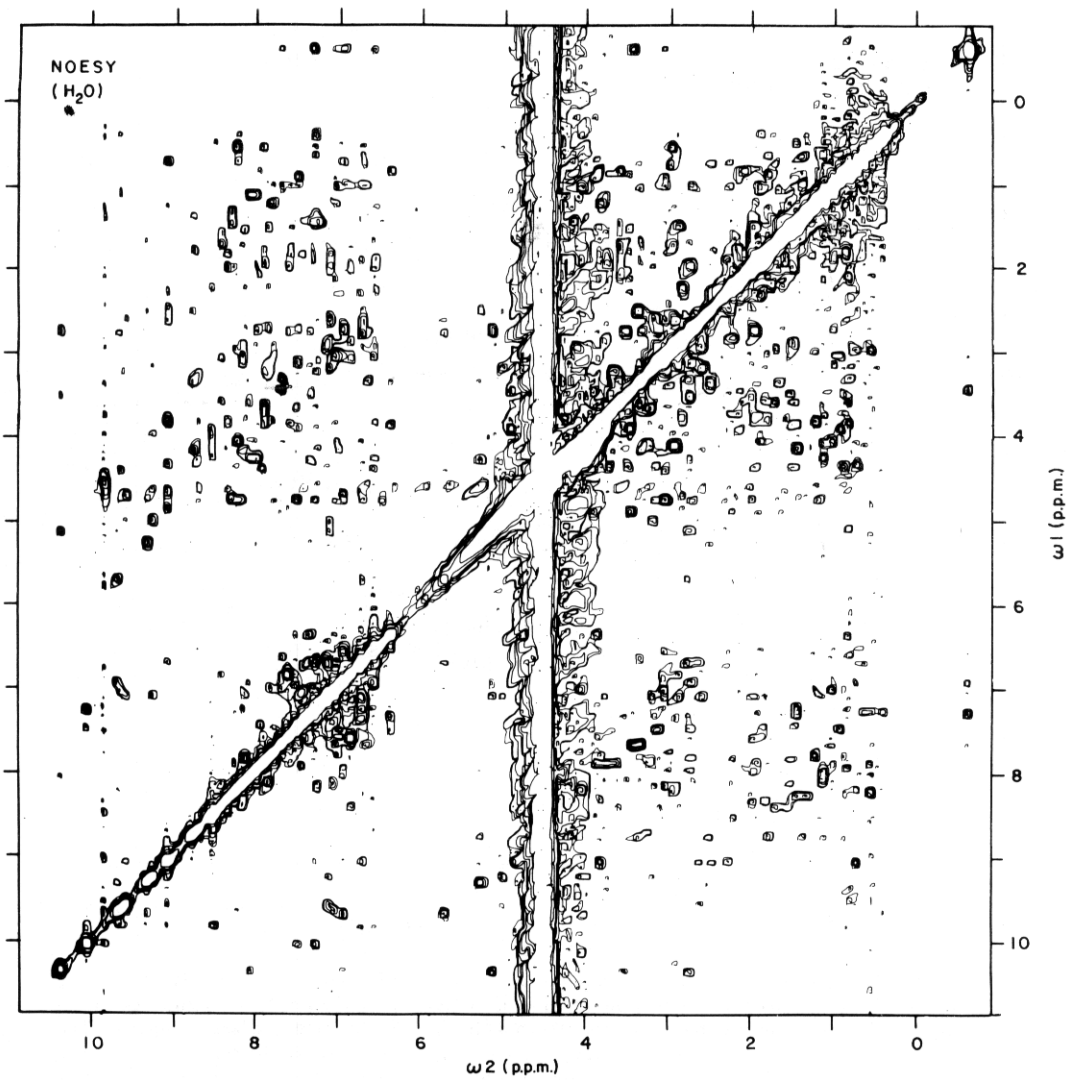

FIG. 6. Contour plot of a 500 MHz ^{1}H NOESY spectrum of a 0·007 M solution of trypsin inhibitor E in H_2O, pH 3·2 at 45°C. The digital resolution is 5·3 Hz/point. The spectrum was recorded in approx. 22 h and a mixing time of 200 ms was used. The strong vertical noise band at 4·5 p.p.m. is at the chemical shift of the water, which was suppressed by selective irradiation as described in Materials and Methods.

(b) *Identification of complete amino acid side-chain spin systems before sequential assignment of the polypeptide backbone protons*

Here, the spin systems of different types of amino acid residues are identified. From this first step of the spectral analysis, assignments to specific positions in the amino acid sequence are obtained only when there are single residues of a given type, such as Val56 in the inhibitor E.

The COSY and SECSY spectra of the inhibitor E in ^2H_2O solution were screened for the cross peak patterns that are characteristic for the different types of amino acid residues (Wider *et al.*, 1982; Nagayama & Wüthrich, 1981). The identifications of the non-aromatic amino acid spin systems are documented in Figures 1, 3 and 7 to 9, and those for the aromatic rings in Figures 2 and 4. In these Figures, the individual assignments obtained as a final result of all the experiments described in this paper are given, and the connectivities are also indicated for the spin systems that were identified only after the sequential assignments of the backbone protons had been obtained.

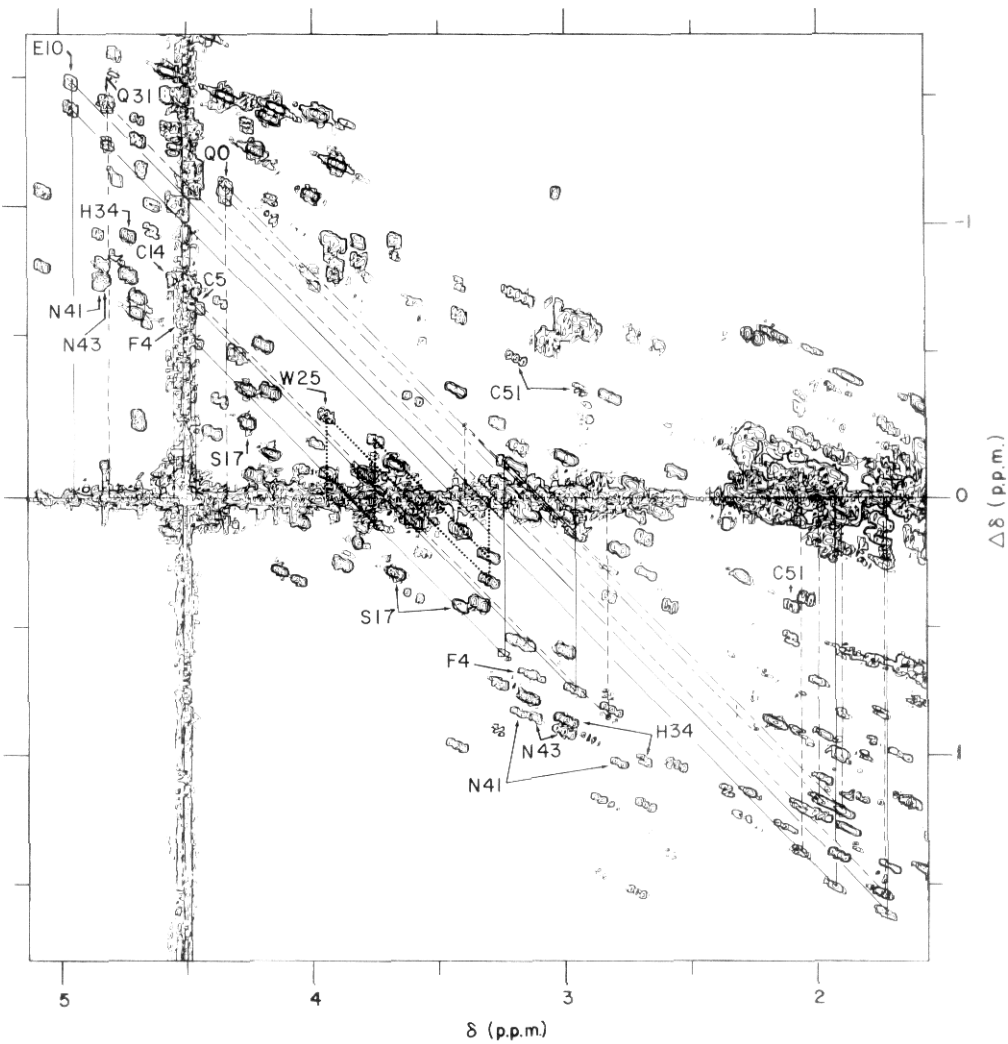

FIG. 7. Contour plot of the region from 1·6 to 5·1 p.p.m. of the same SECSY spectrum as in Fig. 3. The $C^\alpha H$–$C^\beta H_2$ J connectivities are indicated for Gln0 (– · – · –), Glu10 (———), Gln31 (– – –), for these 3 residues the $C^\beta H_2$–$C^\gamma H_2$ connectivities are shown in Fig. 8. Cys5 (———), Cys14 (– – –) and Trp25 (· · ·). In order not to overcrowd the Figure, only the $C^\alpha H$–$C^\beta H_2$ cross peaks of the AMX spin systems of Ser17, His34, Asn41, Asn43 and Cys51, and the A_2X spin system of Phe4 are identified.

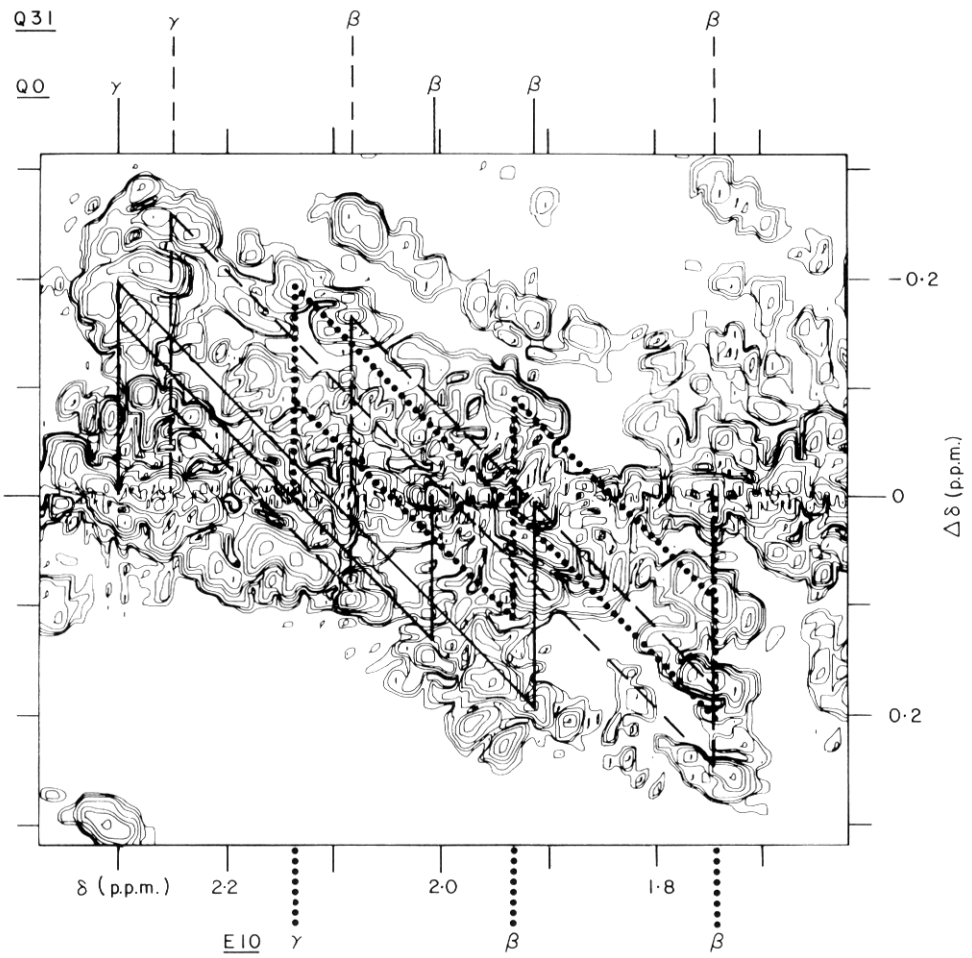

FIG. 8. Contour plot of the region from 1·6 to 2·4 p.p.m. of the same SECSY spectrum as in Fig. 3. The $C^{\beta}H_2$–$C^{\gamma}H_2$ J connectivities are indicated for Gln0 (———), Glu10 (· · ·) and Gln31 (– – –). (See Fig. 7 for the $C^{\alpha}H$–$C^{\beta}H_2$ connectivities of these residues.)

The spin systems that were identified completely in the initial analysis of the COSY and SECSY spectra are those of the five glycyl (Fig. 9), seven alanyl (Fig. 1), and two threonyl residues (Figs 1 and 3), Val56 (Fig. 1), the five-proton system that was later assigned to Glu10 (Figs 7 and 8), and 18 of the 22 $C^{\alpha}H$–$C^{\alpha}H_2$ three-spin systems (Figs 1, 3 and 7), i.e. all except those which were later found and assigned to Phe4, Asn41, Asn43 and Cys51 (Fig. 7). All the aromatic spin systems were identified as indicated in Figures 2 and 4. An intermediate state of mobility (Wagner *et al.*, 1975; Wüthrich, 1976) was observed for Tyr35 at 50°C, where the C3,5H resonances are equivalent but the C2,6H lines are non-equivalent and are too broad to be seen in the spectra of Figures 2 or 4. This spin system (Fig. 4) was unravelled by one-dimensional spin decoupling difference spectra.

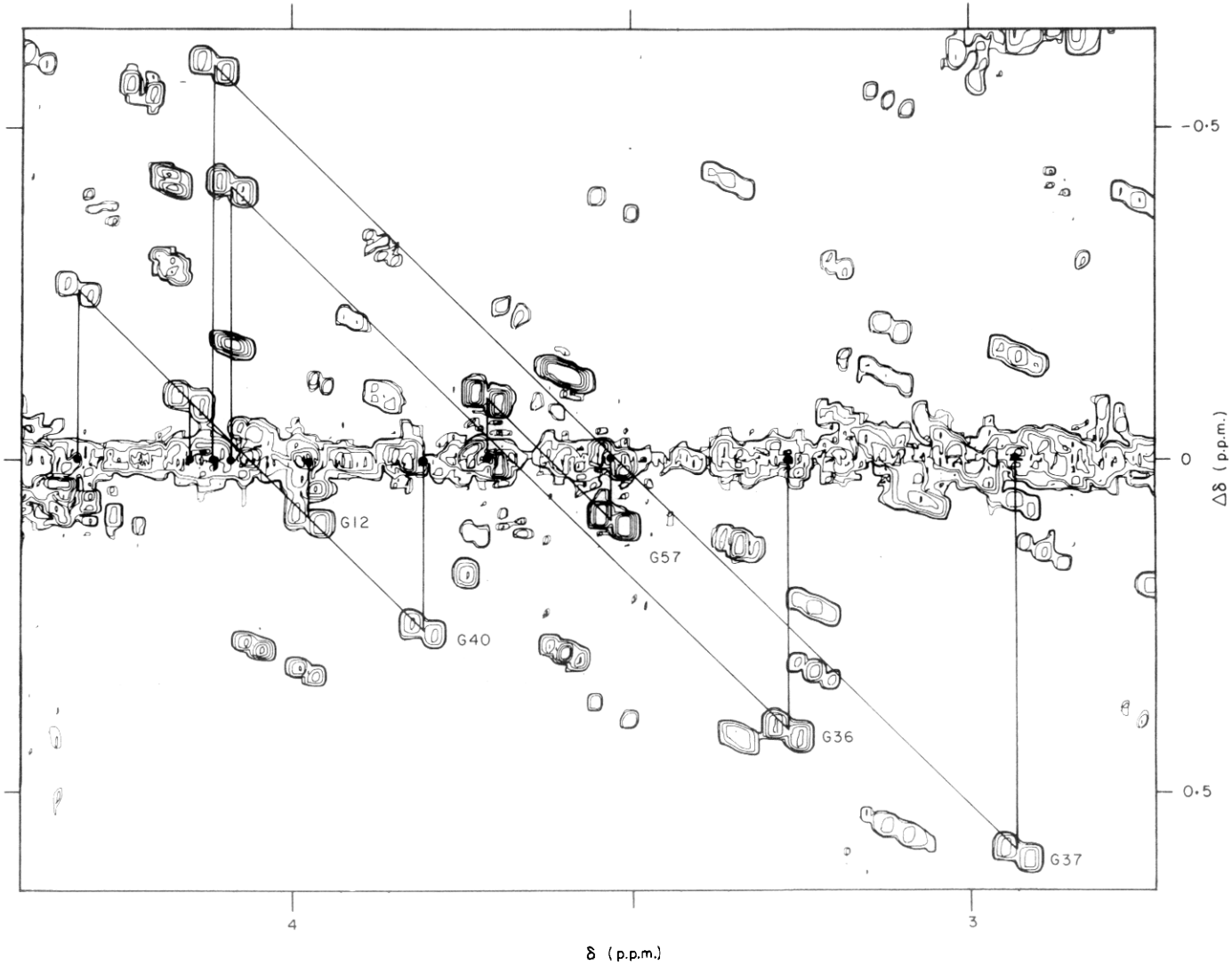

FIG. 9. Contour plot of the region from 2·7 to 4·4 p.p.m. of the same SECSY spectrum as in Fig. 3. The J connectivities are indicated for the AX spin systems of Gly12, Gly36, Gly37, Gly40 and Gly57.

(c)*Sequential resonance assignments for the polypeptide backbone protons*

A general strategy based on stereochemical considerations of polypeptides (Dubs *et al.*, 1979; Billeter *et al.*, 1982) and the use of two-dimensional n.m.r. for the practical realization (Wagner *et al.*, 1981; Wagner & Wüthrich, 1982; Wider *et al.*, 1982) of sequential, individual resonance assignments have been described in detail. Here, we follow closely the procedures outlined in these earlier papers. The crucial quantities for this method of assignment are the distances d_1, d_2 and d_3 between the amide proton of residue $i+1$ and the C^{α} proton, amide proton and C^{β} protons, respectively, of the preceding residue i (Fig. 10). The intrinsic reliability of the identification of sequentially neighbouring amino acid residues is of the order of 80% when it is based on the observation of one of these through-space connectivities, and $\gtrsim$90% when it is based on any two of the three connectivities (Billeter *et al.*, 1982). Additional checks of the sequential assignments are obtained with the use of the previously identified amino acid side-chain spin systems. For example, when starting with the unique valyl residue in position 56 of the inhibitor E, one knows from the amino acid sequence (Fig. 11) that sequential assignments in the direction towards the C terminus must lead to Gly, and in the N-terminal direction to Cys-Ala-His- . . . . Or, once sequential assignments including any of the seven alanyl residues in the inhibitor E extend over two to four residues, these segments can be located uniquely in the sequence, and further sequential assignments can then again be checked against the amino acid sequence.

Figure 11 presents a survey of the sequential assignments obtained in the inhibitor E, and the chemical shifts for the assigned resonances are listed in Table 1. It is seen that nearly complete assignments were obtained, and that most of the sequential connectivities are based on observation of either d_1 and d_3, or d_2 and d_3. The first stretch of continuous assignments extends from residues -1 to 6. Between residues 6 and 7 there are no d_1 and d_3 connectivities, and the d_2 cross peak would be too close to the strong diagonal peaks to be identified unambiguously (see Table 1). The connectivities with Pro8 and Pro13 could not be

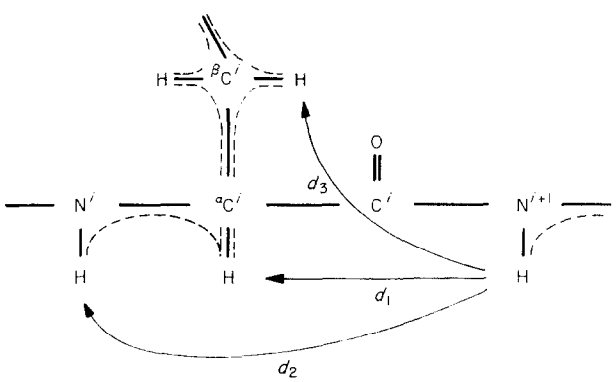

FIG. 10. Polypeptide backbone segment. The through-space distances d_1, d_2 and d_3 are indicated by arrows. The broken lines indicate through-bond J connectivities between hydrogen atoms of the same residue.

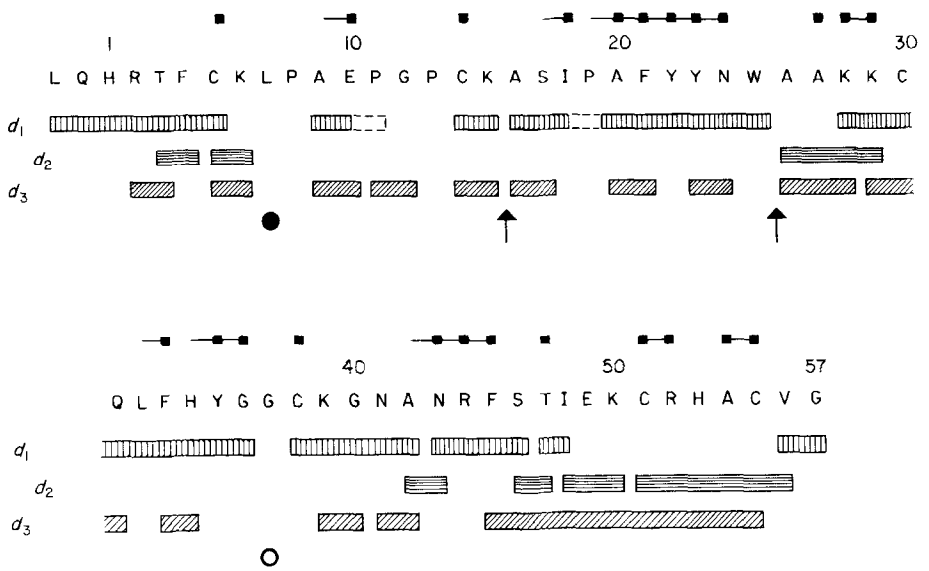

FIG. 11. Amino acid sequence of trypsin inhibitor E from *D. polylepis polylepis* and survey of the sequential connectivities by which the individual resonance assignments listed in Table 1 were obtained. In order to have identical numbers for the homologous Cys residues as in basic pancreatic trypsin inhibitor, the numeration starts with Leu − 1, Gln0, His1. Vertical-lined boxes, sequential assignments *via* d_1 (NOE from NH_{i+1} to $C^{\alpha}H_i$); horizontal-lined boxes, sequential assignments *via* d_2 (NOE from NH_{i+1} to NH_i); hatched boxes, sequential assignments *via* d_3 (NOE from NH_{i+1} to $C^{\beta}H_i$); broken box, sequential assignments *via* NOEs from proline $C^{\delta}H_{i+1}$ to $C^{\alpha}H_i$. ○ and ●, the individual assignment relied primarily on the identification of the complete spin system of the amino acid residues without or with the amide proton, respectively. Once all but one residue of a given type have been assigned, this information is obviously sufficient for the individual assignment in the amino acid sequence, even when none of the connectivities d_1, d_2 or d_3 could be established. The arrows indicate locations where all the resonances were assigned but the connectivity between 2 neighbouring residues was not established. The signs above the sequence indicate the residues for which the amide protons were also observed in ^2H_2O solution (■) and where the sequential connectivities were obtained in ^2H_2O solution (—■). The spectrum in ^2H_2O (Fig. 5) was recorded from 20 to 44 h after the solution was prepared at 25°C and p^{2}H 2·7.

established, but the tetrapeptide segment between these two residues was assigned completely. All backbone NH, $C^{\alpha}H$ and $C^{\beta}H$ lines were assigned for the segment comprising residues 14 to 36, but two connectivities could not be established. Between Lys15 and Ala16, d_3 was absent and d_1 and d_2 would not have been sufficiently well-resolved to provide unambiguous evidence for the connectivity. Between Trp25 and Ala26, d_1 was absent and d_2 and d_3 would not have been sufficiently well-resolved. For Gly37, the amide proton–$C^{\alpha}H_2$ J cross peaks could not be observed. A final stretch of continuous assignments extends from residue 38 to the C-terminal Gly57. The sequential connectivities are documented in Figures 12 to 16, whereby at least one connectivity is shown for each pair of neighbouring residues.

Figures 12 to 14 show d_1 connectivities in the inhibitor E. In the combined COSY–NOESY connectivity diagrams (Wagner *et al.*, 1981), the region of the NH–$C^{\alpha}H$ cross peaks in the upper left triangle of a NOESY spectrum (Fig. 6) is

TABLE 1

Chemical shifts, δ [a], *of the assigned* ^1H *NMR lines of the trypsin inhibitor E from* D. polylepis polylepis *pH 3·2*, t = *50°C*

Amino acid Residue	δ (±0·01 p.p.m.) [a, b]			
	NH	$C^{\alpha}H$	$C^{\beta}H$	Others
Leu − 1		4·00	1·62, 1·62	$C^{\gamma}H$ 1·56 $C^{\delta}H_3$ 0·87, 0·87
Gln0	8·54	4·33	2·00, 1·91	$C^{\gamma}H_2$ 2·29, 2·29 $N^{\varepsilon}H_2$ 8·15, 7·91
His1	8·53	4·68	3·26, 3·14	[c]
Arg2	8·36	4·24	1·15, 1·57	
Thr3	8·13	3·81	4·16	$C^{\gamma}H_3$ 1·20
Phe4	7·16 [d]	4·49	3·13, 3·13	C2,6H 7·11 C3,5H 7·01 or 7·48 [e] C4H 6·81 or 7·40 [e]
Cys5	7·35	4·46	3·24, 2·96	
Lys6	7·04	4·24	1·34	
Leu7	7·18 [d]	4·45	1·98, 2·07	$C^{\gamma}H$ 1·88 $C^{\delta}H_3$ 1·08, 0·96
Pro8				
Ala9	7·31	3·50	−0·56	
Glu10	7·73	4·94	1·93, 1·73	$C^{\gamma}H_2$ 2·14, 2·14
Pro11		4·77	2·90, 2·30	$C^{\gamma}H_2$ 2·34, 2·19 $C^{\delta}H_2$ 3·98, 3·56
Gly12	8·62	4·22, 4·05		
Pro13				
Cys14	9·12 [d]	4·54	3·40, 2·84	
Lys15	7·87	4·46	1·58	
Ala16	8·04	4·34	1·20	
Ser17	8·10	4·25	3·66, 3·41	
Ile18	8·79	4·45	1·87	$C^{\gamma}H_2$ 1·45, 1·06 $C^{\gamma}H_3$ 0·94 $C^{\delta}H_3$ 0·72
Pro19		4·16	1·99, 1·85	$C^{\gamma}H_2$ 2·25, 1·94 $C^{\delta}H_2$ 4·03, 3·80
Ala20	8·25	4·40	0·81	
Phe21	9·10	5·77	2·86, 2·72	C2,6H 6·77 C3,5H 7·34 C4H 7·34 [f]
Tyr22	9·72	5·19	2·86, 2·81	C2,6H 7·02 C3,5H 6·64
Tyr23	10·44	4·36	3·60, 2·83	C2,6H 7·41 C3,5H 6·44
Asn24	8·02	4·46	2·82, 2·04	$N^{\delta}H_2$ 7·63, 6·87
Trp25	7·97	3·94	3·74, 3·30	N1H 10·13 C2H 7·33 C4H 7·57 C5H 7·03 C6H 7·18 C7H 7·54
Ala26	7·82	3·89	1·29	
Ala27	7·05	4·13	1·10	
Lys28	7·72	3·41	1·99, 1·74	
Lys29	6·70	4·57	1·71, 1·46	
Cys30	8·80	5·35	3·43, 2·59	

TABLE 1—*cont.*

Amino acid residue	δ (±0·01 p.p.m.)[a,b]			
	NH	$C^{\alpha}H$	$C^{\beta}H$	Others
Gln31	9·36	4·80	2·07, 1·73	$C^{\gamma}H_2$ 2·24, 2·24 $N^{\varepsilon}H_2$ 7·46, 7·10
Leu32	8·35[d]	4·50	1·52, 1·41	$C^{\gamma}H$ 1·18 $C^{\delta}H_3$ 0·55, 0·49
Phe33	9·69	4·82	3·11, 3·11	C2,6H 7·11 C3,5H 7·48 or 7·01[e] C4H 6·81 or 7·40[e]
His34	8·19	4·72	3·00, 2·69	[c]
Tyr35	9·10	4·62	2·56, 2·35	C2,6H 7·11, 7·02 C3,5H 6·64
Gly36	8·73	4·16, 3·34		
Gly37		4·19, 3·00		
Cys38	7·64	4·91	3·90, 2·86[g]	
Lys39	9·10	3·88	2·07, 1·91	
Gly40	8·39	4·39, 3·88		
Asn41	7·15	4·84	3·18, 2·79	$N^{\delta}H_2$ 8·16, 7·91
Ala42	7·52	3·96	0·98	
Asn43	7·87	4·82	3·12, 3·00	$N^{\delta}H_2$ 8·04
Arg44	6·52	4·77	1·51, 1·33	
Phe45	9·61	5·06	3·27, 2·69	C2,6H 7·17 C3,5H 7·68 C4H 7·42
Ser46	9·27	4·69	4·13, 4·05	
Thr47	7·15	4·82	4·57	$C^{\gamma}H_3$ 1·26
Ile48	8·26	3·03	0·66	$C^{\gamma}H_2$ 0·79, 0·79 $C^{\gamma}H_3$ 0·66 $C^{\delta}H_3$ 0·70
Glu49	8·00	3·81	1·96, 1·81	
Lys50	7·61	3·92	2·07, 1·91	
Cys51	6·89	2·10	3·19, 2·94	
Arg52	8·46	3·66	1·78, 1·67	
His53	8·18	4·30	3·18, 3·18	[c]
Ala54	7·30	4·22	1·52	
Cys55	7·47	4·68	1·92, 1·75	
Val56	7·97	3·91	2·17	$C^{\gamma}H_3$ 0·90, 0·90
Gly57	7·87	3·78, 3·60		

[a] The chemical shifts, δ, are relative to internal sodium 3-trimethyl-silyl-[2,2,3,3-^{2}H]-propionate.

[b] Where no numbers are given in the columns for NH, $C^{\alpha}H$ and $C^{\beta}H$, and where more peripheral side-chain hydrogen atoms are not listed in the last column, no individual resonance assignment was obtained (see the text).

[c] The imidazole proton resonances of the 3 histidyl residues were not assigned individually. Their chemical shifts at pH 3·2 and 50°C are HI: C2H = 8·58, C4H = 7·29; HII: C2H = 8·53, C4H = 7·28; HIII: C2H = 8·54, C4H = 6·79.

[d] Because for these residues the NH–$C^{\alpha}H$ cross peaks in the COSY spectrum at 50°C were bleached out by the solvent irradiation, the amide proton chemical shifts were measured at 25°C for Cys14 and at 45°C for the others.

[e] Since the C2,6 protons of Phe4 and Phe33 have identical chemical shifts, the C3,5H and C4H resonances of these 2 rings could not be assigned individually, even though 2 complete spin systems were identified.

[f] From the multiplicity of the cross peak of C3,5H with C2,6H and the absence of a cross-peak between C3,5H and C4H (Fig. 4), the chemical shifts for C3,5H and C4H of Phe21 must be nearly identical.

[g] The C^{β} proton resonances of Cys38 could not be observed at 50°C and, therefore, the chemical shifts at 25°C are listed.

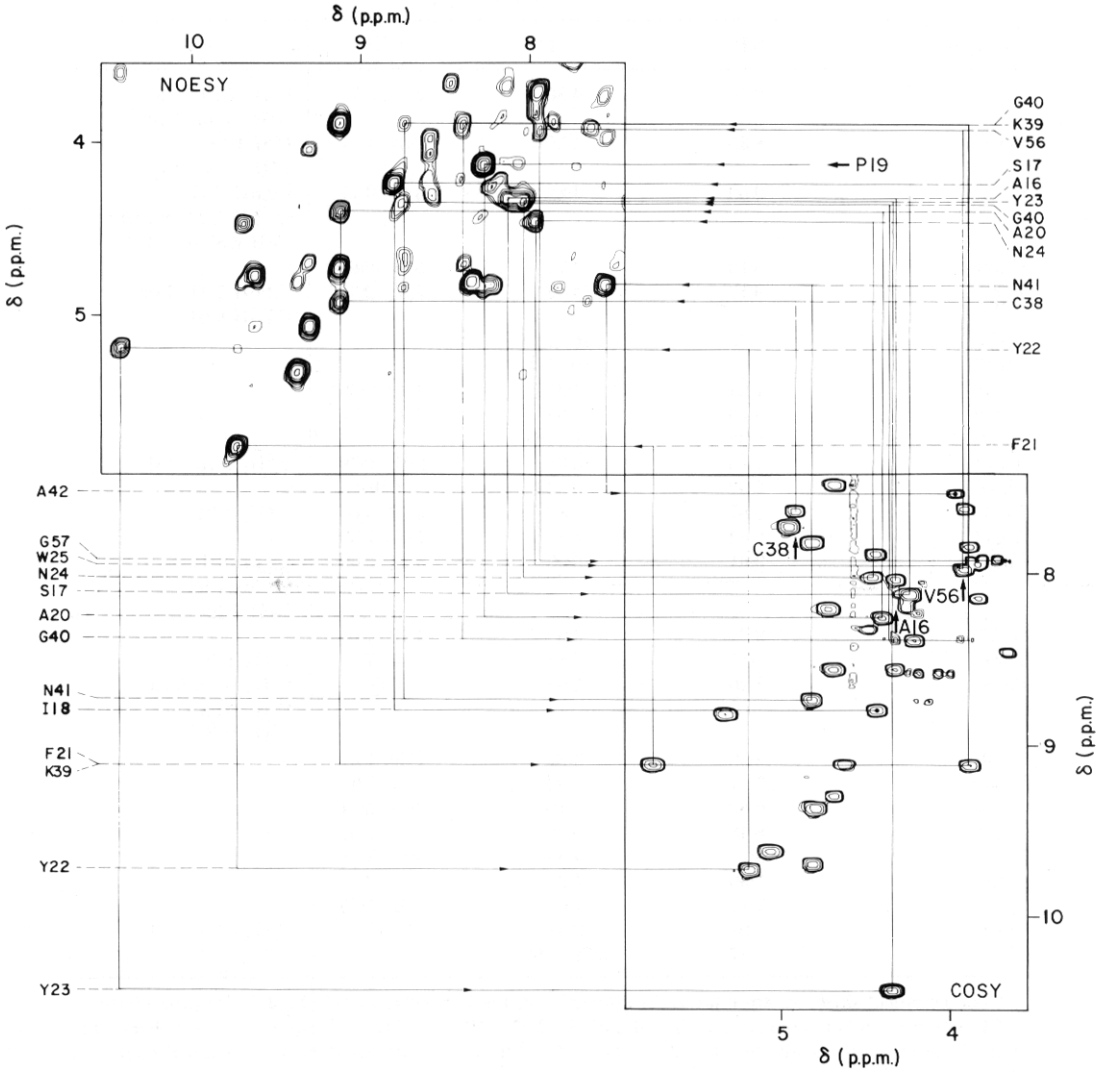

FIG. 12. Combined COSY–NOESY connectivity diagram for sequential resonance assignments *via* NOEs between amide protons and the C^{α} protons of the preceding residue (d_1). In the upper left, the region ($\omega 1 = 3{\cdot}5$ to $5{\cdot}9$ p.p.m., $\omega 2 = 7{\cdot}4$ to $10{\cdot}5$ p.p.m.) of the ^{1}H NOESY spectrum of trypsin inhibitor E in Fig. 6 is presented. In the lower right, the region ($\omega 1 = 7{\cdot}4$ to $10{\cdot}5$ p.p.m., $\omega 2 = 3{\cdot}5$ to $5{\cdot}9$ p.p.m.) from a ^{1}H COSY spectrum recorded from the same sample under identical conditions, i.e. 45°C and pH = 3·2, is shown. The straight lines and arrows indicate the connectivities between neighbouring residues in the segments 16 to 18, 19 to 25, 38 to 42 and 56 to 57. The arrows and the filled circles identify the start and the end for each segment, respectively. The amide proton chemical shifts are indicated by the assignments in the lower left corner, those for the C^{α} protons by the assignments in the upper right corner of the Figure.

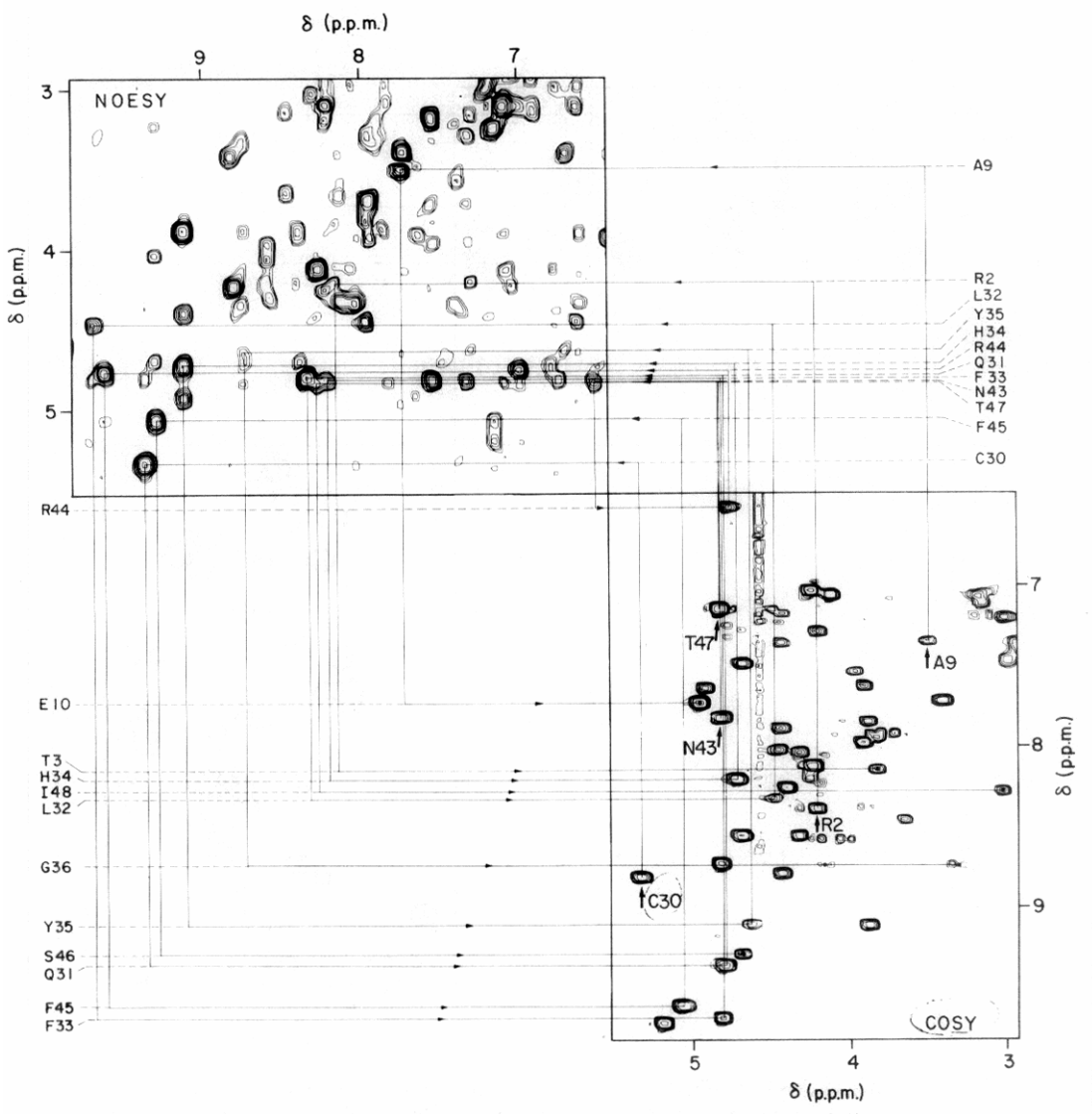

FIG. 13. Combined COSY–NOESY connectivity diagram for sequential resonance assignments *via* NOEs between amide protons and C^{α} protons of the preceding residues (d_1). In the upper left, the region ($\omega 1$ = 2·9 to 5·5 p.p.m., $\omega 2$ = 6·4 to 9·8 p.p.m.) of the ^{1}H NOESY spectrum of trypsin inhibitor E in Fig. 6 is presented. In the lower right, the region ($\omega 1$ = 6·4 to 9·8 p.p.m., $\omega 2$ = 2·9 to 5·5 p.p.m.) from a ^{1}H COSY spectrum recorded from the same sample under identical conditions, i.e. 45°C and pH = 3·2, is shown. The straight lines and arrows indicate the connectivities between neighbouring residues in the segments 2 to 3, 9 to 10, 30 to 36, 43 to 46 and 47 to 48 of trypsin inhibitor E. Arrows and filled circles identify the start and end point of the d_1 connectivity patterns, respectively. The amide proton chemical shifts are indicated by the assignments in the lower left corner and those for the C^{α} protons by the assignments in the upper right corner of the Figure.

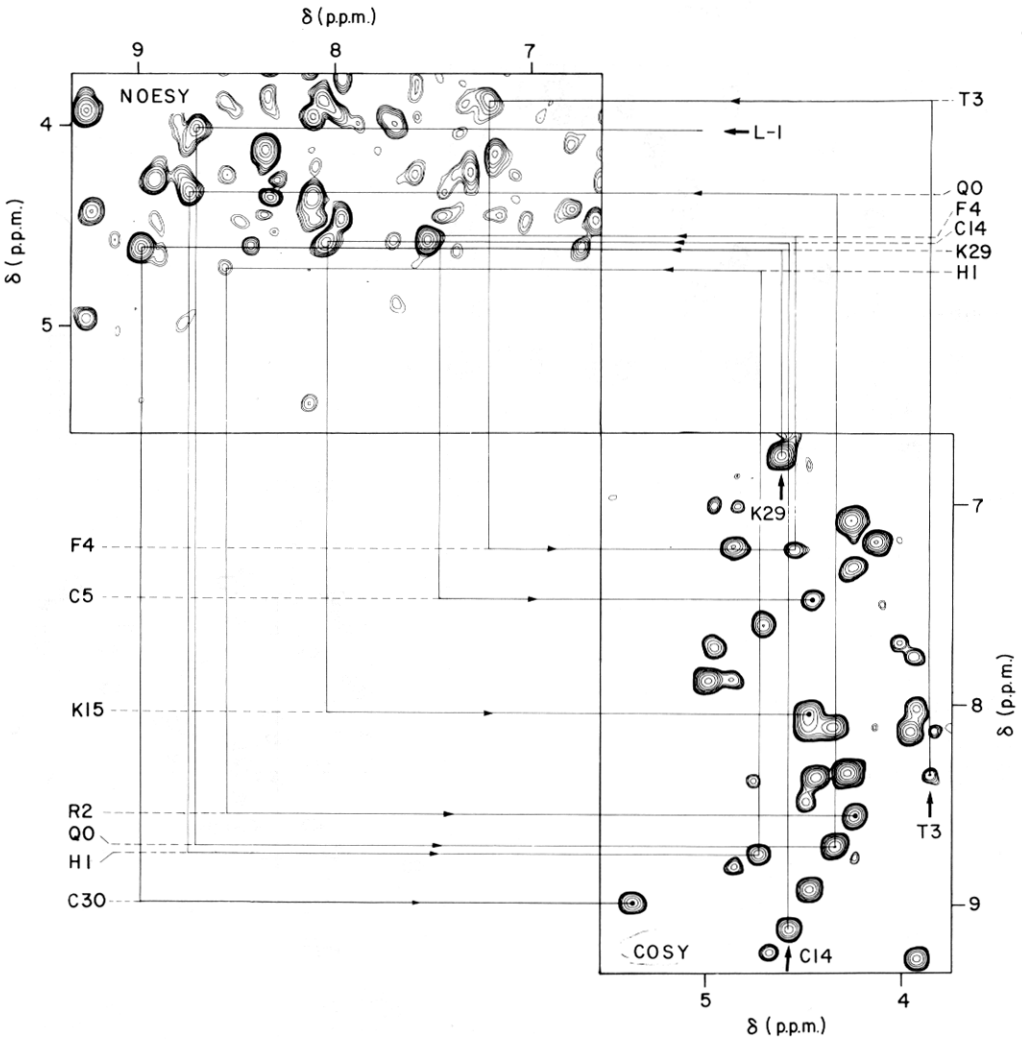

FIG. 14. Combined COSY–NOESY connectivity diagram for sequential resonance assignments *via* NOEs between amide protons and the C^{α} protons of the preceding residue (d_1). In the upper left, the region ($\omega 1$ = 3·7 to 5·5 p.p.m., $\omega 2$ = 6·6 to 9·3 p.p.m.) of a ^1H NOESY spectrum of trypsin inhibitor E in H_2O recorded at pH 3·2 and 25°C with a mixing time of 100 ms and a digital resolution of 5·3 Hz/point is presented. In the lower right, the region ($\omega 1$ = 6·6 to 9·3 p.p.m., $\omega 2$ = 3·7 to 5·5 p.p.m.) of a ^1H COSY spectrum recorded from the same sample, under identical conditions is shown. The straight lines and arrows indicate the connectivities between neighbouring residues in the segments −1 to 2, 3 to 5, 14 to 15 and 29 to 30 of trypsin inhibitor E. Arrows and filled circles identify the start and end point of the connectivity patterns, respectively. The amide proton chemical shifts are indicated by the assignments in the lower left corner, and those for the C^{α} protons by the assignments in the upper right corner of the Figure.

combined with the corresponding spectral region from the lower right triangle of a COSY spectrum recorded under identical conditions.† Such a plot contains both the d_1 connectivities between NH_{i+1} and $C^{\alpha}H_i$ and the J connectivities between $C^{\alpha}H_i$ and NH_i (Fig. 10). When one follows the polypeptide chain in the direction from the N to the C terminus, the lines that connect successive COSY and NOESY cross peaks (Fig. 10) describe a counterclockwise spiral. The data in Figures 12 and 13 are at 45°C, those in Figure 14 at 25°C. To this it may be added that, while the most complete assignments were obtained at 45°C and 50°C (Table 1), nearly complete assignments resulted also at 25°C. For the polypeptide segments in Figure 14, the COSY and NOESY peaks were, as a result of small temperature variations of the chemical shifts, better resolved at 25°C than at the higher temperatures.

d_2 and d_3 connectivities are documented entirely in the NOESY spectrum (Wagner & Wüthrich, 1982). In the presentation of d_2 connectivities (Fig. 15), the diagonal amide proton resonances of neighbouring residues are connected *via* the NOE cross peaks. In the plots of d_3 connectivities (Fig. 16), one uses also information obtained from the COSY and SECSY spectra (Wider *et al.*, 1982). For example, the assignments for the peptide segment Ala26-Ala27-Lys28 started with the observation of the NOE cross peak between Lys28 NH and Ala27 $C^{\beta}H_3$. A connecting line was then drawn to the NOE cross peak between NH and $C^{\beta}H_3$ of Ala27, of which the location was known from the J connectivities between NH, $C^{\alpha}H$ and $C^{\beta}H_3$ of Ala27 observed in the COSY spectra in H_2O. Next, the d_3 connectivity from Ala27 NH to Ala26 $C^{\beta}H_3$ was detected in the NOESY spectrum, and then the location of the NOE cross peak between NH and $C^{\beta}H_3$ of Ala26 was again obtained from COSY. When checking the d_3 connectivities involving residues with non-equivalent C^{β} methylene protons, it is helpful to remember that the NOEs between the two methylene protons and a given amide proton may have different intensities. For example for the two C^{β} methylene protons of Glu49, a strong NOE cross peak with NH of Glu49 is observed only for the resonance at 1·96 p.p.m., and a strong NOESY cross peak with NH of Lys50 only for the resonance at 1·81 p.p.m. (Fig. 16).

The COSY spectrum in Figure 17 provides a final check of the sequential resonance assignments. All the strong $C^{\alpha}H$–NH J cross peaks in the typical $C^{\alpha}H$ chemical shift range from 3·5 to 5 p.p.m. have been assigned and each cross peak has been assigned only once.

Two practical aspects are not documented explicitly. The first is that initial sequential assignments of short peptide segments were obtained in a freshly prepared ^2H_2O solution of the inhibitor E (Fig. 5). Obviously, because of the lesser number of resonances present (e.g. compare Figs 5 and 17), the analysis of this spectrum was somewhat easier. The data obtained in ^2H_2O then served as starting points for the spectral analysis in H_2O. Secondly, once any one of the three connectivities d_1, d_2 or d_3 is established, the locations of the NOESY peaks that manifest the other two connectivities can be predicted from the J connectivities

† In practice, to evade interference with the strong, vertical tail of the water resonance (Fig. 6), the mirror image of the upper left triangle of the COSY spectrum was used.

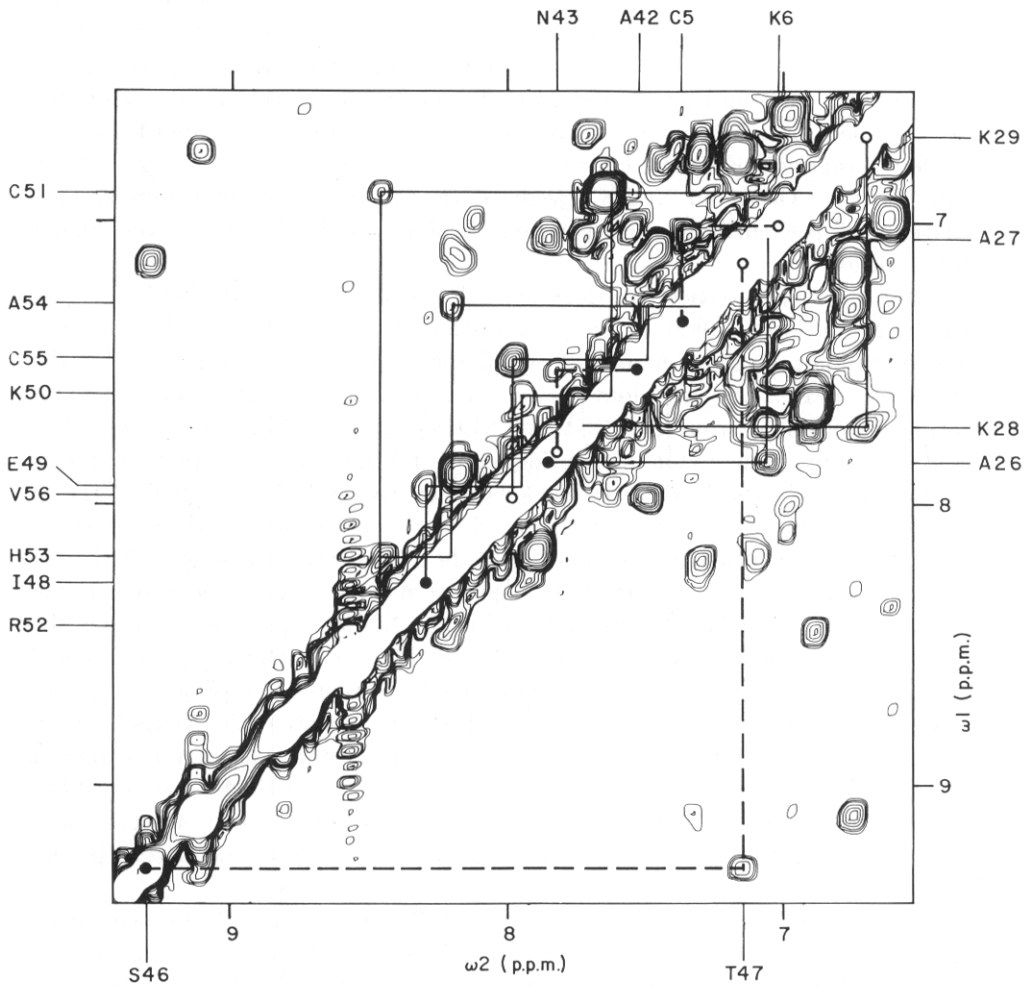

FIG. 15. Contour plot of the spectral region from 6·5 to 9·4 p.p.m. of the 500 MHz ^{1}H NOESY spectrum of Fig. 6, which was recorded from a solution of the inhibitor E in H_2O at pH 3·2 and 45 °C with a mixing time of 200 ms. This spectral region contains the diagonal peaks of the bulk of the backbone amide protons and many cross-peaks manifest NOEs between different amide protons. d_2 connectivities between the amide protons of neighbouring residues are indicated by the unbroken and broken lines, and the resonance positions of the connected amide protons are indicated on the margins of the Figure. The upper left triangle contains the connectivities for the polypeptide segments 5 to 6 and 42 to 43 (– – –, chemical shifts δ indicated at the top of the Figure), and 48 to 56 (———, δ on the left). The lower right triangle contains the sequential resonance assignments for the segments 26 to 29 (———, δ on the right) and 46 to 47 (– – –, δ at the bottom). Filled and empty circles indicate the start and the end for the d_2 connectivity patterns, respectively.

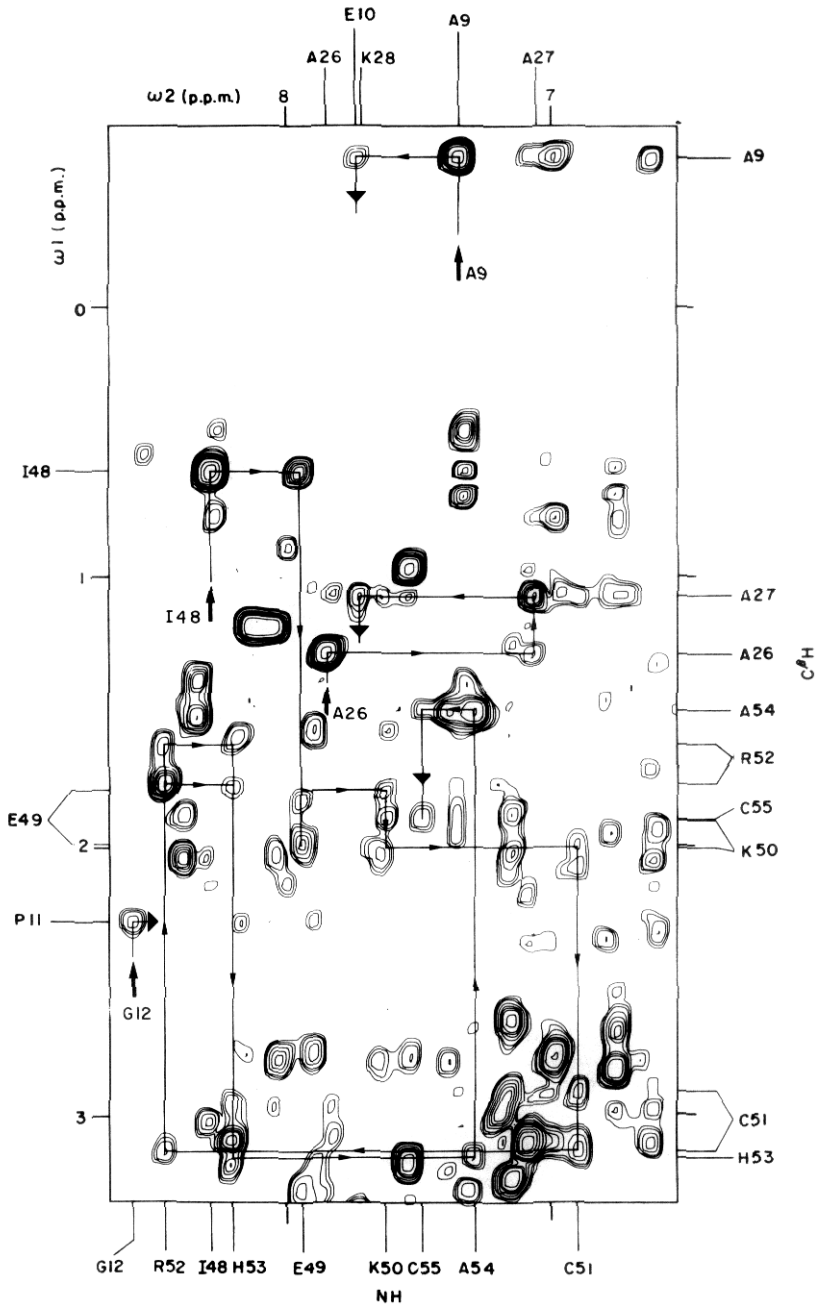

FIG. 16. Contour plot of the spectral region ($\omega 1 = -0{\cdot}7$ to 3·3 p.p.m., $\omega 2 = 6{\cdot}5$ to 8·7 p.p.m.) of the 500 MHz ^{1}H NOESY spectrum in Fig. 6, which was recorded from a solution of the inhibitor E in H_2O at pH 3·2 and 45°C with a mixing time of 200 ms. The unbroken lines with arrows indicate the sequential resonance assignments for the polypeptide segments 9 to 10, 11 to 12, 26 to 28 and 48 to 55, which were obtained from NOEs between amide protons and C^β protons of the preceding residues (d_3). Arrows and arrow heads indicate the start and the end of each d_3 connectivity pattern, respectively. The resonance assignments at the top and at the bottom of the Figure indicate the amide proton chemical shifts; those on the left and the right margins, the C^β proton chemical shifts.

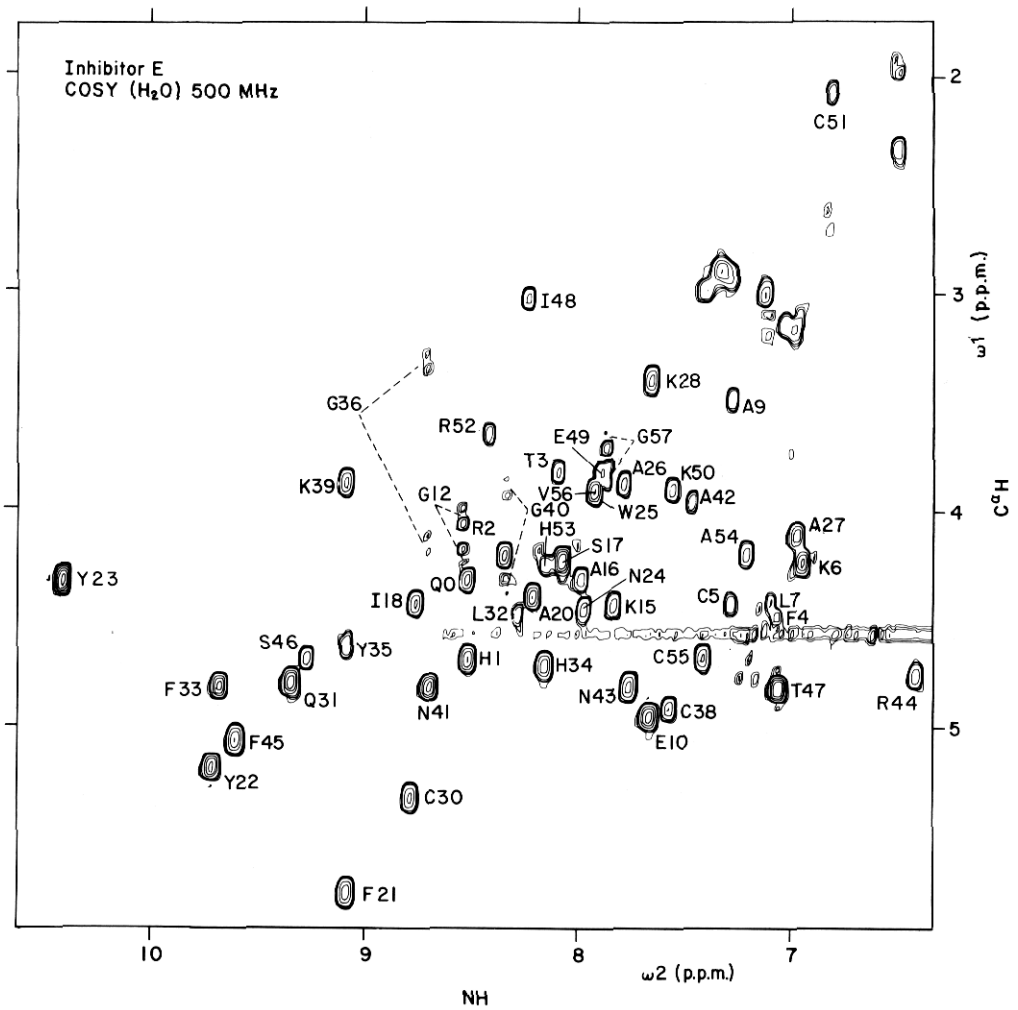

FIG. 17. Spectral region ($\omega 1 = 1{\cdot}7$ to $5{\cdot}9$ p.p.m., $\omega 2 = 6{\cdot}3$ to $10{\cdot}6$ p.p.m.) of a 500 MHz ^{1}H COSY spectrum recorded in a 0·007 M-solution of trypsin inhibitor E in H_2O at pH 3·2 and 45°C. The digital resolution is 5·3 Hz/point, and the spectrum was recorded in approx. 23 h. The letters and numbers indicate the assignments for the $C^{\alpha}H$–NH J cross peaks in the inhibitor E. Peaks are present for all the assigned residues except Leu −1 and Gly37, which have not been observed, and Cys14 and Lys29, which are not visible in this spectrum since their peaks coincide in the $\omega 1$ direction with the water signal and are therefore bleached out by the solvent irradiation (see Table 1). The unassigned peaks in the upper right correspond to arginine and lysine side-chains (Bundi & Wüthrich, 1979).

between NH, $C^{\alpha}H$ and $C^{\beta}H$ (Fig. 10) obtained from COSY and SECSY. Checks for multiple connectivities between neighbouring residues can then be based on comparison of these predicted cross-peak positions with the experimental NOESY spectra. The predicted positions may coincide with a resolved cross peak, which would unambiguously determine a connectivity, fall into an empty region of the spectrum, which would rule out that there is a connectivity, or fall into a spectral region that is too crowded with cross peaks or too close to the diagonal for an

unambiguous identification of individual cross peaks (see Wider *et al.*, 1982, for an illustration. All the connectivities included in Fig. 11 rely on unambiguously identified NOESY cross peaks). Finally, it should be added that, while the NOESY spectra shown in Figures 12, 13, 15 and 16 were recorded with a mixing time of 200 milliseconds, all connectivities indicated in Figure 11 have been confirmed with NOESY spectra recorded with a mixing time of 100 milliseconds.

(d) *Further assignments of amino acid side-chain spin systems after sequential assignments of the polypeptide backbone*

After the sequential assignments of the NH, $C^{\alpha}H$ and $C^{\beta}H_n$ resonances, there remained only relatively few unidentified cross peaks in the COSY and SECSY spectra of Figures 1 to 4. Further, it was known for each $C^{\alpha}H$–$C^{\beta}H_n$ fragment to which type of amino acid residue it belonged. While most of the simple amino acid side-chain spin systems were completely identified before the sequential assignments, many of the more complex side-chains could now, with the availability of this additional information, be assigned. As an illustration, some examples are now described and documented.

The peripheral methyl groups and the $C^{\gamma}H$–$C^{\delta}H_3$ J connectivities were readily identified for all three leucyl residues. For Leu−1 and Leu32, these were then connected to the sequentially assigned $C^{\alpha}H$–$C^{\beta}H_2$ fragments (Fig. 3). For Leu7, no sequential connectivity was available (Fig. 11). However, towards the end of the spectral analysis there was only one unidentified COSY cross peak left in the $C^{\alpha}H$–NH region (Fig. 17). Starting from the $C^{\alpha}H$ chemical shift of this cross peak, the connectivities with the third and last leucine $C^{\gamma}H$–$(C^{\delta}H_3)_2$ fragment could be established in SECSY (Fig. 3).

For Ile18, the J connectivities $C^{\alpha}H$–$C^{\beta}H$, $C^{\beta}H$–$C^{\gamma}H_3$ and $C^{\delta}H_3$–$C^{\gamma}H_2$ were readily apparent, and these three fragments could be connected to the complete spin system (Fig. 18). For Ile48, the $C^{\alpha}H$–$C^{\beta}H$ fragment with the unusual $C^{\alpha}H$ high field chemical shift was sequentially assigned, and $C^{\beta}H$ and $C^{\gamma}H_3$ were found to have identical chemical shifts (Table 1). This degeneracy was confirmed by the observation of a singlet methyl resonance at 0·66 p.p.m. in the conventional, one-dimensional spectrum and in the two-dimensional, J-resolved spectrum. There was only one unidentified methyl triplet left, which thus had to come from $C^{\delta}H_3$ of Ile48. The connectivity of $C^{\gamma}H_2$ was determined by SECSY (Fig. 3). The $C^{\beta}H$–$C^{\gamma}H_2$ connectivity for this residue, however, was not observed.

For Pro11, the $C^{\delta}H_2$ resonances were sequentially assigned *via* NOEs with the C^{α} proton of Glu10 (Fig. 18) and one $C^{\beta}H$ resonance was assigned *via* a NOE with the amide proton of Gly12 (Fig. 16). For Pro19, the $C^{\delta}H_2$ resonances were assigned *via* NOE with $C^{\alpha}H$ of Ile18 (Fig. 18) and the $C^{\alpha}H$ *via* NOE with NH of Ala20 (Fig. 12). As described in detail in the legend to Figure 18, the complete spin systems of these two prolyl residues could then be identified.

Individual assignments for the previously identified (Figs 2 and 4) aromatic spin systems were obtained on the basis of NOE connectivities between aromatic protons and the sequentially assigned $C^{\alpha}H$–$C^{\beta}H_2$ fragments (Fig. 19). The only ambiguity left in the assignments of the aromatics is that between Phe4 and Phe33,

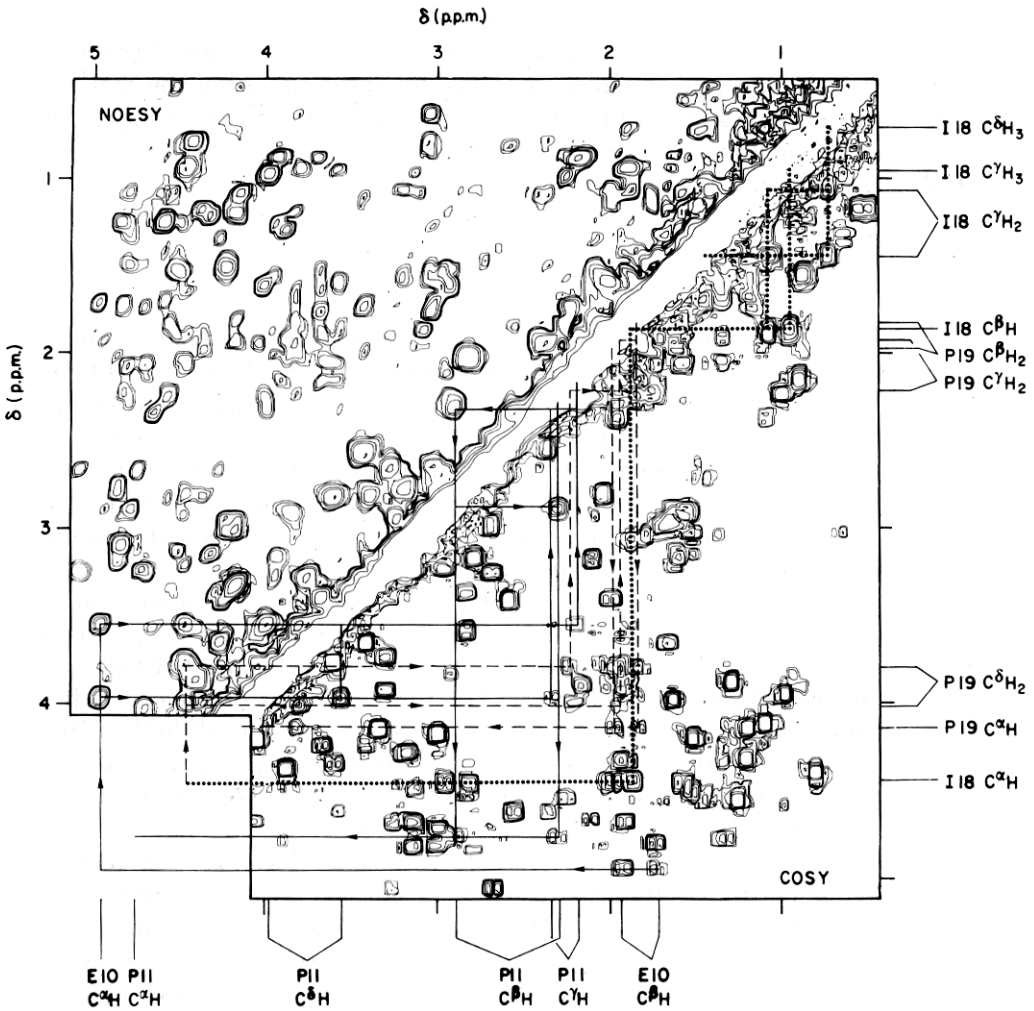

FIG. 18. Resonance assignments for the peptide segments Glu10 to Pro11 and Ile18 to Pro19 in the inhibitor E. A combined plot is shown, which was obtained from the regions 0·5 to 5 p.p.m. of the spectra in Figs 1 and 19 and where the upper left triangle comes from NOESY and the lower right triangle from COSY. The chemical shifts for the residues Glu10 and Pro11 are indicated at the bottom, those for Ile18 and Pro19 on the right of the Figure. The connectivities in the segment Glu10–Pro11 are indicated by unbroken lines. They start from the previously assigned $C^{\beta}H_2$–$C^{\alpha}H$ resonances of Glu10 (Fig. 7), from which the NOE connectivity to $C^{\delta}H_2$ of Pro11 was obtained. Within the Pro11 spin system, the connectivities *via* J-coupling could be established in all cases except from $C^{\gamma}H_2$ to $C^{\beta}H_2$, where a NOE connectivity was used. The J connectivities within the spin system of Ile18 are indicated by dotted lines in the COSY spectrum. The NOE connectivities from $C^{\alpha}H$ of Ile18 to $C^{\delta}H_2$ of Pro19 and the J connectivities within the spin system of Pro19 are indicated with broken lines.

which arose because these two residues have identical C2,6H chemical shifts (Table 1).

Complete assignments were obtained for Gln0, for the previously identified five-spin system of Glu10, and for Gln31. First, the connectivities from the sequentially assigned $C^{\alpha}H$–$C^{\beta}H_2$ fragments to $C^{\gamma}H_2$ were established in SECSY (Figs 7 and 8).

Then the ε amide proton resonances of the two glutamyl residues were assigned from the NOE connectivities with $C^{\gamma}H_2$ (Billeter *et al.*, 1982). Similarly, the δ amide protons of Asn24, Asn41 and Asn43 were assigned from the NOE connectivities with $C^{\beta}H_2$.

4. Conclusions

The inhibitor E is the first protein for which the ^{1}H n.m.r. spectrum was assigned entirely according to the recently proposed strategy of sequential assignments with the use of two-dimensional n.m.r. experiments (Nagayama & Wüthrich, 1981; Wagner *et al.*, 1981; Billeter *et al.*, 1982). In contrast to the previously described BPTI (Wagner & Wüthrich, 1982) and micelle-bound glucagon (Wider *et al.*, 1982), where the two-dimensional n.m.r. studies were a continuation of extensive investigations by conventional one-dimensional n.m.r., this paper describes the first n.m.r. experiments with the inhibitor E. As may be judged from the documentation in this paper, the resonance assignments are based exclusively on the amino acid sequence and the n.m.r. spectra, and they were obtained without reference to the corresponding data on BPTI. With regard to future applications of the presently used assignment techniques, the experience with the inhibitor E is most encouraging. The entire analysis was performed with a single sample of about 18 mg, and we have still about 14 mg of the protein left for further studies of the molecular conformation. Overall, less than one man-year of work was invested to obtain the assignments in Table 1.

Let us briefly consider why a few assignments are not given in Table 1. The missing sequential connectivities are indicated in Figure 11 and have already been discussed. Quite generally, it appears that continuous connectivities across prolyl residues can be established only under favourable conditions. Since the chemical shifts for Pro8 and Pro13 are not known, it is not clear whether for these two residues the absence of NOE connectivities with the neighbouring amino acid residues is due to lack of spectral resolution or to other reasons. It is also unclear at present why the amide proton of Gly37 was not observed, but it seems worth noting that the amide proton resonance of Gly37 was also missing in BPTI (Wagner & Wüthrich, 1982). Besides Pro8 and Pro13, the side-chain assignments beyond $C^{\beta}H_2$ are also missing for Glu49, where the connectivities with $C^{\gamma}H_2$ could not be identified unambiguously, and for the three histidyl, the three arginyl and the six lysyl residues. For the histidyl residues, the NOE connectivities between the sequentially assigned $C^{\alpha}H$–$C^{\beta}H_2$ fragments and the imidazole ring protons were not observed. This was not unexpected, since stereochemical considerations indicate that this method should work reliably for Phe, Tyr and Trp (Fig. 19), but not necessarily for His (Billeter *et al.*, 1982). Intense cross peaks that manifest J connectivities with $C^{\varepsilon}H_2$ of Lys and $C^{\delta}H_2$ of Arg can be recognized readily, e.g. in Figures 1 and 3. However, no special effort was made to obtain individual assignments for the side-chains of Lys and Arg.

The exploitation of the resonance assignments in Table 1 for studies of the solution conformation of the inhibitor E, structural comparisons with homologous

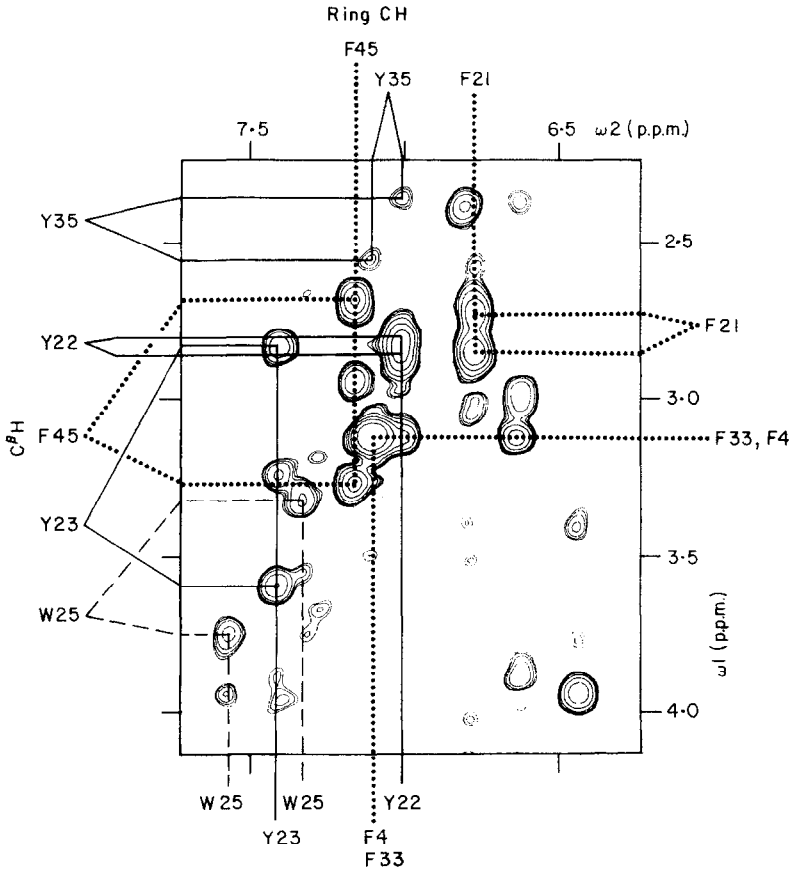

FIG. 19. Spectral region ($\omega 1 = 2{\cdot}2$ to $4{\cdot}1$ p.p.m., $\omega 2 = 6{\cdot}2$ to $7{\cdot}7$ p.p.m.) of a 500 MHz ^{1}H NOESY spectrum recorded in a 0·007 M solution of trypsin inhibitor E in ^2H_2O, p^{2}H 3·2 at 50°C. The mixing time was 100 ms, the digital resolution is 5·3 Hz/point, and the spectrum was recorded in approx. 21 h. The Figure illustrates how the connectivities between $C^{\beta}H_2$ and the rings were obtained for the aromatic side-chains. $C^{\beta}H$ chemical shifts are indicated on the left and right. Chemical shifts for C2,6H of Phe and Tyr and C2H and C4H of Trp25 are indicated at the top and at the bottom.

proteins and investigations of correlations between protein conformation and n.m.r. parameters, as outlined in the Introduction, is in progress in our laboratory.

We thank Dr G. Wagner for helpful discussions, Dr L. Visser for initiating the collaboration between the laboratories in Zürich and in Pretoria, and Mrs E. Huber for the careful preparation of the manuscript. One of us (A.S.A.) was the recipient of a fellowship from the U.S.S.R. Academy of Sciences. We acknowledge financial support by the Schweizerischer Nationalfonds (project 3.528.79) and by a special grant from the Eidgenössische Technische Hochschule, Zürich, for the purchase of the 500 MHz spectrometer.

REFERENCES

Anil Kumar, Ernst, R. R. & Wüthrich, K. (1980*a*). *Biochem. Biophys. Res. Commun.* **95**, 1–6.

Anil Kumar, Wagner, G., Ernst, R. R. & Wüthrich, K. (1980*b*). *Biochem. Biophys. Res. Commun.* **96**, 1156–1163.

Aue, W. P., Bartholdi, E. & Ernst, R. R. (1976). *J. Chem. Phys.* **64**, 2229–2246.
Bax, A. & Freeman, R. (1981). *J. Magn. Reson.* **44**, 542–561.
Billeter, M., Braun, W. & Wüthrich, K. (1982). *J. Mol. Biol.* **155**. 321–346.
Bundi, A. & Wüthrich, K. (1979). *Biopolymers.* **18**, 285–297.
Campbell, I. D., Dobson, C. M. & Williams, R. J. P. (1975). *Proc. Roy. Soc. ser. B*, **189**, 503–509.
Campbell, I. D., Dobson, C. M., Moore, G. R., Perkins, S. J. & Williams, R. J. P. (1976). *FEBS Letters*, **70**, 96–100.
Dubs, A., Wagner, G. & Wüthrich, K. (1979). *Biochim. Biophys. Acta*, **577**, 177–194.
Englander, S. W., Downer, N. W. & Teitelbaum, H. (1972). *Annu. Rev. Biochem.* **41**, 903–924.
Jeener, J., Meier, B. H., Bachmann, P. & Ernst, R. R. (1979). *J. Chem. Phys.* **71**, 4546–4553.
Joubert, F. J. & Strydom, D. J. (1978). *Eur. J. Biochem.* **87**, 191–198.
Macura, S., Huang, Y., Suter, D. & Ernst, R. R. (1981). *J. Magn. Reson.* **43**, 259–281.
Nagayama, K. & Wüthrich, K. (1981). *Eur. J. Biochem.* **114**, 365–374.
Nagayama, K., Wüthrich, K. & Ernst, R. R. (1979). *Biochem. Biophys. Res. Commun.* **90**, 305–311.
Nagayama, K., Anil Kumar, Wüthrich, K. & Ernst, R. R. (1980). *J. Magn. Reson.* **40**, 321–334.
Richarz, R., Sehr, P., Wagner, G. & Wüthrich, K. (1979). *J. Mol. Biol.* **130**, 19–30.
Strydom, D. J. (1973). *Nature New. Biol.* **243**, 88–89.
Strydom, D. J. (1976). *Eur. J. Biochem.* **69**, 169–176.
Tschesche, H. (1974). *Angew. Chemie, Int. Ed. Engl.* **13**, 10–28.
Wagner, G. & Wüthrich, K. (1982). *J. Mol. Biol.* **155**, 347–366.
Wagner, G., DeMarco, A. & Wüthrich, K. (1975). *J. Magn. Reson.* **20**, 565–569.
Wagner, G., DeMarco, A. & Wüthrich, K. (1976). *Biophys. Struct. Mechan.* **2**, 139–158.
Wagner, G., Anil Kumar, & Wüthrich, K. (1981). *Eur. J. Biochem.* **114**, 375–384.
Wider, G., Lee, H. K. & Wüthrich, K. (1982). *J. Mol. Biol.* **155**, 367–388.
Wüthrich, K. (1976). *NMR in Biological Research: Peptides and Proteins*. North-Holland Publishing Company, Amsterdam.
Wüthrich, K. & Wagner, G. (1975). *FEBS Letters*, **50**, 265–268.
Wüthrich, K. & Wagner, G. (1979). *J. Mol. Biol.* **130**, 1–18.
Wüthrich, K., Wider, G., Wagner, G. & Braun, W. (1982). *J. Mol. Biol.* **155**, 311–319.

Edited by V. Luzzati