J. Mol. Biol. (1983) **166**, 641–667

Assignment of the ^1H Nuclear Magnetic Resonance Spectrum of the Proteinase Inhibitor IIA from Bull Seminal Plasma by Two-dimensional Nuclear Magnetic Resonance at 500 MHz

PETR ŠTROP†, GERHARD WIDER‡ AND KURT WÜTHRICH

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

(Received 23 November 1982)

The assignment of the ^1H nuclear magnetic resonance (n.m.r.) spectrum of the protease inhibitor IIA from bull seminal plasma is described and documented. The assignments are based entirely on the amino acid sequence and on two-dimensional n.m.r. experiments at 500 MHz. Individual assignments were obtained at 18°C and 45°C for the backbone protons of all 57 amino acid residues, with the single exception of the N-terminal pyroglutamate amide proton. The amino acid side-chain resonance assignments are complete, with the exception of 17 long side-chains, i.e. Pro13, Met43 and all the Glu, Gln, Lys and Arg, where only one or two resonances of $C^{\beta}H_2$ and in some cases $C^{\gamma}H_2$ could be identified. The sequential assignments showed that the order of the two C-terminal residues in the previously established primary structure had to be changed; this was then confirmed by chemical methods. The chemical shifts for the assigned resonances at 18°C and 45°C are listed for an aqueous solution at pH 4·9. A preliminary characterization of the polypeptide secondary structure was obtained from the observed patterns of sequential connectivities.

1. Introduction

The preceding paper (Štrop & Wüthrich, 1983) provides a qualitative characterization of the ^1H nuclear magnetic resonance spectral properties, the stability and the internal dynamics of the globular conformation of the bull seminal inhibitor BUSI IIA§. The present paper describes and documents the

† Present address: Institute of Organic Chemistry and Biochemistry C.S.A.V., 16610 Prague 6, C.S.S.R.

‡ Present address: Spectrospin AG, Industriestrasse 26, CH-8117 Fällanden, Switzerland.

§ Abbreviations used: BUSI IIA, protease inhibitor IIA from bull seminal plasma; BUSI IIB, protease inhibitor IIB from bull seminal plasma; n.m.r., nuclear magnetic resonance; 2D n.m.r., 2-dimensional n.m.r.; p.p.m., parts per million; COSY, 2D correlated spectroscopy; NOE, nuclear Overhauser enhancement; NOESY, 2D nuclear Overhauser enhancement spectroscopy; SECSY, 2D spin echo correlated spectroscopy.

0022–2836/83/160641–27 $03.00/0

resonance assignments in the ^{1}H n.m.r. spectra of this protein at 18°C and 45°C. These data rely on analysis of 2D n.m.r. experiments at 500 MHz for obtaining sequential, individual assignments for the hydrogen atoms of the polypeptide backbone and most of the amino acid side-chains (Billeter *et al.*, 1982; Wagner & Wüthrich, 1982*a*; Wider *et al.*, 1982). The individual resonance assignments provide a basis for determination of static and dynamic aspects of the protein conformation in solution (Wüthrich *et al.*, 1982), and a preliminary characterization of the secondary structure results directly from the n.m.r. data used to obtain the resonance assignments.

2. Materials and Methods

The starting material, of 4000 ml of bull seminal plasma, was obtained from Dr D. Čechová. Isoinhibitor BUSI IIA was purified according to the method of Čechová *et al.* (1979). The purity of the protein was checked by discontinuous electrophoresis, electrophoresis on cellulose acetate sheets, isoelectric focusing and capillary isotachophoresis.

For the n.m.r. experiments approx. 0·016 M solutions of BUSI IIA were prepared in the following solvents: (1) ^2H_2O, p^{2}H 4·9; 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 4·9; 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 4·9.

The 3 two-dimensional n.m.r. experiments used to obtain the resonance assignments have been described previously in detail. COSY spectra were recorded with the pulse sequence (Aue *et al.*, 1976; Nagayama *et al.*, 1980; Bax & Freeman, 1981):

$$[90° - t_1 - 90° - t_2]_n,$$

where t_1 and t_2 are, respectively, the evolution period and the observation period. To obtain a 2D 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 1·0 s. COSY spectra in ^2H_2O were obtained from 512 measurements, with t_1 values from 0·3 to 59 ms, those in H_2O from 512 measurements, with t_1 from 0·3 to 51 ms. SECSY used the pulse sequence (Nagayama *et al.*, 1979,1980):

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

SECSY spectra were recorded in ^2H_2O solution from 512 measurements, with t_1 values from 0·3 to 262 ms. NOESY used a sequence of 3 pulses (Jeener *et al.*, 1979; Anil Kumar *et al.*, 1980*a*):

$$[90° - t_1 - 90° - \tau_m - 90° - t_2]_n,$$

where τ_m is the so-called mixing time. The NOESY spectra used here were obtained from 512 measurements, with t_1 values from 0·3 to 59 ms in ^2H_2O and from 0·3 to 51 ms in H_2O. Mixing times of 100 and 200 ms were used. To suppress contributions from coherent magnetization transfer to the cross-peak intensities, the 100 ms mixing time was stochastically modulated with a modulation amplitude corresponding to 5% of τ_m (Macura *et al.*, 1981).

The 2D n.m.r. spectra were recorded at 500 MHz on a Bruker WM500 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*; Wider *et al.*, 1983). 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 1024×4096 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 direction 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 1/64 and 1/128 in the t_1 and t_2 directions, respectively. All the spectra are shown in the absolute value representation.

3. Results

The procedures used to assign the ^{1}H n.m.r. spectrum of BUSI IIA are very similar to those employed recently for the assignments of the basic pancreatic trypsin inhibitor and two protease inhibitors isolated from snake venom (Nagayama & Wüthrich, 1981; Wagner & Wüthrich, 1982*a*; Arseniev *et al.*, 1982; Keller *et al.*, 1983). Since BUSI IIA is less stable at low pH values than these three Kunitz-type inhibitors (Štrop & Wüthrich, 1983), the assignments were obtained at pH 4·9. At this pH the amide protons exchange more rapidly than at the pH minimum near 3·5 (Richarz *et al.*, 1979). Therefore, only a small number of sequential assignments could be established in ^2H_2O solution and most of the sequential connectivities had to be obtained from spectra recorded in H_2O. More serious difficulties arose because of the presence of small concentrations of modified BUSI IIA, which were formed in the n.m.r. samples during the recordings of the 2D n.m.r. spectra. This will be discussed further near the end of the paper.

In the following the assignment procedures are described and the assignments are documented with the original spectra. In Table 1, the chemical shifts of the assigned resonances at pH 4·9 are listed at two temperatures, 18°C and 45°C.

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

In this first step the spin systems of the different types of amino acid residues were identified in SECSY spectra of BUSI IIA recorded in ^2H_2O. The SECSY spectra were screened for the characteristic cross-peak patterns of the different amino acid side-chains (Nagayama & Wüthrich, 1981), as is documented in Figures 1 and 2 for the non-aromatic protons and in Figure 3 for the aromatic rings. Table 2 presents a detailed explanation of the interpretation of SECSY spectra. At the digital resolution selected for the SECSY experiments, the J coupling fine structure (Nagayama *et al.*, 1979,1980) is readily apparent in most of the cross peaks. The individual assignments obtained as a final result of all the experiments described in this paper are also indicated in Figures 1 to 3. It is worth noting that well-resolved, high-quality spectra were obtained at 45°C (Figs

TABLE 1

Chemical shifts of the assigned ^1H n.m.r. lines of the bull seminal inhibitor BUSI IIA at pH 4·9, t = 18°C and 45°C

Amino acid residue	Chemical shifts at 45°C‡§				Chemical shifts at 18°C‡§			
	NH	$C^{\alpha}H$	$C^{\beta}H$	Others	NH	$C^{\alpha}H$	$C^{\beta}H$	Others
Pyr1†		4·40	2·59, 2·12			4·41	2·60, 2·12	
Gly2	8·39	4·03, 4·03			8·56	4·03, 4·03		
Ala3	8·29	4·42	1·41		8·51	4·40	1·39	
Gln4	8·32	4·40	2·15, 1·96	$C^{\gamma}H_2$ 2·37	8·45	4·40	2·13, 1·99	$C^{\gamma}H_2$ 2·40
Val5	8·26	4·22	2·17	$C^{\gamma}H_3$ 1·01, 1·01	8·52	4·18	2·17	$C^{\gamma}H_3$ 1·01, 1·01
Asp6	8·58	4·78	2·91, 2·67		8·79	4·80	2·94, 2·67	
Cys7	8·63	5·16	3·45, 2·78		8·83	5·17	3·47, 2·78	
Ala8	8·23	3·93	1·51		8·30	3·93	1·50	
Glu9	8·88	4·00	1·57, 1·41		9·08	3·99	1·56, 1·37	
Phe10	7·64	4·73	3·40, 2·51	C2, 6H 7·16 C3, 5H 7·27 C4H 7·27	7·69	4·75	3·42, 2·48	C2, 6H 7·15 C3, 5H 7·25 C4H 7·25
Lys11	7·36	4·04	1·83,		7·37	4·02	1·82,	
Asp12	7·71	4·75	2·78, 2·57		7·87	4·71	2·77, 2·56	
Pro13		4·41	2·47, 2·02	$C^{\delta}H_2$ 3·80		4·41	2·37, 1·98	$C^{\delta}H_2$ 3·80
Lys14	8·49	4·17	1·79,		8·56	4·12	1·77,	
Val15	7·43	3·92	2·12	$C^{\gamma}H_3$ 1·04, 1·04	7·50	3·89	2·12	$C^{\gamma}H_3$ 1·05, 1·05
Tyr16	8·19	4·78	3·06, 2·93	C2, 6H 7·15 C3, 5H 6·85	8·41	4·78	3·08, 2·87	C2, 6H 7·14 C3, 5H 6·83
Cys17	8·58	5·03	3·35, 2·80		8·83	5·04	3·36, 2·76	
Thr18	8·33	4·57	4·67	$C^{\gamma}H_3$ 1·40	8·53	4·61	4·72	$C^{\gamma}H_3$ 1·40
Arg19	8·28	4·36	2·09, 1·72		8·42	4·38	2·11,1·70	
Glu20	8·05	4·18	2·02,		8·13	4·15	2·02,	
Ser21	8·75	4·75	3·93, 3·82		8·95	4·72	3·92, 3·79	
Asn22	9·05	5·06	3·02, 2·96	$N^{\delta}H_2$ 7·01, 7·76	9·14	5·05	3·02, 2·97	$N^{\delta}H_2$ 7·17, 7·89
Pro23		4·78	2·14, 1·48	$C^{\gamma}H_2$ 2·08, 2·00, $C^{\delta}H_2$ 3·78, 3·67		4·79	2·17, 1·49	$C^{\gamma}H_2$ 2·08, 2·00 $C^{\delta}H_2$ 3·77, 3·67
His24	9·25	4·22	2·61, 2·31	C2H 7·53 C4H 6·12	9·33	4·19	2·63, 2·31	C2H 7·53 C4H 6·09
Cys25	8·29	5·19	2·57, 1·85		8·33	5·17	2·56, 1·78	
Gly26	9·53	4·94, 4·52			9·62	5·18, 4·53		
Ser27	9·43	4·08	4·27, 3·89		9·52	4·06	4·22, 3·93	
Asn28	8·95	4·52	3·05, 2·70		8·89	4·53	3·03, 2·70	
Gly29	8·52	4·14, 3·79			8·57	4·12, 3·81		
Glu30	7·72	4·35	1·94, 1·65		7·75	4·32	1·96, 1·63	

Thr31	8·30	4·65	4·00	$C^{\gamma}H_3$ 1·13	8·47	4·64	3·99	$C^{\gamma}H_3$ 1·13
Tyr32	9·50	4·55	3·02, 2·91	C2, 6H 7·10§	9·59	4·55	2·99, 2·99	C2, 6H 7·16‖
Gly33	9·18	3·99, 3·75			9·32	4·00, 3·74		
Asn34	7·37	5·13	3·55, 3·16	$N^{\delta}H_2$ 6·32, 7·78	7·37	5·13	3·54, 3·14	$N^{\delta}H_2$ 6·26, 7·72
Lys35	8·60	2·93	1·82, 1·72		8·67	2·87	1·82, 1·72	
Cys36	7·92	2·44	2·93,		8·04	2·30	2·88,	
Ala37	8·45	3·91	1·58		8·54	3·92	1·57	
Phe38	7·92	3·40	2·78, 2·72	C2, 6H 6·86	7·98	3·38	2·78, 2·68	C2, 6H 6·88
				C3, 4H 7·03				C3, 5H 7·06
				C4H 6·88				C4H 6·88
Cys39	8·83	3·97	2·79, 2·71		8·92	3·96	2·81, 2·79	
Lys40	8·72	3·95	1·97, 1·77		8·86	3·98	1·95, 1·78	
Ala41	7·23	4·06	1·16		7·26	4·06	1·14	
Val42	8·21	3·12	2·01	$C^{\gamma}H_3$ 0·78, 0·10	8·30	3·10	2·00	$C^{\gamma}H_3$ 0·78, 0·07
Met43	8·13	4·22	2·13,	$C^{\epsilon}H_3$ 2·02	8·18	4·22	2·14,	$C^{\epsilon}H_3$ 2·02
Lys44	8·26	4·14	1·95,		8·30	4·14	1·95,	
Ser45	7·86	4·53	4·23, 4·01		7·88	4·53	4·28, 4·00	
Gly46	8·03	4·03, 3·96			8·12	4·05, 3·96		
Gly47	8·19	4·18, 3·50			8·17	4·15, 3·48		
Lys48	7·29	4·17	1·89, 1·73		7·33	4·15	1·90, 1·73	
Ile49	7·09	3·94	1·32	$C^{\gamma}H_2$ 1·48, 0·88	7·20	3·91	1·29	$C^{\gamma}H_2$ 1·54, 0·88
				$C^{\gamma}H_3$ 0·78				$C^{\gamma}H_3$ 0·88
				$C^{\delta}H_3$ 0·79				$C^{\delta}H_3$ 0·78
Asn50	8·74	5·03	2·87, 2·80		8·89	5·01	2·92, 2·76	$N^{\delta}H_2$ 7·42, 6·84
Leu51	8·95	4·26	1·98, 1·21	$C^{\gamma}H$ 0·97	9·12	4·28	1·96,	$C^{\gamma}H$ 0·90
				$C^{\delta}H$ 0·67, 0·19				$C^{\delta}H_3$ 0·65, 0·19
Lys52	8·98	4·35	1·55, 1·28		9·03	4·32	1·52, 1·28	
His53	7·06	4·58	3·85, 3·21	C2H 8·83	7·14	4·57	3·87, 3·20	C2H 8·82
				C4H 7·08				C4H 7·12
Arg54	9·15	3·83	1·77, 1·65		9·31	3·815	1·76, 1·65	
Gly55	8·00	4·56, 3·56			8·03	4·57, 3·55		
Lys56¶	7·98	4·06	1·79,		8·15	4·03	1·78,	
Cys57¶	8·13	4·41	3·15, 2·56		8·25	4·38	3·12, 2·56	

† The assignments for Pyr1 rely on the observations that only 1 spin system was left unassigned after the resonances of all the other residues had been found, and that this spin system is absent in the isoinhibitor BUSI IIB (Štrop *et al.*, 1983).

‡ The chemical shifts, δ, are relative to internal sodium 3-trimethyl-silyl-[2,2,3,3-$^2H_4$]propionate.

§ 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 assignments were obtained (see the text).

‖ At 18°C and 45°C only one 2-proton doublet Tyr32 was observed, which could be assigned to C2, 6H from the NOE connectivities with $C^{\beta}H_2$ (Fig. 13). At these temperatures, the C3, 5H resonances are broadened by the ring-flip process (Wüthrich, 1976; Campbell *et al.*, 1976) and at higher temperatures they overlap with the C2, 6H resonance (Štrop & Wüthrich, 1983).

¶ Based on evidence from the sequential n.m.r. assignments and subsequent reinvestigation of the C-terminal sequence by conventional methods (see the Appendix), the order of the residues 56 and 57 was inverted as compared to the previously published sequence (Meloun & Čechová, 1979).

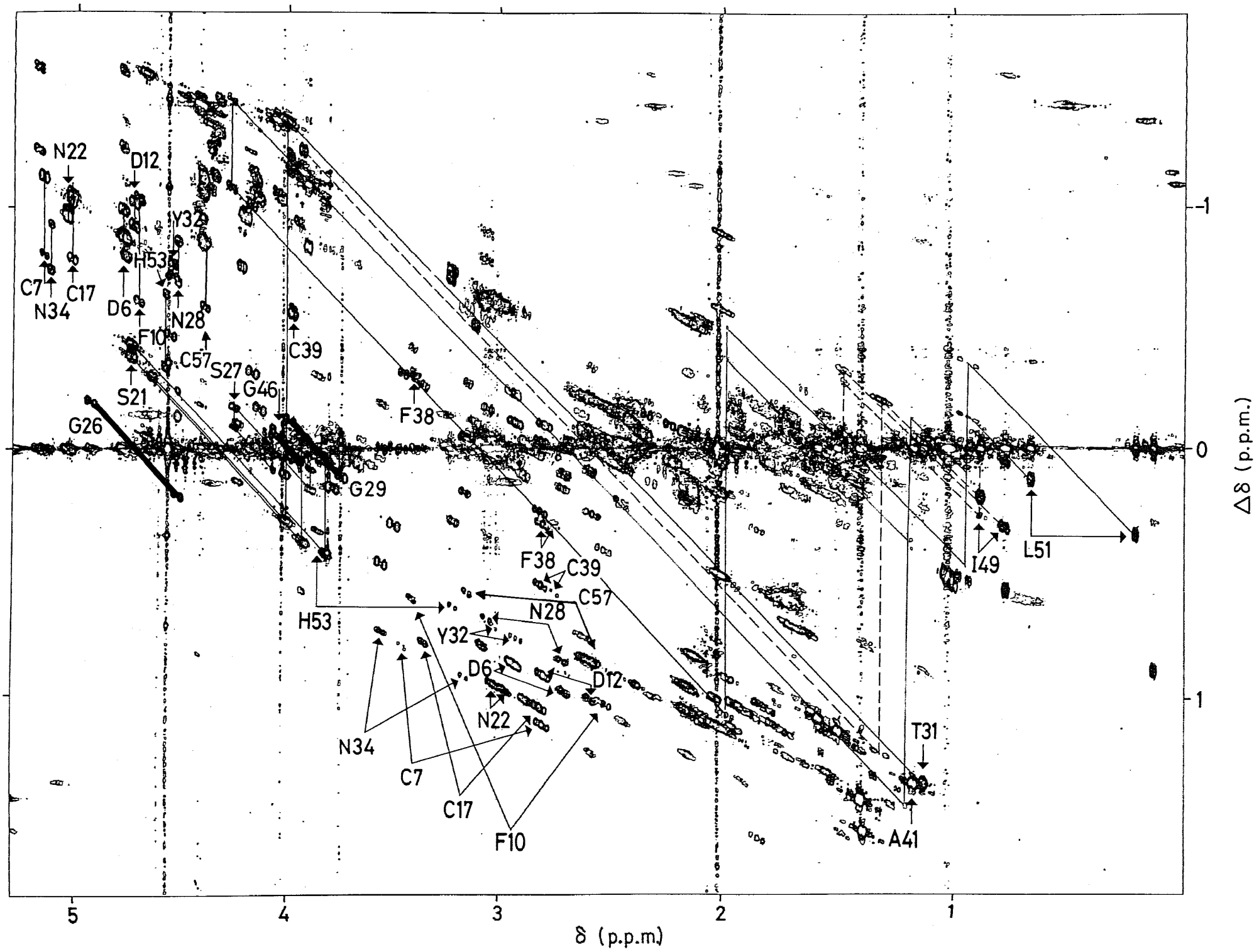

N22
D12
Y32
H53
C7
C17
N34
D6
N28
F10
C57
S27
G46
C39
S21
F38
G26
G29
I49
L51
F38
C39
C57
N28
H53
Y32
D6
D12
N22
T31
N34
C7
C17
F10
A41
5
4
3
2
1
δ (p.p.m.)
−1
0
1
Δδ (p.p.m.)

FIG. 1.

1 and 3) and at 18°C (Fig. 2), where the complete resonance assignments were elucidated.

From the through-bond spin–spin coupling connectivities in SECSY six amino acid types can be uniquely identified from their unique molecular symmetry. These are Gly, Ala, Val, Leu, Ile and Thr. The eight amino acid residues Ser, Asp, Asn, Cys, Phe, Tyr, His and Trp all have $C^{\alpha}H$—$C^{\beta}H_2$ three-proton spin systems and a further distinction between these residues results only when the data of Figures 1 and 2 are combined with sequential assignments of the backbone resonances and/or with NOE connectivities with the aromatic rings (see below). In practice, the complete spin systems of Pro, Glu, Gln, Met, Lys and Arg have been identified only in particularly favorable cases (Nagayama & Wüthrich, 1981; Wagner & Wüthrich, 1982*a*), so that one ends up with a group of "long-side-chain spin systems". These are usually readily distinguished from the three-proton spin systems, since the random-coil chemical shifts of the $C^{\beta}H_2$ resonances in the latter are in all cases at least 1·0 p.p.m. to lower field than in the long side-chains (Bundi & Wüthrich, 1979). Assignments to specific positions in the amino acid sequence are at this point obtained only when the protein contains single residues of one of the unique types, such as Ile49 in BUSI IIA. However, to obtain individual resonance assignments for the entire polypeptide chain eventually, it is crucial that complete identifications of a considerable number of amino acid side-chain spin systems are obtained independently of the sequential assignments (Wüthrich, 1983), which was the case for nearly all the connectivities indicated in Figures 1 to 3.

(b) *Sequential individual resonance assignments for the polypeptide backbone protons*

A general strategy for obtaining sequential, individual resonance assignments on the basis of stereochemical considerations of polypeptides (Dubs *et al.*, 1979; Billeter *et al.*, 1982) and the use of 2D n.m.r. was recently described in detail (Wagner *et al.*, 1981; Wagner & Wüthrich, 1982*a*; Wider *et al.*, 1982). These assignments rely largely on the identification of sequentially neighboring residues by through-space NOE connectivities. It is therefore essential that high-quality NOESY spectra can be obtained from H_2O solutions of the protein. Figure 4 shows that the NOESY spectrum of BUSI IIA at 45°C contains numerous well-resolved cross peaks in all the different spectral regions. Even better NOESY spectra of BUSI IIA were obtained at 18°C.

FIG. 1. Contour plot of the region from 0 to 5·3 p.p.m. of a 500 MHz ^{1}H SECSY spectrum of a 0·016 M solution of protease inhibitor IIA from bull seminal plasma (BUSI IIA), in ^2H_2O, p^{2}H 4·9, $t = 45$°C. Prior to this experiment, all the labile protons were replaced with deuterons. The digital resolution is 1·9 Hz/point along both the $\Delta\delta$ and δ-axes. The spectrum was recorded in approx. 36 h. Proton–proton J connectivities are shown for the following amino acid side-chains: Ile49 (– – – –), Leu51 (———), Thr31 (———), Ser21 (———), Ser27 (———), Gly26, 29 and 46 (▬▬▬). In order not to overcrowd the Figure, the 2 cross peaks of Ala41 have not been connected by a line. A simplified representation is also used for the AMX spin systems of Asp6, Cys7, Phe10, Asp12, Cys17, Asn22, Asn28, Tyr32, Asn34, Phe38, Cys39, His53 and Cys57, whereby the 2 $C^{\alpha}H$–$C^{\beta}H$ cross peaks at the chemical shift position of the C^{α} proton have been connected by a vertical line, the corresponding cross peaks at the chemical-shift positions of the C^{β} protons are indicated by arrows and the $C^{\beta}H$–$C^{\beta}H'$ cross peaks are not indicated (see Table 2 for an explanation of the connectivity patterns).

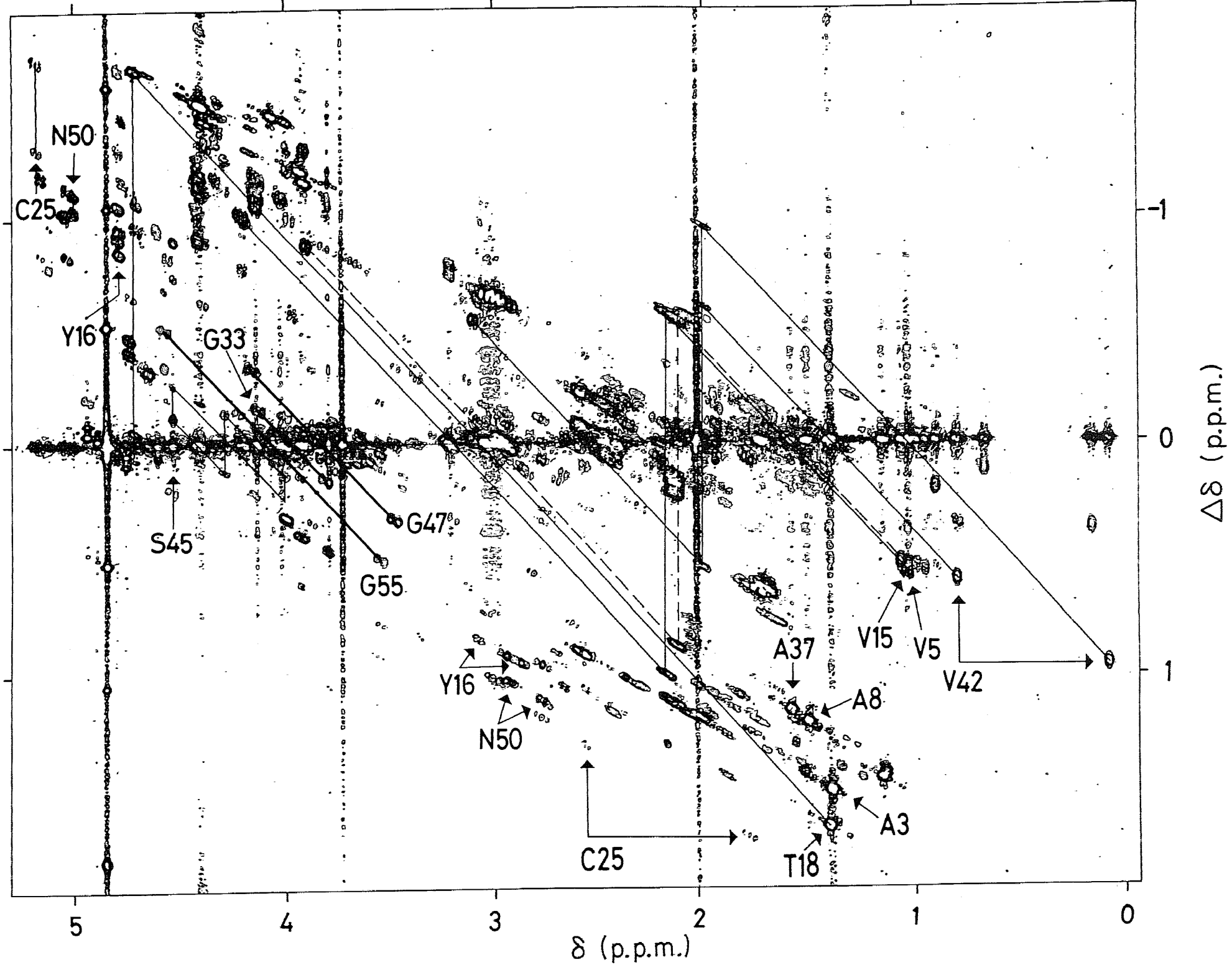

N50
C25
Y16
G33
S45
G47
G55
Y16
N50
C25
A37
V15
V5
V42
A8
A3
T18
5
4
3
2
1
0
δ (p.p.m.)
−1
0
1
Δδ (p.p.m.)

FIG. 2.

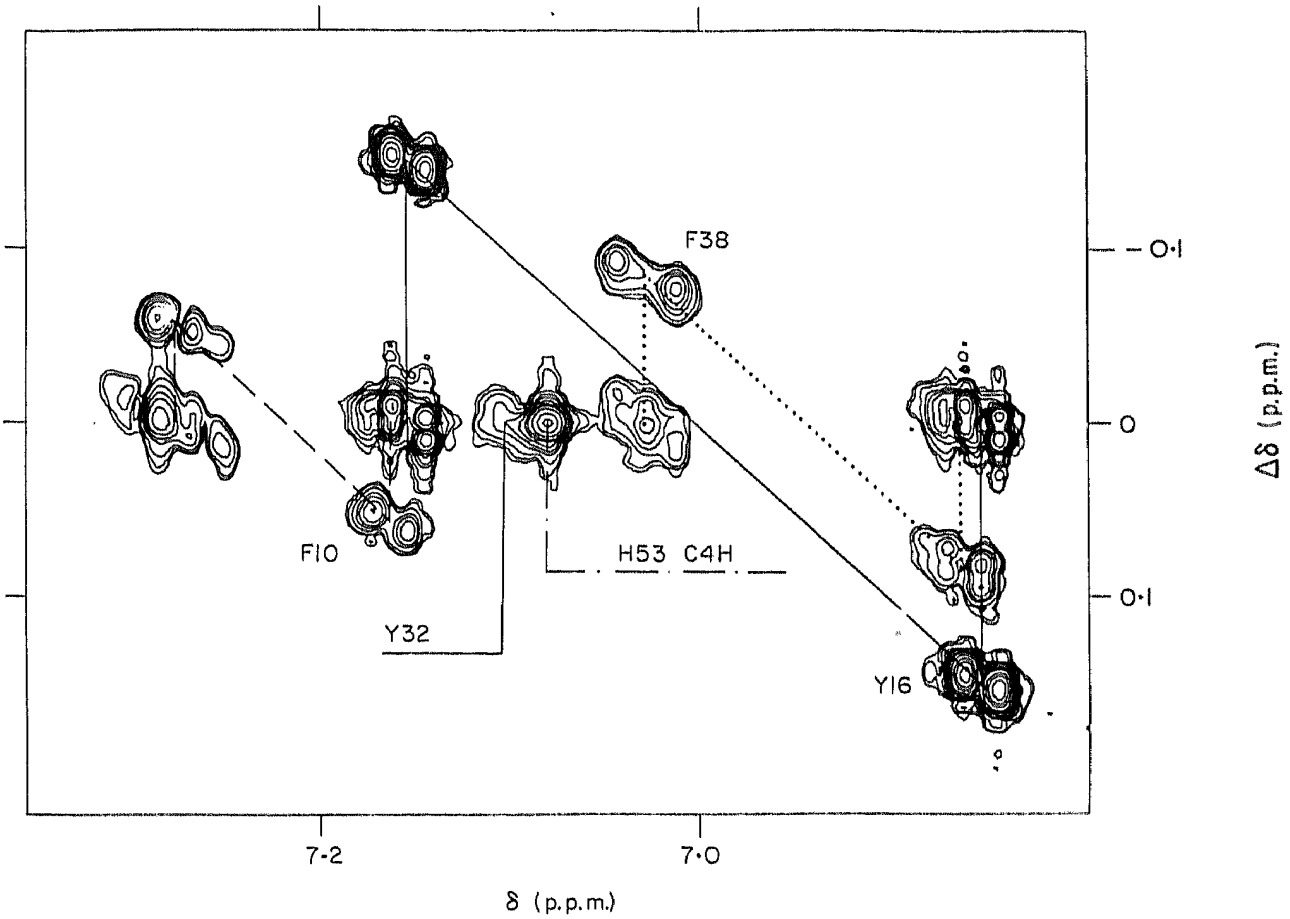


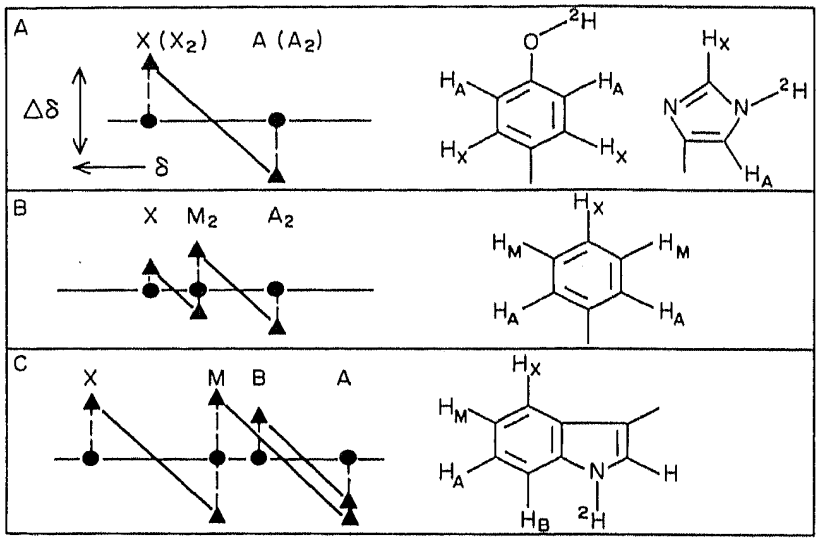
FIG. 3. Contour plot of the region from 6·8 to 7·4 p.p.m. of the same SECSY spectrum as in Fig. 1. The J connectivities are indicated for the following aromatic spin systems: Phe10 (– – –), Tyr16 (———) and Phe38 (· · ·). For Tyr32 only the position of the C2, 6H resonance, which was assigned from the NOE connectivities with $C^{\beta}H_2$ (Fig. 13), is indicated. The absence of J connectivities to the C3, 5 protons is due to the dynamics of this ring (Štrop & Wüthrich, 1983). The position the C4H line of His53, which was determined by a pH titration using 1D n.m.r. (Štrop & Wüthrich, 1983), is also indicated (see Table 2 for an explanation of the connectivity patterns).

The identification of neighboring residues in the primary structure relies on the observation of short proton–proton distances d_1, d_2 and/or d_3 (Fig. 5) in the NOESY spectra. The reliability that short distances between amide protons, $C^{\alpha}H$ and $C^{\beta}H$, of the order of 3·0 Å, are between neighboring residues (as compared to similar connectivities between further distant residues) has been estimated to be 80 to 90%, depending on the connectivities used (Billeter *et al.*, 1982). Typically, stepwise identifications of short polypeptide segments of two to approximately eight amino acids can be obtained by this method. To obtain assignments to

FIG. 2. Contour plot of the region from 0 to 5·3 p.p.m. of a 500 MHz ^{1}H SECSY spectrum of a 0·016 M solution of BUSI IIA in ^2H_2O, at 18°C. Except for the temperature, all the experimental conditions were the same as in Fig. 1. Proton–proton J connectivities are shown for the following amino acid side-chains: Val5 (———), Val15 (– – – –), Val42 (———), Thr18 (———), Gly33, Gly47, Gly55 (▬▬▬) and Ser45 (———). The 2 cross peaks of each of the alanines in positions 3, 8 and 37 have not been connected by lines. The same simplified representation as in Fig. 1 is used for the AMX spin systems of Tyr16, Cys25 and Asn50 (see Table 2 for an explanation of the connectivity patterns).

TABLE 2

SECSY connectivity patterns for the spin systems of the ring protons in the aromatic amino acids



A. Peaks corresponding to the 1D spectrum are on the central axis (●). For His (AX) and Tyr (A_2X_2) a single cross peak (▲) below the central axis at the chemical shift of protons A (δ_A) and a single cross peak above the central axis at δ_X manifest the spin–spin coupling between C2H and C4H in His, and between C2, 6H and C3, 5H in Tyr. The geometric arrangement of the connectivity pattern is defined unambiguously by the requirement that the distance between the central axis and the cross peaks is: $\frac{1}{2}|J\delta| = \frac{1}{2}|\delta_X - \delta_A|$. Because of the small coupling constant the cross peaks linking C2H and C4H of His may be very weak. The symmetric A_2X_2 spin system is for a mobile tyrosine ring and a more complex pattern of 2 separate AX-spin systems for C2H and C3H, and for C5H and C6H, respectively, results when the ring is immobilized in the protein (Wüthrich, 1976).

B. For a mobile phenylalanine ring an A_2M_2X spin system is observed, where the C3, 5H(M) are connected with both the C2, 6 protons and C4H. In real spectra the individual ring protons of Phe often have coinciding chemical shifts, so that incomplete connectivity patterns are observed (Fig. 3).

C. The indole ring of Trp gives rise to an AMPX 4-proton spin system and a 1-proton singlet. In the AMPX system, C5H(M) and C6H(A) have connectivities to the hydrogen atoms of both neighboring ring carbons. (There are no SECSY or COSY connectivities between the singlet resonance of C2H and the AMPX system, but connectivities can be established either by NOEs *via* $C^{\beta}H_2$ observed in NOESY spectra recorded in ^2H_2O or by NOEs *via* N1H in NOESY spectra recorded in H_2O (Wider *et al.*, 1982; Arseniev *et al.*, 1982.)

Gly (AX) and Ala (A_3X) correspond to the connectivity pattern A (see Figs 1 and 2). Connectivity patterns for the aliphatic protons of the other common amino acid residues have been presented by Nagayama & Wüthrich (1981).

specific locations in the amino acid sequence, the chemical shifts of the C^{α} and C^{β} protons in the sequentially identified backbone segments are then matched with corresponding chemical shifts in the side-chain spin systems identified with SECSY (Figs 1 and 2). The amino acid sequences for the sequentially connected polypeptide segments can thus be established in sufficient detail to locate them in the primary structure (Wüthrich, 1983; Wagner & Wüthrich, 1982*a*; Arseniev *et al.*, 1982). For example, in BUSI IIA, assignments to specific locations in the sequence (Fig. 6) resulted when the dipeptide segments Gly–Ala, Ala–Val and Gly–Gly were connected and found to fit into the positions 2–3, 41–42 and 46–47, respectively. Obviously, once the location in the primary structure is known, each

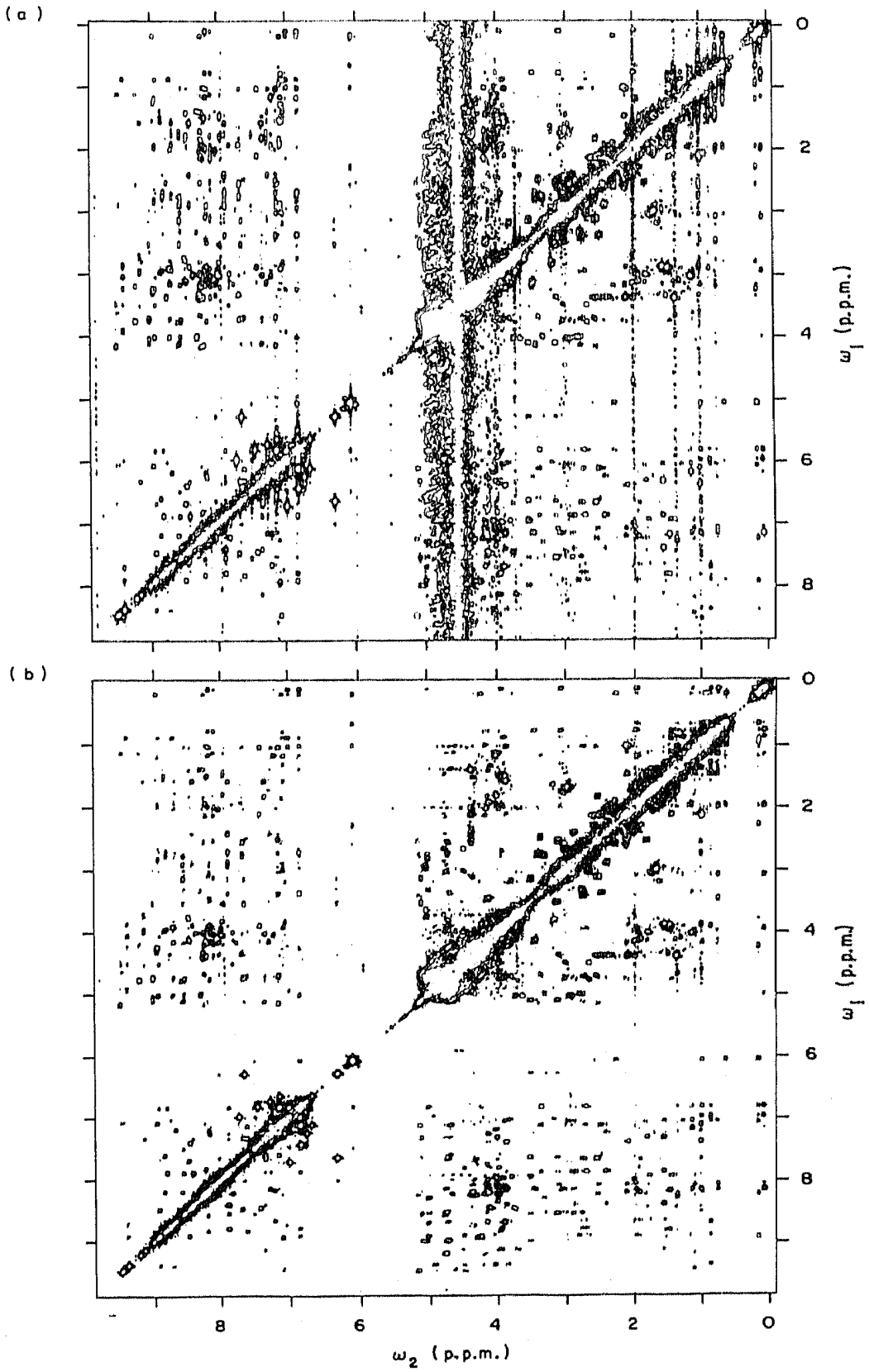


FIG. 4. Contour plot of a 500 MHz ^{1}H NOESY spectrum of a 0·016 M solution of BUSI IIA in H_2O, pH 4·9, $t = 45°C$. The digital resolution is 4·8 Hz/point. The spectrum was recorded in approx. 25 h and a mixing time of 200 ms was used. The strong vertical noise band at 4·6 p.p.m. is at the chemical shift of the water resonance, which was suppressed by selective irradiation as described in Materials and Methods. (a) Unsymmetrized spectrum; (b) spectrum after symmetrization (Baumann *et al.*, 1981).

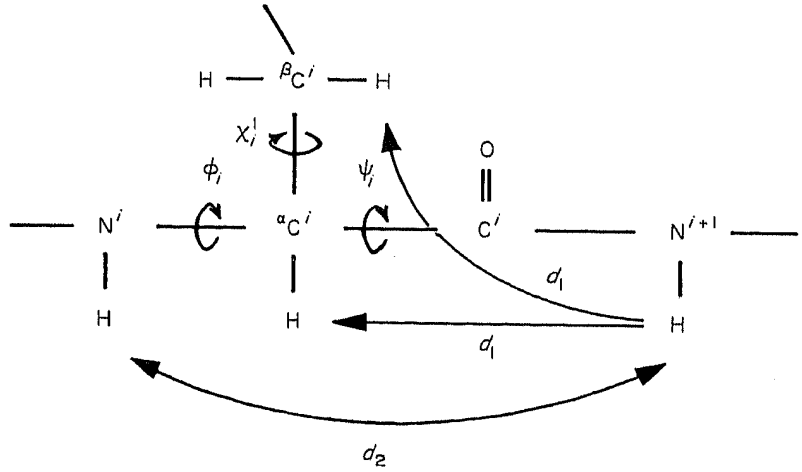


FIG. 5. Polypeptide backbone segment. The through-space distances d_1, d_2 and d_3 used for the sequential resonance assignments are indicated by arrows. The relationships between d_1, d_2 and d_3 and the dihedral angles ψ_i, ϕ_i and χ_i^1 have been discussed previously (Billeter *et al.*, 1982).

successive sequential assignment can be checked against the amino acid sequence. While this helps in most instances to confirm the resonance assignments, evaluation of the n.m.r. observations against the primary structure may also reveal errors in the amino acid sequence, as will be discussed further, below.

Figure 6 presents a survey of the sequential assignments obtained for BUSI IIA

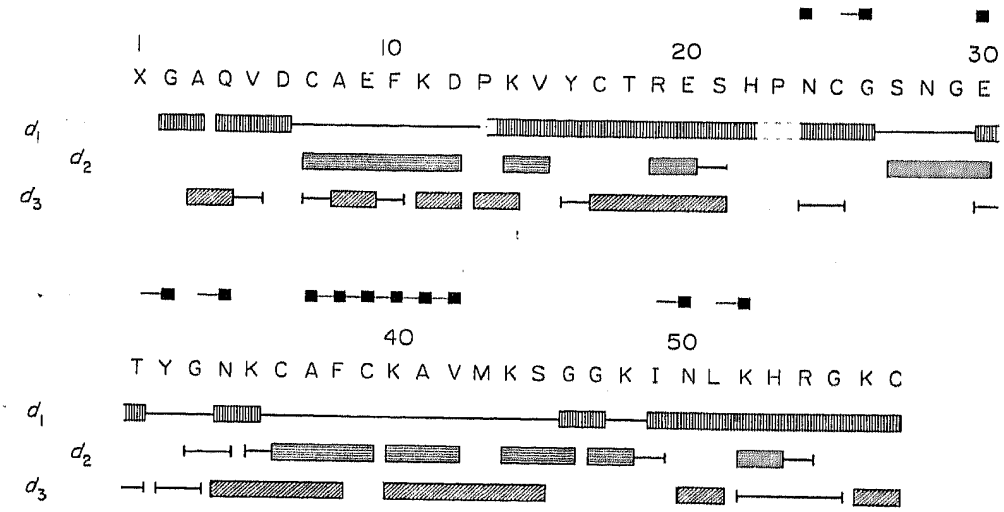


FIG. 6. Amino acid sequence of protease inhibitor IIA from bull seminal plasma (BUSI IIA), where X stands for pyroglutamic acid, and survey of the sequential connectivities by which the individual resonance assignments listed in Table 1 were obtained. (▭) Sequential assignments *via* d_1 (NOE from NH_{i+1} to $C^{\alpha}H_i$); (▭) sequential assignments *via* d_2 (NOE from NH_{i+1} to NH_i); (▭) sequential assignments *via* d_3 (NOE from NH_{i+1} to $C^{\beta}H_i$); (⊏ = = = ⊐) sequential assignments *via* NOEs from proline $C^{\delta}H_{i+1}$ to $C^{\alpha}H_i$. Single lines indicate additional d_1, d_2 or d_3 connectivities, which were manifested by less intense NOEs (see the text). The signs above the sequence indicate the residues for which the amide protons were also observed in ^2H_2O solution (■) and for which the sequential connectivities were obtained in ^2H_2O solution (—■). The 2D spectra in ^2H_2O were recorded from 0·5 to 24 h after the solution was prepared at 25°C and p^{2}H 4·9. All the connectivities shown here could be observed at 18°C, and with few exceptions they were also manifested at 45°C. Compared to the previously published sequence (Meloun & Čechová, 1979), the order of residues 56 and 57 was inverted on the basis of evidence obtained from the sequential assignments (see also Note added in proof).

at 18°C. It is seen that nearly complete assignments were obtained and that many sequential connectivities are based on observation of two connectivities, mostly d_1 and d_3, or d_2 and d_3. In NOESY spectra recorded with a mixing time of 100 ms weak d_1 connectivities were observed between all residues except Pyr1 and Gly2, and Ala3 and Gln4. While this is not *a priori* unexpected, since $d_1(\psi)$ (Fig. 5) varies only between 2·2 and 3·6 Å (Leach *et al.*, 1977; Kuo & Gibbons, 1979; Billeter *et al.*, 1982), it is in contrast to the experience gained previously with proteins of similar size (Wagner & Wüthrich, 1982*a*; Arseniev *et al.*, 1982; Keller *et al.*, 1982), and is probably a consequence of the superior quality of the NOESY spectra obtained with BUSI IIA (Figs 4 and 7 to 10). For each pair of sequentially neighboring residues one of the strong connectivities is documented, whereby Figures 7 to 9 show d_1 connectivities and Figure 10 shows d_2 connectivities. To demonstrate the quality of the spectra obtained at both temperatures used, some of these illustrations are at 18°C and others at 45°C. Detailed explanations of how the sequential connectivities can be found in the 2D n.m.r. spectra are given in the legends to Figures 7 and 10 (see also Wagner *et al.*, 1981; Wagner & Wüthrich, 1982*a*).

Once the sequential assignments are completed, they must be compatible with the "n.m.r. fingerprint" of the protein in the $C^{\alpha}H$–NH region of the COSY spectrum (Wagner & Wüthrich, 1982*a*). The fingerprint contains one peak for each amino acid residue except for the prolines, which give no peak, and the glycines, which may give two peaks. The BUSI IIA fingerprints in Figure 11 contain peaks corresponding to all the residues in the protein, but there are also additional, mostly weak, peaks in this spectral region, which have not been assigned. This will be followed up further after the determination of additional assignments for the amino acid side-chains.

(c) *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 one knows to which type of amino acid residue each $C^{\alpha}H$–$C^{\beta}H_n$ fragment belongs. With this additional information, the assignments of some of the more complex side-chain spin systems could be completed by renewed inspection of the SECSY and COSY spectra (Table 1). As an illustration, Figure 12 documents the assignment of the spin system of Pro23.

Individual assignments of the previously identified (Fig. 3 of Štrop & Wüthrich, 1983) 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 observed in ^2H_2O solution of the protein (Fig. 13). At 45°C such NOEs were observed for all aromatic residues, and thus no ambiguities remained. It is worth noting that this is our first observation of NOE connectivities between $C^{\beta}H_2$ and the imidazole ring protons of His. Stereochemical considerations indicated that this connectivity should be more evasive than the corresponding NOEs in Phe, Tyr and Trp (Billeter *et al.*, 1982).

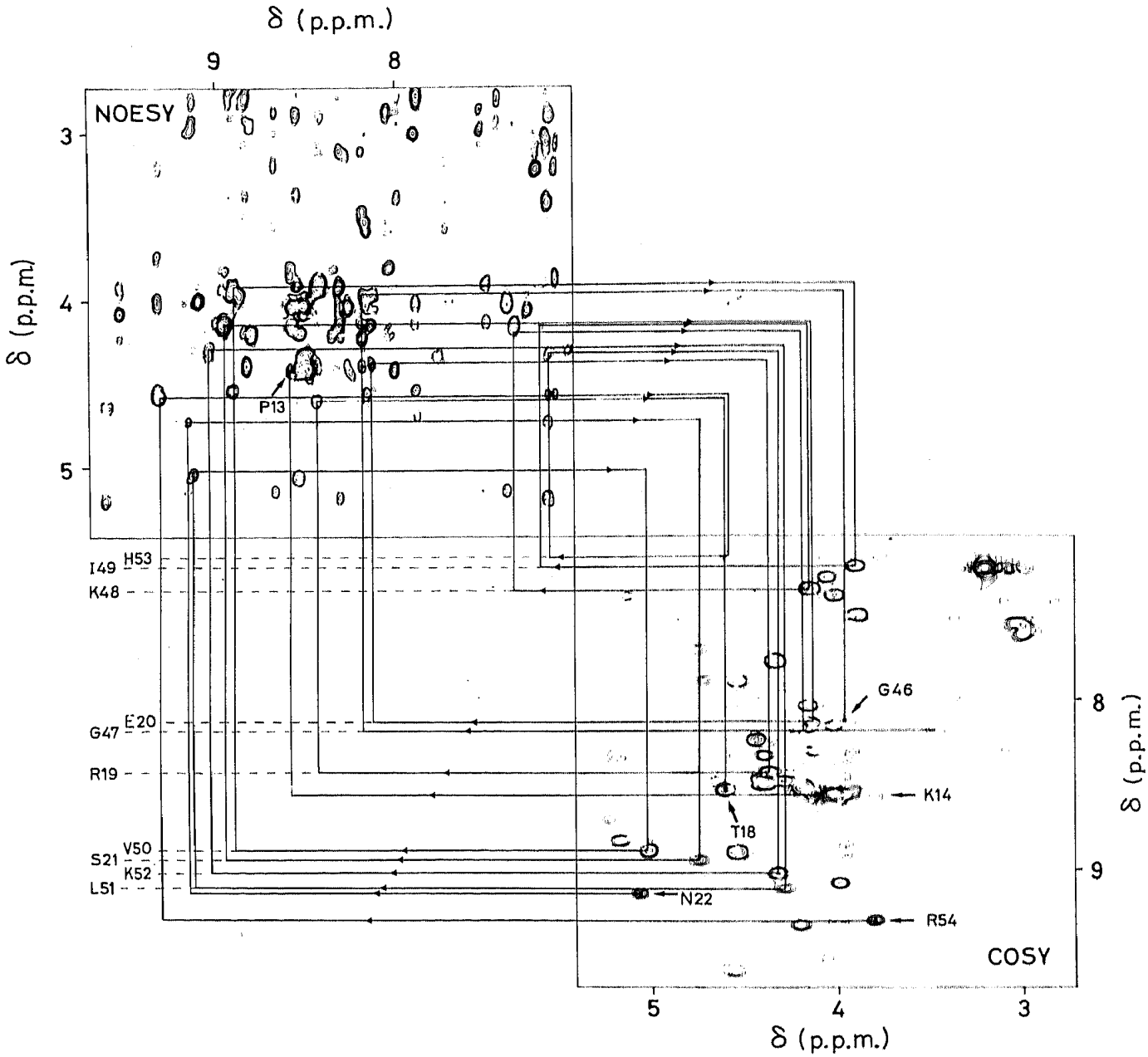


FIG. 7. Combined COSY–NOESY connectivity diagram for sequential resonance assignments *via* NOEs between amide protons and the C^{α} protons of the preceding residues (d_1) (Wagner *et al.*, 1981). In the upper left, the region ($\omega_1 = 2{\cdot}7$ to $5{\cdot}4$ p.p.m., $\omega_2 = 7{\cdot}0$ to $9{\cdot}7$ p.p.m.) of a ^{1}H NOESY spectrum of BUSI IIA recorded with a mixing time of 200 ms is presented. In the lower right, the region ($\omega_1 = 7{\cdot}0$ to $9{\cdot}7$ p.p.m., $\omega_2 = 2{\cdot}7$ to $5{\cdot}4$ p.p.m.) of a ^{1}H COSY spectrum recorded from the same sample under identical conditions, i.e. $t = 18°C$ and pH = 4·9, is shown. 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. 5). When one follows the polypeptide chain in the direction from the C to the N terminus, the lines that link successive COSY and NOESY connectivities describe a clockwise spiral (Wagner *et al.*, 1981; Wagner & Wüthrich, 1982*a*). For example, the connectivities from residues 22 to 18 were obtained from the following considerations. Starting from a (at that stage not yet assigned) $C^{\alpha}H$–NH cross peak in COSY a horizontal line leads to the chemical-shift position of the amide proton on the virtual diagonal, which substitutes for the actual diagonal peaks seen, e.g. in Fig. 4. If present, the d_1 connectivity to $C^{\alpha}H$ of the preceding residue (Fig. 5) must be manifested by a NOESY cross peak located on the vertical line through the diagonal position of NH. Such a connectivity between residues 22 and 21 is indeed observed, and from this cross peak a horizontal line leads to the chemical shift position of $C^{\alpha}H$ of Ser21

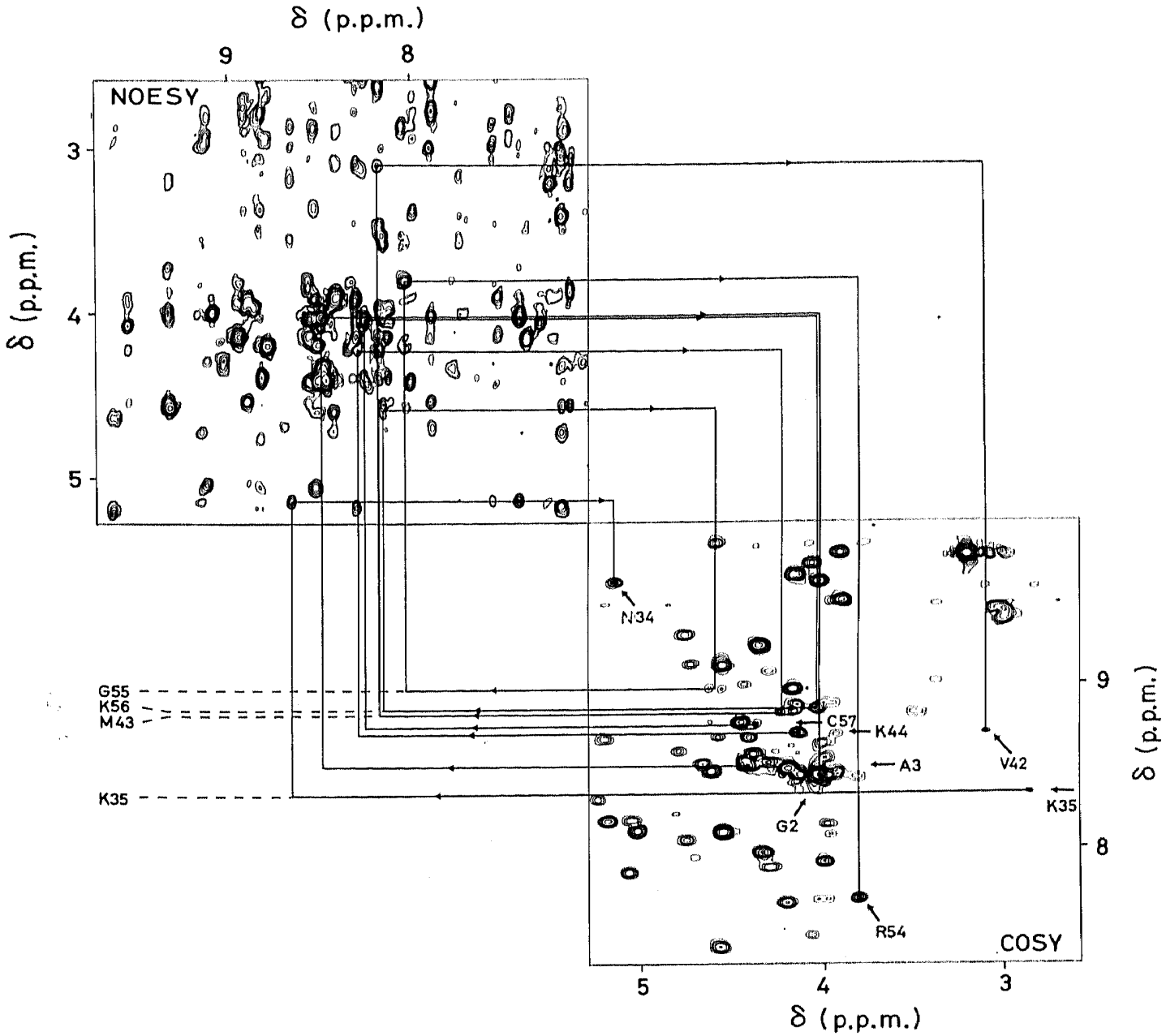


FIG. 8. Combined COSY–NOESY d_1 connectivity diagram, same spectra and same presentation as in Fig. 7. The connectivity patterns are indicated by unbroken lines for the following segments of the BUSI IIA polypeptide chain: 57 to 54, 44 to 42, 35 to 34 and 3 to 2.

---

on the diagonal. From there a vertical line connects to the C$^\alpha$H–NH cross peak of Ser21 in COSY. The search for the connectivity with C$^\alpha$H of Glu20 then starts again with a horizontal line to the diagonal location of the amide proton of Ser21. The connectivities thus found between neighboring residues in the segments 54 to 46, 22 to 18 and 14 to 13 are indicated by solid horizontal and vertical lines. For each segment, the start and the end of the connectivity pattern are indicated by (●←—), and by (●) and identification of the last residue, respectively. To identify the connected cross peaks, assignments are indicated in the lower left at the amide proton chemical-shift positions. For the segment 14 to 13 the terminal connectivity is to C$^\alpha$H of proline in the NOESY spectrum.

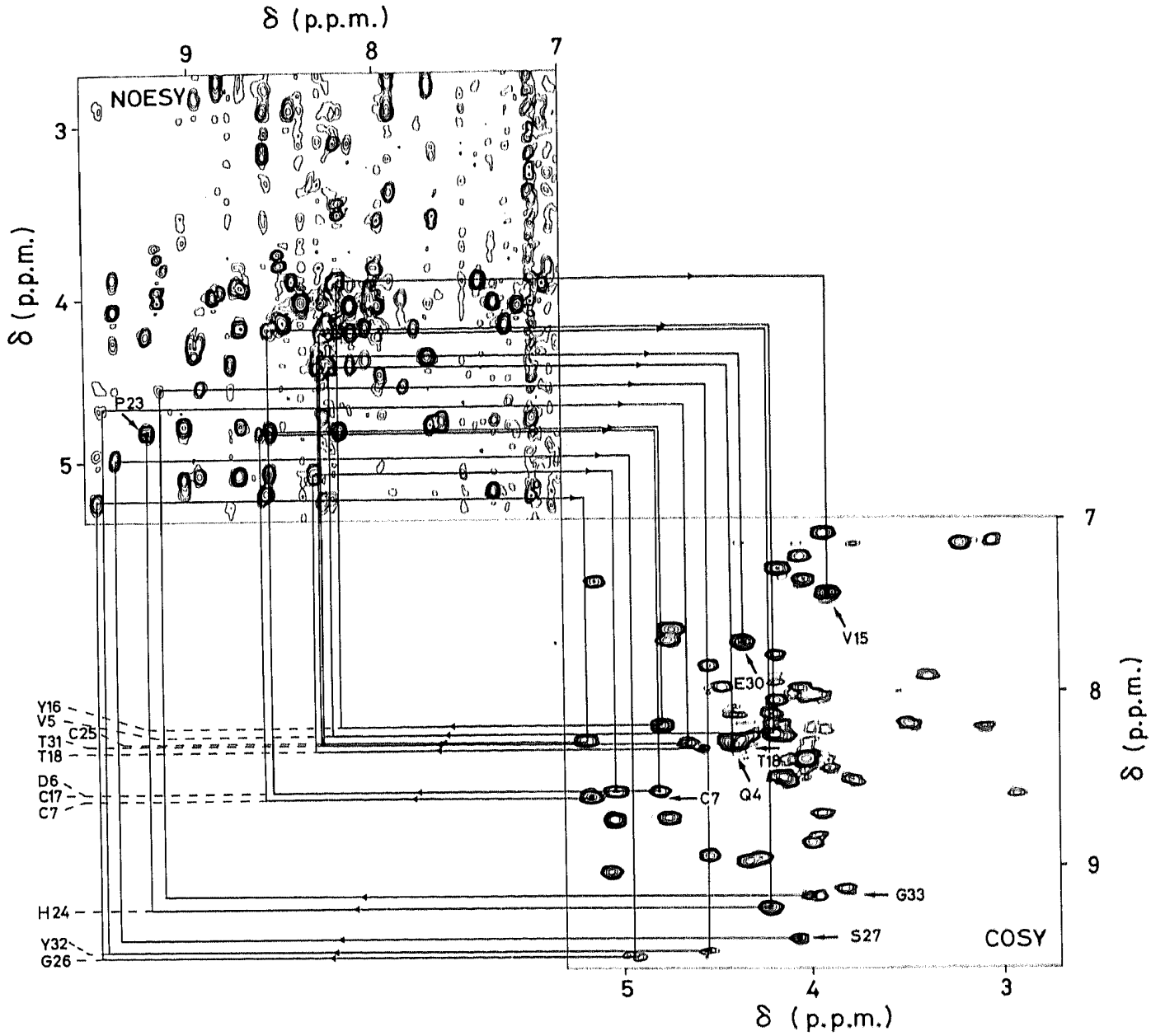


FIG. 9. Combined COSY–NOESY d_1 connectivity diagram, same presentation as in Fig. 7. In the upper left, the region ($\omega_1 = 2{\cdot}7$ to $5{\cdot}3$ p.p.m., $\omega_2 = 7{\cdot}0$ to $9{\cdot}6$ p.p.m.) of a ^{1}H NOESY spectrum of BUSI IIA recorded with a mixing time of 100 ms, and at $t = 45$°C and pH 4·9, is presented. In the lower right, the region ($\omega_1 = 7{\cdot}0$ to $9{\cdot}6$ p.p.m., $\omega_2 = 2{\cdot}7$ to $5{\cdot}3$ p.p.m.) of a ^{1}H COSY spectrum recorded from the same sample, under identical conditions, is shown. The connectivities between neighboring residues in the segments 7 to 4, 18 to 15, 27 to 23 and 33 to 30 are indicated.

(d) *Effects of partial hydrolysis of Asn and Gln side-chains during the n.m.r. experiments*

An initial count of the cross peaks in the COSY fingerprint of BUSI IIA (Fig. 11) showed that there was a larger number of cross peaks than one would have expected from the amino acid sequence. Once the resonance assignments were completed it was quite obvious that these extra peaks could not be

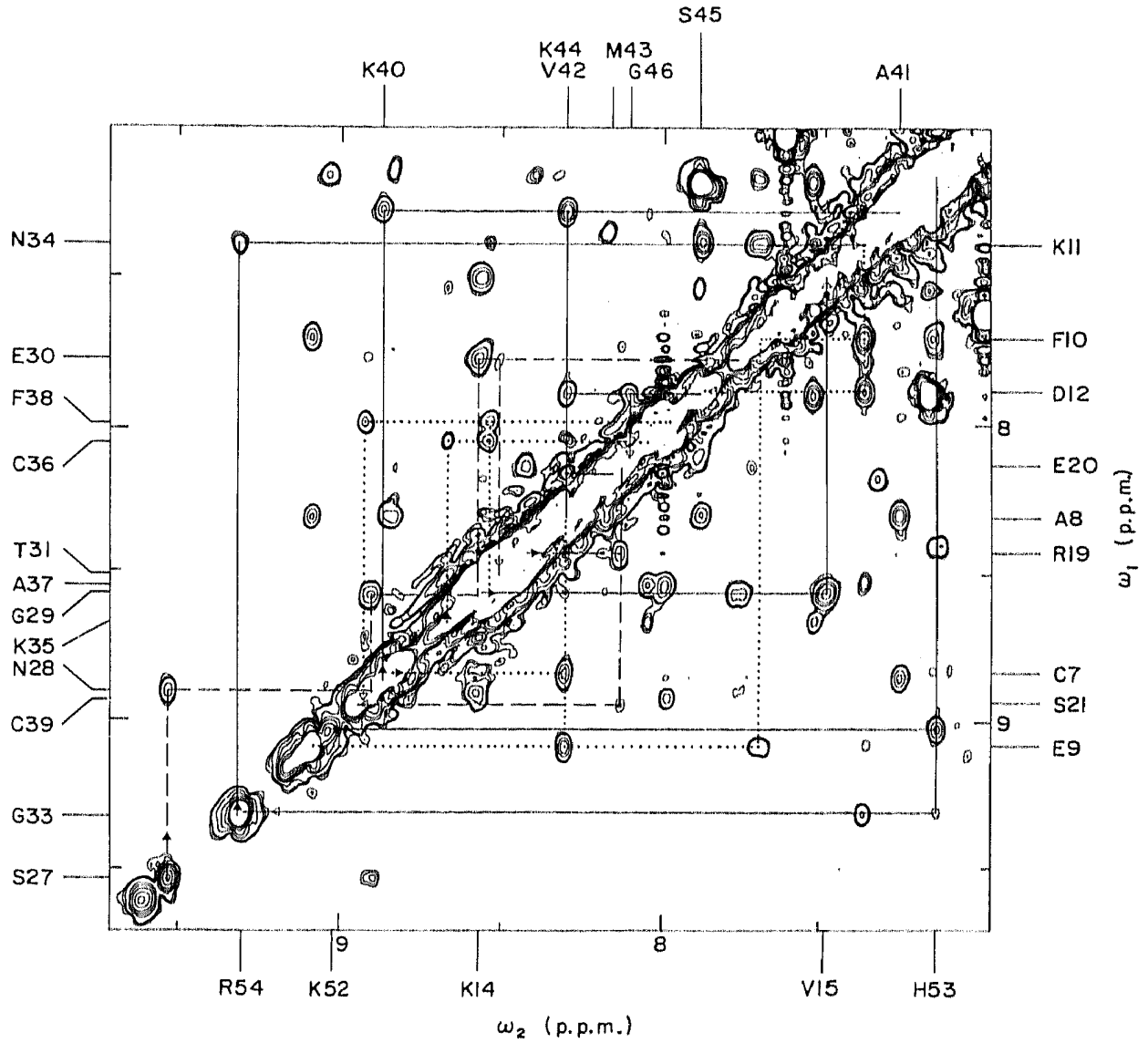


FIG. 10. Identification of sequentially neighboring residues by d_2 connectivities in the spectral region from 7·0 to 9·7 p.p.m. of a 500 MHz ^{1}H NOESY spectrum, which was recorded from a 0·016 M solution of BUSI IIA in H_2O at pH 4·9 and $t = 18°C$, with a mixing time of 200 ms and a digital resolution of 4·8 Hz/point. This spectral region contains the diagonal peaks of the bulk of the backbone amide protons, and NOEs between different amide protons are manifested by numerous cross peaks. d_2 connectivities between the amide protons of neighboring residues are indicated by 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 27 to 31 (– – – –), 33 to 34 (———), 35 to 39 (·····) (positions of the diagonal peaks indicated on the left) and 40 to 46 (———) (diagonal peak positions indicated at the top of the Figure; note that the cross peaks 42 → 43 and 43 → 44 are overlapped). The lower right triangle contains the connectivities for the segments 7 to 12 (·····) and 19 to 21 (– – – –; assignments on the right), and 14 to 15 and 52 to 54 (———; assignments at the bottom). For each segment (▶) and (▷), respectively, indicate the start and the end of the d_2 connectivity pattern. When searching for these patterns it is important that d_2 connectivities are symmetrical with respect to the direction of the polypeptide chain (Billeter *et al.*, 1982) and that a NH–NH NOE connectivity may thus equally well be observed with either of the 2 neighboring residues in the sequence (Fig. 5). For example, the amide proton of Asn28 in BUSI IIA has d_2 connectivities to both Ser27 and Gly29. Since the NH resonance of Ser27 is at lower field than that of Asn28, the connecting cross peak is towards the left side on a horizontal line through the diagonal peak of NH of Asn28. Gly29 NH, however, is at higher field and therefore the d_2 cross peak is observed on a vertical line through the diagonal peak of the Asn28 amide proton.

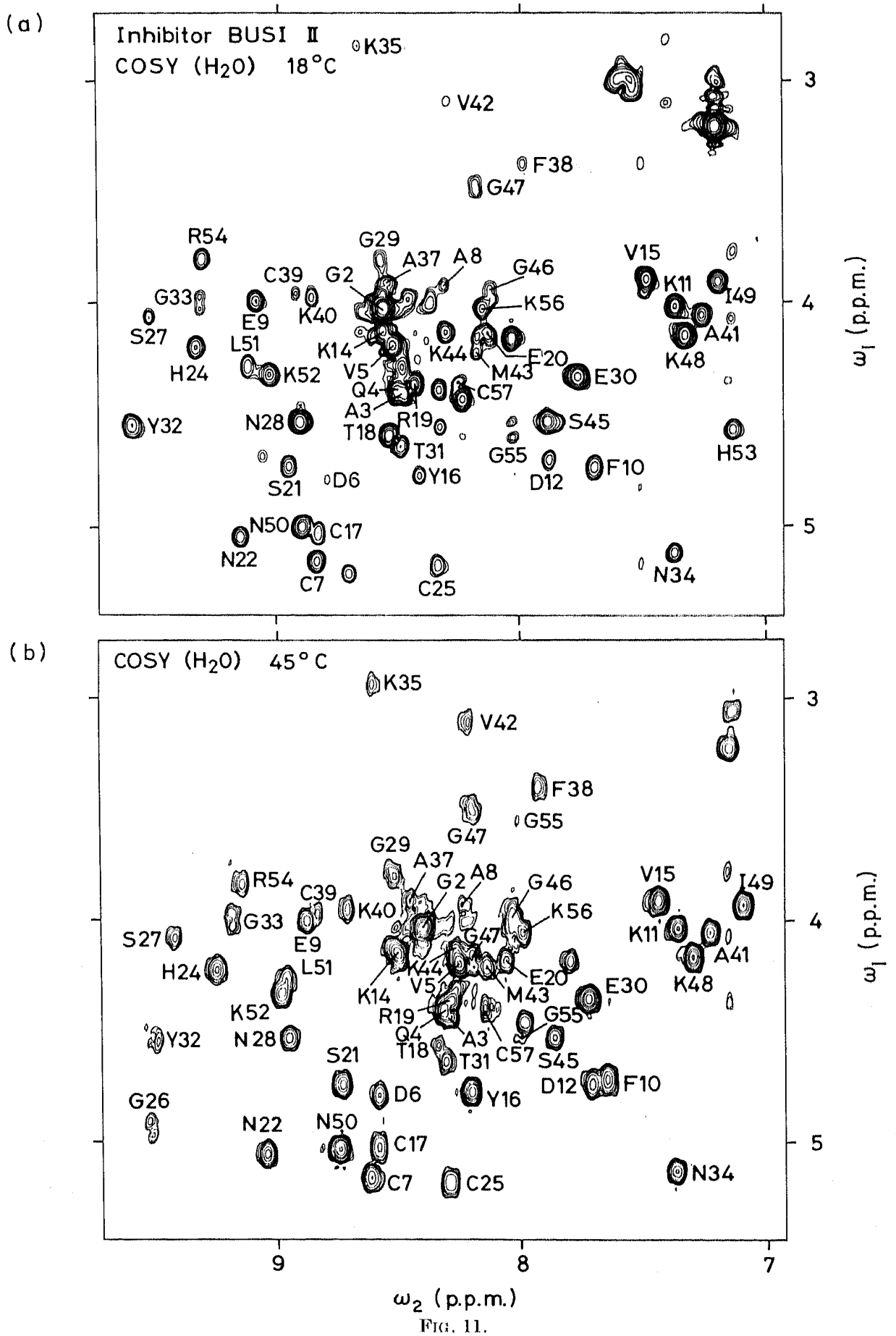

(a)
Inhibitor BUSI II
COSY (H_2O) 18°C
K35
V42
F38
G47
R54
G29
A37
A8
G46
V15
C39
G2
K11
I49
G33
E9
K40
K56
S27
L51
K14
K44
A41
H24
K52
V5
M43
E20
K48
Q4
C57
E30
A3
Y32
N28
R19
S45
T18
T31
G55
H53
S21
D6
Y16
D12
F10
N50
C17
N22
C7
C25
N34
(b)
COSY (H_2O) 45° C
K35
V42
F38
G55
G47
G29
A37
R54
G2
A8
G46
V15
I49
C39
K40
K56
G33
S27
E9
G47
K11
A41
H24
L51
K44
E20
K48
K14
V5
M43
E30
K52
R19
G55
Y32
N28
Q4
A3
S21
T18
T31
C57
S45
D12
F10
G26
D6
Y16
N22
N50
C17
C7
C25
N34
3
4
5
9
8
7
ω_1 (p.p.m.)
ω_2 (p.p.m.)

FIG. 11.

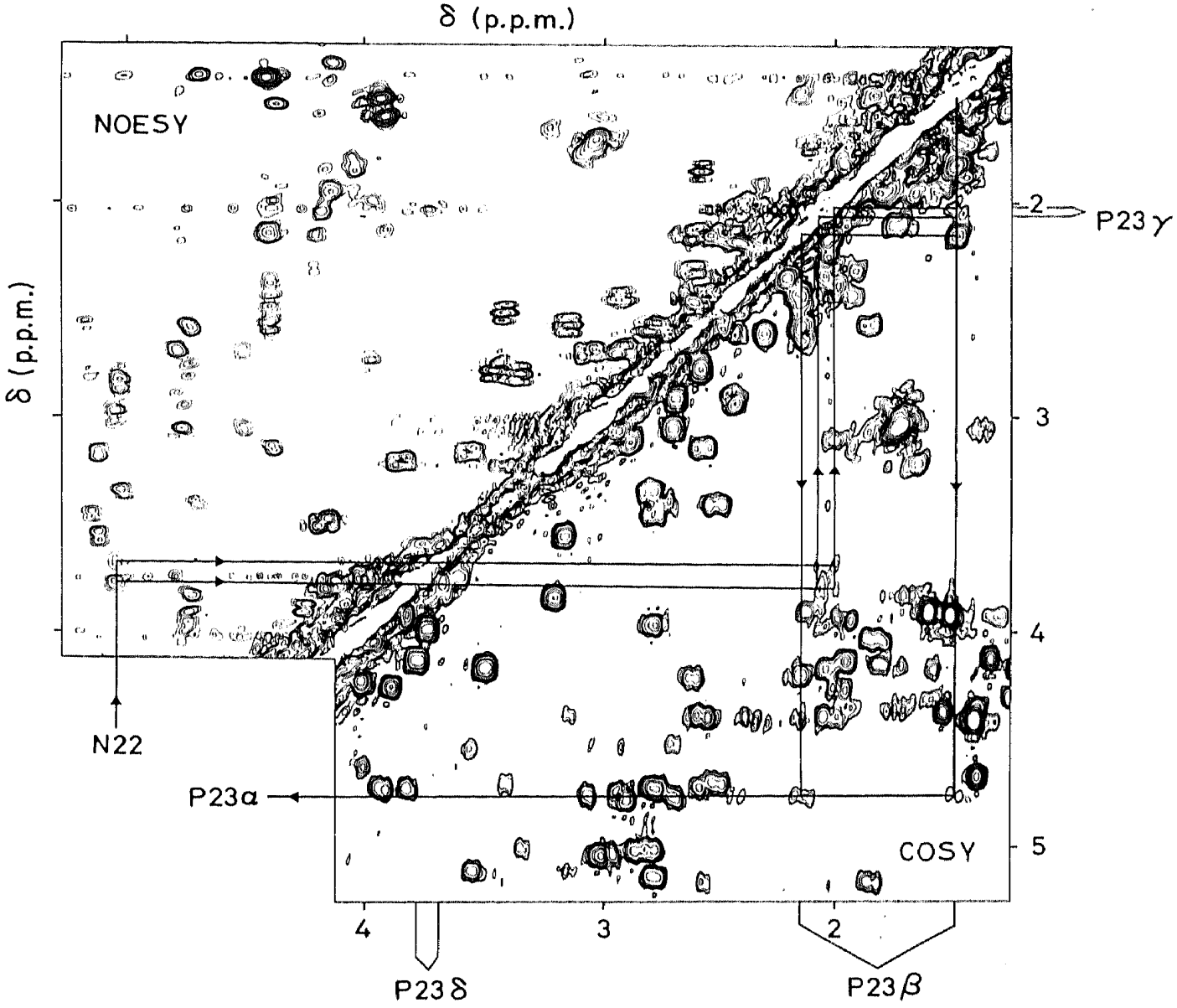


FIG. 12. Resonance assignments for the peptide segment Asn22 to Pro23. A combined plot is shown, where the upper left triangle comes from the NOESY spectrum of BUSI IIA in Fig. 4(b) and the lower right triangle from a COSY spectrum recorded under identical conditions, i.e. pH 4·9 and $t = 45$°C. The connectivities in the segment Asn22–Pro23 start from the previously assigned $C^{\beta}H_2$–$C^{\alpha}H$ resonances of Asn22 (Figs 1 and 6), from which the NOE connectivities to $C^{\delta}H_2$ of Pro23 were obtained. Within the proline spin system the connectivities could be established *via* J couplings in COSY. The NOE from the amide proton of His24 to $C^{\alpha}H$ of Pro23 is shown in Fig. 9. The chemical shifts for the resonances of Pro23 are indicated on the margins of the Figure.

FIG. 11. Spectral region ($\omega_1 = 2{\cdot}6$ to $5{\cdot}4$ p.p.m., $\omega_2 = 6{\cdot}9$ to $9{\cdot}7$ p.p.m.) of a 500 MHz ^{1}H COSY spectrum recorded in a 0·016 M solution of BUSI IIA in H_2O (pH 4·9) at (a) 18°C and (b) 45°C. The digital resolution is 4·8 Hz/point; the spectra were recorded in approx. 22 h. 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. The unassigned peaks in the upper right correspond to arginine and lysine side-chains (Bundi & Wüthrich, 1979). The high-field-shifted cross peak of Cys36 (see Table 1 and Fig. 15) is outside of the spectral region shown in this Figure. Because of the water irradiation the following peaks are bleached out (Anil Kumar *et al.*, 1980*b*; Wider *et al.*, 1983). At 18°C Gly26 and Asp6 are very weak; at 45°C His53 is very weak.

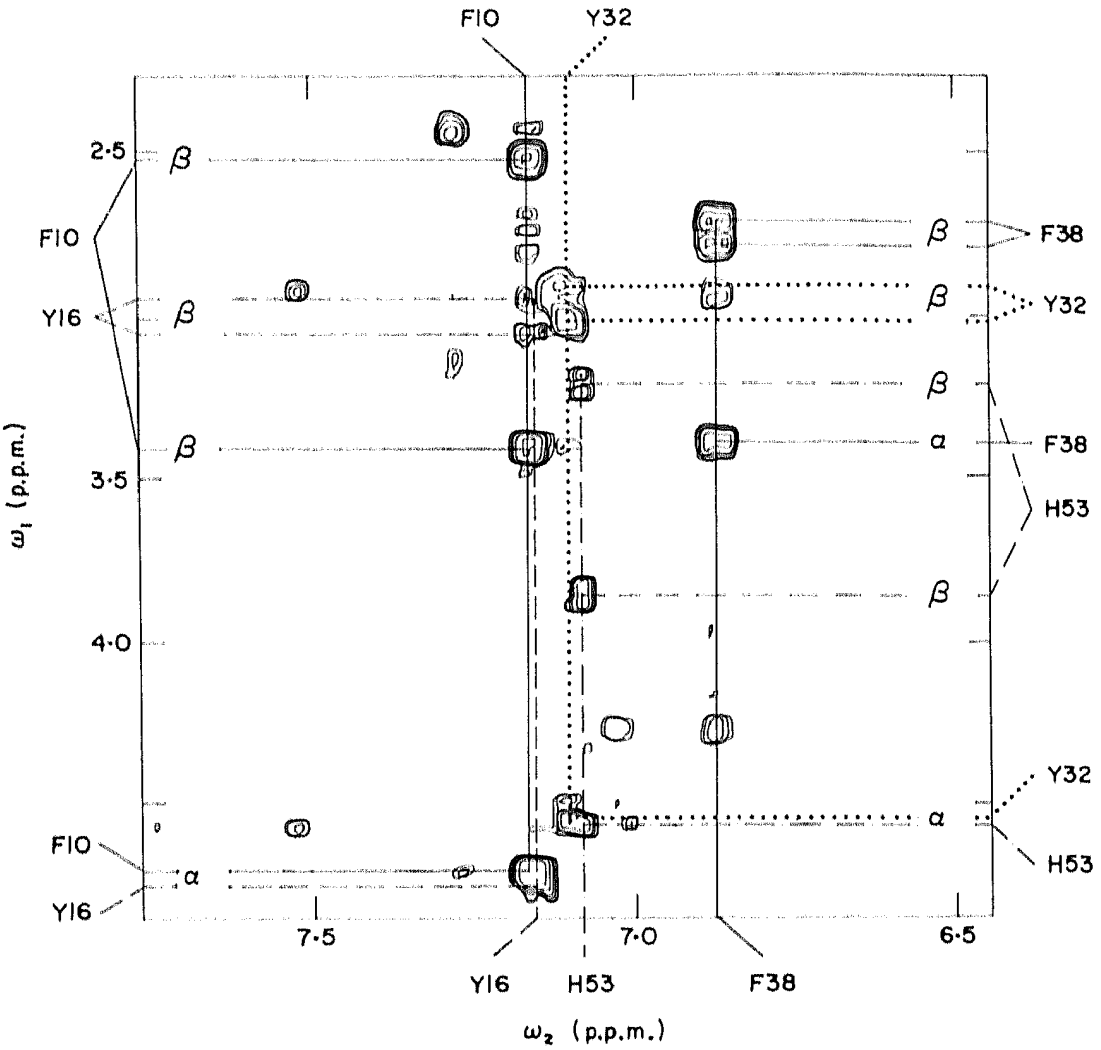


FIG. 13. Spectral region (ω_1 = 2·3 to 4·8 p.p.m., ω_2 = 6·4 to 7·8 p.p.m.) of a 500 MHz ^{1}H NOESY spectrum recorded in a 0·016 M solution of BUSI IIA in ^2H_2O, p^{2}H 4·9 at 45°C. The mixing time was 200 ms, the digital resolution is 3·7 Hz/point and the spectrum was recorded in approx. 25 h. The Figure illustrates how the connectivities between $C^\beta H_2$ and the rings were obtained for the aromatic side-chains of Phe10, Tyr16, Tyr32, Phe38 and His53. The chemical shifts of the ring protons nearest to $C^\beta H_2$ are indicated at the top and the bottom of the Figure, those for $C^\alpha H$ and $C^\beta H_2$ on the right and left margins.

attributed to the major protein species in solution. Similarly, unassigned extra peaks remained in the SECSY spectra of the aliphatic side-chains (Figs 1 and 2).

2D n.m.r. spectra of BUSI IIA were recorded from the same samples over a period of approximately 12 months. When the spectra recorded early during this period were compared with those obtained later, we found that the number of extra peaks and their intensities increased with time after the protein was dissolved in water. Subsequently, several of the old n.m.r. samples were pooled and the resulting batch of approximately 100 mg BUSI IIA was repurified on a 39 cm × 27 cm column of CM-Sephadex-C25 equilibrated with 0·1 M-Tris·HCl buffer (pH 7·4) (Čechová *et al.*, 1979). Elution was performed by the linear salt gradient formed using 600 ml of the starting buffer and 600 ml of 0·1 M-Tris·HCl

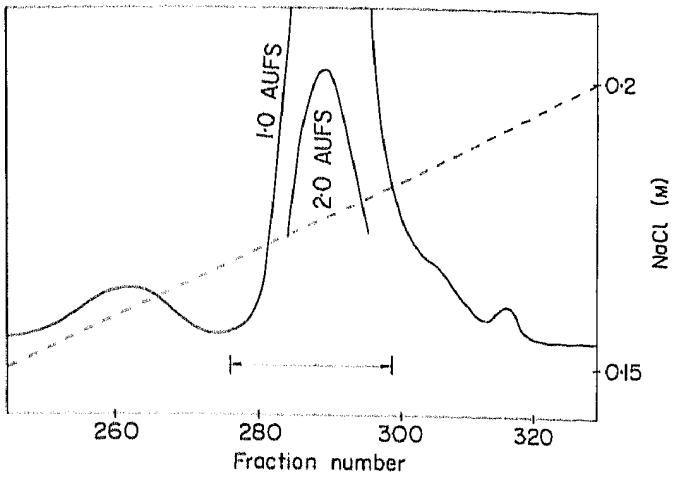


FIG. 14. Rechromatography of isoinhibitor BUSI IIA after prolonged exposure of the solution to 45°C during the 2D n.m.r. experiments. The unbroken curve indicates the absorbance at 280 nm (1·0 AUFS and 2·0 AUFS, respectively, as indicated in the different regions of the plot). The scale on the right indicates the NaCl concentration; the broken line the linear salt gradient used for the elution. The arrow extends over those fractions that were pooled to prepare the sample for the 2D n.m.r. experiment in Fig. 15.

buffer (pH 7·4) with 0·25 M-NaCl. The flow-rate was 26 ml per hour, the size of the fractions 2·9 ml. The elution profile for fractions 245 to 330 (Fig. 14) shows that at least three modifications of BUSI IIA could be separated from the native protein. Fractions 277 to 299 (indicated by the arrow) were pooled, dialyzed against H_2O, concentrated by ultrafiltration and studied by n.m.r. The COSY fingerprints of this sample contained all the cross peaks that had been assigned to BUSI IIA, but none of the "extra peaks" was observed (Fig. 15). The sample was then lyophilized, dissolved in ^2H_2O, heated to 50°C for one hour to exchange all the labile protons, lyophilized again and redissolved in ^2H_2O. Most of the extra SECSY cross peaks in the aliphatic region were not present in the rechromatographed sample and a few peaks persisted with only very small intensity. This clearly shows that the resonances listed in Table I correspond to native BUSI IIA and that in the course of the investigation modified proteins had accumulated in the n.m.r. samples.

(e) *Verification of the amino acid sequence by sequential resonance assignments*

The amino acid sequence of BUSI IIA had been found to terminate with the fragment –Gly55–Cys56–Lys57–OH (Meloun & Čechová, 1979). The ^{1}H n.m.r. assignments in this region (Figs 1, 6 and 8), however, indicated that there was a long side-chain in position 56 and that residue 57 had a $C^{\alpha}H$–$C^{\beta}H_2$ three-proton spin system. Since there was no other corresponding discrepancy between the published sequence and the n.m.r. assignments, the most likely explanation, which would also be compatible with the amino acid composition (Čechová *et al.*, 1979), was that the terminal tripeptide is actually –Gly55–Lys56–Cys57–OH. This was subsequently confirmed by chemical methods, as described in the Appendix.

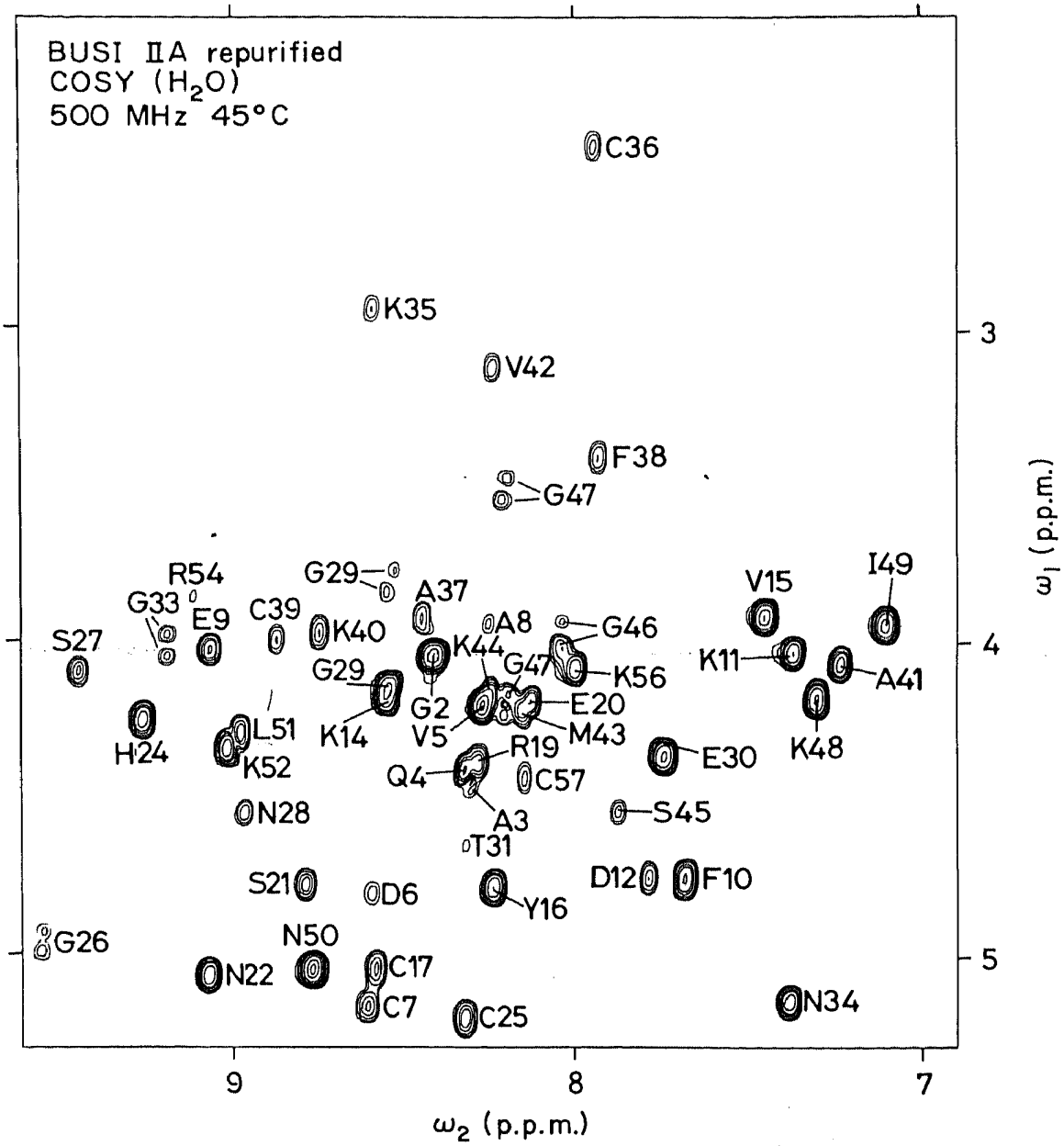


FIG. 15. Spectral region ($\omega_1 = 2{\cdot}0$ to $5{\cdot}3$ p.p.m., $\omega_2 = 6{\cdot}9$ to $9{\cdot}7$ p.p.m.) of a 500 MHz ^{1}H COSY spectrum recorded in a 0·012 M solution of the repurified isoinhibitor BUSI IIA (Fig. 14) in H_2O, pH 5·5 at 45°C. The digital resolution is 4·8 Hz/point and the spectrum was recorded in approx. 17 h. 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. The cross peaks of Thr18, Tyr32, His53 and Gly55 are missing and the Thr31 peak is very weak in this spectrum, since they were bleached out by the water irradiation (Anil Kumar *et al.*, 1980*b*).

4. Discussion

The general conditions for the elucidation of the resonance assignments in BUSI IIA were rather favorable. Large amounts of the protein were available, so that good quality spectra could be obtained for different conditions. Major difficulties arose at certain stages of the project because of the presence of

modified protein in the n.m.r. samples (Figs 14 and 15, and the text). The treatment of the protein during n.m.r. investigations of the type discussed here involves, for all the samples, prolonged standing in preferably slightly acidic, aqueous solution; and for the ^2H_2O solutions, further repeated heating to higher temperatures to exchange the labile protons, followed by lyophilization to minimize the residual solvent protons. These conditions favor modifications of the covalent polypeptide structure, such as hydrolysis of the side-chain amide groups of Asn and Gln. This was previously also observed in solutions of the basic pancreatic trypsin inhibitor, when these were studied over periods of several months for measurements of the amide proton exchange rates (Richarz *et al.*, 1979). This and the present experience with BUSI IIA indicate that much care will have to be used when handling certain proteins, and the n.m.r. recordings should preferably be made with freshly purified material.

The extent to which individual assignments for BUSI IIA were obtained (Fig. 6, Table 1) coincides closely with the results for other proteins of similar size (Wagner & Wüthrich, 1982*a*; Arseniev *et al.*, 1982; Keller *et al.*, 1983). As in these other proteins, the assignments are complete for the entire polypeptide chain, with the exception of most of the long side-chains, where the assignments extend only to the C^β or possibly the C^γ hydrogen atoms (Table 1).

From the fundamental considerations that underlie the methodology of sequential resonance assignments (Billeter *et al.*, 1982; Wagner & Wüthrich, 1982*a*; Wüthrich, 1983) it is readily apparent that the same procedures could in principle, with certain limitations, also be used to obtain information on the primary structure. BUSI IIA presents the first example in which the sequential assignments lead also to a modification of an amino acid sequence that had been previously determined by conventional chemical techniques (see also Note added in proof).

Much further work, based on the use of the resonance assignments described here, will be needed to obtain data on the spatial structure of BUSI IIA (Wüthrich *et al.*, 1982). However, from the patterns of d_1, d_2 and d_3 connectivities (Fig. 6) some preliminary conclusions on the secondary structure can be drawn (Dubs *et al.*, 1979; Leach *et al.*, 1977; Kuo & Gibbons, 1979; Billeter *et al.*, 1982; Wagner *et al.*, 1981; Wagner & Wüthrich, 1982*a*). Fundamentally, stretches of d_1 connectivity are indicative of extended polypeptide chain, such as that found in β structures, d_2 connectivities extending over four or more residues are indicative of α-helices, and short stretches of d_2 connectivity are expected in tight turns. The connectivity data should be combined with observations on the exchange of the backbone amide protons, which are indicated in Figure 6. For example, slow exchange for a continuous stretch of residues is indicative of a central strand in a β structure or an α-helix, and slow exchange for each second residue is typical for peripheral strands in β structures (Wagner *et al.*, 1981; Wagner & Wüthrich, 1982*b*). The data in Figure 6 thus indicate that the long stretch of extended structure from residues 14 to 26 ends as part of a β structure near the residues 23 to 26. β structure is also indicated in the regions of residues 30 to 35 and 49 to 53. Evidence for the formation of an α-helix is present for the peptide segment of residues 36 to 42. While we want to emphasize again that these conclusions on the

secondary structure are preliminary and that it is, in particular, not possible from Figure 6 alone to determine the residues that limit the segments of regular secondary structure, it is quite interesting that the above conclusions coincide rather closely with the X-ray structure of a homologous protein, the Japanese quail ovomucoid domain 3 (Papamokos *et al.*, 1982). In this protein residues 21 to 32 (which are homologous to residues 22 to 33 in BUSI IIA) form an antiparallel β hairpin, to which the segment 49 to 54 (corresponding to 50 to 55 in BUSI IIA) is added as a third antiparallel strand, and residues 33 to 44 (corresponding to 34 to 45 in BUSI IIA) are in an α-helix.

Financial support by the Schweizerischer Nationalfonds (project 3.528.79) and by a special grant of the Eidgenössische Technische Hochschule, Zürich is gratefully acknowledged. We thank Mrs E. Huber and Mrs E. H. Hunziker for careful preparation of the manuscript and the illustrations, Dr D. Marion for helpful discussions and Dr D. Čechová for providing us with the starting material for the preparation of BUSI IIA and for valuable advice on the isolation procedures.

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.

Arseniev, A. S., Wider, G., Joubert, F. J. & Wüthrich, K. (1982). *J. Mol. Biol.* **159**, 323–352.

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

Baumann, R., Wider, G., Ernst, R. R. & Wüthrich, K. (1981). *J. Magn. Reson.* **44**, 402–406.

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**, 285–298.

Campbell, I. D., Dobson, C. M., Moore, G. R., Perkins, S. J. & Williams, R. J. P. (1976). *FEBS Letters*, **70**, 96–100.

Čechová, D., Jonáková, V., Sedláková, E. & Mach, O. (1979). *Hoppe-Seyler's Z. Physiol. Chem.* **360**, 1753–1758.

Dubs, A., Wagner, G. & Wüthrich, K. (1979). *Biochim. Biophys. Acta*, **577**, 177–194.

Jeener, J., Meier, B. H., Bachmann, P. & Ernst, R. R. (1979). *J. Chem. Phys.* **71**, 4546–4553.

Keller, R. M., Baumann, R., Hunziker-Kwik, E. H., Joubert, F. J. & Wüthrich, K. (1983). *J. Mol. Biol.* **163**, 623–646.

Kuo, M. & Gibbons, W. A. (1979). In *Peptides: Structure and Biological Function, Proc. Sixth Amer. Peptide Symp.* (Gross, E. & Meienhofer, J., eds), pp. 229–232, Pierce Chem. Co., Rockford.

Leach, S. J., Nemethy, G. & Scheraga, H. A. (1977). *Biochem. Biophys. Res. Commun.* **75**, 207–215.

Macura, S., Huang, Y., Suter, D. & Ernst, R. R. (1981). *J. Magn. Reson.* **43**, 259–281.

Meloun, B. & Čechová, D. (1979). *Coll. Czechoslov. Chem. Commun.* **44**, 2710–2720.

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.

Papamokos, E., Weber, E., Bode, W., Huber, R., Empie, M. W., Kato, I. & Laskowski, M. (1982). *J. Mol. Biol.* **158**, 515–537.

Richarz, R., Sehr, P., Wagner, G. & Wüthrich, K. (1979). *J. Mol. Biol.* **130**, 19–30.
Štrop, P. & Wüthrich, K. (1983). *J. Mol. Biol.* **166**, 631–640.
Štrop, P., Čechová, D. & Wüthrich, K. (1983). *J. Mol. Biol.* **166**, 669–676.
Wagner, G. & Wüthrich, K. (1982*a*), *J. Mol. Biol.* **155**, 347–366.
Wagner, G. & Wüthrich, K. (1982*b*). *J. Mol. Biol.* **160**, 343–361.
Wagner, G., Anil Kumar & Wüthrich, K. (1981). *Eur. J. Biochem.* **114**, 375–384.
Wider, G., Lee, K. H. & Wüthrich, K. (1982). *J. Mol. Biol.* **155**, 367–388.
Wider, G., Hosur, R. V. & Wüthrich, K. (1983). *J. Magn. Reson.*, in the press.
Wüthrich, K. (1976). *NMR in Biological Research: Peptides and Proteins*, North-Holland Publishing Co., Amsterdam.
Wüthrich, K. (1983). *Biopolymers*, **22**, 131–138.
Wüthrich, K., Wider, G., Wagner, G. & Braun, W. (1982). *J. Mol. Biol.* **155**, 311–319.

Edited by J. C. Kendrew

APPENDIX

Reinvestigation of the C-terminal End in the Amino Acid Sequence of the Proteinase Inhibitor IIA from Bull Seminal Plasma

GERHARD FRANK

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

Chemical sequence determination of BUSI IIA by Meloun & Čechová (1979) resulted in the C-terminal sequence –GCK, whereas ^{1}H nuclear magnetic resonance studies resulted in –GKC. In view of this apparent discrepancy we decided to reinvestigate this part of the sequence by means of chemical methods.

In order to rule out differences that might be due to different origins of BUSI IIA, this investigation was performed with a sample of the same batch that had been used for the n.m.r. studies.

(a) *The chymotryptic fragment from the C terminus of BUSI IIA*

To avoid interference with other regions of the sequence a C-terminal hexapeptide of the inhibitor was prepared in the following manner: the disulfide bridges in BUSI IIA were reductively cleaved with dithiothreitol and all sulfhydryl groups were alkylated with iodoacetamide following a standard procedure. A sample (5 mg) of the alkylated protein was then digested in 0·1 M-NH_4HCO_3 for five hours with 100 μg of chymotrypsin (EC 3.4.21.1) at room temperature. In high-voltage paper electrophoresis at pH 6·5 the C-terminal