Štrop, P., Wider, G., & Wüthrich, K. (1983) *J. Mol. Biol. 166*, 641–667.
Vliegenthart, J. F. G., Dorland, L., & Van Habeek, H. (1983) *Adv. Carbohydr. Chem. Biochem. 41*, 209–374.
Wagner, G. (1983a) *J. Magn. Reson. 55*, 151–156.
Wagner, G. (1983b) *Q. Rev. Biophys. 16*, 1–57.
Wagner, G., & Wüthrich, K. (1982) *J. Mol. Biol. 155*, 347–366.
Wagner, G., Neuhaus, D., Wörgötter, E., Vasák, M., Kägi, J. H. R., & Wüthrich, K. (1986) *Eur. J. Biochem. 157*, 275–289.
Watanabe, K., Matsuda, T., & Sato, Y. (1981) *Biochim. Biophys. Acta 667*, 242–250.
Westler, W. M., Ortiz-Polo, G., & Markley, J. L. (1984) *J. Magn. Reson. 58*, 354–357.
Wider, G., Macura, S., Kumar, A., Ernst, R. R., & Wüthrich, K. (1984) *J. Magn. Reson. 56*, 207–234.
Williamson, M. P., Havel, T. F., & Wüthrich, K. (1985) *J. Mol. Biol. 182*, 295–315.
Wlodawer, A., & Sjölin, L. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 1418–1422.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.
Wüthrich, K., Billeter, M., & Braun, W. (1984) *J. Mol. Biol. 180*, 715–740.

Two-Dimensional NMR Studies of Kazal Proteinase Inhibitors. 2. Sequence-Specific Assignments and Secondary Structure of Reactive Site Modified Turkey Ovomucoid Third Domain†

Gyung Ihm Rhyu and John L. Markley*
Department of Biochemistry, College of Agricultural and Life Sciences, University of Wisconsin, Madison, Wisconsin 53706
Received August 11, 1987; Revised Manuscript Received November 24, 1987

ABSTRACT: The solution structure of modified turkey ovomucoid third domain (OMTKY3*) was investigated by high-resolution proton NMR techniques. OMTKY3* was obtained by enzymatic hydrolysis of the scissile reactive site peptide bond (Leu[18]–Glu[19]) in turkey ovomucoid third domain (OMTKY3). All of the backbone proton resonances were assigned to sequence-specific residues except the N*H*'s of Leu[1] and Glu[19], which were not observed. Over 80% of the side-chain protons also were assigned. The secondary structure of OMTKY3*, as determined from assigned NOESY cross-peaks and identification of slowly exchanging amide protons, contains antiparallel β-sheet consisting of three strands (residues 21–25, 28–32, and 49–54), one α-helix (residues 33–44), and one reverse turn (residues 26–28). This secondary structure closely resembles that of OMTKY3 in solution [Robertson, A. D., Westler, W. M., & Markley, J. L. (1988) *Biochemistry* (preceding paper in this issue)]. On the other hand, changes in the tertiary structure of the protein near to and remote from the cleavage site are indicated by differences in the chemical shifts of numerous backbone protons of OMTKY3 and OMTKY3*.

Ovomucoids are glycoproteins present in high concentration in avian egg white. They inhibit serine proteinases according to the standard mechanism (Laskowski & Kato, 1980) given by

$$E + I \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} C \underset{k^*_{on}}{\overset{k^*_{off}}{\rightleftharpoons}} E + I^* \qquad (1)$$

where E is the proteinase, I and I* are the virgin (reactive site peptide bond intact) and modified (reactive site peptide bond hydrolyzed) inhibitor, respectively, and C is the stable complex. Avian ovomucoids consist of three tandem domains. Each domain is homologous to single-domain pancreatic secretory trypsin inhibitors (Kazal). Among those domains, third domains have been the most extensively studied because they can be easily isolated, free of carbohydrate, following limited proteolysis of ovomucoid. Third domains from 101 avian species have been sequenced (Laskowski et al., 1987).

Turkey ovomucoid third domain (OMTKY3)[1] (M_r 6060) consists of a single chain of 56 amino acid residues having three disulfide bridges. Modified OMTKY3 (OMTKY3*) is obtained by clipping the scissile reactive site peptide bond, Leu[18]–Glu[19]. Proteinases that are inhibited by OMTKY3 also are inhibited by OMTKY3*. The rate constants for the as-

†Supported by NIH Grant GM35976. This study made use of the National Magnetic Resonance Facility at Madison, which is supported in part by NIH Grant RR02301 from the Biomedical Research Technology Program, Division of Research Resources. Equipment in the facility was purchased with funds from the University of Wisconsin, the NSF Biological Instrumentation Program (Grant PCM-845048), the NIH Biomedical Research Technology Program (Grant RR02301), the NIH Shared Instrumentation Program (Grant RR02781), and the U.S. Department of Agriculture.

[1] Abbreviations: COSY, two-dimensional homonuclear correlated spectroscopy; 2D, two dimensional; des(1–3)OMTKY3*, turkey ovomucoid third domain which lacks the N-terminal three residues (Leu[1]-Ala[2]-Ala[3]-) and is cleaved at the reactive site (Leu[18]-Glu[19]); FID, free induction decay; HOHAHA, two-dimensional homonuclear Hartmann–Hahn spectroscopy; NOE, nuclear Overhauser effect; NOESY, two-dimensional NOE spectroscopy; NMR, nuclear magnetic resonance; OMSVP3*, modified silver pheasant ovomucoid third domain (reactive site peptide bond, Met[18]–Glu[19], cleaved); OMTKY3, turkey ovomucoid third domain; OMTKY3*, modified turkey ovomucoid third domain (reactive site peptide bond, Leu[18]–Glu[19], cleaved); pH*, uncorrected pH meter reading of a ^2H_2O solution measured with a glass electrode standardized with H_2O buffer; RELAY, two-dimensional relayed coherence transfer spectroscopy; [^{2}H]TSP, sodium 3-(trimethylsilyl)-[2,2,3,3-$^2H_4$]propionate; Tris-HCl, tris(hydroxymethyl)aminomethane hydrochloride. The amino acid protons are designated according to IUPAC–IUB conventions (IUPAC–IUB Commission on Biochemical Nomenclature, 1970). For example, $C_\alpha H$ is the proton attached to the main-chain α-carbon.

0006-2960/88/0427-2529$01.50/0

sociation and dissociation of serine proteinases with modified inhibitor are generally much lower than those with virgin inhibitor (Ardelt & Laskowski, 1985); e.g., the ratios k_{on}/k^*_{on} and k_{off}/k^*_{off} with chymotrypsin are about 10^6. As shown by X-ray crystallographic structures of serine proteinase–protein proteinase inhibitor complexes [see, for example, Read et al. (1983), Bode et al. (1986), and Fujinaga et al. (1987)] and NMR studies of such complexes (Baillargeon et al., 1980; Richarz et al., 1980; Hunkapiller et al., 1979; Kainosho, 1985), the reactive site peptide bond of the inhibitor is intact. The stable complex, C, appears to be a Michaelis E·I complex; the reactive site peptide bond is resynthesized following association of E with I* (Laskowski & Kato, 1980).

Even though OMTKY3 inhibits proteinases by acting as a "good substrate", the equilibrium constant for the hydrolysis K_{hyd} = [OMTKY3*]/[OMTKY3], unlike normal substrates, is close to unity at neutral pH (Ardelt & Laskowski, 1983). The pH dependence of K_{hyd} suggests that the free energies of OMTKY3 and OMTKY3* are essentially the same. Thus the newly formed amino- and carboxy-terminal fragments in OMTKY3* do not appear to gain significant rotational freedom and may be relatively close to one another. The interesting question is whether or not the hydrogen bonds that stabilize residues in the neighborhood of the reactive site bond in Kazal-type inhibitors (Read et al., 1983; Papamokos et al., 1982; Bode et al., 1985) still exist after the reactive site peptide bond has been cleaved. No refined crystallographic data have been published for OMTKY3* or for any other modified inhibitor. Crystal structures have been determined for OMTKY3 in complex with *S. griseus* proteinase B (Read et al., 1983), human leukocyte elastase (Bode et al., 1986), and chymotrypsin (Fujinaga et al., 1987). Crystal structures also are available for the homologous Japanese quail ovomucoid third domain (Papamokos et al., 1982) and silver pheasant ovomucoid third domain (Bode et al., 1985).

In this paper we discuss the results of our investigation of the solution conformation of modified turkey ovomucoid third domain. High-resolution ^{1}H NMR techniques have been employed including phase-sensitive COSY, phase-sensitive NOESY, RELAY, and HOHAHA. We present a comparison of the backbone proton chemical shifts and secondary structures of OMTKY3* with those of OMTKY3 (Robertson et al., 1988).

Materials and Methods

Proteins. Ovomucoid was isolated from turkey egg whites and purified by a modification of the method of Lineweaver and Murray (1947) as described by Bogard et al. (1980). Carbohydrate-free third domain was produced by limited proteolysis of the ovomucoid with the proteinase from *Staphylococcus aureus* and purified chromatographically (Ortiz-Polo, 1986). Pure OMTKY3 was then converted to OMTKY3* by the method of Ardelt and Laskowski (1985) with minor modifications. OMTKY3 (2.5×10^{-3} M) was dissolved in 1.5×10^{-2} M HCl, and the pH of the solution was adjusted to 1.5 with 1 M HCl. Pronase was then added to a final concentration of 1 mg/mL, and the mixture was stirred gently at room temperature until about 90% of the OMTKY3 was converted to OMTKY3*. For the first preparation, the extent of the reaction was determined by analytical anion-exchange chromatography (Ardelt & Laskowski, 1982); later on, the modification reaction was stopped after 43–50 h. The digest was then chromatographed, to remove Pronase, on Sephadex G-50f (5×90 cm) equilibrated with 0.02 M HCl. The product was lyophilized and then desalted through a Sephadex G-25f column (5×16 cm) equilibrated with 15 mM ammonium bicarbonate. OMTKY3* was separated on DEAE-Sepharose CL-6B (3×19 cm) from the remaining OMTKY3 and other contaminants. The anion-exchange column was developed with a 0–0.1 M linear NaCl gradient in 0.04 M Tris-HCl buffer, pH 8.8. The inhibitor was desalted and lyophilized. OMSVP3* and des(1–3)OMTKY3* were prepared in the same way as OMTKY3*. Pronase was purchased from Calbiochem, ^2H_2O (99.98% isotope pure) was from Stohler Isotope Chemicals, and staphyococcal proteinase was purchased from Miles Laboratories, Inc.

All the NMR samples were prepared as 15 mM protein in 0.2 M KCl, pH 4.0. Samples in H_2O were prepared by dissolving 45 mg of protein in 0.5 mL of solvent mixture consisting of 90% H_2O and 10% ^2H_2O (v/v) containing 0.2 M KCl. For non-preexchanged protein samples in ^2H_2O, 45 mg of protein was dissolved in 1 mL of distilled water, and the pH* was adjusted to 3.65. The protein was lyophilized and redissolved in 0.5 mL KCl solution in ^2H_2O, and the pH* was adjusted to 4.0, if necessary. Protein to be preexchanged was dissolved in ^2H_2O, and the pH* was adjusted to ~9.5. The resulting solution was held for 1 h at room temperature; then, the pH* was readjusted to ~4, and the sample was lyophilized 3 times from ^2H_2O.

NMR Spectroscopy. NMR spectra were recorded on Bruker AM-500 and AM-400 NMR spectrometers equipped with digital phase shifters and an ASPECT 3000 computer. The water signal was suppressed by selective irradiation during the relaxation delay and, in the case of NOESY, during the mixing time as well. Typically, 512 blocks were collected, containing 32 transients for ^2H_2O samples and 64 transients for H_2O samples, each accumulated into 2048 data points. The digital resolution was adjusted to 5.88 Hz per point in both dimensions by zero filling once in the t_1 dimension only. COSY data were acquired in the absolute value mode with a $90°-t_1-90°-t_2$ pulse sequence with the second 90° pulse replaced by a 60° or 45° pulse to optimize cross-peak intensity and sensitivity (Bax & Freeman, 1981). The delay, t_1, was incremented from 3 μs in steps of 166 μs. Phase cycling was as described by Wider et al. (1984). The FIDs were apodized by multiplication with a sine bell function before Fourier transformation. RELAY spectra were recorded in the absolute value mode (Wagner, 1983), with phase cycling as described by Bax and Drobny (1985), and processed in the same way as COSY data in the absolute value mode. Double quantum filtered COSY data were obtained in the pure absorption mode (Rance et al., 1983) with quadrature detection in ω_1 achieved by incrementing the phase of the first pulse by 90° in successive t_1 values (Marion & Wüthrich, 1983). NOESY data were recorded in pure absorption mode as described by Williamson et al. (1984) with a mixing time of 100 ms, which is short enough to avoid occurrence of excessive spin-diffusion (Kumar et al., 1981). To reduce contributions from coherent magnetization transfer to the cross-peak intensities, the mixing time was modulated randomly with an amplitude of ±10 ms (Macura et al., 1981). HOHAHA spectra (Bax & Davis, 1985) were recorded on a Bruker AM-400 spectrometer in pure absorption mode at two mixing times (50 and 100 ms). For every spectrum, additional phase cycling (CYCLOPS) was employed to reduce quadrature images in ω_2 and to remove axial peaks (Hoult & Richards, 1975). Absorption mode spectra were acquired with the same conditions as those for absolute value mode spectra, except that, for the former, the t_1 increment was 83 μs. The resulting data were processed by real Fourier transformation. A nonshifted sine bell window function in the ω_2 dimension and a Gaussian window function

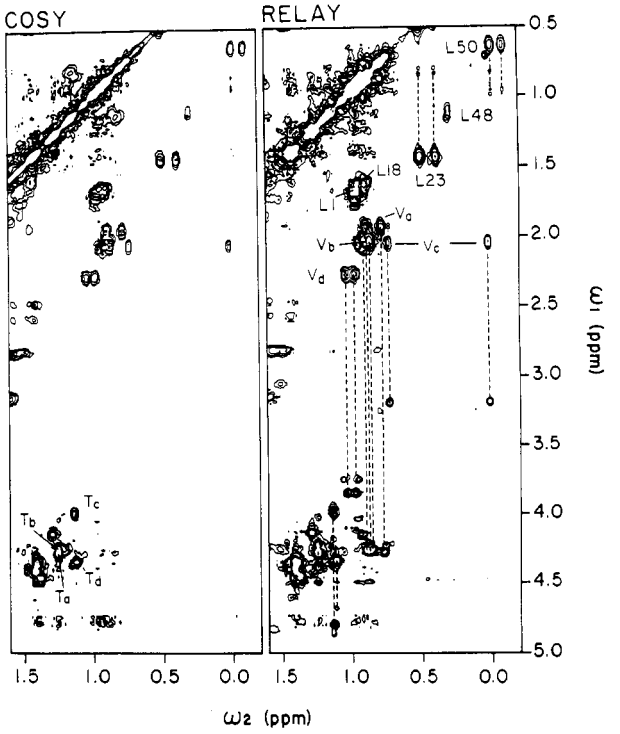

FIGURE 1: Identification of the spin systems of valine and threonine residues. The relevant cross-peaks are labeled by amino acid type. The spectra were obtained at 25 °C. Protein concentration was 15 mM in 0.2 M KCl in ^2H_2O, pH* 4.0. The RELAY spectrum was obtained with a delay of 23 ms.

in the ω_1 dimension with a line-broadening factor of −20 and maximum at $0.3t_1$ were used to apodize FIDs for NOESY and COSY. Relaxation delays were 2 s for NOESY and 1.5 s for the other experiments. In some of the final spectra, the background noise was eliminated by the method of Zolnai et al. (1986). Other details are given in the figure legends.

Chemical shifts are referenced to internal [^{2}H]TSP. Conventional symbolic notation is used for sequential NOEs such as d_{NN}, $d_{\alpha N}$, or $d_{\beta N}$ [definitions are given in Robertson et al. (1988)].

RESULTS

Identification of Signals from Methyl-Containing Amino Acids. The methyl region of the COSY map of OMTKY3* in ^2H_2O is shown in Figure 1. OMTKY3* contains four valine, five leucine, four alanine, and four threonine residues. COSY cross-peaks in which one active coupling partner is a methyl resonance tend to be very intense and fall in a characteristic region of the spectrum: e.g., the methyl cross-peaks from alanine or threonine fall in the region ω_2 = 1.0–1.6 ppm and ω_1 = 3.6–4.8 ppm; the methyl cross-peaks from valine or leucine fall in the region ω_2 = −0.2–1.2 ppm and ω_1 = 0.5–2.2 ppm (Figure 1). In the valine and leucine methyl region, nine pairs of methyl groups were observed. Among those, two pairs overlapped partially at about (0.97, 1.68) ppm, and the partner of a methyl cross-peak at (0.31, 1.14) ppm was not resolved well from the diagonal. Among those nine pairs, four methyl pairs at ω_1 = 1.94, 2.04, 2.07, and 2.28 ppm had RELAY cross-peaks to resonances at 4.27, 4.25, 3.21, and 3.84 ppm, respectively (Figure 1). Since only methyl pairs from valine residues can have RELAY cross-peak to their own $C_\alpha H$, these resonances were identified as $C_\alpha H$'s of Val^a, Val^b, Val^c, and Val^d, respectively. Thus, the remaining methyl groups had to belong to leucine residues. Some of the RELAY cross-peaks between $C_\delta H$ and $C_\beta H$ from leucine residues can be seen in Figure 1. The weak pairs of RELAY cross-peaks at ω_1 = 3.25 and 3.75 ppm close to Val^c and Val^d RELAY cross-peaks come from extra valine peaks.

In Figure 2 are shown COSY maps of OMTKY3*, OMSVP3*, and des(1–3)OMTKY3*. Leu^{18} in OMTKY3* is replaced by Met^{18} in OMSVP3*. The $C_\alpha H$ chemical shift of Leu^{18} was assigned to 4.23 ppm on the basis of the single amino acid replacement. Leu^{18} $C_\beta H$, $C_\gamma H$, and $C_\delta H$ chemical

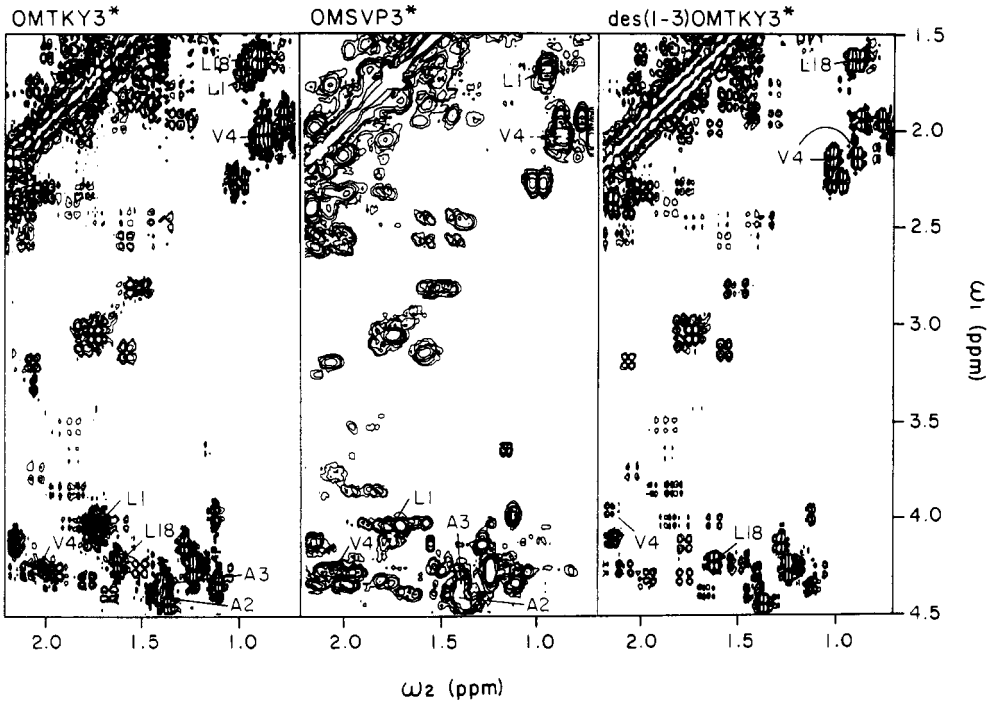

FIGURE 2: Assignments of Leu^{18} and N-terminal residues by comparison of COSY maps of ovomucoid third domains from turkey and silver pheasant. OMTKY3* and OMSVP3* differ by a single substitution: Leu^1 → Met^{18}. Des(1–3)OMTKY3* differs from OMTKY3* by the lack of the three N-terminal residues (Leu^1-Ala^2-Ala^3). COSY of OMTKY3* and des(1–3)OMTKY3* were recorded in the pure absorption mode, and COSY of OMSVP3* was recorded in the absolute value mode. The background noise was eliminated from the pure absorption mode spectra by the method of Zolnai et al. (1986).

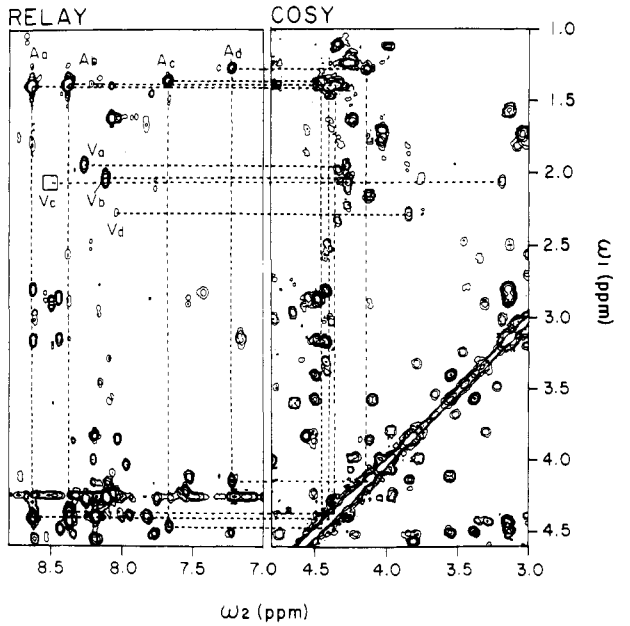

FIGURE 3: Identification of the spin systems of alanine residues. Alanine ($C_\alpha H$, NH) and ($C_\alpha H$, $C_\beta H_3$) COSY cross-peaks and ($C_\beta H$, NH) RELAY cross-peaks are connected by dashed lines. Valine spin systems are labeled but not connected by boxes because of crowding. The small square indicates the expected position of the (NH, $C_\beta H$) RELAY cross-peak of Val[41]. The RELAY spectrum was obtained with a delay of 40 ms.

shifts were observed at 1.62, 1.64, 0.89, and 0.92 ppm. In the COSY map of des(1–3)OMTKY3*, the methyl cross-peaks of Val[b] were shifted downfield (labeled as Val[4] in Figure 2). The corresponding ($C_\alpha H$, $C_\beta H$) cross-peak also appeared at a different position. Since Val[4] is the N-terminal residue in des(1–3)OMTKY3*, Val[b] was assigned to Val[4], and this assignment was confirmed by sequential assignment via NOE connectivities (see below). The methyl resonances of Leu[1] were assigned as 0.95 and 0.97 ppm by comparison of COSY maps of OMTKY3* and des(1–3)OMTKY3*, which lacked the peaks. The ($C_\alpha H$, $C_\beta H$) cross-peak from Leu[1] at (4.03, 1.74) ppm also was missing in the des(1–3)OMTKY3* COSY map, but it was assigned by sequential assignment (see below). The region for alanine and threonine methyl cross-peaks also was simplified in the COSY spectrum of des(1–3)OMTKY3* due to the lack of signals from Ala[2] and Ala[3].

Eight intense cross-peaks were observed in the region of the COSY map dominated by alanine and threonine methyl resonances (Figure 1). The four methyl resonances at 1.24, 1.24, 1.12, and 1.10 ppm had RELAY cross-peaks to resonances at 4.38, 4.22, 4.79, and 4.68 ppm, respectively (Figure 1). These resonances were identified as $C_\alpha H$'s of four threonine residues: a, b, c, and d, respectively. The two partially overlapped methyl resonances at 1.24 ppm could be distinguished by means of a ($C_\alpha H$, $C_\beta H$) COSY cross-peak which lines up with only one part of the pattern at (4.38, 4.28) ppm. The RELAY cross-peak at 4.22 ppm was not well resolved from the COSY peak at (4.25, 1.24) ppm; thus, classification of this spin system as a threonine residue required confirmation by sequential assignment (see below). The remaining peaks thus arise from Ala ($C_\alpha H$, $C_\beta H$). As expected for alanines, strong RELAY cross-peaks were observed from amide protons to each of the methyl resonances of the remaining cross-peaks (Figure 3). The alanine COSY cross-peaks [($C_\alpha H$, $C_\beta H$) and ($C_\alpha H$, NH)] and RELAY cross-peak (NH, $C_\beta H$) are connected to the corresponding diagonal peaks in Figure 3. The COSY cross-peaks at (4.40, 1.42), (4.36, 1.40), (4.45, 1.37), and (4.13, 1.28) ppm are labeled as Ala[a], Ala[b], Ala[c], and Ala[d], respectively. From the RELAY spectrum, the chemical shifts of the NH's of Ala[a], Ala[b], Ala[c], and Ala[d] were assigned to 8.54, 8.40, 7.70, and 7.24 ppm, respectively. Two of these are missing in the spectrum of des(1–3)OMTKY3* and must correspond to Ala[2] and Ala[3].

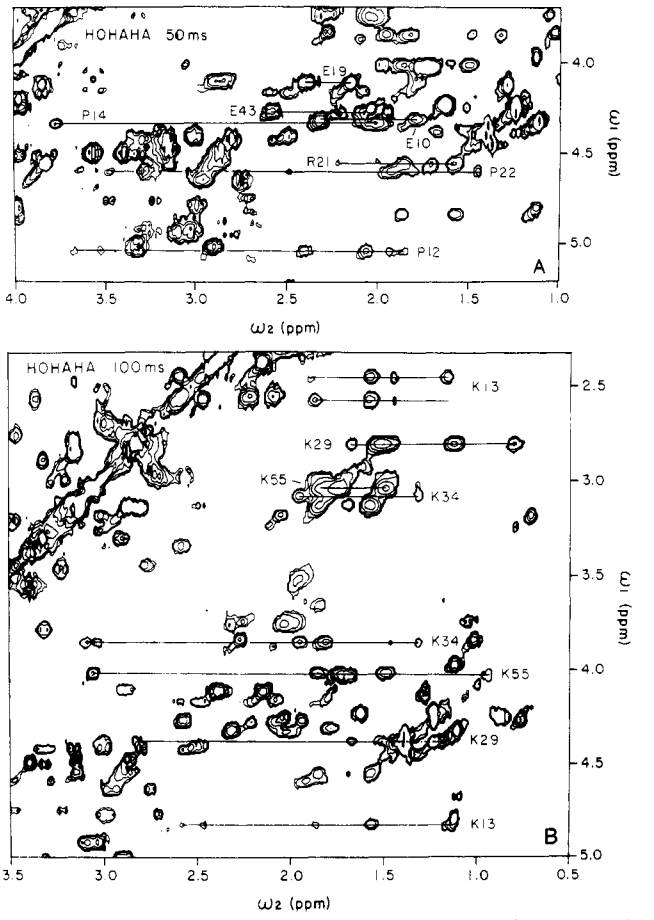

FIGURE 4: Pure-phase absorption HOHAHA spectra of OMTKY3* recorded at 400 MHz: (A) 50-ms mixing time; (B) 100-ms mixing time. With a mixing time of 50 ms, multiple relayed connectivities are observed mostly for prolines and arginine. With a mixing time of 100 ms, multiple relayed connectivities are seen mostly for lysine residues. The FIDs were apodized by a shifted sine bell window function in both dimensions. Assignments shown here were obtained only after completion of the backbone assignments. Background noise was eliminated by the method of Zolnai et al. (1986).

RELAY peaks at 8.13, 8.25, and 8.06 ppm served to identify the NH's of Val[a], Val[b] (Val[4]), and Val[d], respectively. The RELAY cross-peak for Val[c] (Val[41]) was not observed, but the expected position is marked by a square in Figure 3. The NH chemical shift for Val[c] was found to be 8.50 ppm after sequential assignment (see below).

Identification of Long Side Chain Amino Acids. HOHAHA data (Figure 4) permitted identification of most of the long side chain amino acids except leucines. Two leucine spin systems were identified by comparison with variants lacking Leu[1] and Leu[18] (Figure 2); others were identified in the RELAY map (Figure 1). With a short mixing time (50 ms; Figure 4A), most of multiple relayed connectivities were observed for arginine and proline residues including $C_\delta H$'s. Multiple relayed connectivities from $C_\alpha H$ to $C_\epsilon H$ for lysines were obtained only with a long mixing time (100 ms; Figure 4B), under conditions where the multiple relayed connectivities for prolines disappear. This made it easy to discriminate between prolines and lysines. Glutamate connectivities could be observed in spectra with either mixing time.

Sequential Assignment. The fingerprint region of the

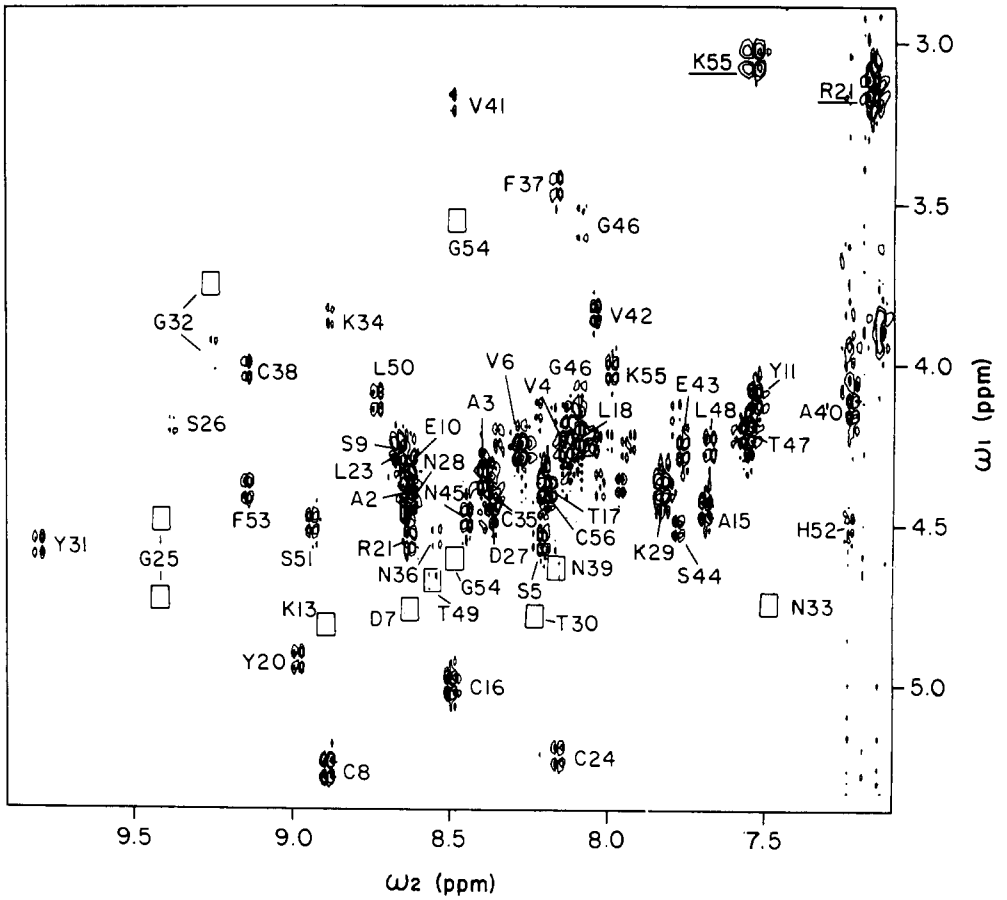

FIGURE 5: COSY contour plot of OMTKY3* dissolved in H_2O. The identity of each cross-peak is labeled with its assignments. Some of the cross-peaks were bleached out by water irradiation. The peaks near the water peak were observed in later experiments; their positions are indicated by labeled boxes. The spectrum contained a number of extra peaks that have not been assigned (see Discussion). The FIDs were apodized in ω_2 by an unshifted sine bell window function and in ω_1 by a Gaussian window function.

COSY map of OMTKY3* is shown in Figure 5. Cross-peaks from all expected amide protons, except those of the two N-terminal residues (Leu[1] and Glu[19]) which appear to exchange rapidly with solvent, were observed and assigned to sequence-specific residues: 56 residues minus 3 prolines and 2 N-terminal residues. Backbone amide protons from a reduced set of 25 residues were detected in COSY data obtained with non-preexchanged protein dissolved in ^2H_2O (data not shown). These data facilitated the analysis by resolving some COSY or NOESY cross-peak overlaps that otherwise occur. For displaying the sequential assignments via $d_{\alpha N}$, we have adopted a presentation recently used by Wagner et al. (1986) in place of the more familiar NOESY–COSY connectivity diagrams. The information from the COSY fingerprint (Figure 5) and the (N*H*, C$_\alpha$*H*) region of the NOESY map was combined as described in the legend to Figure 6.

Analysis of a leucine spin system provided a starting point for sequential assignments. In the COSY map obtained in ^2H_2O, a pair of methyl cross-peaks was observed at (0.49, 1.44) ppm and (0.38, 1.44) ppm (Leu[23] in Figure 1). Both of these methyls have RELAY cross-peaks to 0.84 and 1.09 ppm (observed only at lower contour levels), which were assigned to the C$_\beta$*H*'s of the leucine residue. COSY cross-peaks were observed at (1.09, 4.27) ppm and (0.84, 4.27) ppm. In the COSY data shown in Figure 1, these cross-peaks are very weak; in other experiments, they turned out to be stronger and better defined. The N*H* of this spin system resonates at 8.68 ppm as evidenced by a COSY cross-peak observed with non-preexchanged protein sample in ^2H_2O (data not shown). Additional evidence for this assignment came from NOEs from the N*H* to protons resonating at 1.09, 1.44, and 0.49 ppm. The

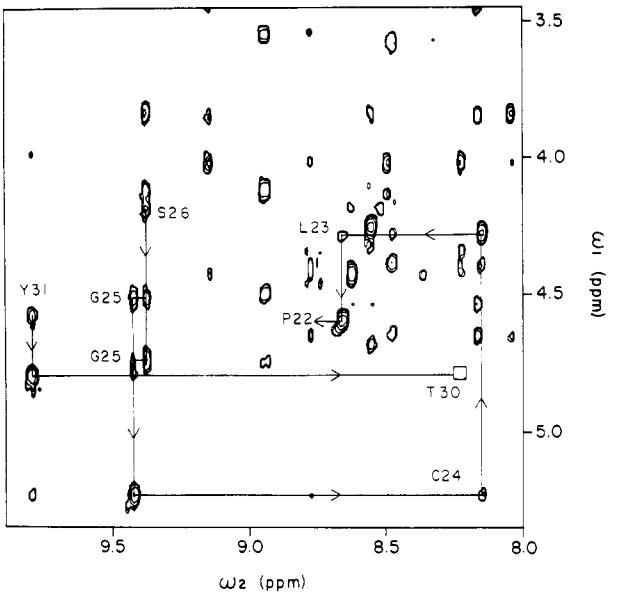

FIGURE 6: Sequential $d_{\alpha N}$ connectivities for OMTKY3*. NOESY data were recorded with non-preexchanged protein dissolved in ^2H_2O. Information contained in the COSY fingerprint (not shown, but see Figure 5) and in the (N*H*, C$_\alpha$*H*) cross-peak region of the NOESY spectrum is combined in this figure. The positions of the relevant (N*H*, C$_\alpha$*H*) COSY cross-peaks are indicated by labeling of intraresidue $d_{\alpha N}$ cross-peaks or by labeling of a box enclosing the position of the peak if the intraresiduel $d_{\alpha N}$ was not observed. Apodization functions used were as in Figure 5.

C$_\alpha$*H* of this leucine residue showed a strong NOE at 8.20 ppm (Figure 6). The C$_\alpha$*H* of this presumed adjacent residue was

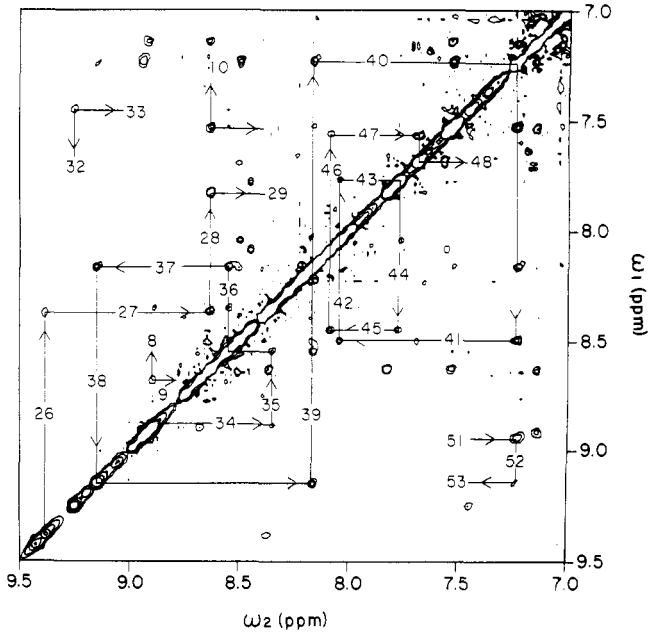

FIGURE 7: Sequential d_{NN} connectivities observed in the NOESY map of OMTKY3* obtained in H_2O. The mixing time was 100 ms. Shifted sine bell window functions were used in both dimensions to apodize the FIDs. The final spectrum was obtained after polar symmetrization.

then located in the COSY contour map at 5.23 ppm. This $C_\alpha H$ showed an NOE to a resonance at 9.40 ppm, which in turn had cross-peaks to two $C_\alpha H$ resonances located at 4.50 and 4.75 ppm. Hence, the spin system (9.40, 4.50, 4.75) ppm was identified as that of a glycine residue. The unique Leu-X-Gly sequence is located at Leu[23]-Cys[24]-Gly[25] in the primary sequence of OMTKY3*. This assignment is consistent with the results presented in Figure 2, which indicated that this leucine spin system corresponds to Leu[23], Leu[48], or Leu[50]. The cross-peak at (8.20, 5.23) ppm thus belongs to (NH, $C_\alpha H$) from Cys[24]. The two $C_\alpha H$ protons of Gly[25] showed an NOE to a proton resonating at 9.37 ppm, which was assigned to the NH of Ser[26]. A series of (NH, NH) connectivities (d_{NN}) extended from Ser[26] through Asp[27], Asn[28], and Lys[29] (Figure 7). A strong NOE was observed from the NH of Leu[23] to a resonance at 4.59 ppm, which was assigned to Pro[22] $C_\alpha H$.

Two of the four valine $C_\alpha H$'s, Val[c] at 3.21 ppm and Val[d] at 3.84 ppm, showed well-resolved COSY connectivities to 8.50 and 8.06 ppm, respectively, in spectra obtained in ^2H_2O (data not shown, but see Figure 5). These two NH resonances also were found to be connected by a d_{NN}-type NOE (Figure 7). The OMTKY3* sequence has only one Val-Val dipeptide; thus, Val[c] and Val[d] were assigned pairwise to Val[41] and Val[42]. Since resonances of Val[b], but not those of Val[a], were affected by deletion of three N-terminal residues from OMTKY3* (Figure 2), Val[b] was assigned tentatively to Val[4] and Val[a] to Val[6]. (This was confirmed by sequential assignment described below.) The NH at 8.50 ppm (Val[c]) also showed an NOE connectivity to a resonance at 7.24 ppm, which in turn showed a strong RELAY cross-peak at 1.28 ppm and a COSY cross-peak at 4.13 ppm. The (4.13, 1.28) ppm cross-peak is labeled Ala[d] (see above and Figure 3). In the primary sequence of OMTKY3*, only Val[41] and Val[4] have alanine as a neighbor. Since Val[c] had been shown to have another valine residue as a neighbor (Val[d]), Val[c] was assigned to Val[41], Val[d] to Val[42], and Ala[d] to Ala[40]. A chain of d_{NN} connectivities was followed from Ala[40] to Lys[34]. In this sequence, it was found that the NH's of Asn[39] and Phe[37] overlap at 8.18 ppm. The ambiguity caused by this overlap was resolved in the following way.

The NH of Ala[40] showed an NOE to a resonance at 8.18 ppm. Two COSY cross-peaks were observed, one at (8.18, 4.64) ppm and the other at (8.18, 3.45) ppm. The NH and $C_\alpha H$ of the latter residue showed NOEs to an NH proton resonating at 9.17 ppm, indicating that the cross-peak at (8.18, 3.45) ppm did not arise from Asn[39], and the cross-peak at (8.18, 4.64) ppm was assigned to Asn[39] by elimination. Two d_{NN} connectivities were observed at (8.18, 9.17) ppm and (8.18, 8.54) ppm. These lined up with COSY cross-peaks at (4.01, 9.17) ppm and (4.54, 8.54) ppm, one of which had to correspond to Cys[38]. But $d_{\beta N}$ connectivities were observed from both residues to 8.18 ppm, so these could not be used to assign the Cys[38] resonances. The two d_{NN} spirals were continued to the next residue, Phe[37]. The phenylalanine spin system can often be identified through NOEs from the ring $C_\delta H$ to the $C_\beta H$. We were able to determine that the cross-peak at (4.01, 9.17) ppm corresponded to Cys[38] and that at (8.18, 3.45) ppm belonged to Phe[37]. As a result of the overlap of Phe[37] and Asn[39] NH protons, their d_{NN} cross-peaks with Cys[38] overlap. The COSY cross-peak at (4.54, 8.54) ppm was assigned to Asn[36]. A further d_{NN} connectivity led to assignment of Cys[35]. The NH of Lys[34] gave an NOE ($d_{\beta N}$) to resonances at 3.45 and 3.21 ppm (data not shown); the $C_\alpha H$ resonance of this spin system (Asn[33]) was located at 4.75 ppm and the NH resonance at 7.48 ppm. An intraresidue NOE ($d_{\beta N}$) for Asn[33] was not observed. The NH of Asn[33] exhibited NOEs to resonances at 3.76 and 3.97 ppm which form an AB spin system and are assigned to Gly[32]. The NH of Gly[32] was located at 9.25 ppm in the COSY fingerprint region. Only one COSY cross-peak was observed (3.97 ppm); the other $C_\alpha H$ chemical shift was assigned to 3.76 ppm via a strong NOE to its NH. A sequential d_{NN} connectivity confirmed the assignments of the NH resonances of Asn[33] and Gly[32] (Figure 7). The backbone protons of Tyr[31] and Thr[30] were then assigned via $d_{\alpha N}$ connectivities (Figure 6).

The chain of d_{NN} connectivities could be followed from 8.50 ppm (Val[42]) to 7.76, 8.43, 8.08, 7.55, and then 7.68 ppm (Figure 7). Two COSY cross-peaks were observed at (8.08, 3.58) ppm and (8.08, 4.10) ppm, which included two mutually coupled $C_\alpha H$ resonances. Hence the spin system (8.08, 3.58, 4.10) ppm was assigned to Gly[46]. The remaining two resonances, at 7.76 and 8.43 ppm, thus must represent three NH protons (Glu[43], Ser[44], and Asn[45]), with two of them overlapping. Two COSY cross-peaks were observed at (7.78, 4.50) ppm and (7.76, 4.25) ppm. The $C_\alpha H$ at 4.50 ppm belonged to an AMX spin system, with the other spins at 4.20 ppm and 3.95 ppm. These chemical shifts are close to those expected for $C_\beta H$'s of serine residues. Therefore, the cross-peak at (7.78, 4.50) ppm was assigned to the backbone protons of Ser[44]. Consequently, the NH resonances at 7.76, 8.43, 8.08, 7.55, and 7.68 ppm were assigned to Glu[43], Asn[45], Gly[46], Thr[47], and Leu[48], and their respective $C_\alpha H$ resonances were located on the fingerprint region at 4.25, 4.48, (3.18, 4.10), 4.20, and 4.25 ppm, respectively. The assignments for residues 45, 46, 47, and 48 were confirmed by $d_{\alpha N}$ connectivities (Figure 8A). Further $d_{\alpha N}$ connectivities permitted assignment of backbone protons from Thr[49], Leu[50], and Ser[51] (Figure 8A). In Figure 8B, sequential $d_{\alpha N}$ connectivities are shown from His[52] to Cys[56]. Overlap of $C_\alpha H$'s of Ser[51] (4.49 ppm) and His[52] (4.50 ppm) was resolved by reference to d_{NN}-type NOEs between Ser[51], His[52], and Phe[53] (Figure 7).

A spiral was constructed for four consecutive residues, starting with Ala[c], whose COSY fingerprint cross-peaks are located at (7.70, 4.45), (8.50, 4.99), (8.20, 4.38), and (8.09, 4.23) ppm (Figure 8B). Among these, the $C_\alpha H$ resonance at

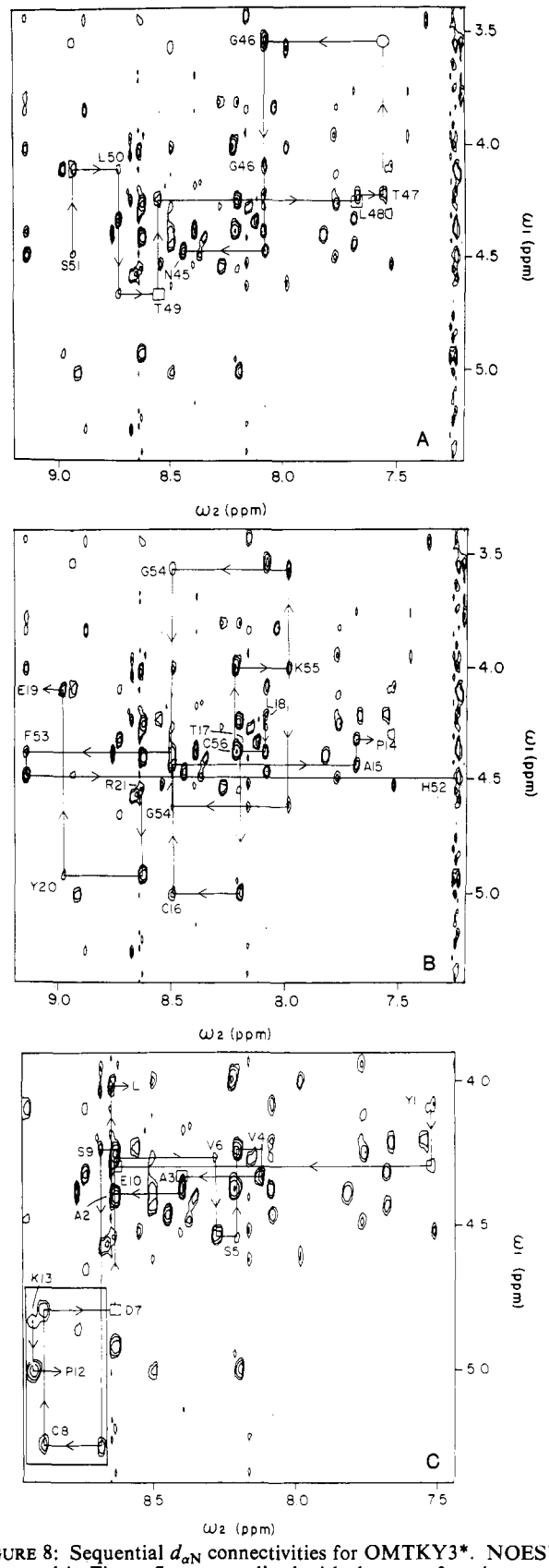

FIGURE 8: Sequential $d_{\alpha N}$ connectivities for OMTKY3*. NOESY data used in Figure 7 were apodized with the same functions used for COSY data in Figure 5. The information contained in the COSY fingerprint (Figure 5) and in the (N*H*, C$_\alpha$*H*) cross-peak region of the NOESY spectrum was combined as described in the legend to Figure 6. The NOE from the N*H* of Thr^{47} to the C$_\alpha$*H* of Glu^{46} [indicated by the circle in (A)] was observed in a NOESY spectrum (not shown) recorded with a mixing time of 200 ms. The box inserted in the lower left corner in (C) contains data obtained with a newer ^{1}H probe; the mixing time was 200 ms. The signals close to the water signal (Asp^7 and Lys^{13}) were better determined. Assignments for the following polypeptide segments are shown: (A) 45–51; (B) 14–18, 19–21 and 52–56; (C) 1–11 and 12–13.

4.38 ppm coincides with the C$_\alpha$*H* chemical shift of Thr^a, and the resonance at 4.23 coincides with the C$_\alpha$*H* chemical shift of Leu^{18}. The Ala-X-Thr-Leu sequence corresponded to Ala^{15}-Cys^{16}-Thr^{17}-Leu^{18} in the primary sequence of the protein. The chemical shift of the C$_\alpha$*H* of Pro^{14} was located at 4.33 ppm through a strong NOE from the N*H* of Ala^{15} (Figure 8B).

A $d_{\alpha N}$ spiral between five consecutive amino acid residues was constructed from Val^a (4.27, 8.25) through (4.50, 8.20), Val^b (4.25, 8.13), Ala^b (4.36, 8.40), to Ala^a (4.40, 8.54) ppm (Figure 8C). The cross-peaks were assigned to Val^6, Ser^5, Val^4, Ala^3, and Ala^2, respectively. The C$_\alpha$*H* chemical shift of Leu^1 was located at 4.03 ppm by the NOE from the N*H* of Ala^2. The $d_{\alpha N}$ connectivities were extended further from Val^6 to Tyr^{11}. Connectivities (d_{NN}) between Tyr^{11} and Glu^{10} and between Ser^9 and Cys^8 also were observed (Figure 7). No further NOE connectivity was observed from the C$_\alpha$*H* of Tyr^{11}, as was to be expected with a proline residue at position 12.

The remaining unassigned backbone proton chemical shifts at this point were those from the peptide segments Tyr^{11}-(Pro^{12}-Lys^{13})-Pro^{14} and Pro^{22}-(Arg^{21}-Tyr^{20}-Glu^{19}-N*H*$_2$). A spiral between three consecutive amino acid residues was constructed from (8.62, 4.55) ppm through (8.98, 4.93) ppm and 4.13 ppm (Figure 8B). By elimination of peptide segments assigned so far, this spiral was assigned to the peptide segment Pro^{22}-(Arg^{21}-Tyr^{20}-Glu^{19}). Proline does not have an amide proton, but the NOE from the C$_\delta$*H* of proline to the C$_\alpha$*H* of the preceding residue frequently can be used for sequential assignment. An NOE was indeed observed between the C$_\alpha$*H* resonance at 4.55 ppm (Arg^{21}) and the C$_\delta$*H* of Pro^{22} (data not shown). A residue with an extended side chain (Lys^{13} in Figure 4B) showed a COSY cross-peak at (8.89, 4.82) ppm (Figure 8C); its N*H* gave an NOE at 5.02 ppm. In addition, an NOE was observed between the resonance at 4.82 ppm and the C$_\delta$*H* of Pro^{14} (data not shown). Hence, the resonances at 8.89 and 4.82 ppm were assigned to the N*H* and C$_\alpha$*H* of Lys^{13}. The resonance at 5.02 ppm could be assigned to C$_\alpha$*H* of Pro^{12}; a strong NOE observed between this proton and the C$_\alpha$*H* of Tyr^{11} indicated that the peptide bond between these residues is in the cis conformation (Wüthrich et al., 1984).

Side-Chain Proton Assignments. With the sequence-specific backbone proton assignments in hand, C$_\beta$*H*'s could be assigned from (C$_\alpha$*H*, C$_\beta$*H*) and (C$_\beta$*H*^{1}, C$_\beta$*H*^{2}) COSY cross-peaks. These assignments were confirmed by RELAY cross-peaks from N*H* resonances and/or by intraresidue $d_{\beta N}$-type NOEs. Other long side chain protons were assigned by analysis of cross-peak connectivities in HOHAHA (Figure 4) and COSY maps. NOEs from C$_\alpha$*H* or N*H* to the side-chain protons also aided these assignments. The side-chain N$_\delta$*H* resonances of four of the five asparagine residues in OMTKY3* were assigned via the intraresidue NOEs from the C$_\beta$*H*'s.

Aromatic Ring Proton Assignments. Short distances between C$_\delta$*H* on aromatic rings and C$_\beta$*H* or C$_\alpha$*H* of the same residue ensure reliable NOE connectivities (Billeter et al., 1982). Hence, the C$_\delta$*H*'s from Tyr^{11}, Tyr^{20}, Tyr^{31}, Phe^{37}, and Phe^{53} were assigned by intraresidue NOEs from their C$_\beta$*H*'s. The assignments were extended to C$_\epsilon$*H* and C$_\zeta$*H* (in phenylalanines) via COSY cross-peaks. The C$_\epsilon$*H* resonance of His^{52} (the single histidine) was assigned by virtue of its unique downfield chemical shift; the C$_\delta$*H* was then assigned by a COSY cross-peak to C$_\epsilon$*H*.

*Secondary Structure in OMTKY3**. Once sequence-specific resonance assignments have been obtained, further NMR measurements, such as NOEs (Billeter et al., 1982), spin–spin coupling constants (Pardi et al., 1984), or amide–proton exchange rates, can be attributed to specified locations in the sequence. The occurrence of certain patterns of NMR pa-

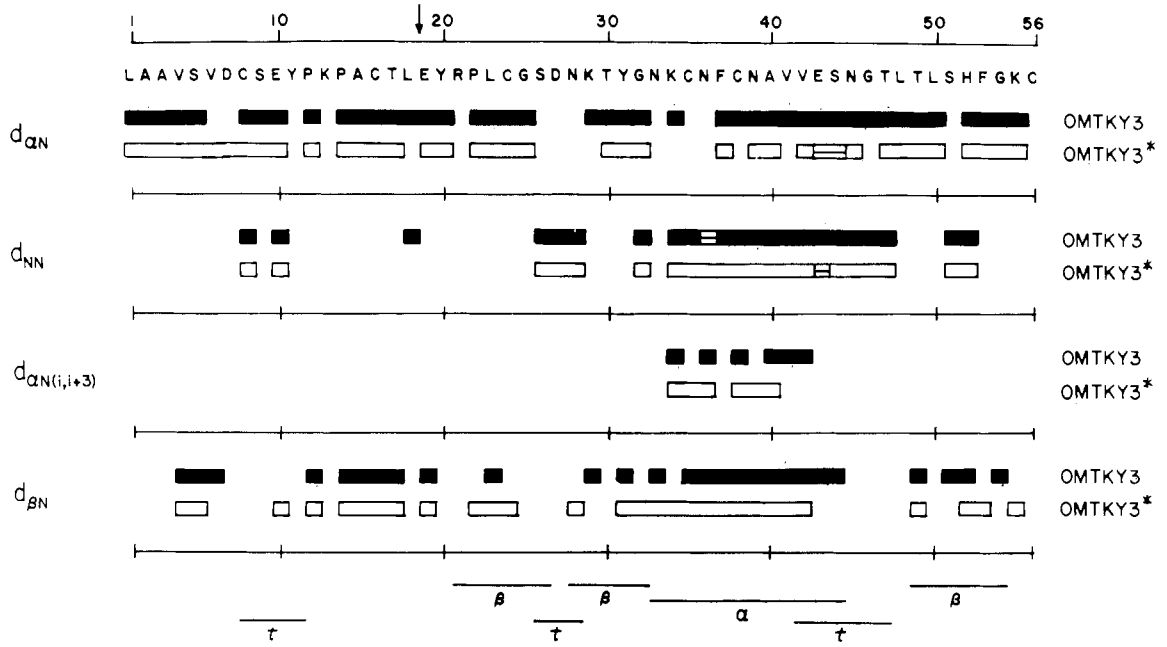

FIGURE 9: Summary of short-range NOEs observed in OMTKY3*. Corresponding NOEs observed in OMTKY3 are shown for comparison (Robertson et al., 1988). Open bars with a line in the middle represent results derived from overlapping NOESY cross-peaks; NOEs at these positions are uncertain. Secondary structural elements of OMTKY3 (in complex with *S. griseus* proteinase B) determined from single-crystal diffraction data (Fujinaga et al., 1982; Papamokos et al., 1982) are shown at the bottom of the figure.

rameters along the polypeptide chain is then indicative of particular secondary structures. For example, $d_{\alpha N}$ connectivities can be used to identify extended structure in the polypeptide chain (even though the uniqueness is very poor; Wüthrich et al., 1984). α-Helices are characterized by short NH_i–NH_{i+1} distances (<3 Å), medium-length $C_\alpha H_i$–NH_{i+3} and $C_\alpha H$–$C_\beta H_{i+3}$ distances (<3 Å), and $C_\alpha H_i$–NH_{i+1} distances of ~3.5 Å. Extended strands are characterized by very short $C_\alpha H_i$–NH_{i+1} distances (<3.0 Å). The short-range NOEs observed in OMTKY3* are summarized in Figure 9, which also presents a comparison with the results from studies of OMTKY3 [adapted from Robertson et al. (1988)]. The data indicate that residues 1–7, 13–18, and 19–21 most likely are in an extended-chain conformation, as evidenced by strong $d_{\alpha N}$ connectivities and the lack of other NOE connectivities. The NOE map also indicates a turn at residues 8–11. The peptide segment from residue 33 to residue 44 is α-helical. Slowed exchange rates for amide protons also indicate hydrogen-bonded structure (amide protons from residues 33 and 35–44 were observed in COSY spectra recorded in ^2H_2O). As shown in Figure 10, antiparallel β-structure was constructed through observed short- and medium-range NOEs. Individual strands were aligned by use of short interstrand distances between backbone protons. The β-sheet structure consists of three strands, residues 22–25, 28–32, and 49–54, and one reverse turn at residues 26–28. There is irregularity on one strand, around His^{52}, in the NOE connectivities between the strands, indicating a β-bulge. The NOE between the N*H* of His^{52} and Cys^{24} was not identified unambiguously, as the NOE may overlap with the d_{NN}-type NOE between Asn^{39} and Ala^{40} (see Table I). This type of β-bulge structure was detected in the crystal structure of OMTKY3 in complex with *S. griseus* proteinase B (Read et al., 1983) as well as in the solution structure of OMTKY3 (Robertson et al., 1988).

Cis peptide bonds can be identified by $d_{\alpha\alpha}(i,i+1)$ connectivities (Wüthrich et al., 1984). In OMTKY3*, one sequential $d_{\alpha\alpha}$ was observed between Pro^{12} and Tyr^{11}. This cis bond was also detected in OMTKY3 in the solid state (Fujinaga et al., 1982) and in solution (Robertson et al., 1988).

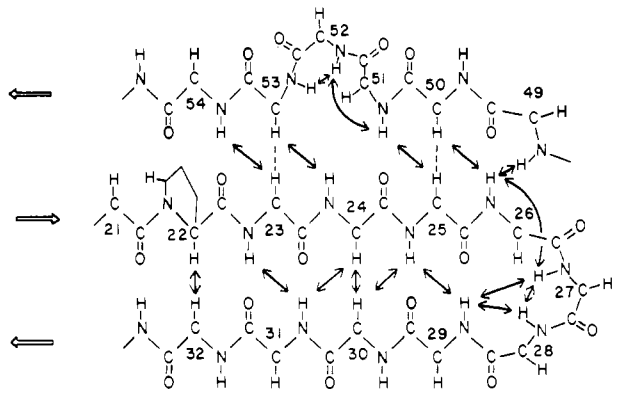

FIGURE 10: Schematic representation of antiparallel β-structure in OMTKY3* derived from the present NMR results. Observed interstrand backbone NOEs as well as the short-range d_{NN} in the turn are indicated by arrows. Note the β-bulge at residue 52 (histidine). NOEs expected from the X-ray structure (Fujinaga et al., 1982), but not detected, are indicated as dashed lines.

Extra Peaks. Some extra peaks were observed in the fingerprint region (Figure 5) as well as in the aliphatic region of the 2D spectra. Some of the extra peaks observed in COSY and NOESY spectra obtained in ^2H_2O with non-preexchanged protein were assigned to Leu^{23}, Cys^{24}, Ala^{40}, and Thr^{49} (data not shown). Those extra fingerprint peaks detected only in H_2O could not be assigned to specific residues because they lacked NOE connectivities. The origin of these extra peaks is not known. Possible explanations that need to be investigated include a second conformational state of the protein or a protein impurity as might result from an additional nick in the protein during hydrolysis.

DISCUSSION

Structure determination by NMR relies on sequence-specific assignment of a large number of resonances. Recent techniques make it possible to obtain such assignments reliably

Table I: Chemical Shifts of the Assigned ^{1}H NMR Resonances of OMTKY3*, in 0.2 M KCl, pH 4.0, 25 °C

	chemical shift in ppm from [^{2}H]TSP[a]			
residue	N*H*	$C_\alpha H$	$C_\beta H$	other
Leu[1]		4.03	1.74	$C_\gamma H$ 1.74; $C_\delta H$ 0.95, 0.97
Ala[2]	8.64	4.40	1.42	
Ala[3]	8.40	4.36	1.40	
Val[4]	8.13	4.25	2.04	$C_\gamma H$ 0.86, 0.90
Ser[5]	8.20	4.54	3.83, 3.98	
Val[6]	8.25	4.27	1.94	$C_\gamma H$ 0.77, 0.88
Asp[7]	8.62	4.77	2.70, 2.98	
Cys[8]	8.87	5.27	2.57, 3.36	
Ser[9]	8.68	4.26	3.96, 4.05	
Glu[10]	8.60	4.30	1.77, 2.04	$C_\gamma H$ 2.35
Tyr[11]	7.53	4.11	2.83, 2.86	$C_\delta H$ 7.15; $C_\epsilon H$ 7.03
Pro[12]		5.02	2.06, 2.41	$C_\gamma H$ 1.85, 1.94; $C_\delta H$ 3.52, 3.67
Lys[13]	8.89	4.82	1.54, 1.87	$C_\gamma H$ 1.16; $C_\delta H$ 1.43, 1.58; $C_\epsilon H$ 2.46, 2.58
Pro[14]		4.33	1.96, 2.30	$C_\gamma H$ 1.88, 2.04; $C_\delta H$ 3.75
Ala[15]	7.70	4.45	1.37	
Cys[16]	8.50	5.00	2.89, 3.30	
Thr[17]	8.20	4.38	4.28	$C_\gamma H$ 1.24
Leu[18]	8.09	4.23	1.62	$C_\gamma H$ 1.64; $C_\delta H$ 0.89, 0.92
Glu[19]	7.33	4.13	2.16	$C_\gamma H$ 2.40
Tyr[20]	8.98	4.93	3.05, 3.10	
Arg[21]	8.62	4.55	1.69, 1.85	$C_\gamma H$ 1.56, 2.22; $C_\delta H$ 3.15
Pro[22]		4.59	1.44, 1.94	$C_\gamma H$ 2.46; $C_\epsilon H$ 3.45
Leu[23]	8.66	4.27	0.84, 1.09	$C_\gamma H$ 1.44; $C_\delta H$ 0.38, 0.49
Cys[24]	8.16	5.23	1.37, 2.48	
Gly[25]	9.42	4.50, 4.75		
Ser[26]	9.38	4.18	3.84, 4.13	
Asp[27]	8.35	4.43	2.57, 3.00	
Asn[28]	8.60	4.42	2.81, 3.15	$N_\delta H$ 6.83, 7.49
Lys[29]	7.83	4.40	1.42, 1.68	$C_\gamma H$ 0.85, 1.15; $C_\delta H$ 1.52; $C_\epsilon H$ 2.81
Thr[30]	8.23	4.79	3.98	$C_\gamma H$ 1.12
Tyr[31]	9.80	4.58	2.90	$C_\delta H$ 7.07
Gly[32]	9.25	3.76, 3.97		
Asn[33]	7.48	4.75	3.21, 3.45	$N_\delta H$ 6.63, 7.35
Lys[34]	8.87	3.85	1.83, 1.95	$C_\gamma H$ 1.33, 1.47; $C_\delta H$ 1.80; $C_\epsilon H$ 3.06
Cys[35]	8.33	4.42	3.27, 3.36	
Asn[36]	8.54	4.54	2.87, 3.14	
Phe[37]	8.18	3.45	2.77	$C_\delta H$ 6.52; $C_\epsilon H$ 6.80; $C_\zeta H$, 6.56
Cys[38]	9.17	4.02	1.61, 1.97	
Asn[39]	8.18	4.64	2.75, 2.95	$N_\delta H$ 7.52
Ala[40]	7.24	4.13	1.28	
Val[41]	8.50	3.21	2.07	$C_\gamma H$ 0.02, 0.73
Val[42]	8.06	3.84	2.28	$C_\gamma H$ 0.96, 1.03
Glu[43]	7.76	4.25	2.09, 2.23	$C_\gamma H$ 2.57
Ser[44]	7.78	4.50	3.95, 4.20	
Asn[45]	8.43	4.48	2.87, 3.14	$N_\delta H$ 6.85, 7.54
Gly[46]	8.08	3.58, 4.10		
Thr[47]	7.55	4.20	4.28	$C_\gamma H$ 1.24
Leu[48]	7.68	4.25	1.21, 1.51	$C_\gamma H$ 1.10; $C_\delta H$ 0.29, 1.16
Thr[49]	8.56	4.68	4.33	$C_\gamma H$ 1.12
Leu[50]	8.76	4.12	0.98, 1.75	$C_\gamma H$ 0.66; $C_\delta H$ −0.02, −0.10
Ser[51]	8.93	4.49	3.40, 3.55	
His[52]	7.26	4.50	3.30, 3.78	$C_\delta H$ 7.22; $C_\epsilon H$ 8.78
Phe[53]	9.13	4.40	3.03, 3.19	$C_\delta H$ 7.23; $C_\epsilon H$ 7.15; $C_\zeta H$ 6.80
Gly[54]	8.48	3.57, 4.62		
Lys[55]	7.99	4.02	1.76, 1.84	$C_\gamma H$ 0.92, 1.50; $C_\delta H$ 1.72; $C_\epsilon H$ 3.05
Cys[56]	8.24	4.40	2.49, 3.14	

[a] Chemical shifts are measured to ±0.02 ppm.

in small proteins. Once sequence-specific resonance assignments have been obtained for the backbone protons of a polypeptide chain, the secondary structure can be identified from characteristic patterns of ^{1}H–^{1}H short distances manifested in the NOESY spectrum. In practice, this approach has been applied successfully to identify antiparallel β-sheet and α-helices in several small proteins (Wemmer & Kellenbach, 1983; Zuiderweg et al., 1983; Williamson et al., 1984; Kline & Wüthrich, 1985; Klevit & Waygood, 1986; Van de Ven & Hilbers, 1986; Clore et al., 1986a; Sukumaran et al., 1987). In this paper, we have used various 2D NMR techniques to achieve sequence-specific resonance assignments of the ^{1}H NMR spectrum and to determine the secondary structure of modified turkey ovomucoid third domain (OMTKY3*).

The assigned backbone and side-chain proton resonances in OMTKY3* are tabulated in Table I. The assignments include all backbone proton resonances except the two N-terminal amide protons and over 80% of the side-chain proton resonances. Under the conditions of the data collection (0.2 M KCl, pH 4.0), all the N*H* and $C_\alpha H$ resonances were separated well enough for facile assignment. The $C_\beta H$ chemical shifts are close to those obtained with model peptides (Bundi & Wüthrich, 1979) with a few exceptions, e.g., $C_\beta H$'s of Pro[22], Cys[24], Cys[38], and Ser[51]. The observed deviations in chemical shift probably result from ring current anisotropy, since these residues are sequentially adjacent to aromatic residues, Tyr[21], Phe[37], and His[52]. Similar deviations were observed in the chemical shifts of OMTKY3 (Robertson et al., 1988).

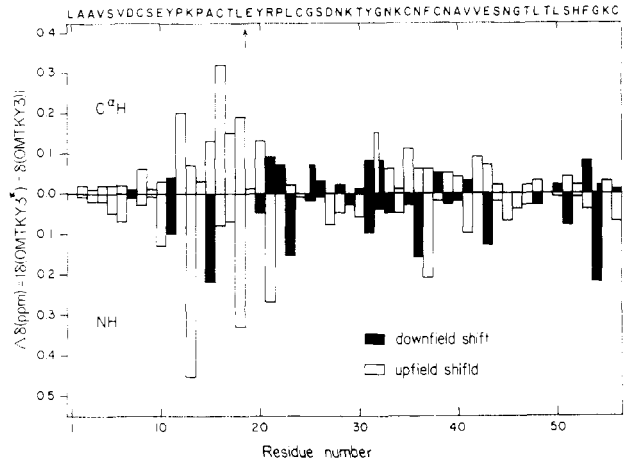

FIGURE 11: Backbone chemical shift differences between virgin and modified OMTKY3. Filled bars represent downfield shifts occurring upon modification; open bars represent upfield shifts upon modification. Since the chemical shifts were measured with an accuracy of ±0.02 ppm, differences of ±0.04 ppm represent experimental error. Chemical shifts of the two $C_{\alpha}H$'s in glycine residues are compared separately by use of two bars of half-width.

Chemical shifts of backbone protons in OMTKY3* are compared with those in OMTKY3 in Figure 11. Whereas ^1H chemical shifts are of limited use for investigations of unknown protein conformation, they are very sensitive probes of conformational changes (Pardi et al., 1983). The largest chemical shift differences occur mainly in residues that comprise the loop (Fujinaga et al., 1982) at the modification site. Significant chemical shift differences occur, however, in regions of the protein remote from the modification site. Chemical shift changes outside the reactive site loop can be rationalized by reference to the X-ray structure of virgin third domain (Fujinaga et al., 1982). Residues 31, 32, and 33 are hydrogen bonded to the loop, which may explain the chemical shift changes in this region. In particular, the side-chain N*H*'s of Asn^{33} are hydrogen bonded to the carbonyl oxygens of Glu^{19} and Thr^{17}. Cys^{35} is linked to the loop through a disulfide bond to Cys^{16}, a residue on the loop near the reactive site. Moreover, the N*H* of Asn^{36} is hydrogen bonded to the O^{δ} of Asn^{33}. Chemical shift changes around Val^{41}, Val^{42}, and Glu^{43} might result from alteration in the ring current effect of Phe^{37} resulting from a slight conformational change of the ring. In fact, an NOE was observed from Phe^{37} ring protons to the $C_{\gamma}H$'s of Val^{41} (data not shown). Ser^{51}, Phe^{53}, and Gly^{54} are hydrogen bonded to the central β-strand from Leu^{23} to Gly^{25} (see Figure 10). His^{52}, which forms the β-bulge (Figure 10), does not show a significant chemical shift difference upon modification. All the observed chemical shift changes cannot be explained, however, in the absence of a highly refined three-dimensional structure of the modified third domain.

By contrast, the NOE map (Figure 9) shows that virtually identical short-range NOEs were observed in OMTKY3* and OMTKY3. This indicates close similarity of the local backbone conformations along the entire sequence in the two proteins. Also, the secondary structure of OMTKY3 (and OMTKY3*), determined by NMR data (Robertson et al., 1988), is in close agreement with that observed by crystallography (secondary structural elements observed in single crystals are given at the bottom of Figure 9). In some respects, the solution structures of OMTKY3* and OMTKY3 appear more similar to each other than to the crystal structures. The amide protons of Asp^{27}, Asn^{33}, and Lys^{55} in OMTKY3 and OMTKY3* are in slow exchange in solution, but no possibility of hydrogen bonding of these protons is indicated in the crystal data (Fujinaga et al., 1982; Papamokos et al., 1982; Bode et al., 1985).

Only slowly exchanging amide protons can be detected in ^2H_2O solution, i.e., those from buried or hydrogen-bonded residues (Wlodawer & Sjölin, 1982). With non-preexchanged protein in ^2H_2O, backbone amides were observed from residues 23–31, 33, 36–44, 48–49, 51, and 54–56 of OMTKY3*. These slowly exchanging amide protons are located in turns, α-helix, and antiparallel β-strands (Figure 9). In OMTKY3, the amide protons from a slightly different set of residues (11, 13, 21, 23–29, 31, 33, 35–43, 49, 51) were observed in ^2H_2O (Robertson et al., 1988). Among these, the amide protons from residues 11, 13, 21, and 35 in OMTKY3 (those not observed in OMTKY3* in ^2H_2O) exchange faster than the others (A. D. Robertson and J. L. Markley, unpublished results). It is thus interesting that the N*H*'s that shift the farthest upon conversion of OMTKY3 to OMTKY3* are those of Lys^{13} and Arg^{21} (Figure 11). Moreover, the direction in which they shift is consistent with loss or weakening of hydrogen bonds involving these N*H* groups upon modification, although the possibility of ring current shifts by Tyr^{11} and Tyr^{20} cannot be ruled out entirely at present. These results suggest that hydrolysis of the reactive site peptide bond destabilizes the rigid hydrogen-bonded reactive site conformation. The N*H* of Arg^{21} participates in the construction of the rigid reactive site conformation of the virgin inhibitor (Read et al., 1983) by forming a hydrogen bond to the carbonyl oxygen of Gly^{32}. On the pathway of reverse hydrolysis, the nucleophilic Ser^{195} O^{γ} of the serine proteinase attacks the carboxyl carbon of Leu^{18} to form the tetrahedral intermediate. In OMTKY3, hydrogen bonding of the side chain of Asn^{33} to the carbonyl oxygen atoms of Thr^{17} and Glu^{19} (Read et al., 1983) would oppose movement of the carbon of Leu^{18} toward Ser^{195} O^{γ} on the enzyme. In fact, there is an indication that the hydrogen-bonded reactive site conformation involving Asn^{33} is modified by the cleavage at the reactive site peptide bond: (1) the backbone N*H* of Asn^{33} exchanges much faster with solvent molecule in OMTKY3* than in OMTKY3 (G. I. Rhyu, A. D. Robertson, and J. L. Markley, unpublished results), and (2) the side-chain $N_{\delta}H$'s of Asn^{33} shift from 6.32 and 8.11 ppm before cleavage to 6.63 and 7.35 ppm after cleavage. Hydrogen bonding around the reactive site is being investigated further through more detailed studies of amide proton exchange rates. If the rigid structure at the reactive site of the inhibitor obtains local freedom by loosening or breaking the hydrogen bonds after cleavage, the carboxyl carbon of Leu^{18} can move toward Ser^{195} O^{γ} rather easily to synthesize the acyl-enzyme intermediate which then reacts with the amino group on Glu^{19} to resynthesize the peptide bond.

Even though the secondary structure of OMTKY3 does not change after reactive site peptide bond cleavage, widespread changes in backbone proton chemical shifts indicate differences in the tertiary structure. The magnitude of the changes, in terms of atom displacements, cannot be gauged from chemical shift data. With the present assignments in hand, we plan to measure long-range NOEs in order to compute the three-dimensional structures of OMTKY3 and OMTKY3* by using a combination of distance geometry (Havel & Wüthrich, 1985) and restrained molecular dynamics calculations (Clore et al., 1986b). This comparison may further elucidate structure–function relationships in Kazal-type proteinase inhibitors.

ACKNOWLEDGMENTS

We are grateful to Dr. W. Ardelt for assaying the protein modification reaction by analytical anion-exchange column chromatography, and we thank Dr. W. M. Westler and A. D.

Robertson for helpful discussions.

REFERENCES

Ardelt, W., & Laskowski, M., Jr. (1982) *Anal. Biochim. 120*, 198–203.

Ardelt, W., & Laskowski, M., Jr. (1983) *Acta Biochim. Pol. 30*, 115–126.

Ardelt, W., & Laskowski, M., Jr. (1985) *Biochemistry 24*, 5313–5320.

Baillargeon, M. W., Laskowski, M., Jr., Neves, D. F., Porubcan, M. A., Santini, R. E., & Markley, J. L. (1980) *Biochemistry 19*, 5703–5710.

Bax, A., & Freeman, R. (1981) *J. Magn. Reson. 44*, 542–561.

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

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

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

Bode, W., Epp, O., Huber, R., Laskowski, M., Jr., & Ardelt, W. (1985) *Eur. J. Biochem. 147*, 387–395.

Bode, W., Wei, A.-Z., Huber, R., Meyer, E., Travis, J., & Neumann, S. (1986) *EMBO J. 5*, 2453–2458.

Bogard, W. C., Jr., Kato, I., & Laskowski, M., Jr. (1980) *J. Biol. Chem. 255*, 6569–6574.

Bundi, A., & Wüthrich, K. (1979) *Biopolymers 18*, 285–279.

Clore, A. M., Martin, S. R., & Gronenborn, A. M. (1986a) *J. Mol. Biol. 191*, 553–561.

Clore, A. M., Brünger, A. T., Karplus, M., & Gronenborn, A. M. (1986b) *J. Mol. Biol. 191*, 523–551.

Fujinaga, M., Read, R. J., Sielecki, A., Ardelt, W., Laskowski, M., Jr., & James, M. N. G. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 4868–4872.

Fujinaga, M., Sielecki, A. R., Read, R. J., Ardelt, W., Laskowski, M., Jr., & James, M. N. G. (1987) *J. Mol. Biol. 195*, 397–418.

Havel, T. F., & Wüthrich, K. (1985) *J. Mol. Biol. 182*, 281–294.

Hoult, D. I., & Richards, R. E. (1975) *Proc. R. Soc. London, A 34*, 311–340.

Hunkapiller, M. W., Forgae, M. D., Yu, E. H., & Richards, J. H. (1979) *Biochem. Biophys. Res. Commun. 87*, 25–31.

Kainosho, M. (1985) in *Protein Protease Inhibitor—The Case of Streptomyces Subtilisin Inhibitor (SSI)* (Hiromi, K., et al., Eds.) pp 326–328, Elsevier, New York.

Klevit, R. E., & Waygood, E. B. (1986) *Biochemistry 25*, 7774–7781.

Kline, A. D., & Wüthrich, K. (1985) *J. Mol. Biol. 183*, 503–507.

Kumar, A., Wagner, G., Ernst, R. R., & Wüthrich, K. (1981) *J. Am. Chem. Soc. 103*, 3654–3658.

Laskowski, M., Jr., & Kato, I. (1980) *Annu. Rev. Biochem. 49*, 593–627.

Laskowski, M., Jr., Kato, I., Ardelt, W., Cook, J., Denton, A., Empie, M. W., Kohr, W. J., Park, S. J., Parks, K., Schatzley, B. L., Schoenberger, O. L., Tashiro, M., Vichot, G., Whatley, H. E., Wieczorek, A., & Wieczorek, M. (1987) *Biochemistry 26*, 202–221.

Lineweaver, H., & Murray, C. W. (1947) *J. Biol. Chem. 171*, 565–581.

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

Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967–974.

Ortiz-Polo, G. (1986) M.S. Thesis, Purdue University.

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

Pardi, A., Wagner, G., & Wüthrich, K. (1983) *Eur. J. Biochem. 137*, 445–454.

Pardi, A., Billeter, M., & Wüthrich, K. (1984) *J. Mol. Biol. 180*, 741–751.

Rance, M., Sørensen, O. W., Bodenhausen, G., Ernst, R. R., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 117*, 479–485.

Read, R. J., Fujinaga, M., Sielecki, A. R., & James, M. N. G. (1983) *Biochemistry 22*, 4420–4433.

Richarz, R., Tschesche, H., & Wüthrich, K. (1980) *Biochemistry 19*, 5711–5715.

Robertson, A. D., Westler, W. M., & Markley, J. L. (1988) *Biochemistry* (preceding paper in this issue).

Sukumaran, D. K., Clore, M., Press, A., Zarbock, J., & Gronenborn, A. M. (1987) *Biochemistry 26*, 333–338.

Van de Ven, F. J. M., & Hilbers, C. W. (1986) *J. Mol. Biol. 192*, 419–441.

Wagner, G. (1983) *J. Magn. Reson. 55*, 151–156.

Wagner, G., Neuhaus, D., Wörgötter, E., Vasák, M., Kägi, J. H. R., & Wüthrich, K. (1986) *Eur. J. Biochem. 157*, 275–289.

Wemmer, D., & Kallenbach, N. R. (1983) *Biochemistry 22*, 1901–1906.

Wider, G., Macura, S., Kumar, A., Ernst, R. R., & Wüthrich, K. (1984) *J. Magn. Reson. 56*, 207–234.

Williamson, M. P., Marion, D., & Wüthrich, K. (1984) *J. Mol. Biol. 173*, 341–360.

Wlodawer, A., & Sjölin, L. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 1418–1422.

Wüthrich, K., Billeter, M., & Braun, W. (1984) *J. Mol. Biol. 180*, 715–740.

Zolnai, Zs., Macura, S., & Markley, J. L. (1986) *Comput. Enhanced Spectrosc. 3*, 141–145.

Zuiderweg, E. R. P., Kaptein, R., & Wüthrich, K. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 5837–5841.