Two-Dimensional NMR Studies of Kazal Proteinase Inhibitors. 1. Sequence-Specific Assignments and Secondary Structure of Turkey Ovomucoid Third Domain†

Andrew D. Robertson, William M. Westler, 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: Two-dimensional proton NMR experiments have been used to sequentially assign resonances to all of the peptide backbone protons of turkey ovomucoid third domain (OMTKY3) except those of the N-terminal α-amino group whose signal was not resolved owing to exchange with the solvent. Assignments also have been made for more than 80% of the side-chain protons. Two-dimensional chemical shift correlated spectroscopy (COSY), relayed coherence transfer spectroscopy (RELAY), and two-dimensional homonuclear Hartmann–Hahn spectroscopy (HOHAHA) were used to identify the spin systems of almost half of the residues prior to sequential assignment. Assignments were based on two-dimensional nuclear Overhauser enhancements observed between adjacent residues. The secondary structure of OMTKY3 in solution was determined from additional assigned NOESY cross-peaks; it closely resembles the secondary structure determined by single-crystal X-ray diffraction of OMTKY3 in complex with *Streptomyces griseus* proteinase B [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]. The NMR data provide evidence for three slowly exchanging amide protons that were not identified as hydrogen-bond donors in the crystal structure.

Ovomucoid, the second most abundant protein in egg white, is a Kazal-type serine proteinase inhibitor. It is made up of three homologous domains, each of which is a potential inhibitor. The three domains can be cut apart by limited proteolysis of the linker peptides and isolated. The third domains are the most extensively studied of the three since they are easily isolated, generally in both glycosylated and nonglycosylated forms (Laskowski et al., 1987). Investigations of avian ovomucoid third domains have encompassed assays of inhibitory activity (Empie & Laskowski, 1982; Ardelt & Laskowski, 1983), determination of crystal structures (Papamakos et al., 1982; Fujinaga et al., 1982, 1987; Bode et al., 1985, 1986), and measurements of physical properties in solution (Matsuda et al., 1981; Watanabe et al., 1981; Ogino et al., 1982). Sequences for over 100 third domains have been determined, and 65 of these sequences are unique (Laskowski et al., 1987). Of these, 37 pairs differ by single amino acid substitutions. Its abundance, small size, well-characterized function, and naturally occurring variety of sequences have made ovomucoid third domain an excellent subject for solution-state NMR[1] studies of the effects of amino acid substitutions on protein structure and dynamics (Markley et al., 1986). As our starting point, we have chosen ovomucoid third domain from the domestic turkey (OMTKY3). Its amino acid sequence places it in the midst of a family of 12 naturally occurring third domains related by single amino acid substitutions [see Figure 5 of Laskowski et al. (1987)]. Its X-ray structure has been determined in complexes with proteinase B from *Streptomyces griseus* (Fujinaga et al., 1982), human leukocyte elastase (Bode et al., 1986), and chymotrypsin (Fujinaga et al., 1987).

Detailed study of protein structure in solution requires extensive assignment of proton resonances. To this end we have employed the sequential assignment technique introduced by Wüthrich and co-workers (Wagner & Wüthrich, 1982). We present here results from 2D NMR experiments that have allowed us to classify proton resonances by amino acid types and, subsequently, to assign them to specific residues in the amino acid sequence of OMTKY3. In this way, we have been able to assign peptide backbone proton resonances to all 56 residues. These NMR data also permit identification of β-sheet and α-helix in OMTKY3. We compare these secondary structural elements with those determined from X-ray data on OMTKY3 complexed with serine proteinases (Fujinaga et al., 1982, 1987; Bode et al., 1986).

The assignments described herein were achieved independently of those presented previously for OMTKY3 (Markley et al., 1984, 1986; Westler et al., 1984; Ortiz-Polo, 1985; Ortiz-Polo et al., 1986), which were based on comparison

†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: BUSI IIA, bull seminal plasma inhibitor IIA; BPTI, bovine pancreatic trypsin inhibitor; COSY, two-dimensional correlated spectroscopy; 2D, two dimensional; HOHAHA, two-dimensional homonuclear Hartmann–Hahn spectroscopy; DQF-COSY, double quantum filtered COSY; GlcNAc, *N*-acetylglucosamine; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, two-dimensional NOE spectroscopy; OMTKY3, turkey ovomucoid third domain; OMTKY3*, modified turkey ovomucoid third domain (reactive site peptide bond, Leu[18]–Glu[19], cleaved); pH*, glass electrode pH reading uncorrected for deuterium isotope effect; RELAY, two-dimensional relayed coherence transfer spectroscopy; [^{2}H]TSP, sodium 3-(trimethylsilyl)[2,2,3,3-$^2H_4$]propionate. The amino acid protons are designated according to IUPAC–IUB convention (IUPAC–IUB Commission on Biochemical Nomenclature, 1970). For example, $C_\alpha H$ is the proton attached to the main-chain α-carbon. The chemical shift positions of cross-peaks in two-dimensional NMR spectra are reported as (ω_2, ω_1) ppm in the text, unless otherwise indicated.

0006-2960/88/0427-2519$01.50/0

of spectra of variant ovomucoid third domains. The present assignments have been made in a more complete and rigorous fashion. All assignments in the methyl region (Westler et al., 1984; Ortiz-Polo, 1985) have been confirmed. We have corrected a number of previous assignments in other regions of the OMTKY3 spectrum that appear to have been incorrect.

A further objective of our work, in collaboration with Laskowski and co-workers, is a more complete understanding of the mechanism of inhibition of serine proteinases by ovomucoid third domain. Each third domain has a reactive site peptide bond which is a potential substrate for serine proteinases. Enzymatic cleavage of this site is reversible. This suggests, along with the observed K_{hyd} near unity for most third domains, that third domain structure may be largely retained upon reactive site cleavage (Ardelt & Laskowski, 1983). There is, however, little direct physical evidence for this. The following paper (Rhyu & Markley, 1988) describes ^1H NMR resonance assignments and secondary structure determination for modified turkey ovomucoid third domain (OMTKY3*), the form having a cleaved reactive site peptide bond. The results indicate that the secondary structure is preserved on reactive site peptide bond hydrolysis but that minor conformational changes are propagated throughout the protein molecule.

Materials and Methods

Materials. Domestic turkey (*Maleagris gallopavo*) eggs were obtained from Wolff's Bellefontaine Farms (Belleville, WI). *Staphylococcus aureus* V8 proteinase was purchased from Miles Laboratories (Naperville, IL) as the lyophilized powder. It was used without further purification. Ultrapure KCl was from Alfa Products (Danver, MA). Deuterium oxide was from Aldrich Chemical Co. (Milwaukee, WI) (99.8 atom % ^2H) and from Stohler Isotope Chemicals (Waltham, MA) (99.996 atom % ^2H). ^2HCl (99 atom % ^2H, 35% w/w in ^2H_2O), KO^2H (99 atom % ^2H, 40% w/w in ^2H_2O), and [^2H]TSP were purchased from MSD Isotopes (St. Louis, MO). All other chemicals were of reagent grade or better.

Protein Purification. Nonglycosylated OMTKY3 was purified by the method of Bogard et al. (1980) with modifications as described by Ortiz-Polo (1985).

NMR Spectroscopy. All samples were 10–15 mM in protein and 0.2 M in KCl, with the pH adjusted to 4.0–4.2 with dilute ^2HCl, KO^2H, HCl, or KOH. Samples dissolved in H_2O contained 10% (v/v) ^2H_2O for the lock. All pH values for samples in ^2H_2O are reported without correction. Data were acquired on samples maintained at 25 °C, and chemical shifts are reported relative to [^2H]TSP.

Absolute value mode COSY (Aue et al., 1976) data were acquired with a 90°–t_1–60°–acquire pulse sequence to minimize diagonal peaks while maximizing cross-peak intensities (Bax & Freeman, 1981) and with phase cycling for quadrature detection and suppression of axial and P-type peaks (Wider et al., 1984). Absolute value RELAY experiments (Wagner, 1983a) were performed with the eight-step phase cycling described by Bax and Drobny (1985) plus additional CYCLOPS cycling (Hoult & Richards, 1975), which gave a total of 32 steps. DQF-COSY (Rance et al., 1983), NOESY (Kumar et al., 1980), and HOHAHA (Bax & Davis, 1985) data were acquired in pure absorption mode with time-proportional phase incrementation (Marion & Wüthrich, 1983). Phase cycling for the DQF-COSY experiment was as described by Rance et al. (1983). The NOESY mixing time, usually 100 ms, was varied randomly by ±10% to suppress scalar coupling effects (Macura et al., 1982), and the phase cycling was that described by Bodenhausen et al. (1984). CYCLOPS phase cycling was used for the HOHAHA experiments. The mixing times used in RELAY and HOHAHA experiments are indicated in the figures.

Data were acquired on Bruker AM-500 and AM-400 spectrometers equipped with Aspect 3000 computers and digital phase shifters. HOHAHA data were acquired with the proton decoupler channel for proton excitation. The decoupler channel and the receiver were referenced to a common frequency. In all experiments the carrier frequency was set in the middle of the spectrum. Sweep widths were 6024 Hz on the AM-500 and 4808 Hz on the AM-400. Typical acquisitions in ^2H_2O consisted of 256–512 blocks of 32 summed transients, each transient being 2048 time-domain data points in size; 64 transients were collected for experiments in H_2O. Solvent suppression was achieved by irradiation at the solvent frequency during all times except acquisition. The relaxation delay was 1.4–2.0 s.

All data were processed on a Bruker offline data processing station equipped with an Aspect 3000 computer and an array processor. Most data sets were zero filled to 2048 × 2048 or 4096 × 2048 ($t_2 \times t_1$) data points before transformation. Absolute value mode COSY and RELAY and the pure-phase HOHAHA maps were symmetrized after transformation. Apodization before Fourier transformation in the t_2 dimension consisted of a sine bell function with a phase shift of $\pi/64$ for absolute value COSY, $\pi/32$ for pure-phase COSY, $\pi/8$ for NOESY, and $\pi/8$ for HOHAHA and with no phase shift for the RELAY experiment. Sine bell multiplication was performed in t_1 on absolute value COSY, pure-phase NOESY, RELAY, and HOHAHA data with phase shifts of $\pi/32$, $\pi/4$, 0, and $\pi/8$, respectively. A Gaussian weighting function, with a maximum at 0.2 times the time-domain data size and a line broadening of −20 Hz, was applied to the t_1 dimension of the pure-phase COSY data.

Calculation of Protein Static Solvent Accessibility. The solvent accessibility of atoms in OMTKY3 was calculated according to the algorithm of Connolly (1981) with the molecular surface (ms) utility of the MIDAS molecular graphics system (Langridge et al., 1981) installed on a Silicon Graphics, Inc., IRIS 2400 computer. Protons were indirectly included in the calculation by expanding the van der Waals radii of the carbon, nitrogen, and oxygen atoms. The water probe radius was 1.4 Å. The X-ray crystallographic coordinate data, obtained from the Brookhaven National Laboratory Protein Data Bank, were those for OMTKY3 in complex with *Streptomyces griseus* proteinase B (Read et al., 1983).

Results and Discussion

Sequential Assignment Strategy. The first stage of the sequential assignment method is the identification of scalar coupling networks, or spin systems, corresponding to different amino acid types (Wagner & Wüthrich, 1982). Particular emphasis is placed on identifying the backbone N*H* and $C_\alpha H$ resonances, as these are used for sequence-specific assignment in the second stage of assignment. In addition, these resonances and the NOESY data can be used to determine protein secondary and tertiary structure.

Many amino acids have characteristic spin systems owing to both the chemical structure of the side chain (bonded arrangement of protons) and the chemical shifts of these protons. For proteins with overlapping proton resonances, the COSY experiment facilitates identification of protons coupled through two and three bonds. The COSY map contains a diagonal, representing the one-dimensional spectrum, and off-diagonal cross-peaks, which result from scalar coupling of protons whose chemical shifts describe the position of the cross-peak. When

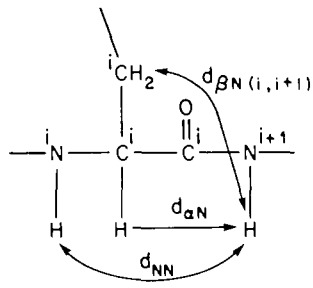

FIGURE 1: NOESY connectivities used for sequential assignment and secondary structure determination. Additional connectivities that are not shown but that are used for determination of secondary structure are those between the $C_\alpha H$ of residue i and the NH's of residues $i + 2$, $i + 3$, and $i + 4$. Shorthand notations for these connectivities are $d_{\alpha N}(i, i + 2)$, $d_{\alpha N}(i,i + 3)$, and $d_{\alpha N}(i,i + 4)$, respectively.

the COSY map is complicated by overlap, other 2D NMR experiments such as RELAY and HOHAHA may allow resolution of ambiguities. However, none of these experiments permits connection of sequentially neighboring residues, since each residue acts as an isolated spin system. This is accomplished in the next stage of assignment, in which NOESY data, first-stage assignments, and protein sequence data are utilized to deduce sequence-specific backbone NH and $C_\alpha H$ assignments. The NOESY experiment relies on dipolar coupling connectivities, which are a function of through-space interproton distance. Proteins can exhibit a number of interresidue NOESY connectivities, some of which are diagrammed in Figure 1, that can link elements of the spin systems of sequentially adjacent residues. Valid links must connect spin systems consistent with the primary amino acid sequence.

Identification of Amino Acid Side-Chain Spin Systems. (*1*) *Glycine Residues.* Two of the four glycine spin systems of OMTKY3 were identified by noting the characteristic pairs of ($C_\alpha H$, NH) cross-peaks in the COSY map. One was observed only in ^2H_2O (NH = 9.40 ppm), since its coupled $C_\alpha H$ resonances (4.68 and 4.50 ppm) were bleached during irradiation of the nearby water peak in H_2O. The other glycine spin system was observed in H_2O (NH = 8.12, $C_\alpha H$ = 4.10 and 3.58 ppm).

(*2*) *Methyl-Bearing Amino Acids.* Amino acids with side-chain methyl groups are, aside from glycine residues, perhaps the easiest to identify in the first stage of assignment. Each of these amino acid types has a unique spin system that may be resolved.

OMTKY3 contains four alanine, five leucine, four threonine, and four valine residues, having a total of 26 methyl groups. The eight intense alanine ($C_\alpha H$, $C_\beta H$) and threonine ($C_\beta H$, $C_\gamma H$) cross-peaks are located between 1.00 and 1.50 ppm in the ω_2 dimension and 4.00 and 5.00 ppm in the ω_1 dimension (Figure 2A). The protons of the two methyl groups of leucine and valine residues are coupled to a single methine proton. These should yield a total of nine pairs of intense methyl cross-peaks, with each pair aligned along a methine proton resonance frequency. Eight of the nine pairs of leucine and valine methyl cross-peaks are observed in the high-field region of the spectrum in Figure 2A. Two of these pairs, at approximately (0.9, 1.7) ppm, overlap partially. The ninth pair was located indirectly through the following reasoning.

The COSY cross-peak at (0.30, 1.14) ppm (Figure 2A) was assigned tentatively to a valine or leucine spin system since alanine and threonine cross-peaks usually do not fall in this region and since OMTKY3 does not contain an isoleucine residue. The signal at 0.30 ppm is a well-resolved doublet in the one-dimensional ^{1}H NMR spectrum with an intensity equivalent to that of three protons, indicative of a methyl

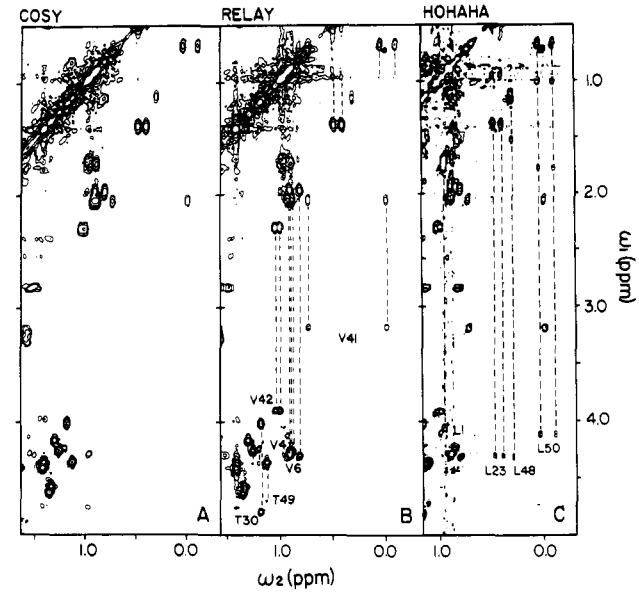

FIGURE 2: (A) Expanded region of the COSY map of OMTKY3. (B) Expansion of the analogous region of the RELAY map. The mixing time was set at 25 ms. (C) Expansion of the HOHAHA map, with a mixing time of 101 ms. All samples were 15 mM OMTKY3–0.2 M KCl, dissolved in ^2H_2O. The pH* was adjusted to 4.0, and the temperature was maintained at 25 °C.

resonance. It follows then that the resonance at 1.14 ppm is that of the methine proton and that the second methyl resonance is close to or overlaps with this signal. This brings our methyl resonance count to 26.

RELAY and HOHAHA experiments were used in the next step toward identification of methyl-containing residues. Additional peaks in the RELAY map arise from pairs of protons that share scalar coupled protons but are not themselves coupled. For example, ($C_\alpha H$, $C_\gamma H$) and (NH, $C_\beta H$) cross-peaks are expected in the RELAY map, the connectivities being mediated by the $C_\beta H$ and $C_\alpha H$, respectively. The data in Figure 2B suffice to identify all the valine $C_\alpha H$, $C_\beta H$, and $C_\gamma H$ resonances along with ($C_\alpha H$, $C_\gamma H$) connectivities in the region around (1.00, 4.50) ppm from two threonine residues. There is, however, a weak RELAY cross-peak at (1.20, 4.36) ppm in Figure 2B which appears to originate from a fifth threonine residue. Further evidence for this identification is discussed below.

Because leucine methyl resonances have RELAY connectivities only to $C_\beta H$ resonances, $C_\alpha H$ resonances could not be identified directly with this experiment. This information was provided by the HOHAHA experiment in which magnetization is transferred from one spin to all other spins within a scalar coupling network, often resulting in a map with strings of cross-peaks representing the entire spin system. In the HOHAHA map of OMTKY3, the $C_\delta H$'s for four of the five leucines show connectivities to their $C_\alpha H$ resonances and to some $C_\beta H$ resonances.

Up to this point we had identified $C_\alpha H$ resonances for all four valine residues, four of the five leucine residues, two of the four expected threonine residues, and the minor "extra" threonine residue. To identify the NH resonances for these spin systems and to distinguish the remaining unidentified threonine spin system from those of the alanines, we used the COSY and RELAY data for OMTKY3 dissolved in H_2O and focused on the NH region of the latter map. As indicated in Figure 3, all alanine and valine NH resonances were linked easily via their RELAY connectivities to $C_\beta H$ resonances. The valine NH resonating at lowest field shows a weak connectivity to the $C_\beta H$ in this map, and the ($C_\alpha H$, NH) COSY-type

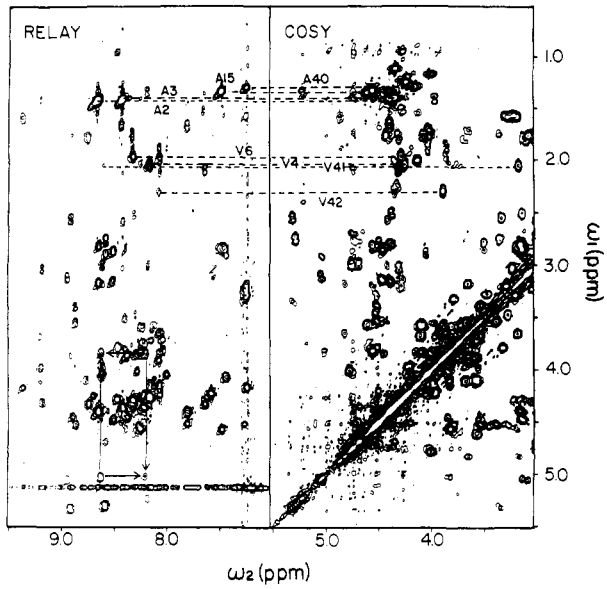

FIGURE 3: Expanded regions of RELAY and COSY maps. RELAY data were acquired on OMTKY3 dissolved in H_2O. The mixing time was 40 ms. The COSY data were acquired under the conditions described for Figure 2.

cross-peak in the RELAY map at (8.60, 3.18) ppm confirms this identification. The same reasoning was used in the case of all ambiguities in the RELAY map, as the COSY peaks usually were not suppressed completely in this experiment.

Because of the strong RELAY ($C_\beta H$, NH) connectivities for alanine, we were able to identify unambiguously all four alanine spin systems in OMTKY3. By elimination, we could identify all remaining threonine spin systems (including that containing the extra methyl peak), or at least their $C_\beta H$ and $C_\gamma H$ components. The putative extra threonine shows ($C_\beta H$, $C_\gamma H$) connectivities, at (1.20, 4.24) ppm (Figure 2A) that are of lower intensity than the others. It clearly shows a RELAY connectivity to 4.36 ppm, indicating that it is indeed a threonine residue (Figure 2B). Two of the four other putative threonine spin systems did not show RELAY connectivities, as can result if components of the spin system overlap. The sequential assignments (to be discussed later) confirmed that the $C_\alpha H$ and $C_\beta H$ resonances for these residues do overlap.

We were now left with identifying the leucine and threonine NH resonances. From the RELAY and HOHAHA data discussed earlier, we identified $C_\beta H$ resonances for three of the five leucine residues. Using this information and focusing again on the NH region of the RELAY map, we were able to identify NH resonances for two of these leucine residues, at 8.48 and 7.65 ppm, via ($C_\beta H$, NH) connectivities. The third NH resonance was located through a well-resolved COSY ($C_\alpha H$, NH) cross-peak, at (8.77, 4.10) ppm. The fourth leucine, for which a $C_\alpha H$ resonance was identified at 4.05 ppm, does not have an observable ($C_\alpha H$, NH) COSY cross-peak.

Identification of three of the five threonine NH resonances was made with the COSY and RELAY data. An NH resonance at 8.56 ppm was identified as a threonine through a ($C_\alpha H$, NH) cross-peak, at (8.56, 4.68) ppm, in the COSY map of third domain dissolved in ^2H_2O. A second threonine NH resonance at 8.29 ppm was located in the RELAY map via a ($C_\beta H$, NH) connectivity at (8.29, 4.02) ppm. A third NH resonance (assigned below to the putative extra threonine) was identified with COSY and RELAY data, in which the ($C_\alpha H$, NH) and ($C_\beta H$, NH) connectivities are observed at 8.22 ppm.

We thus had identified $C_\alpha H$ and NH resonances for all eight valine and alanine residues, three of the five leucine residues,

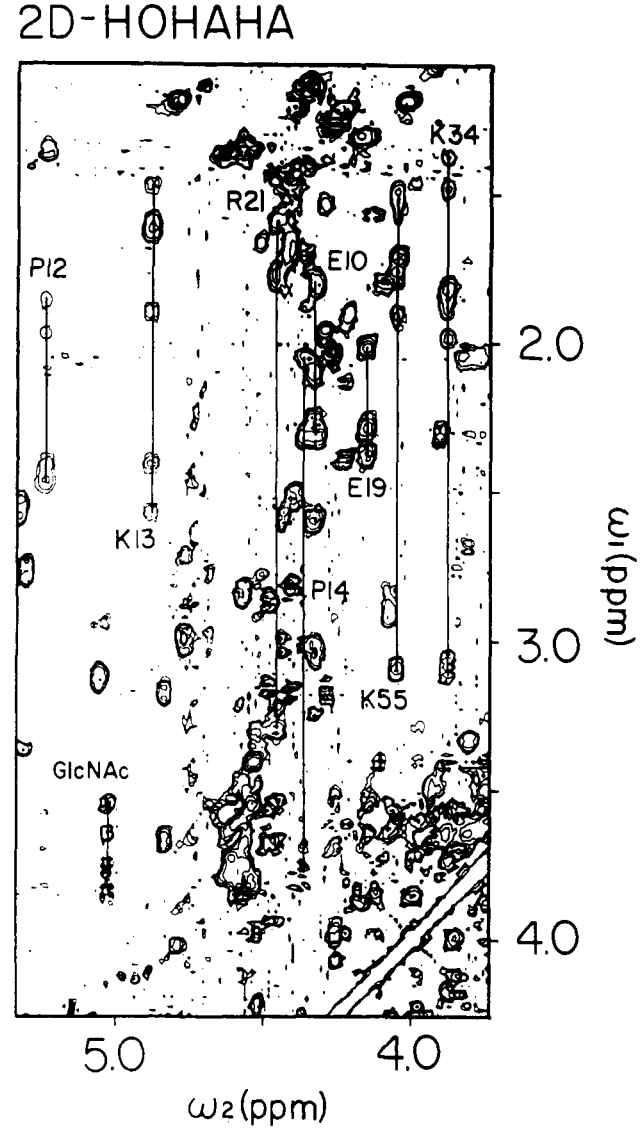

FIGURE 4: Expanded region of the HOHAHA map of OMTKY3. Experimental conditions were as described for Figure 2.

and three of the five threonine residues in OMTKY3 prior to sequence-specific assignments. A $C_\alpha H$ resonance for a fourth leucine residue was identified as well.

(*3*) *Long Side Chain Amino Acids.* In favorable cases, the long side chain amino acid spin systems can be elucidated with COSY and RELAY data alone, but extensive overlap in the high-field region of the OMTKY3 spectra made this difficult. We chose to rely on the HOHAHA experiment as a simpler way to make these identifications. Because we had characterized the spin systems of most of the methyl-containing amino acids, the only remaining amino acid spin systems with more than two scalar-coupled side-chain protons in OMTKY3 were those belonging to one arginine, three glutamate, four lysine, and three proline residues. Thus, if more than two connectivities to a $C_\alpha H$ resonance were observed in the aliphatic region of the HOHAHA map, the spin system had to belong to one of the above. The $C_\beta H$ protons of the aromatic amino acids (Phe, Tyr, and His in OMTKY3) do not have observable coupling with the aromatic ring protons.

Figure 4 is part of a HOHAHA map obtained with a mixing time of 101 ms. Of the 11 long-chain amino acids in OMTKY3, spin systems for 8 were elucidated at least partially prior to sequential assignment. Three of the spin systems contained NH resonances coupled to side-chain aliphatic

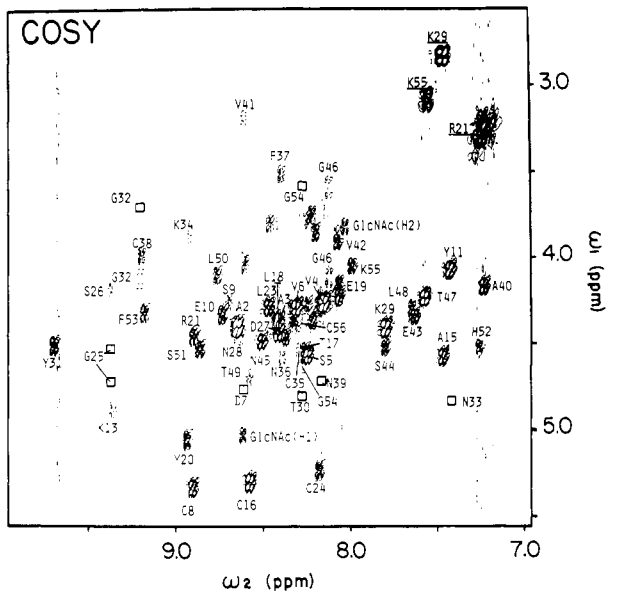

FIGURE 5: Fingerprint region of the pure-phase COSY map of OMTKY3 (15 mM at pH 4.0, 0.2 M KCl, 25 °C) in H_2O. Assignments are indicated. Boxes indicate cross-peaks that are not observed in H_2O and that have been located through other evidence. Underlined assignments are for side-chain resonances.

protons, which identified them as belonging to arginine or lysine residues. A slowly exchanging backbone N*H*, at 7.80 ppm, and additional aliphatic proton resonances were linked to a lysine spin system through HOHAHA connectivities (data not shown). Two spin systems were attributed to glutamate residues ($C_\alpha H$ = 4.14 and 4.33 ppm) and the remaining three to arginine, lysine, or proline residues ($C_\alpha H$ = 5.22, 4.89, and 4.46 ppm).

(4) Linkage of NH Resonances to Side-Chain Spin Systems. A total of 32 backbone NH resonances were linked to side chains via RELAY data; some of these are described above. An additional 26 N*H* resonances were linked to their spin systems via COSY data (Figure 5) or intraresidue (N*H*, $C_\beta H$) NOESY connectivities (data not shown).

Sequential Assignment of Peptide Backbone Resonances. Identification of unambiguous starting points is essential to the sequential assignment of $C_\alpha H$ and N*H* resonances. By combining the first-stage assignments, described above, with NOESY data and primary sequence data, we were able to locate seven such sites in the sequence for OMTKY3 (Figure 8).

Two good starting sites were provided by the Ala^2-Ala^3-Val^4 and Ala^{40}-Val^{41}-Val^{42} sequences (Figure 8). Each of these tripeptides is unique in OMTKY3. We could link the spin systems of residues 2, 3, and 4 through $d_{\alpha N}$ connectivities between previously identified alanine and valine residues (Figure 6A). Another set of $d_{\alpha N}$ connectivities between an alanine residue and two valine residues located residues 40–42 (Figure 6C).

Given the information at hand, the sequential assignments can be built up in several different ways. Here we will first extend the assignments made for residues 2–4 (Figure 6A). The Leu^1 $C_\alpha H$ resonance (4.05 ppm) was assigned through a $d_{\alpha N}$ connectivity to Ala^2 N*H*. This assignment agreed with a previously identified leucine spin system. The absence of an observable ($C_\alpha H$, N*H*) or "fingerprint" cross-peak in the COSY map at 4.05 ppm is probably the result of rapid exchange of the N-terminal α-amino protons with the solvent. Moving in the C-terminal direction from Val^4 N*H*, we located $C_\alpha H$ and N*H* resonances for Ser^5 and Val^6 through $d_{\alpha N}$ connectivities. The spin system of residue 6 had been identified above as valine. A possible $d_{\alpha N}$ connectivity from Val^6 to Asp^7 was observed in the NOESY map. This assignment was confirmed by a $d_{\beta N}$ connectivity between Val^6 $C_\beta H$ and Asp^7 N*H* (8.61 ppm). The only COSY cross-peak at this N*H* resonance belongs to a valine residue ($C_\alpha H$ = 3.18 ppm), but the evidence cited below suggests that the Asp^7 $C_\alpha H$ is not observed because of overlap with the solvent resonance.

Assignment of a putative Asp^7 $C_\beta H$ resonance (2.98 ppm) was made through a RELAY connectivity to an N*H* resonating at 8.61 ppm. Inspection of the aliphatic region of the COSY map of OMTKY3 dissolved in ^2H_2O revealed a cross-peak at (4.76, 2.98) ppm. This signal at 4.76 ppm overlaps with that of the solvent. A second COSY cross-peak appears at (4.76, 2.72) ppm and can be linked to a NOESY cross-peak at (8.61, 2.72) ppm. The latter corresponds to an intraresidue NOE between Asp^7 N*H* and $C_\beta H$.

Starting at residues 40–42, we extended the assignments in the N-terminal direction by constructing a $d_{\alpha N}$ spiral from Ala^{40} to Phe^{37} $C_\alpha H$ (Figure 6C). The N*H* resonance for residue 37 was assigned via a COSY cross-peak at (8.39, 3.51) ppm (Figure 5). The assignments from this spiral were confirmed by d_{NN} and $d_{\beta N}$ connectivities (Figures 7 and 8). Moving in the other direction from Val^{42} $C_\alpha H$, a $d_{\alpha N}$ spiral was constructed to residue 46, whose spin system had been identified earlier as glycine. This spiral was continued to a putative Gly^{54} N*H* resonance (Figure 6D). The spin systems assigned to residues 48–50 had been identified previously as leucine, threonine, and leucine, respectively. The $C_\alpha H$ resonance for Thr^{47} identified in this spiral (4.22 ppm) was found to be very close to a threonine $C_\beta H$ resonance located earlier (4.26 ppm). The proximity of these resonances explains why we were unable to resolve the ($C_\alpha H$, $C_\gamma H$) cross-peak for this spin system in the RELAY map. The $C_\alpha H$ resonances for Gly^{54} are not seen in the fingerprint region of the COSY map for OMTKY3 dissolved in H_2O, but one, at 4.61 ppm, is seen in the map of COSY OMTKY3 dissolved in ^2H_2O (figure 6E). Most of these assignments were corroborated by d_{NN} and $d_{\beta N}$ connectivities (Figure 8).

We could now assign the remaining alanine spin system to residue 15 by elimination. From this new start site, we first assigned the Pro^{14} $C_\alpha H$ resonance through a $d_{\alpha N}$ connectivity to Ala^{15} (Figure 6B). From Ala^{15}, a sequential spiral to Arg^{21} $C_\alpha H$ was constructed. The assignment of this last residue is consistent with a previously identified long-chain residue with a side-chain nitrogen-bound proton. Glu^{19} assignments were confirmed in the same way. A NOESY connectivity was observed between a threonine $C_\gamma H$ resonance, identified earlier, and the newly assigned Thr^{17} N*H* resonance.

Assignments to this point left a single leucine spin system which could now be assigned by elimination to Leu^{23}. A $d_{\alpha N}$ connectivity to Leu^{23} N*H* permitted assignment of Pro^{22} $C_\alpha H$ (Figure 6B). The spiral was continued to Gly^{25} N*H*, part of a glycine spin system identified previously. A $d_{\alpha N}$ connectivity from Gly^{25} $C_\alpha H$ to Ser^{26} N*H* is observed in ^2H_2O (Figure 6E), and the $C_\alpha H$ was assigned through a well-resolved COSY fingerprint cross-peak. These assignments were extended through d_{NN} connectivities to Lys^{29} N*H* (Figure 7); the spin system containing the NH resonance was characterized as a lysine in the first stage of assignment. The $C_\alpha H$ resonances for Asp^{27} and Asn^{28} were assigned by reference to COSY fingerprint cross-peaks. The Asp^{27} cross-peak is obscured somewhat in the map of the sample dissolved in H_2O, but it is well resolved in ^2H_2O (data not shown). The Asn^{28} cross-peak completely overlaps the Ala^2 cross-peak. Two lines of

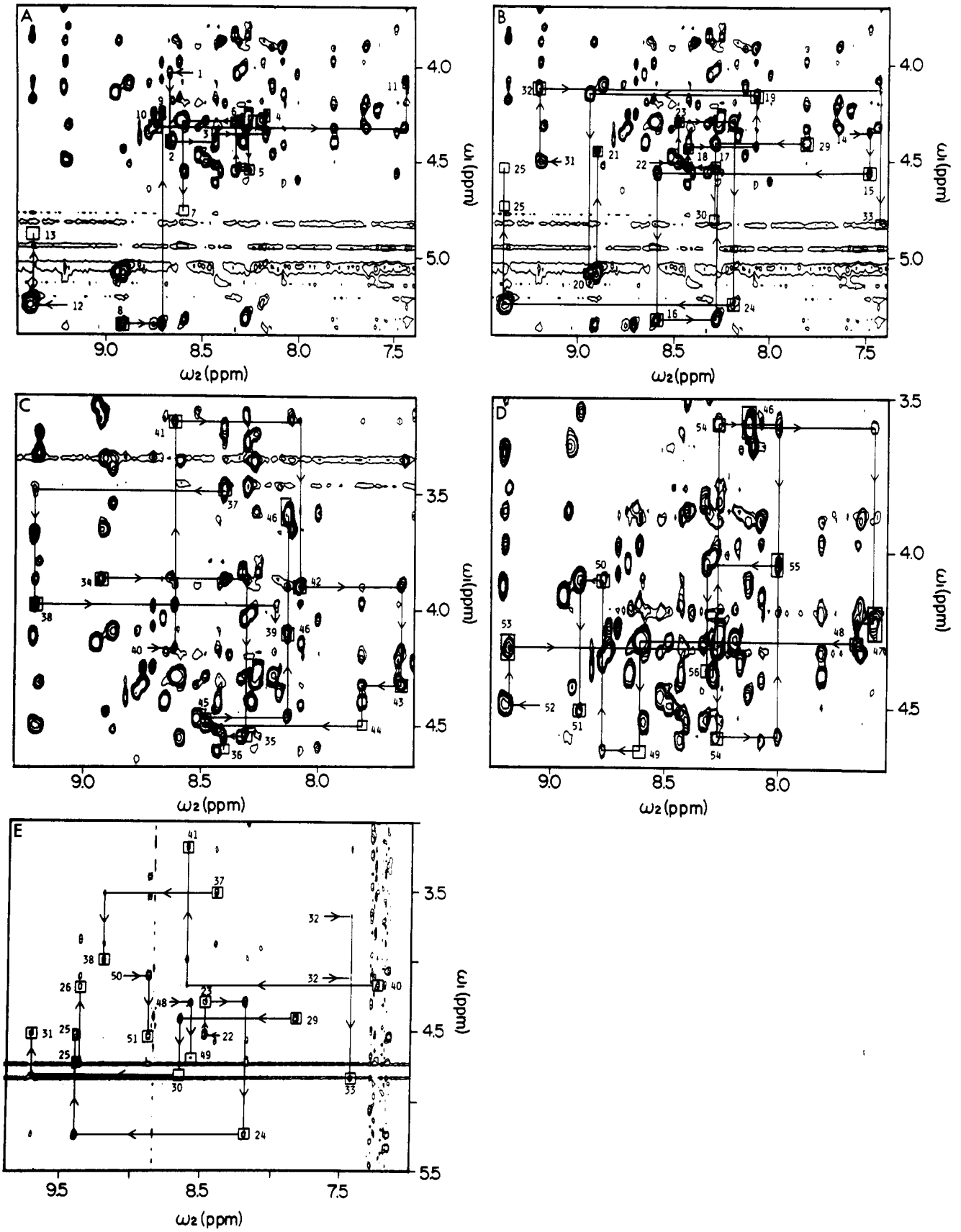

FIGURE 6: Assignment spirals constructed from $d_{\alpha N}$ connectivities. The connectivities are represented with the convention suggested by Wagner and co-workers (Wagner et al., 1986) in which NOESY data alone are presented. The COSY-type cross-peaks are boxed, and the assignments are indicated by the numbers next to the boxes. Horizontal lines represent the NOESY $d_{\alpha N}$ connectivities used to connect residue i to residue $i + 1$. The vertical line is used then to locate the COSY fingerprint ($C_{\alpha}H$, NH) cross-peak for residue $i + 1$. (A) Assignment spirals for residues 1–7, 8–11, and 12–13. (B) Assignment spirals for residues 14–21, 22–25, and 31–33. (C) Assignment spirals for residues 34–36, 37–39, and 40–46. (D) Assignment spirals for residues 46–51 and 52–56. (E) Assignment spirals constructed with data acquired in ^2H_2O.

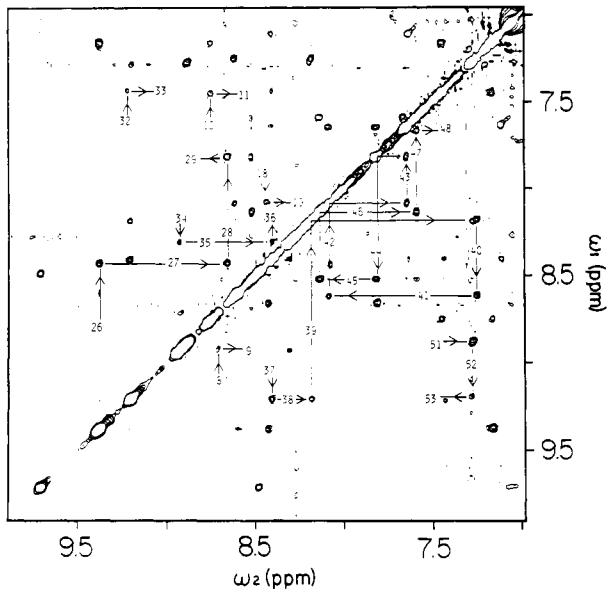

FIGURE 7: Expanded region of a pure-phase NOESY map in H_2O, pH 4.0, 25 °C. Sequential d_{NN} connectivities are indicated by the lines and arrows.

evidence confirm this. First, in the aliphatic region of the COSY map of OMTKY3 in H_2O we observed three cross-peaks at 4.42 ppm, one of which corresponds to Ala^2 (1.42 ppm) and the other two of which presumably belong to Asn^{28} (3.17 and 2.82 ppm). This can be confirmed by reference to the low-field region of the RELAY map of samples dissolved in H_2O or ^2H_2O. Second, OMTKY3 dissolved in H_2O shows three RELAY cross-peaks to 8.70 ppm from the aforementioned 3.17, 2.82, and 1.42 ppm (Figure 3), whereas OMTKY3 dissolved in ^2H_2O shows RELAY cross-peaks to 8.70 ppm only at 3.17 and 2.82 ppm (data not shown). Apparently the Ala^2 N*H* proton exchanged out before the start of the RELAY experiment while Asn^{28} N*H* exchanged slowly enough to be observed. Moreover, the d_{NN} connectivity between residues 28 and 29 is still observed.

Thr^{30} N*H* (8.29 ppm) was located through a $d_{\alpha N}$ connectivity to Lys^{29} (Figure 6B). This agreed with a previously identified threonine spin system. The Tyr^{31} N*H* assignment was made through a $d_{\alpha N}$ connectivity observed in ^2H_2O (Figure 6E). The Thr^{30} C$_\alpha$*H* resonance (4.78 ppm) was bleached during solvent irradiation in H_2O. Tyr^{31} C$_\alpha$*H* was assigned by reference to the fingerprint region of the COSY map, where a well-resolved cross-peak occurs at (9.70, 4.50) ppm (Figure 5). Gly^{32} N*H* and C$_\alpha$*H* resonances were assigned through $d_{\alpha N}$ connectivities, a single COSY fingerprint cross-peak (4.12 ppm), and a RELAY cross-peak (9.21, 3.68 ppm). Asn^{33} N*H* was located through a weak $d_{\alpha N}$ connectivity and a d_{NN} connectivity from Gly^{32}. A well-resolved COSY cross-peak (7.43, 4.81 ppm) observed in OMTKY3 dissolved in ^2H_2O permitted assignment of Asn^{33} C$_\alpha$*H* (Figure 6E).

A $d_{\alpha N}$ connectivity was observed between two long side chain residues at (5.22, 9.37) ppm (Figure 6A). These had to be Pro^{12} and Lys^{13} as these were the only two unassigned neighboring amino acids with long side chains. A $d_{\alpha\alpha}$ connectivity, at (5.22, 4.07) ppm, from Pro^{12} to Tyr^{11} intiated a $d_{\alpha N}$ spiral back to Cys^8 (Figure 6A). In addition, d_{NN} connectivities linked residues 8 and 9 and residues 10 and 11 (Figure 7).

Two three-residue stretches, 34–36 and 54–56, remained to be assigned. Each included one lysine residue, which would account for the two long-chain spin systems that remained to be assigned. We had previously been able to identify tentatively the Gly^{54} N*H* resonance. We could locate two NOESY connectivities to the putative Gly^{54} N*H* at 4.61 and 3.57 ppm (Figure 6D). A COSY cross-peak was observed at (4.61, 3.57) ppm. These were likely to be the C$_\alpha$*H* resonances of Gly^{54}. Additional evidence for this assignment was provided by $d_{\alpha N}$ connectivities to a lysine N*H* at 7.99 ppm, presumably that of Lys^{55}. A further $d_{\alpha N}$ connectivity from Lys^{55} C$_\alpha$*H* (4.05 ppm) led to assignment of the Cys^{56} spin system.

By elimination, the remaining long-chain spin system could be assigned to Lys^{34}. A $d_{\alpha N}$ connectivity to Cys^{35} N*H* and the COSY cross-peak at (8.30, 4.53) ppm led to assignment of this residue (Figure 6C). A d_{NN} connectivity from Cys^{35} N*H* indicated that the N*H* resonance of Asn^{36} (8.39 ppm) overlapped with that of Phe^{37}. We could assign Asn^{36} C$_\alpha$*H* through a COSY cross-peak at (8.39, 4.58) ppm.

Side-Chain Assignments. We were able to assign a number of aromatic proton resonances on the basis of NOESY correlations between C$_\beta$*H* and C$_\delta$*H* resonances. Two assignment criteria were used here. One was that the C$_\beta$*H* resonances be identified unambiguously from the sequential assignments. The second criterion was that NOESY correlations from C$_\delta$*H* resonances be observed to both C$_\beta$*H* resonances identified above. The C$_\delta$*H* for Phe^{37} and Phe^{53} and for Tyr^{20} were assigned easily in this way. The Phe^{37} C$_\epsilon$*H* resonance was assigned through the single COSY cross-peak involving Phe^{37} C$_\delta$*H*. The Phe^{53} and Tyr^{20} C$_\delta$*H* resonances overlap, but assignment of the C$_\epsilon$*H* resonances could be made on the basis

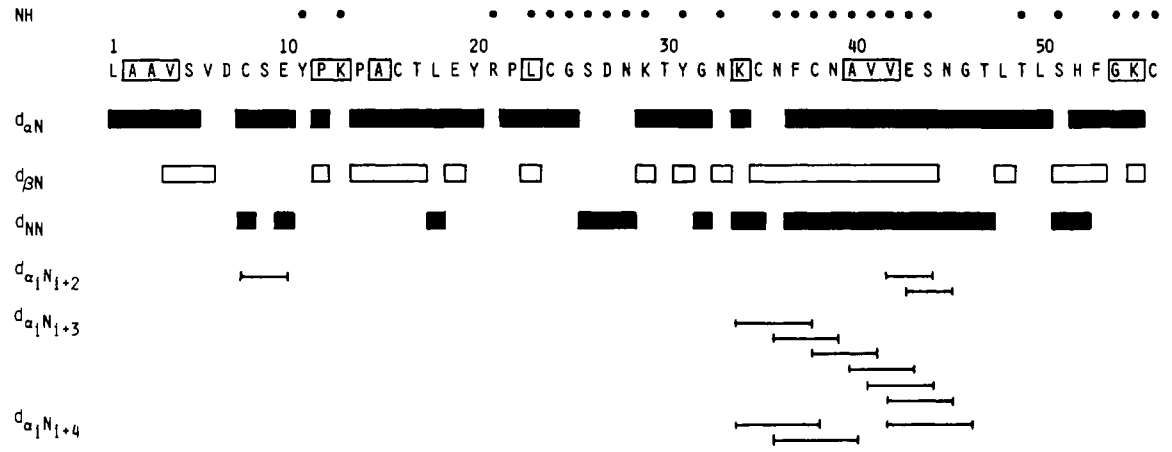

FIGURE 8: Sequence of OMTKY3, with sequential assignment starting sites (boxed). NOESY connectivities are indicated below the sequence. The dots above the sequence denote slowly exchanging N*H*'s as defined by observation of a cross-peak in the fingerprint region of the COSY map (8-h acquisition time) for OMTKY3 freshly dissolved in ^2H_2O at pH 4, 25 °C.

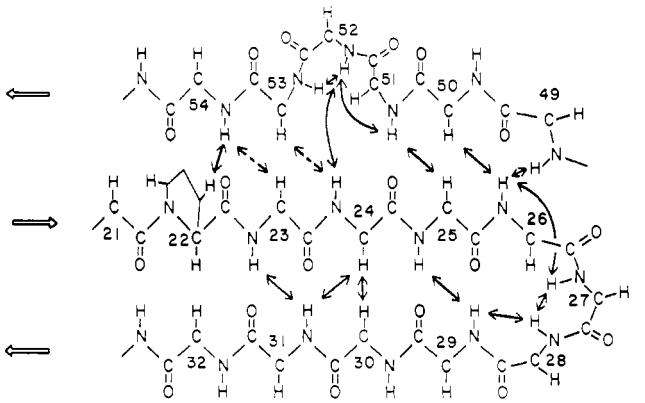

FIGURE 9: NOESY connectivities used to establish the relative orientation of the strands of β-sheet. The dashed lines indicate possible NOE connectivities that are obscured by overlapping cross-peaks.

of relaxation experiments, as the Tyr[20] ring protons have T_1 relaxation times which differ from those of the other aromatic protons (data not shown). The Tyr[11] and Tyr[31] $C_\beta H$ resonances are very similar. On the basis of a NOESY correlation to Tyr[11] $C_\alpha H$, the Tyr[11] $C_\delta H$ resonances were assigned tentatively to 7.15 ppm. Side-chain $N_\delta H$ resonances for four of the five asparagine residues in OMTKY3 could be assigned through NOESY connectivities to sequentially assigned $C_\beta H$ resonances. The $N_\delta H$ resonances for Asn[36] were assigned by elimination.

Identification of N-Linked N-Acetylglucosamine. Earlier samples of OMTKY3 gave rise to extra peaks in the fingerprint region of the COSY spectrum that were not assigned to backbone or side-chain resonances. Two of the extra cross-peaks, (8.61, 5.01) ppm and (8.20, 3.82) ppm, were assigned to protons at positions 1 and 2 of N-linked *N*-acetylglucosamine. These assignments were based on agreement of these chemical shifts with those of GlcNAc in model glycopeptides (Vliegenthart et al., 1983), on RELAY cross-peaks at (8.61, 3.82) ppm and (8.20, 5.01) ppm (Figure 3), and on the four cross-peaks in the aliphatic region of the HOHAHA map at 5.01 ppm (Figure 4). These cross-peaks correspond to the protons at positions 2–5 of GlcNAc. All of the GlcNAc cross-peaks are absent or of lower intensity in the COSY maps of more recent OMTKY3 preparations. Five extra fingerprint cross-peaks (unidentified) also were absent in more recent preparations of OMTKY3.

Determination of Secondary Structure. The α-helix and triple-stranded β-sheet of OMTKY3 were identified through NOESY correlations (Figures 8 and 9), by use of rules developed by Wüthrich and co-workers (Wüthrich et al., 1984). Residues 1–7 are most likely in an extended-chain conformation, as determined by a stretch of $d_{\alpha N}$ connectivities with attendant lack of other NOESY connectivities. Residues 8–10 are part of a small kink or turn, as evidenced by the d_{NN} and $d_{\alpha N}(i,i+2)$ connectivities. A $d_{\alpha\alpha}$ connectivity between Tyr[11] and Pro[12] indicates a cis peptide bond between these residues. The region from residue 13 to residue 21 is an extended chain, as there is little other than $d_{\alpha N}$ connectivities. Residues 23–26 are the central strand of a triple-stranded β-sheet. The other strands include residues 28–31 and 50 and 51, as evidenced by the interstrand NOESY connectivities indicated in Figure 9. The β-turn, which is made up of residues 26–29, is well defined by d_{NN} connectivities and the interstrand $d_{NN}(i,i+4)$ connectivity between residues 25 and 29. The α-helix extends from residue 33 to residue 46, as defined by the d_{NN} and $d_{\alpha N}(i,i+2)$, $d_{\alpha N}(i,i+3)$, and $d_{\alpha N}(i,i+4)$ connectivities. The latter connectivities are particularly strong indicators of α-helix (Wüthrich et al., 1984).

CONCLUSIONS

Sequential Assignments. Use of the RELAY and HOHAHA experiments greatly facilitated identification of the side-chain spin systems of OMTKY3. Probably the best example of the power of these experiments was our ability to identify the spin system of the unexpected N-linked GlcNAc. Complete assignment of the backbone proton resonances has been accomplished, and over 80% of the side-chain proton resonances have been assigned as well (Table I).

A limited number of assignments in OMTKY3 had been reported earlier by our group (Markley et al., 1984, 1986; Westler et al., 1984; Ortiz-Polo, 1985; Ortiz-Polo et al., 1986). The assignments made by Ortiz-Polo (1985) for the methyl and aromatic ^{1}H NMR resonances are in complete accord with the assignments described here. In addition, we have been able to distinguish the Thr[47] and Thr[49] methyl and the Phe[37] and Phe[53] aromatic resonances. The present results indicate that some of the assignments presented elsewhere (Markley et al., 1984, 1986; Westler et al., 1984) were incorrect. We can identify two sources of error in these: (1) assignment based on spectral differences for proteins differing by more than one or two amino acid substitutions and (2) sequential assignments based on identification of a limited number of side-chain spin systems. The present assignments do not rely on amino acid substitutions and are much more complete. They have been assisted by the availability of more sophisticated spectrometers and pulse sequences such as HOHAHA and RELAY.

The chemical shifts of the "extra" threonine spin system are close to the values observed for threonine in small peptides (Wüthrich, 1986), but the slow exchange rate of the amide suggests that the threonine is not present as the free amino acid, but part of a peptide or protein. The resonance intensities for this spin system are lower than those for the other methyl-containing amino acids. Its relative intensity, however, has been equivalent in COSY maps of three different OMTKY3 preparations. Because of its low intensity, we were unable to link this threonine spin system to any of the other spin systems in OMTKY3. Polymorphisms have been detected in the sequences of avian ovomucoid third domains, but not in OMTKY3 (Laskowski et al., 1987). Threonine is a difficult amino acid to detect at minor occupancy in a sequencing run. Additional efforts will be made to place this threonine in the sequence.

Secondary Structure. The secondary structure of OMTKY3 in solution closely resembles that determined through X-ray crystallographic studies (Fujinaga et al., 1982). All but three of the slowly exchanging N*H*'s have been identified as hydrogen-bond donors in the X-ray studies. The exceptions are the N*H*'s of residues 27, 33, and 55. Determination of the solvent accessibilities of atoms in the OMTKY3 crystal structure indicates that the Asp[27] and Asn[33] N*H*'s are not exposed to solvent. The Asp[27] NH may hydrogen bond with an Asp[27] side-chain oxygen. We are unable to identify a potential hydrogen-bond acceptor for the Asn[33] backbone NH.

The third slowly exchanging N*H*, that of Lys[55], appears to be accessible to solvent and has no obvious potential hydrogen-bonding partner in the crystal structure. One possible explanation for retarded exchange is aggregation of OMTKY3 at the high protein concentrations (10–15 mM) used in these studies. Slowed N*H* exchange due to aggregation was observed in BPTI (Wagner, 1983b). In recent experiments with OMTKY3, slowed exchange still was observed at 5 mM protein. Exchange experiments at lower concentrations are to be conducted.

Table I: Chemical Shifts of the Assigned ^{1}H NMR Resonances of OMTKY3 (15 mM) in 0.2 M KCl, pH 4.0, at 25 °C[a]

residue	NH	$C_\alpha H$	$C_\beta H$	other
Leu[1]		4.05		$C_\gamma H$ 1.70; $C_\delta H$ 1.00, 0.96
Ala[2]	8.65	4.42	1.42	
Ala[3]	8.42	4.37	1.41	
Val[4]	8.15	4.27	2.05	$C_\gamma H$ 0.93, 0.88
Ser[5]	8.25	4.56	3.84	
Val[6]	8.32	4.29	1.97	$C_\gamma H$ 0.92, 0.82
Asp[7]	8.61	4.76	2.98, 2.72	
Cys[8]	8.90	5.33	3.36, 2.55	
Ser[9]	8.69	4.27	4.08, 3.98	
Glu[10]	8.73	4.33	2.08, 1.78	$C_\gamma H$ 2.31
Tyr[11]	7.43	4.07	2.90, 2.78	$C_\delta H$ 7.15
Pro[12]		5.22	2.47, 2.08	$C_\gamma H$ 1.98, 1.87; $C_\delta H$ 3.68, 3.56
Lys[13]	9.37	4.89	1.90, 1.59	$C_\gamma H$ 1.14, 0.91; $C_\delta H$ 1.60, 1.43; C_ϵ 2.57, 2.40
Pro[14]		4.36	2.31	$C_\gamma H$ 2.05; $C_\delta H$ 3.82, 3.76
Ala[15]	7.48	4.58	1.33	
Cys[16]	8.58	5.32	3.37, 2.79	
Thr[17]	8.27	4.53	4.63	$C_\gamma H$ 1.36
Leu[18]	8.42	4.42	1.66	$C_\gamma H$ 1.73; $C_\delta H$ 0.96, 0.90
Glu[19]	8.06	4.14	2.00	$C_\gamma H$ 2.37, 2.28
Tyr[20]	8.93	5.06	3.11, 2.92	$C_\delta H$ 7.28; $C_\epsilon H$ 6.82
Arg[21]	8.89	4.46	1.74	$C_\gamma H$ 1.59; $C_\delta H$ 3.29, 3.20; $N_\delta H$ 7.25
Pro[22]		4.52	1.67	$C_\gamma H$ 1.73; $C_\delta H$ 3.36
Leu[23]	8.48	4.28	0.98	$C_\gamma H$ 1.39; $C_\delta H$ 0.48, 0.41
Cys[24]	8.17	5.23	2.40, 1.35	
Gly[25]	9.40	4.68, 4.50		
Ser[26]	9.37	4.15	4.10, 3.82	
Asp[27]	8.43	4.43	3.01, 2.59	
Asn[28]	8.65	4.40	3.17, 2.82	$N_\delta H$ 7.53, 6.85
Lys[29]	7.80	4.40	1.66, 1.44	$C_\gamma H$ 1.14, 0.81; $C_\delta H$ 1.54; 1.43; $C_\epsilon H$ 2.82; $N_\epsilon H$ 7.45
Thr[30]	8.29	4.78	4.02	$C_\gamma H$ 1.18
Tyr[31]	9.70	4.50	2.90, 2.79	
Gly[32]	9.21	4.12, 3.68		
Asn[33]	7.43	4.81	3.67, 3. 19	$N_\delta H$ 6.32, 8.11
Lys[34]	8.92	3.86	1.98, 1.86	$C_\gamma H$ 1.46, 1.34; $C_\delta H$ 1.78; $C_\epsilon H$ 3.10, 3.04
Cys[35]	8.30	4.53	3.47, 3.25	
Asn[36]	8.39	4.58	3.22, 2.84	$N_\delta H$ 7.61, 7.11
Phe[37]	8.39	3.51	2.848 2.78	$C_\delta H$ 6.55; $C_\epsilon H$ 6.82; $C_\zeta H$ 6.55
Cys[38]	9.19	3.97	1.77, 1.38	
Asn[39]	8.17	4.69	2.98, 2.78	$N_\delta H$ 7.52, 6.81
Ala[40]	7.22	4.17	1.30	
Val[41]	8.60	3.18	2.06	$C_\gamma H$ 0.72, −0.02
Val[42]	8.06	3.93	2.31	$C_\gamma H$ 1.05, 1.00
Glu[43]	7.63	4.32	2.25, 2.10	$C_\gamma H$ 2.58
Ser[44]	7.80	4.52	4.22 3.98	
Asn[45]	8.50	4.48	3.15, 2.85	$N_\delta H$ 7.58, 6.85
Gly[46]	8.12	4.10, 3.58		
Thr[47]	7.58	4.22	4.26	$C_\gamma H$ 1.26
Leu[48]	7.65	4.28	1.51	$C_\gamma H$ 1.12; $C_\delta H$ 1.17, 0.30
Thr[49]	8.56	4.68	4.36	$C_\gamma H$ 1.13
Leu[50]	8.77	4.10	1.78, 1.00	$C_\gamma H$ 0.65; $C_\delta H$ 0.04, −0.12
Ser[51]	8.85	4.53	3.53, 3.38	
His[52]	7.27	4.52	3.79, 3.32	$C_\delta H$ 7.26; $C_\epsilon H$ 8.80
Phe[53]	9.17	4.32	3.24, 3.04	$C_\delta H$ 7.28; $C_\epsilon H$ 7.17
Gly[54]	8.26	4.61, 3.57		
Lys[55]	7.99	4.05	1.88, 1.74	$C_\gamma H$ 1.55, 1.49; $C_\delta H$ 1.75; $C_\epsilon H$ 3.09; $N_\epsilon H$ 7.54
Cys[56]	8.31	4.39	3.15, 2.52	

[a] Chemical shifts are reported in ppm (±0.02 ppm) relative to [^{2}H]TSP.

Comparison of OMTKY3 and BUSI IIA. Little is known of the relationship between peptide backbone conformation and the chemical shifts of backbone protons. One approach to an understanding of this relationship has been correlation of proton chemical shifts and structure in homologous proteins for which detailed structure information is available (Pardi et al., 1983). OMTKY3 and BUSI IIA share approximately 60% sequence homology. Similiarities in the tertiary structures (Williamson et al., 1985) and COSY maps of these two proteins [see Figure 11 of Štrop et al. (1983) and Figure 5 of the present work] led us to compare the chemical shifts of backbone protons as a function of sequence position. To compare resonances of nonhomologous residues, the small-peptide or random-coil chemical shifts (Wüthrich, 1986) were substracted from each of the backbone chemical shift values for OMTKY3 and BUSI IIA. The remaining chemical shift values are "conformation-dependent" values (Pardi et al., 1983).

The conformation-dependent backbone chemical shifts are, overall, very similiar in OMTKY3 and BUSI IIA (Figure 10). This suggests that the contribution of conformation to backbone chemical shift is fairly independent of amino acid type. The similiarity between OMTKY3 and BUSI IIA is more apparent for the $C_\alpha H$ resonances than for the N*H* resonances. Pardi et al. (1983) reached the same conclusion in their comparison of chemical shifts from BPTI and two homologous proteins. Some of the scatter seen for the N*H* resonances may be due to the different pH and temperatures used during sequential assignment in OMTKY3 (pH 4.0 and 25 °C) and BUSI IIA (pH 4.9 and 18 °C). For both $C_\alpha H$ and N*H* resonances, the largest differences are located in two regions

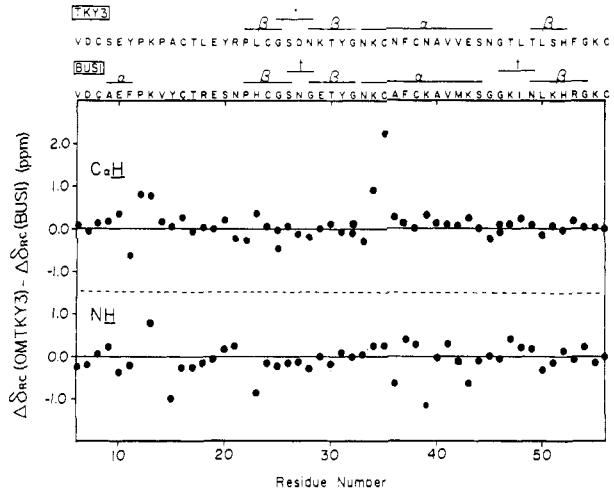

FIGURE 10: Comparison of the conformation-dependent chemical shifts of OMTKY3 and BUSI IIA. Secondary structure is indicated above the sequences (α-helix = α; β-sheet = β; turn = t). To obtain conformation-dependent chemical shift values ($\Delta\delta_{RC}$) for each of the proteins, random-coil chemical shift values (Wüthrich, 1986) were subtracted from the observed chemical shifts. Sequence numbering is that for OMTKY3, and the sequences are aligned to maximize homology. OMTKY3 has a two-residue deletion relative to BUSI IIA between residues 11 and 12.

of the primary sequence: between residues 10–15 and 33–40. BUSI IIA has two additional residues between positions 11 and 12 (OMTKY3 numbering), and residues 7–10 participate in a small α-helix. In OMTKY3, these residues are part of a kink or bulge. Residues 34 and 35 are at the N-terminal end of the large α-helix and are close, in the three-dimensional structures, to residues 10–15. Ring current effects arising from either His^{24} or Phe^{10} in BUSI IIA may be responsible for the very large upfield shifts for the $C_\alpha H$ of Lys^{34} and Cys^{35}. Despite the differences, it is interesting to note that the conformation-dependent backbone chemical shift differences between OMTKY3 and BUSI IIA, although larger, are more localized than those between OMTKY3 and OMTKY3* (Rhyu & Markley, 1988).

The ^1H NMR assignments presented here will be used to derive further, more detailed, structural information for OMTKY3 in solution. The exchange of backbone amide protons is being characterized to determine the relative stability of the elements of secondary structure in OMTKY3. Detailed studies of third domains differing from OMTKY3 by single amino acid substitutions are under way. Ongoing studies address the mechanism of serine proteinase inhibition by ovomucoid third domains and, more generally, the relationship between amino acid sequence and the structure and dynamics of proteins.

ACKNOWLEDGMENTS

We are grateful to Drs. Thomas Farrar and Bruce Adams of the Department of Chemistry at the University of Wisconsin—Madison for providing access to high-field NMR spectrometers in the early stages of our work. A.D.R. thanks Dr. Slobodan Macura for initial training in NMR spectroscopy and Dr. Gyung Ihm Rhyu for many helpful discussions.

REFERENCES

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

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

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

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

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

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., Mayer, E., Travis, J., & Neumann, S. (1986) *EMBO J. 5*, 2453–2458.

Bodenhausen, G., Kogler, H., & Ernst, R. R. (1984) *J. Magn. Reson. 58*, 370–388.

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

Connolly, M. L. (1981) Ph.D. Thesis, University of California, Berkeley.

Empie, M. W., & Laskowski, M., Jr. (1982) *Biochemistry 21*, 2274–2284.

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. (1984) *Bull. Math. Biol. 46*, 673–698.

Hoult, D. I., & Richards, R. E. (1975) *Proc. R. Soc. London, A 344*, 311–340.

IUPAC–IUB Commission on Biochemical Nomenclature (1970) *J. Mol. Biol. 52*, 1–17.

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

Langridge, R., Ferrin, T. E., Kuntz, I. D., & Connolly, M. L. (1981) *Science (Washington, D.C.) 211*, 661–666.

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, M. T., Tashiro, M., Vichot, G., Watley, H. E., Wieczorak, A., & Wieczorak, M. (1987) *Biochemistry 26*, 202–221.

Macura, S., Wüthrich, K., & Ernst, R. R. (1982) *J. Magn. Reson. 46*, 269–282.

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

Markley, J. L., Westler, W. M., Chan, T. M., Kojiro, C. L., & Ulrich, E. L. (1984) *Fed. Proc., Fed. Am. Soc. Exp. Biol. 43*, 2648–2656.

Markley, J. L., Croll, D. H., Krishnamoorthi, R., Ortiz-Polo, G., Westler, W. M., Bogard, W. C., & Laskowski, M., Jr. (1986) *J. Cell. Biochem. 30*, 291–309.

Matsuda, T., Watanabe, K., & Sato, Y. (1981) *FEBS Lett. 124*, 185–188.

Ogino, T., Croll, D. H., Kato, I., & Markley, J. L. (1982) *Biochemistry 21*, 3452–3460.

Ortiz-Polo, G. (1985) M.S. Thesis, Purdue University.

Ortiz-Polo, G., Krishnamoorthi, R., Markley, J. L., Live, D. H., Davis, D. G., & Cowburn, D. (1986) *J. Magn. Reson. 68*, 303–310.

Papamakos, E., Weber, 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.

Rance, M., Sørensen, O. W., Bodenhausen, G., Wagner, 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.

Rhyu, G. I., & Markley, J. L. (1988) *Biochemistry* (following paper in this issue).

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