Biochemistry

 Volume 32, Number 31 August 10, 1993

Articles

An Improved Strategy for Determining Resonance Assignments for Isotopically Enriched Proteins and Its Application to an Engineered Domain of Staphylococcal Protein A†

Barbara A. Lyons,‡ Mitsuru Tashiro,‡,§ Lena Cedergren,∥,⊥ Björn Nilsson,∥,⊥ and Gaetano T. Montelione*,‡,§

Center for Advanced Biotechnology and Medicine, Rutgers University, Piscataway, New Jersey 08854-5638, Graduate Program in Biochemistry and Molecular Biology, UMDNJ-Robert Wood Johnson Medical School, Piscataway, New Jersey 08854-5638, Department of Structural Biology, Kabi Pharmacia Bioscience Center, S-112 87 Stockholm, Sweden, and Department of Biochemistry, Royal Institute of Technology, S-100 44 Stockholm, Sweden

Received March 9, 1993; Revised Manuscript Received May 21, 1993

ABSTRACT: Sequence-specific resonance assignments provide the basis for interpreting multidimensional NMR spectra and for determining 3D structures of proteins from these data. We have developed an improved strategy for determining these sequence-specific NMR assignments in small proteins and applied this method in determining proton and nitrogen resonance assignments for an 8.2-kDa engineered domain (the Z-domain) of the cell wall protein A of *Staphylococcus aureus*. First, HCCNH-TOCSY [Lyons, B. A. & Montelione, G. T. (1993) *J. Magn. Reson. 101B*, 206] data were used together with 2D 2QF-COSY, TOCSY, and ^{15}N-HSQC data to identify amino acid spin systems. Most asparagine and glutamine spin systems were also identified uniquely from these triple-resonance data. Next, complementary HCC(CO)-NH-TOCSY [Montelione, G. T., et al. (1992) *J. Am. Chem. Soc. 114*, 10975] data were used to identify sequential connections from the aliphatic H^{α}, H^{β}, H^{γ}, H^{δ}, and H^{ϵ} resonances of residue *i* to the amide and nitrogen resonances of residue *i* + 1. By combined analysis of HCCNH-TOCSY and HCC(CO)NH-TOCSY spectra we have determined most of the proton and nitrogen resonance assignments for the Z-domain. This represents the first example of the use of this triple-resonance technique to determine extensive resonance assignments in a small protein.

Staphylococcal protein A (SPA)[1] is a cell wall constituent of *Staphylococcus aureus* with the ability to bind the F_c portion of IgG antibodies from various mammalian species (Forsgren & Sjöquist, 1966). The SPA gene sequence (Uhlén et al., 1984) exhibits internal homologies involving an approximately 58 amino acid unit, repeated five times, corresponding to IgG-binding domains A, B, C, D, and E (Sjödahl, 1977; Uhlén et al., 1984; Moks et al., 1986). Other regions of the gene encode portions of SPA which do not bind IgGs, including regions responsible for the binding of SPA to the cell wall, an N-terminal signal peptide, and a C-terminal transmembrane region. The three-dimensional structure of the B-domain has been determined both in solution using NMR spectroscopy

† This work was supported by The National Science Foundation (DIR-9019313) and The National Institutes of Health (GM-47014). L.C. is supported by a research grant from NUTEK, Sweden.

‡ Rutgers University.
§ UMDNJ-Robert Wood Johnson Medical School.
∥ Kabi Pharmacia Bioscience Center.
⊥ Royal Institute of Technology.

[1] Abbreviations: SPA, staphylococcal protein A; 2D, two-dimensional; 3D, three-dimensional; MD, multidimensional; 2QF-COSY, 2D two-quantum filtered correlated spectroscopy; HSQC, 2D ^{15}N–^{1}H heteronuclear single quantum coherence spectroscopy; INEPT, insensitive nuclear enhancement by polarization transfer; NOE, nuclear Overhauser effect; NOESY, 2D NOE spectroscopy; TOCSY, 2D ^{1}H–^{1}H total correlation spectroscopy; HSQC–TOCSY, MD NMR experiment combining HSQC and TOCSY; HCCNH-TOCSY (or CA-TOCSY) and HCC(CO)NH-TOCSY (or CO-TOCSY), MD triple resonance experiments using ^{13}C–^{13}C TOCSY.

0006-2960/93/0432-7839$04.00/0

(Gouda et al., 1992; Torigoe et al., 1990) and in complex with an Fc fragment of human IgG antibodies by X-ray crystallography (Deisenhofer, 1981).

In order to facilitate protein engineering of multivalent SPA-like molecules for use in high-level protein expression vectors, a synthetic IgG-binding domain the Z-domain, has also been designed and expressed in *Escherichia coli* (Nilsson et al., 1987; Cedergren et al., 1993). The DNA sequence of the Z-domain can be polymerized at the DNA level in an obligate head-to-tail fashion using nonpalindromic restriction sites, allowing the generation of repetitive genes containing mutagenized SPA-like domains. The protein sequence of the Z-domain is highly homologous to SPA domains A–E (Nilsson et al., 1987), but it is more resistant to chemical treatments useful for sequence-specific chain cleavage, as it lacks methionine residues and Asn-Gly dipeptide sequences. Uni- and divalent forms of Z-domain fusion proteins are used in high-level expression vectors for producing several mammalian proteins [for a review, see Nilsson et al. (1991)] which can be purified using IgG–Sepharose chromatography.

We have begun studies of the Z-domain by multinuclear NMR with the aim of using solution structural information in combination with protein engineering efforts (Cedergren et al., 1993) to understand its interactions with the constant part (Fc) of class G antibodies. Our Z-domain construct includes a 14 amino acid N-terminal leader sequence which does not change the biochemical properties of its interactions with Fc (Cedergren et al., 1993). Initial NMR work was hindered by the severe chemical shift degeneracy exhibited by this molecule. The spectral overlap is especially severe because of the largely disordered N-terminal leader sequence. In order to overcome this degeneracy problem, we have developed a powerful triple-resonance NMR strategy useful for rapid determination of sequence-specific resonance assignments. This strategy is based on recently described multidimensional HCCNH-TOCSY (Logan et al., 1992; Lyons & Montelione, 1993) and HCC(CO)NH-TOCSY experiments (Logan et al., 1992; Montelione et al., 1992). While initial reports (Logan et al., 1992; Montelione et al., 1992; Lyons & Montelione, 1993) have suggested that these new triple-resonance experiments can provide all of the information needed to determine most sequence-specific resonance assignments in small proteins, this proposal has yet to be validated for a complete protein sequence. In this paper we present the first demonstration of the combined use of these experiments in determining nearly complete proton and nitrogen-15 resonance assignments for this 8.2-kDa Z-domain. These results demonstrate the power of this approach even for apparently disordered regions of the polypeptide chain which exhibit little or no sequential NOE data.

MATERIALS AND METHODS

Production and Purification of Uniformly ^{15}N,^{13}C-Enriched Z-Domain. *E. coli* strain RV308 (Maurer et al., 1980) was transformed with a previously described Z-domain expression vector (Altman et al., 1991; Cedergren et al., 1993) and then grown overnight in 2× YT medium supplemented with 15 μg/mL tetracycline (Tc). The overnight culture was diluted 1:25 in 1× MJ minimal growth media (details to be published elsewhere) containing 0.4% [$^{13}C_6$]glucose (Isotec), 0.2 % ($^{15}NH_4$)$_2$SO$_4$ (Isotec), and 15 μg/mL tetracycline. Induction was performed by addition of 25 μg/mL β-indoleacrylic acid at midlog growth phase. The cells were grown for an additional 12 h and harvested by centrifugation at 4000*g* for 10 min. The cell pellet was resuspended in TST (50 mM Tris-HCl, pH 7.4, 150 mM NaCl, and 0.05% Tween) and subsequently disrupted by sonication for 5 × 30 s using a Sonifer B15 cell disruptor (Branson) equipped with a microtip. After centrifugation at 10000*g* for 10 min, the supernatant was passed over an IgG–Sepharose FF (Pharmacia Biosystems) column equilibrated with TST. The column was washed with 10 column volumes of TST and 2 column volumes of 0.5 mM ammonium acetate, pH 5.5. The protein was then eluted with 0.3 M acetic acid titrated to pH 3.2 using ammonium acetate. The yield of purified [^{13}C,^{15}N]Z-domain was approximately 40 mg/L of growth volume, as determined by quantitative amino acid analysis. The homogeneity of this sample (>95%) was further confirmed by polyacrylamide gel electrophoresis in sodium dodecyl sulfate, and the isotopic enrichment (>99%) was verified by determining the molecular mass using a JEOL SX102 mass spectrometer equipped with an electrospray unit.

NMR Spectroscopy. Samples for NMR spectroscopy were prepared in aqueous solutions of 2–3 mM protein concentration at pH 6.5 ± 0.1 containing 10 mM K_2HPO_4 and 0.2 mM NaN_3. NMR spectra were obtained with a Varian Unity 500 spectrometer system using a sample temperature of 30 ± 0.1 °C. Two-quantum filtered correlation spectroscopy (2QF-COSY) (Piatini et al., 1982), NOE spectroscopy (NOESY) (Jeener et al., 1979), total correlation spectroscopy (TOCSY) (Braunschweiler & Ernst, 1983), ^{15}N-heteronuclear single quantum coherence spectroscopy (HSQC) (Bodenhausen & Ruben, 1980), HSQC–TOCSY (Moy et al., 1993), HCCNH-TOCSY (Lyons & Montelione, 1993), and HCC(CO)NH-TOCSY (Montelione et al., 1992) were recorded using pulse sequences described elsewhere. All 2D and 3D NMR experiments were run for 24–36 spectrometer hours. Carbon isotropic mixing was carried out using the DIPSI-3 multipulse sequence (Shaka, et al., 1988) with a mixing time of 24 ms. Collection of 3D data included 64 points in the t_1 and t_2 dimensions and 1024 points in the t_3 dimension. These data were then zero-filled prior to Fourier transformation. ^{1}H chemical shifts are reported in parts per million (ppm) relative to the methyl resonance of 2,2-dimethyl-2-silapentane-5-sulfonate (DSS), assigned as 0.0 ppm. ^{15}N chemical shifts are reported in parts per million (ppm) relative to external $^{15}NH_4Cl$, assigned as 24.9 ppm.

RESULTS

The Z-domain consists of a highly helical 58-residue domain of SPA plus a largely disordered N-terminal 14-residue leader sequence (B. A. Lyons and G. T. Montelione, unpublished results). Analysis using sequential NOEs was limited by the lack of resolved data in the N-terminal polypeptide leader sequence and by significant overlap of these resonances with those of the remainder of the protein. In addition, there are internal repeats of spin system motifs. For example, polypeptide segments Asn-6–Phe-13 (NKEQQNAF) and Asn-23–Phe-30 (NEEQRNAF) have similar AMX-Lng-Lng-Lng-Lng-AMX-Ala-AMX spin system sequences. These features of the sequence and structure of the Z-domain result in high degeneracy among H^α and H^β proton resonances, which complicates the analysis by standard NOE methods (Wüthrich, 1986). Even sophisticated MD triple-resonance experiments (Montelione & Wagner, 1989, 1990; Clore et al., 1990; Ikura et al., 1990; Kay et al., 1990, 1991, 1992; Boucher et al., 1992; Clubb & Wagner, 1992) which provide connections between sequential amino acid spin systems by heteronuclear related coherence transfer through the peptide bond are challenged by this system because of a high degree of degeneracy among the H^α and C^α resonances. For these reasons, we have

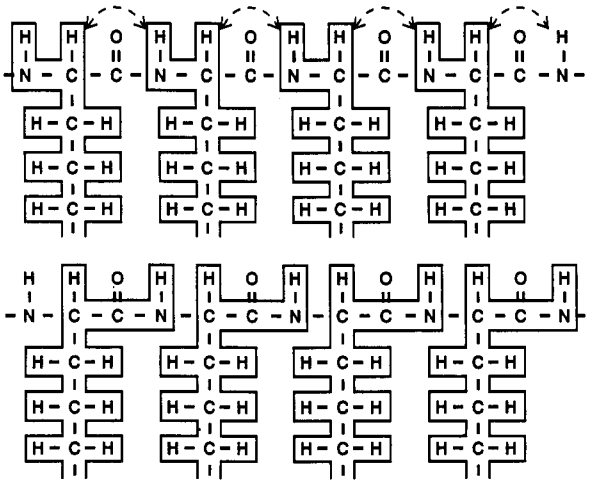

FIGURE 1: Schematic diagram showing magnetization pathways used to establish sequential relationships between spin systems in CA-TOCSY (top) and CO-TOCSY experiments (bottom). The CA-TOCSY experiment provides intraresidue correlations from backbone and side-chain aliphatic resonances to the amide nitrogen and proton resonances of the same residue, while CO-TOCSY data provide sequential correlations from the backbone and aliphatic resonances of residue i to the amide nitrogen and proton resonances of residue $i + 1$. In addition, the CA-TOCSY experiment results in weak sequential correlations, indicated by the dashed double-headed arrows.

developed a more powerful strategy for establishing sequential resonance assignments in highly degenerate small protein systems like the Z-domain.

Our strategy, outlined in Figure 1, is based on recently developed multidimensional HCC(CO)NH-TOCSY (Logan et al., 1992; Montelione et al., 1992) and HCCNH-TOCSY experiments (Logan et al., 1992; Lyons & Montelione, 1993). For convenience, we refer to these two triple-resonance experiments in the remainder of this paper as CO-TOCSY and CA-TOCSY, respectively. For each amino acid spin system, "CA-ladders" are first identified using 3D CA-TOCSY spectra. A CA-ladder consists of a string of 3D cross peaks for a single amino acid residue corresponding to the aliphatic proton (or carbon-13) resonances in the ω_1 dimension, the backbone amide nitrogen resonance in the ω_2 dimension, and the backbone amide proton resonance in the ω_3 dimension. Next, for each amino acid residue a second "ladder", a "CO-ladder", is identified using 3D CO-TOCSY. A CO-ladder consists of a string of 3D cross peaks representing a link between amino acid residues (Figure 1). These correspond to the aliphatic (or carbon-13) resonances of residue i in the ω_1 dimension and the backbone amide nitrogen and proton resonances of residue $i + 1$ in the ω_2 and ω_3 dimensions, respectively. These CO-TOCSY data provide connections exclusively between sequential amino acid spin systems (Montelione et al., 1992), except for some intraresidue connections to side-chain amides discussed below, and are therefore easier to interpret than NOESY spectra.

The CO-TOCSY and CA-TOCSY pulse sequences have been reported previously (Montelione et al., 1992; Lyons & Montelione, 1993). Similar pulse sequences have also been described recently by the Fesik laboratory (Logan et al., 1992). While it is not necessary to repeat the details of these multipulse experiments here, it is useful to understand how the experiments work in order to apprciate the strategy outlined in Figure 1. In both experiments, the magnetization pathway begins on the aliphatic protons of residue i. After this proton magnetization is frequency-labeled in the t_1 period, it is transferred to attached ^{13}C nuclei using refocused INEPT (Morris, 1980). This carbon magnetization is then transferred

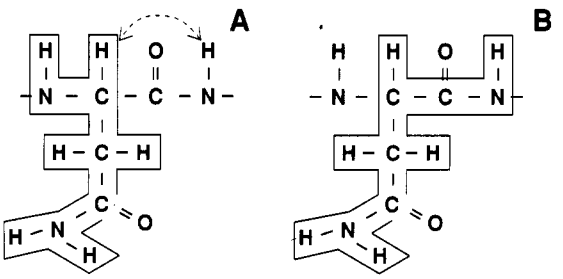

FIGURE 2: Schematic representation showing magnetization pathways used to identify asparagine spin systems in (A) CA-TOCSY and (B) CO-TOCSY experiments. Magnetization beginning on H^{α} and H^{β} resonances is transferred first to aliphatic carbon-13 nuclei, then between carbon-13 nuclei using isotropic mixing, to the side-chain and backbone amide nitrogen nuclei, and finally to these amide proton nuclei for detection. Three-dimensional cross peaks at $\omega_1 = H^{\alpha}$ and/or H^{β}, $\omega_2 = N^{\delta}$, and $\omega_3 = H^{\delta 1}$ and $H^{\delta 2}$ allow unique identification of some AMX-type spin systems with asparagine residues. A similar strategy was used to identify some long-type spin systems as glutamine residues.

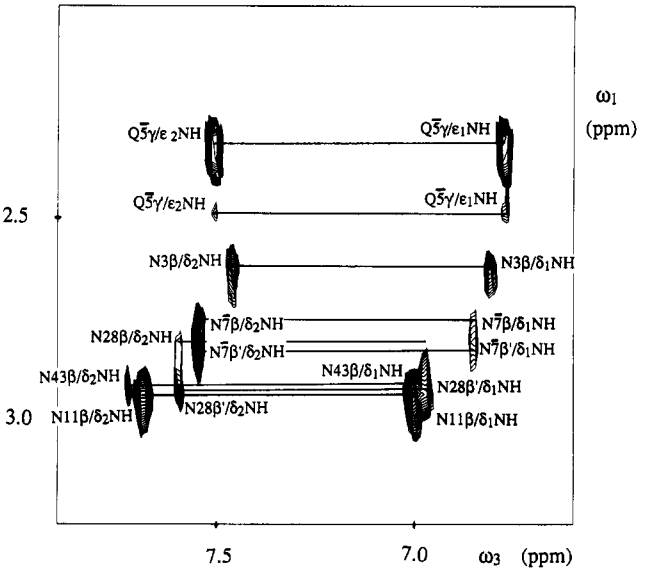

FIGURE 3: A ^{15}N-plane from the 3D CA-TOCSY spectrum of the Z-domain showing cross peaks used to identify several asparagine and glutamine spin systems. Each H^{β}–N^{δ}–H^{δ} (or H^{γ}–N^{ϵ}–H^{ϵ}) 3D cross peak of these asparagine (and glutamine) spin systems is labeled with the corresponding sequence-specific resonance assignment using the one-letter code for amino acid residues. In our notation, the residue numbers corresponding to the N-terminal 14 amino acid leader sequence of the Z-domain are designated with a bar over the sequence number (e.g., the asparagine residue at position −7 is designated $N\bar{7}$).

throughout the aliphatic carbon skeleton using ^{13}C isotropic mixing (Shaka et al., 1988). At this point, the CA-TOCSY and CO-TOCSY experiments follow different pathways (Figure 1). In CA-TOCSY, magnetization originating on peripheral aliphatic nuclei of residue i which ends up on the carbon-α nucleus at the end of the isotropic mixing step is next transferred to the nitrogen nucleus of residue i using a concatenated INEPT sequence. This nitrogen magnetization is frequency-labeled during the t_2 period and then transferred to the amide proton of residue i, which is detected in the t_3 period. In CO-TOCSY, the C^{α} magnetization originating on peripheral aliphatic nuclei of residue i is transferred to the carbonyl carbon of residue i and then to the nitrogen nucleus of residue $i + 1$. This nitrogen magnetization is frequency-labeled during the t_2 period and then transferred to the amide proton of residue $i + 1$ for detection during the t_3 period.

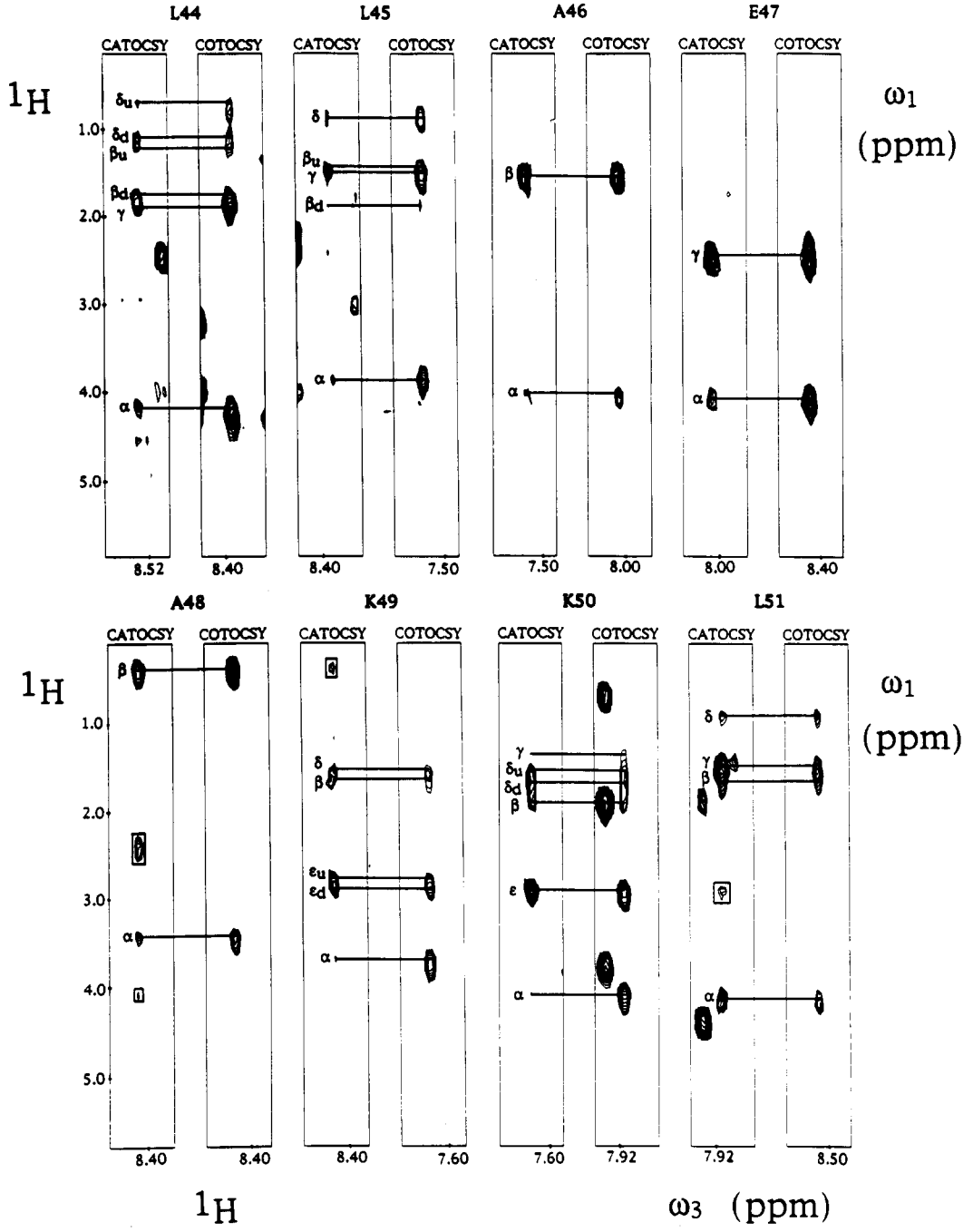

FIGURE 4: Experimental data showing slices of 3D CA-TOCSY and CO-TOCSY data used to race out the sequential connections between amino acid spin systems of the Z-domain. Shown in the figure are slices at different nitrogen frequencies providing intraresidue and sequential connections for the polypeptide fragment Leu-44–Leu-51. In each slice are 3D cross peaks between the aliphatic proton resonances of residue i (CA-TOCSY) or residue $i-1$ (CO-TOCSY) (ω_1 dimension), the ^{15}N resonance of residue i (ω_2 dimension), and the H^N resonance of residue i (ω_3 dimension). These $\omega_2 = {}^{15}$N slices are labeled at the top with the name of residue i (CA-TOCSY) or $i-1$ (CO-TOCSY). Each pair of cross peaks indicating a sequential connection is identified with a horizontal line at the ω_1 frequency of the aliphatic proton resonance. As stereospecific resonance assignments are not available yet for this protein, the stereochemically distinct resonances of methylene and isopropylmethyl groups are designated with u (upfield) and d (downfield) labels. Cross peaks shown in rectangular boxes result from sequential transfer in the CA-TOCSY data from the aliphatic resonances of residue i to the amide group of residue $i+1$, which were suppressed as much as possible by appropriate tuning of the coherence-transfer delays (Lyons & Montelione, 1993).

In principle, 3D CA-TOCSY data provide sufficient information to characterize many spin system types, particularly when ^{13}C frequency information is included in the analysis. For the Z-domain, however, proton spin systems were first characterized using 2D 2QF-COSY and TOCSY data recorded on an unenriched Z-domain sample dissolved in ^2H_2O. Nitrogen resonances were then connected to the spin systems using ^{13}C-decoupled ^{15}N-HSQC–TOCSY together with 2D and 3D CA-TOCSY data using a sample of [^{15}N,^{13}C]Z-domain. Identification of spin system types generally followed the classic method (Wüthrich, 1986). However, it was also possible to use the CA- and CO-TOCSY data to uniquely identify some AMX spin systems as asparagines and some long-type spin systems as glutamines (Montelione et al., 1992). This strategy is outlined in Figure 2, and a representative slice showing the relevant connections between aliphatic and side-chain amide resonances in a 3D CA-TOCSY spectrum is shown in Figure 3. In this way AMX-type spin systems later assigned to residues Asn-(–7), Asn-3, Asn-11, Asn-28, Asn-43, and Asn-52 were identified as

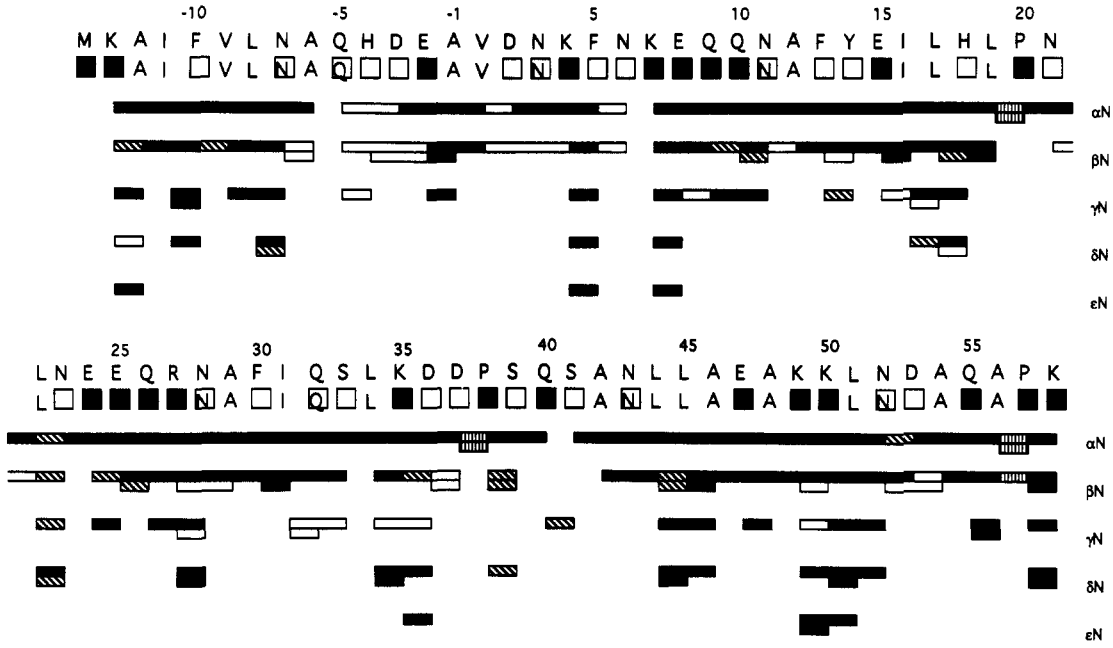

FIGURE 5: Survey of sequential CO-TOCSY connections used in determining resonance assignments for the Z-domain. The amino acid sequence of the Z-domain is shown along with the corresponding sequence of spin system types which are either unique, AMX-type (□), or long-type (■). The AMX and long spin systems which could be connected to side-chain amide protons of asparagine or glutamine with CA-TOCSY or CO-TOCSY data are designated as N and Q spin system types, respectively. Cross peaks from the H^{α}, H^{β}, H^{γ}, H^{δ}, and H^{ϵ} resonances of residue i to the H^{N} resonance of residue $i + 1$ in the CO-TOCSY spectrum are indicated by horizontal bars labeled αN, βN, γN, δN, and ϵN, respectively. Open, diagonally hatched, and solid horizontal bars indicate that the corresponding sequential connection was identified in 2D CO-TOCSY only, in 3D CO-TOCSY only, or in both spectra. As proline residues do not have an amide proton, Xxx-Pro dipeptides cannot show connectivities in CO-TOCSY experiments, and the corresponding vertically hatched links indicate sequential H^{α}_{i}–H^{δ}_{i+1} and H^{β}_{i}–H^{δ}_{i+1} connections established by analysis of 2D NOESY data.

asparagine spin systems, and long-type spin systems later assigned to residues Gln-(–5) and Gln-32 were identified as glutamine spin systems.

The resulting spin system list was next used to identify CA-ladders in the 3D CA-TOCSY spectrum of the Z-domain. These ladders were then used to search for corresponding CO-ladders in the 3D CO-TOCSY spectrum. This combined analysis is shown for the polypeptide segment Leu-44 – Leu-51 of the Z-domain in Figure 4, and the complete set of connectivities used in establishing sequence-specific resonance assignments for all but the N-terminal Met-(–14) residue are summarized in Figure 5. These resonance assignments are tabulated in Table I. In most cases, multiple sequential connections were established between each amino acid spin system in the sequence, including many H^{δ}_{i}–H^{N}_{i+1} and H^{ϵ}_{i}–H^{N}_{i+1} connections (Figure 5). In the polypeptide segment Ala-(–5)–Glu(–1), the amide proton intensities are very weak due to saturation-transfer effects, and these sequential connections could only be established using 2D CO-TOCSY spectra, which exhibit higher signal-to-noise ratios. Connections in this region of the sequence have also been confirmed using pulsed-field gradient CO-TOCSY experiments (L. Wang, B. A. Lyons, and G. T. Montelione, manuscript in preparation).

DISCUSSION

A key step in the process of determining protein structures and dynamics by NMR spectroscopy is the assignment of proton, nitrogen, and carbon resonances to their sequence-specific positions in the primary amino acid sequence. These sequence-specific resonance assignments provide the basis for interpreting multidimensional NMR spectra and for determining 3D structures of proteins from these data (Wüthrich, 1986). Traditional methods (Wüthrich, 1986) in use in most protein NMR laboratories for determining these resonance assignments involve two key steps. First, MD NMR experiments are used to identify spin systems of aliphatic and backbone amide resonances for each amino acid. Next, MD NOESY data (Jeener et al., 1979) are used to identify sequential relationships between these spin systems and to order them according to the amino acid sequence. Sequential connections between amino acid spin systems are established using short distances (<4 Å) between the H^{N}, H^{α}, and H^{β} atoms of residue i and the H^{N} atom of the next residue in the sequence, residue $i + 1$ (Dubs et al., 1979; Billeter et al., 1982). These sequential distances are conformation dependent and are sometimes too long to result in an observable NOESY cross peak. However, sequential connections can be established with high statistical certainty when two or more such sequential NOE interactions are identified between a pair of spin systems (Billeter et al., 1982). The analysis of sequential connections from NOE data is further complicated by the fact that, besides the intraresidue and sequential interactions in NOESY spectra, there are additional longer range interactions that arise from the folding of the protein structure. These complexities slow the analysis of sequence-specific resonance assignments from NOESY spectra and represent one of the primary bottlenecks in the process of determining protein structures from NMR data.

In order to overcome these shortcomings of the NOESY experiment in establishing sequential connections between amino acid spin systems, several groups have developed multidimensional triple-resonance pulse sequences which use "through-bond" scalar coupling interactions between ^{1}H, ^{13}C, and ^{15}N nuclei to transfer magnetization from one amino acid spin system to the next directly through the peptide bonds of doubly ^{15}N,^{13}C-enriched proteins. The first generation of these triple-resonance experiments (Montelione & Wagner, 1989,

Table I: Proton and Nitrogen-15 Resonance Assignments for the Z-Domain of Staphylococcal Protein A at 30 °C and pH 6.5

amino acid residue	chemical shifts (ppm)				
	N	HN	Hα	Hβ	others
K(−13)			4.35	1.78	Hγ 1.44; Hδ 1.68; Hε 2.98
A(−12)	123.4	8.35	4.30	1.30	
I(−11)	116.9	7.97	4.12	1.75	Hγ 1.10, 1.36; $H_3\gamma$ 0.80; $H_3\delta$ 0.80
F(−10)	121.2	8.17	4.68	3.06, 2.98	Hδ 7.18
V(−9)	120.0	7.94	4.04	1.97	$H_3\gamma$ 0.86
L(−8)	123.0	8.18	4.25	1.97	Hγ 1.57; $H_3\delta$ 0.91, 0.86
N(−7)	116.3	8.37	4.66	2.82, 2.76	Hδ 7.56, 6.86; Nδ 117.4
A(−6)	120.9	8.15	4.26	1.35	
Q(−5)			4.27	2.06, 1.97	Hγ 2.49, 2.32; Hε 7.50, 6.77; Nε 117.1
H(−4)	119.3	8.32	4.67	3.21, 3.13	Hδ4 8.30; Hε2 7.17
D(−3)		8.30	4.61	2.67, 2.59	
E(−2)	119.2	8.47	4.22	2.05, 1.96	Hγ 2.27
A(−1)	121.1	8.23	4.26	1.35	
V1	115.4	7.79	3.78	1.95	$H_3\gamma$ 0.74, 0.76
D2	119.1	7.98	4.42	2.47	
N3	116.0	8.08	4.57	2.59	Hδ 7.46, 6.82; Nδ 117.3
K4	116.7	8.16	4.21	1.61	Hγ 1.24; Hδ 1.69; Hε 2.90
F5	115.9	7.84	5.07	3.41, 3.13	Hδ 7.07; Hε 6.95; Hζ 7.25
N6	117.8	8.42	4.78	3.35, 2.95	Hδ 7.44, 6.87; Nδ 115.0
K7	115.9	8.30	3.99	1.88	Hγ 1.45; Hδ 1.72; Hε 3.02
E8	116.8	8.20	4.11	2.1	Hγ 2.32
Q9	118.8	8.49	3.88	2.50, 2.25	Hγ 1.55; Hε 7.21, 6.91; Nε 115.5
Q10	116.2	8.72	3.93	2.18, 2.57	Hγ 2.42; Hε 7.22, 6.84; Nε 116.3
N11	114.8	8.29	4.62	2.92	Hδ 7.68, 7.00; Nδ 117.1
A12	119.5	7.86	4.10	1.47	
F13	115.3	8.10	3.81	3.32, 2.95	Hδ 6.92; Hε 7.06; Hζ 7.25
Y14	113.9	8.14	3.95	3.16	Hδ 7.17; Hε 6.78
E15	116.3	8.48	4.03	2.15	Hγ 2.42
I16	116.7	8.37	3.39	1.78	Hγ 1.78, 1.88; $H_3\gamma$ 0.75; $H_3\delta$ 0.53
L17	114.9	7.85	3.68	1.45, 1.13	Hγ 1.32; $H_3\delta$ 0.63, 0.53
H18	109.0	7.18	4.52	3.45, 2.84	Hδ4 7.09; Hε2 8.21
L19	121.7	7.21	4.49	1.75, 1.38	Hγ 2.23; $H_3\delta$ 0.86, 0.68
P20			4.42		Hδ 4.08, 3.82
N21	111.0	8.86	5.03	2.90	Hδ 7.37, 6.94; Nδ 119.3
L22	114.6	6.48	4.43	1.69	Hγ 1.69; $H_3\delta$ 0.96, 0.88
N23	116.3	8.51	4.94	3.28, 2.83	Hδ 7.46, 7.00; Nδ 116.7
E24	115.6	8.60	3.98	2.05	Hγ 2.37
E25	117.0	8.23	4.03	2.26, 2.04	Hγ 1.94
Q26	116.9	8.46	3.97	2.48	Hγ 2.35; Hε 8.21, 7.60; Nε 117.6
R27	116.1	8.53	3.78	1.85, 1.75	Hγ 1.73, 1.46; Hδ 3.43, 3.23; Hε 7.63
N28	112.3	8.43	4.40	2.90, 2.79	Hδ 6.97, 7.58; Nδ 117.0
A29	120.4	7.83	4.17	1.33	
F30	114.2	7.96	4.37	3.10, 2.99	Hδ 7.31; Hε 7.20; Hζ 7.14
I31	115.6	8.25	3.75	2.11	Hγ 1.59, 1.35; $H_3\gamma$ 0.97; $H_3\delta$ 0.62
Q32	116.7	8.41	3.94	2.26	Hε 7.82, 6.89; Nε 120.9
S33	112.3	8.01	4.28	4.10, 3.98	
L34	121.4	8.08	3.78	1.95	Hγ 1.64; $H_3\delta$ 0.78, 0.76
K35	112.8	7.97	4.00	1.98	Hδ 1.68, 1.40; Hε 2.88
D36	115.3	8.10	4.43	2.78	
D37	111.0	7.55	4.93	2.99, 2.57	
P38			4.51	2.16, 1.91	Hγ 2.24; Hδ 3.87, 3.70
S39	109.9	8.00	4.36	4.04, 3.98	
Q40	117.6	7.83	4.62	2.66, 2.00	Hγ 2.46, 2.32; Hε 7.58, 6.81; Nε 118.8
S41	112.8	7.75	3.74	4.00	
A42	120.2	8.44	4.15	1.42	
N43	115.7	7.87	4.54	2.91	Hδ 7.73, 6.98; Nδ 117.7
L44	118.7	8.56	4.18	1.74, 1.27	Hγ 1.87; $H_3\delta$ 1.11, 0.79
L45	116.0	8.38	3.83	1.92, 1.46	Hγ 1.53; $H_3\delta$ 0.91, 0.91[a]
A46	116.7	7.56	4.05	1.54	
E47	116.3	8.02	4.04	2.23	Hγ 2.46
A48	121.2	8.43	3.49	0.50	
K49	115.3	8.45	3.78	1.95, 1.65	Hγ 1.34; Hδ 1.61; Hε 2.94, 2.84
K50	117.4	7.66	4.13	1.96	Hγ 1.41; Hδ 1.72, 1.62; Hε 2.96
L51	119.3	7.90	4.19	1.70	Hγ 1.57; $H_3\delta$ 1.00
N52	114.2	8.53	3.97	3.11, 2.39	Hδ 7.91, 6.80; Nδ 120.9
D53	115.9	8.20	4.48	2.77, 2.72	
A54	120.1	8.01	4.24	1.58	
Q55	112.2	7.52	4.40	2.32, 1.82	Hγ 2.67, 2.46; Hε 8.70, 7.25; Nε 115.8
A56	121.6	7.09	4.36	1.45	
P57			4.43	2.32, 1.97	Hγ 2.08; Hδ 3.80, 3.64
K58	124.5	8.00	4.20	1.86	Hγ 1.44; Hδ 1.70; Hε 3.04

[a] The two degenerate methyl resonances of Leu-45 have distinct chemical shifts at pH 6.5 and a temperature of 10 °C.

1990; Clore et al., 1990; Ikura et al., 1990; Kay et al., 1990, 1991, 1992; Boucher et al., 1992; Clubb & Wagner, 1992) provide magnetization transfer between H^{α}, $^{13}C^{\alpha}$, and/or $^{13}C'$ carbonyl nuclei of residue *i* and H^N and ^{15}N amide nuclei of

residue $i + 1$. These methods have proven successful in determining sequence-specific resonance assignments in several proteins. However, even when these experiments are carried out with four-dimensional frequency labeling, degeneracies among H^{α}, $^{13}C^{\alpha}$, and $^{13}C'$ resonances of several amino acid spin systems within a protein can prevent the identification of unambiguous sequential connections. More recently, a second generation of triple-resonance experiments have been developed, including CBCANH (Grzesiek & Bax, 1992a), CBCA(CO)NH (Grzesiek & Bax, 1992b), CA-TOCSY (Logan et al., 1992; Lyons & Montelione, 1993), and CO-TOCSY (Logan et al., 1992; Montelione et al., 1992), which overcome degeneracies among backbone H^{α}, $^{13}C^{\alpha}$, and $^{13}C'$ atoms by also frequency-labeling proton and carbon resonances of the amino acid side chains. This work on the Z-domain demonstrates that these new experiments provide a powerful approach for determining resonance assignments for complete sequences of small proteins. The resulting spectra are significantly simpler to analyze than NOESY spectra and provide a more extensive set of sequential connectivity information than first-generation triple-resonance experiments. In addition, MD CA- and CO-TOCSY spectra are highly amenable to automated analysis with artificial intelligence (Zimmerman, et al., 1993).

One disadvantage of our approach to sequential assignments is that the strategy requires uniformly ^{13}C,^{15}N-enriched proteins. For this reason, high-level expression systems are essential for successful implementation of these experiments. The recombinant production system utilized here for isotope enrichment of Z-domain (Cedergren et al., 1993), for example, yielded as much as 40 mg of purified protein per liter of growth medium containing 4 g/L of [$^{13}C_6$]glucose. A second shortcoming of this triple-resonance approach is that the magnetization-transfer pathways in both CA- and CO-TOCSY involve significant time periods (10–50 ms) during which they are transverse on C^{α} nuclei that in proteins have relatively short transverse relaxation times (T_2's). In larger proteins (>20 kDa) these $^{13}C^{\alpha}$ T_2's may be too short to allow efficient magnetization transfer. In this regard, the CO-TOCSY experiment generally exhibits higher cross-peak intensities than CA-TOCSY because it is less sensitive to the effects of $^{13}C^{\alpha}$ relaxation. The higher signal-to-noise of CO-TOCSY data is apparent even in the spectra of the Z-domain shown in Figure 4. Despite these limitations, we anticipate that the general strategy outlined here and demonstrated for the first time in determining nearly complete resonance assignments for the Z-domain will become a standard approach for determining resonance assignments in small isotope-enriched proteins.

NOTE ADDED IN PROOF

Since submitting this work for publication, we have learned that a similar strategy has also been used to determine many resonance assignments for the FK506 binding protein denatured in urea. [Logan, T. M., et al. (1993) *J. Biomol. NMR 3*, 225–231].

ACKNOWLEDGMENT

We thank Drs. P.-O. Edlund and B. Norén for performing the mass spectrometry analysis and Dr. G. Palm for performing the amino acid analyses. We also thank Ms. D. Zimmerman for useful discussions.

REFERENCES

Altman, J. D. Henner, D., Nilsson, B., Anderson, S., & Kuntz, I. D. (1991) *Protein Eng. 4*, 593–600.

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

Bodenhausen, G., & Ruben, D. J. (1980) *Chem. Phys. Lett. 69*, 185–189.

Boucher, W., Laue, E. D., Campbell-Burk, S., & Domaille, P. J. (1992) *J. Am. Chem. Soc. 114*, 2262–2264.

Braunschweiler, L., & Ernst, R. R. (1983) *J. Magn. Reson. 53*, 521–528.

Cedergren, L., Andersson, R., Jansson, B., Uhlén, M., & Nilsson, B. (1993) *Protein Eng.* (in press).

Clore, G. M., Bax, A., Driscoll, P. C., Wingfield, P. T., & Gronenborn, A. M. (1990) *Biochemistry 29*, 8172–8184.

Clubb, R. T., & Wagner, G. (1992) *J. Biomol. NMR 2*, 389–394.

Deisenhofer, J. (1981) *Biochemistry 20*, 2361–2370.

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

Forsgren, A., & Sjöquist, J. (1966) *J. Immunol. 97*, 822–827.

Gouda, H., Torigoe, H., Saito, A., Sato, M., Arata, Y., & Shimada, I. (1992) *Biochemistry 31*, 9665–9672.

Grzesiek, S., & Bax, A. (1992a) *J. Magn. Reson. 99*, 201–207.

Grzesiek, S., & Bax, A. (1992b) *J. Am. Chem. Soc. 114*, 6291–6293.

Ikura, M., Kay, L. E., & Bax, A. (1990) *Biochemistry 29*, 4659–4667.

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

Kay, L. E., Ikura, M., Tschudin, R., & Bax, A. (1990) *J. Magn. Reson. 89*, 496–514.

Kay, L. E., Ikura, M., Zhu, G., & Bax, A. (1991) *J. Magn. Reson. 91*, 422–428.

Kay, L. E., Wittekind, M., McCoy, M., Friedrichs, M. S., & Müller, L. (1992) *J. Magn. Reson. 98*, 443–450.

Logan, T. M., Olejniczak, E. T., Xu, R. X., & Fesik, S. W. (1992) *FEBS Lett. 314*, 413–418.

Lyons, B. A., & Montelione, G. T. (1993) *J. Magn. Reson. 101B*, 206–209.

Maurer, R., Meyer, B. J., & Ptashne, M. (1980) *J. Mol. Biol. 139*, 147–161.

Moks, T., Abrahmsén, L., Nilsson, B., Hellman, U., Sjöquist, J., & Uhlén, M. (1986) *Eur. J. Biochem. 156*, 637–643.

Montelione, G. T., & Wagner, G. (1989) *J. Am. Chem. Soc. 111*, 5474–5475.

Montelione, G. T., & Wagner, G. (1990) *J. Magn. Reson. 87*, 183–188.

Montelione, G. T., Lyons, B. A., Emerson, S. D., & Tashiro, M. (1992) *J. Am. Chem. Soc. 114*, 10974–10975.

Morris, G. (1980) *J. Am. Chem. Soc. 102*, 428–429.

Moy, F. J., Li, Y.-C., Rauenbuehler, P., Winkler, M. J., Scheraga, H. A., & Montelione, G. T. (1993) *Biochemistry* (in press).

Nilsson, B., Moks, T., Jansson, B., Abrahmsén, L., Elmblad, A., Holmgren, E., Henrichson, C., Jones, T. A., & Uhlén, M. (1987) *Protein Eng. 1*, 107–113.

Nilsson, B., Forsberg, G., & Hartmanis, M. (1991) *Methods Enzymol. 198*, 3–16.

Piatini, U., Sørensen, O. W., & Ernst, R. R. (1982) *J. Am. Chem. Soc. 104*, 6800–6801.

Shaka, A. J., Lee, C. J., & Pines, A. (1988) *J. Magn. Reson. 77*, 274–293.

Sjödahl, J. (1977) *Eur. J. Biochem. 73*, 343–351.

Torigoe, H., Shimada, I., Saito, A., Sato, M., & Arata, Y. (1990) *Biochemistry 29*, 8787–8793.

Uhlén, M., Guss, B., Nilsson, B., Gatenbeck, S., Philipson, L., & Lindberg, M. (1984) *J. Biol. Chem. 259*, 1695–1702.

Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, New York.

Zimmerman, D., Kulikowski, C. & Montelione, G. T. (1993) *Proceedings of the First International Conference of Intelligent Systems for Molecular Biology*, Washington, DC (in press).