Sequence-Specific Assignments of the Backbone ^{1}H, ^{13}C, and ^{15}N Resonances of the MutT Enzyme by Heteronuclear Multidimensional NMR†

Chitrananda Abeygunawardana,‡ David J. Weber,‡,§ David N. Frick,‖,⊥ Maurice J. Bessman,‖ and Albert S. Mildvan*,‡

Department of Biological Chemistry, The Johns Hopkins School of Medicine, 725 North Wolfe Street, Baltimore, Maryland 21205, and Department of Biology, The Johns Hopkins University, Charles and 34th Streets, Baltimore, Maryland 21218

Received July 15, 1993; Revised Manuscript Received September 17, 1993®

ABSTRACT: The MutT protein, a 129-residue enzyme from *Escherichia coli* which prevents A·T → C·G mutations, catalyzes the hydrolysis of nucleoside triphosphates (NTP) to nucleoside monophosphates (NMP) and pyrophosphate [Bhatnagar, S. K., Bullions, L. C., & Bessman, M. J. (1991) *J. Biol. Chem. 266*, 9050–9054], by a mechanism involving nucleophilic substitution at the rarely attacked β-phosphorus of NTP [Weber, D. J., Bhatnagar, S. K., Bullions, L. C., Bessman, M. J., & Mildvan, A. S. (1992a) *J. Biol. Chem. 267*, 16939–16942]. The bacterial MutT gene was inserted into the plasmid pET-11b under control of the T7 promoter and overexpressed in minimal media, permitting labeling of MutT with ^{13}C and/or ^{15}N. The yield after purification of the soluble fraction was ~35 mg of homogeneous MutT/L with physical and enzymatic properties indistinguishable from those of the originally isolated enzyme. Essentially complete sequence-specific assignments of the backbone HN, N, Cα, Hα, and CO resonances of the free enzyme (1.5 mM) were made at pH 7.4 and 32 °C, by heteronuclear double- and triple-resonance experiments using a modified Bruker AM 600 NMR spectrometer. Specifically, $^{1}H[^{15}N]$HSQC, $^{1}H[^{15}N]$TOCSY-HMQC, and $^{1}H[^{15}N]$NOESY-HMQC experiments were done with uniformly ^{15}N-labeled enzyme. A $^{1}H[^{15}N]$HSQC experiment was done with selective [α-^{15}N]Lys-labeled enzyme. Also HNCA, HN(CO)CA, HNCO, constant time $^{1}H[^{13}C]$HSQC, HCACO, and HCA(CO)N experiments were done with uniformly ^{13}C- and ^{15}N-labeled enzyme. Sequence-specific assignments were initiated from HN and ^{15}N chemical shifts of Gly residues and of selectively labeled Lys residues in $^{1}H[^{15}N]$HSQC experiments. They were confirmed by Cα chemical shifts of Ala residues uniquely identified by residual coupling to Cβ resonances in constant time $^{1}H[^{13}C]$HSQC experiments. The sequence-specific assignments proceeded bidirectionally, terminating at Pro residues and at residues with undetectable NH signals, and the segments were linked to complete the backbone assignments. The backbone assignments reported here have permitted the interpretation of NOEs in the elucidation of the solution secondary structure of MutT, and the Cα and Hα chemical shifts have provided an independent approach to identifying secondary structural elements and to define their extent [Weber, D. J., Abeygunawardana, C., Bessman, M. J., & Mildvan, A. S. (1993) *Biochemistry* (following paper in this issue)].

The MutT protein is one of a class of enzymes which may be involved in "sanitizing" the nucleotide pool by selectively hydrolyzing nucleotides deleterious to DNA replication (Bhatnagar *et al.*, 1991). Deficiency of this protein in its active form results in a profound increase in AT to CG transversions in *Escherichia coli* (Treffers *et al.*, 1954; Yanofsky *et al.*, 1966; Cox, 1973; Bhatnagar *et al.*, 1990; Bullions, 1993). The reaction catalyzed by the MutT enzyme in the presence of divalent cations is the unusual hydrolysis of nucleoside or deoxynucleoside triphosphates (NTP) to yield nucleotides (NMP) and pyrophosphate (Bhatnagar *et al.*, 1991) by nucleophilic substitution at the rarely attacked β-phosphorus of NTP (Weber *et al.*, 1992a). Like other enzymes which catalyze substitution at Pβ of NTP, the MutT enzyme requires two divalent cations for activity (Frick *et al.*, 1993). Although MutT hydrolyzes all of the canonical nucleoside triphosphates at measurable rates, it has a strong preference for dGTP, and especially for purine dNTP's with a bulky substituent at C8 (Bhatnagar *et al.*, 1991). One of these, 8-oxo-dGTP, has been reported by Maki and Sekiguchi (1992) to be an excellent substrate for the MutT protein, and

† This work supported by the National Institutes of Health Grant DK28616 (to A.S.M.) and Grant GM18649 (to M.J.B.).

* To whom correspondence should be addressed. Phone: 410-955-2038; FAX: 410-955-5759.

‡ The Johns Hopkins School of Medicine.

§ Present address: Department of Biological Chemistry, University of Maryland School of Medicine, 108 N. Greene Street, Baltimore, MD 21201.

‖ The Johns Hopkins University.

⊥ D.N.F. is supported by National Institutes of Health Training Grant 5-T32-GM07231.

® Abstract published in *Advance ACS Abstracts*, November 15, 1993.

[1] Abbreviations: COSY, correlation spectroscopy; CT, constant time; DANTE, delays alternating with nutation for tailored excitation; GARP, globally optimized alternating-phase rectangular pulses; HCACO, proton to α-carbon to carbonyl correlation; HCA(CO)N, proton to α-carbon to nitrogen (via carbonyl) correlation; HMQC, heteronuclear multiple-quantum correlation; HNCA, amide proton to nitrogen to α-carbon correlation; HNCO, amide proton to nitrogen to carbonyl correlation; HN(CO)CA, amide proton to nitrogen to α-carbon (via carbonyl) correlation; HOHAHA, homonuclear Hartman–Hahn spectroscopy; HSQC, heteronuclear single-quantum correlation; INEPT, insensitive nuclei enhancement by polarization transfer; IPTG, isopropyl thio-β-D-galactoside; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; RF, radio frequency; TOCSY, total correlation spectroscopy; 2D, two dimensional; 3D, three dimensional; TPPI, time-proportional phase incrementation; Tris-HCl, tris(hydroxymethyl)aminomethane hydrochloride; TSP, 3-(trimethylsilyl)propionate-*2,2,3,3-d*$_4$.

0006-2960/93/0432-13071$04.00/0

they have proposed that MutT exerts its antimutagenic activity by removing the potentially mutagenic 8-oxo-dGTP from the nucleotide pool. Recently, an activity related to that of MutT has been reported to be present in extracts of a human cell line (Mo *et al.*, 1992).

Because of its small size (129 residues) and profound mechanistic interest, we have undertaken a determination of the solution structure of the MutT enzyme by NMR[1] methods. As expected for a protein of this size, the classical approach to the sequential assignment of resonances based solely on 2D proton NMR (Wüthrich, 1986) was unsuccessful due to severe resonance overlap and broadening even at 600 MHz. For this reason, we have greatly increased the yield of active enzyme by placing the *mutT* gene under the control of the T7 promoter, permitting ^{13}C and ^{15}N enrichment, and have assigned the backbone resonances of the enzyme using the recently developed triple-resonance experiments in which sequential connectivities are based on scalar coupling between ^{15}N- and ^{13}C-enriched backbone nuclei (Ikura *et al.*, 1990a,b). This strategy has been applied to several proteins including calmodulin (Ikura *et al.*, 1990b, 1991), the RNase H domain of HIV-1 reverse transcriptase (Powers *et al.*, 1991), and *E. coli* enzyme III^{Glc} (Pelton *et al.*, 1991), where the resolution afforded by these experiments was sufficient to obtain complete backbone assignments with minimal identification of side chain spin systems. An abstract of this work has been published (Abeygunawardana *et al.*, 1993).

MATERIALS AND METHODS

MutT Preparation

For ^{15}N and ^{13}C labeling of the MutT enzyme, large yields of the purified MutT protein in minimal media were required. Accordingly, the gene for MutT was ligated into a pET11b plasmid (Novagen, Inc.) containing a T7 promoter. In this procedure, a segment of DNA containing the MutT gene was prepared utilizing the polymerase chain reaction with oligonucleotides which anneal to the 3′ and 5′ ends of the MutT gene and which contain *Bam*I and *Nde*I restriction sites, respectively. The resulting segment of DNA containing the MutT gene and the plasmid pET11b were then both hydrolyzed with the restriction enzymes *Nde*I and *Bam*I, purified, mixed, and ligated to give the plasmid required to produce the complete MutT enzyme with no additional residues. The plasmid was transformed into HB101 cells and sequenced to verify the presence of the intact, wild-type MutT gene. The purified MutT-containing plasmid was then transformed into *E. coli* strain HMS174 (DE3) cells (Novagen, Inc.) grown at 37 °C in M9 (Maniatis *et al.*, 1982) or MOPS minimal media (Weber *et al.*, 1992b) until the OD at 600 nm was ~1.0. At that point IPTG was added, and the cells were grown for an additional 2 h prior to harvest. Longer growth periods resulted in a loss of MutT to inclusion bodies. After opening the cells with a French press, soluble MutT was found to account for at least 30% of the total cell protein to give 35 mg of fully active MutT enzyme per liter of minimal media.

Purification of MutT from the cell lysates to >98% homogeneity was performed as previously described (Bhatnagar & Bessman, 1988) except that the final two DEAE-cellulose columns were substituted with a single DEAE-Sepharose fast-flow column at pH 7.5, eluting with 50 mM Tris-HCl, pH 7.5, with a stepwise gradient from 50 to 150 mM NaCl. At this stage, the protein was judged to be more than 98% pure by SDS gel electrophoresis. The MutT protein was then concentrated by lyophilization and redissolving in 50 mM NaCl and 1 mM d_{11}-Tris-HCl pH 7.5, passed down a Sephadex G-25 column, dialyzed against 1 mM d_{11}-Tris-HCl containing a 0.01% (v/v) suspension of Chelex 100 to remove trace metal contaminants, and again concentrated by lyophilization. The homogeneous MutT protein had a specific activity of 70 ± 5 units/mg under standard assay conditions (Bhatnagar et al., 1991), and a 1% solution of this pure MutT gave an $A_{280}^{1\%}$ = 22.0 at pH 7.5. In cases where labeled MutT was prepared, $^{15}NH_4Cl$ (99 atom %) and/or [^{13}C]-glucose (99 atom %) was substituted for the unlabeled compounds in the growth medium. For optimal MutT production, the most cost-effective amounts of $^{15}NH_4Cl$ and [^{13}C]glucose required were found to be 0.5 and 2.0 g/L, respectively. For preparations of [^{15}N]Lys labeled MutT, the cells were grown in media containing 0.1 g/L of ^{15}N-labeled Lys with the other amino acids unlabeled as described previously (Weber *et al.*, 1992).

Uniformly (99%) deuterated Tris (d_{11}), $^{15}NH_4Cl$ (99% ^{15}N), [α-^{15}N]-L-Lys (95–99% ^{15}N), uniformly ^{13}C-labeled glucose (99% ^{13}C), and D_2O (99.996% D) were purchased from Cambridge Isotope Labs (Woburn, MA). All other reagents for the enzymatic assay and for the NMR samples were of the highest purity commercially available.

NMR samples contained 1.5 mM MutT, 0.1 mM EDTA, 0.34 mM NaN_3, 5–8 mM d_{11}-Tris-HCl, and enough NaCl (19–21 mM) to give an ionic strength equal to 25 mM. The pH was adjusted to 7.4 in H_2O (10% D_2O) and in D_2O (uncorrected for isotope effects). It was necessary to work at this pH due to the decreased stability of the free enzyme at lower pH values. Enzyme preparations were assayed before and after NMR experiments (Bhatnagar *et al.*, 1991), and in all cases at least 85% of the activity remained after prolonged studies at 32 °C (3–6 days).

NMR Instrumentation and General Methods

All NMR experiments were recorded at 32 °C on a modified Bruker AM 600 NMR spectrometer equipped with an external timer device (AM TIMER, Tschudin Associates, Kensington, MD) to reduce disk read/write time at the end of each (t_1, t_2) increment. The heteronuclear correlation data sets were recorded with either a 5-mm inverse broad-band probe or a triple-resonance probe using Bruker reverse electronics, i.e., protons were pulsed with the decoupler, while other nuclei were pulsed with the transmitter operating in the low-power mode. For triple-resonance experiments, two additional heteronuclear RF channels (third and fourth) were constructed using external synthesizers (Model 310, PTS Inc., Littleton, MA) phase locked to the 10-MHz spectrometer master clock as described by Kay *et al.* (1990). High-speed RF gates/level switches, GARP/WALTZ modulators, and phase-shifters for the PTS310 frequency synthesizers are from Tschudin Associates. External components, including the phase-shifters were controlled by up to seven separate transistor–transistor logic (TTL) lines under pulse programmer control. The RF band-stop filters and ^1H band pass filter are from Bruker, while the ^{13}C and ^{15}N band pass filters are from K & L Microwave (Salisbury, MD). RF signals from all of the heteronuclear channels in the various hardware configurations used in the experiments described below were amplified by a 50-W class A linear amplifer (Model 525 LA, ENI, Rochester, NH), with the exception of the ^{15}N channel in the triple-resonance probe in which the signal was first amplified (<5 W) by a 25-W amplifier (ENI, 325 LA) followed by a 13 dB booster (Model MM200, Kalmus Engineering Inc., Woodinville, WA).

The NMR data sets were transferred to a Personal IRIS 4D/35 workstation (Silicon Graphics, Inc., Mountain View, CA) and processed with the software package FELIX (Hare Research, Inc., Woodinville, WA). For all indirectly detected dimensions, acquired time domain data points were extended by one-third of the original size by the forward linear prediction routine in FELIX. However, these predicted data points were greatly attenuated, typically by a cosine squared bell function, and zero-filled before Fourier transformation.

The observed ^{1}H chemical shifts are reported with respect to the H_2O or HOD signal, which is taken as 4.706 ppm downfield from external sodium 3-(trimethylsilyl)propionate-*2,2,3,3-d*$_4$ (TSP) in D_2O (0.0 ppm) at 32 °C. The carbon chemical shifts are reported with respect to external TSP in D_2O (0.0 ppm). The nitrogen chemical shifts are reported with respect to external $^{15}NH_4Cl$ (2.9 mM in 1 M HCl) at 20 °C, which is at 24.93 ppm downfield from liquid NH_3 (Levy & Lichter, 1979).

2D Heteronuclear NMR Spectroscopy

^{15}N–^{1}H Correlation Spectra. 2D ^{1}H–^{15}N HSQC spectra (Bodenhousen & Ruben, 1980) were recorded in H_2O using the pulse sequence described by Bax *et al.* (1990b) with a 1.5-ms water suppression purge pulse at the end of the first INEPT transfer (Messerle *et al.*, 1989) for uniformly ^{15}N-labeled MutT, as well as for samples selectively labeled with ^{15}N in lysine residues. ^{1}H decoupling was achieved with a 180° pulse in the middle of the t_1 period, and ^{15}N decoupling during acquisition was achieved with GARP-1 modulation of a 1.25-kHz RF field (Shaka *et al.*, 1985). The delay times of 1.5 s and 2.8 ms were used for relaxation and INEPT delays, respectively. In both experiments, 768 t_1 values were recorded using TPPI (Marion & Wuthrich, 1983) with either 16 or 32 ([^{15}N]Lys-labeled sample) scans per t_1 value. Acquisition times were 160 (t_1) and 127 ms (t_2), respectively. The data were processed with a cosine squared bell filter in both dimensions and zero-filled to obtain final data matrices of 4K × 2K real points. The digital resolutions in ^{1}H and ^{15}N dimensions were 0.003 and 0.02 ppm/pt, respectively.

^{13}C–^{1}H Correlation Spectra. The 2D constant-time ^{1}H–^{13}C HSQC (CT-HSQC) spectrum was recorded on uniformly ^{13}C- and ^{15}N-enriched MutT in D_2O using the pulse sequence and the phase cycle of Vuister and Bax (1992). Due to hardware limitations, carbonyl 180° decoupling pulses were applied as hard pulses using an off-resonance DANTE sequence. The ^{13}C carrier frequency was set at 56.00 ppm, and the pulse length (55 ms) was adjusted such that $^{13}C\alpha$ pulses caused minimal excitation at the carbonyl region and vice versa. The experiments then yielded optimum correlations at the Hα–Cα region. During t_1, modulation by ^{13}C–^{13}C coupling was avoided by the constant time, and ^{15}N decoupling was carried out by GARP-1 modulation using a 1.25-kHz RF field. The residual HOD signal was suppressed by a 2-ms spin lock pulse at the end of first INEPT transfer. The relaxation delay, INEPT delay, and constant time period were 0.9 s, 1.7 ms, and 53.2 ms, respectively. A total of 256 t_1 values were recorded using the TPPI mode with 16 scans per t_1 value. Acquisition times were 32.8 ms (t_1) and 85 ms (t_2), respectively. Decoupling of ^{13}C from protons during the data acquisition was accomplished with asynchronous GARP-1 modulation, using a 2.3-kHz RF field. The data were processed with a 45°-shifted sine bell filter in t_2 and a cosine bell filter in t_1 and zero-filled to obtain the final data matrix of 4K × 2K real points with digital resolutions of 0.002 and 0.013 ppm/pt in the ^{1}H and ^{13}C dimensions, respectively.

3D Heteronuclear NMR Spectroscopy

In all of the heteronuclear 3D experiments described below, quadrature data were collected in t_1 and t_2 dimensions according to the TPPI–States method (Marion *et al.*, 1989a). The ^{1}H carrier frequency was set to 4.71 ppm, and the ^{15}N carrier frequency was set at 121.00 ppm. For spectra recorded in H_2O, a weak solvent presaturation pulse with an effective field strength of ~25 Hz was used during the relaxation period. The residual solvent signal was further suppressed by convolution of the time domain data (Marion *et al.*, 1989b) during the data processing. To obtain the highest possible digital resolution in the f_3 (^{1}H) dimension, the aliphatic half of the proton spectrum was discarded after zero-filling and Fourier transformation. Generally, the cosine bell squared filter was applied in t_1 and t_2 while the sine-bell filter with phase shift ranging from 60° to 90° was applied in the t_3 dimension. The data matrices were zero-filled in all three dimensions prior to Fourier transformation to obtain the desired digital resolution.

^{15}N–^{1}H Double Resonance Spectra. The 3D ^{1}H[^{15}N]-TOCSY-HMQC experiments (Marion *et al.*, 1989c) were carried out on uniformly ^{15}N-labeled MutT in H_2O using the pulse sequence and phase cycle of Pelton *et al.* (1991). The experiment was carried out with a single purge pulse (1 ms) preceding the 34-ms spin lock time, the latter using the DIPSI-2 sequence (Shaka *et al.*, 1988). A delay of 17 ms was used immediately following the spin lock to remove rotating frame NOE effects. The delay for the ^{1}H–^{15}N HMQC portion of the ^{1}H[^{15}N]TOCSY-HMQC was set to 4.5 ms, which is slightly less than $1/2J_{NH}$. ^{15}N decoupling during the t_1 and t_3 periods was achieved by a 180° pulse in the middle of the t_1 period and by asynchronous GARP decoupling (1.25-kHz) during the acquisition, respectively. Low-power presaturation of water was used during the 1.2-s relaxation period. The spectrum was derived from a 64 (complex) × 32 (complex) × 1024 (real) data matrix with acquisition times of 9.1 ms (^{1}H, t_1), 17.54 ms (^{15}N, t_2), and 63.5 ms (^{1}H, t_3), respectively. A total of 16 scans were acquired per (t_1, t_2) point, resulting in an experimental time of 50 h. The final data matrix consisted of 256 × 128 × 512 real points with digital resolutions of 0.05, 0.24, and 0.14 ppm/pt in the f_1, f_2, and f_3 dimensions, respectively.

The 3D ^{1}H[^{15}N]NOESY-HMQC spectrum (Zuiderweg & Fesik, 1989; Marion *et al.*, 1989) was acquired with ^{15}N decoupling during the t_1 and t_3 periods as described for the TOCSY-HMQC experiment. The water signal was presaturated during the relaxation delay (1–2 s) as well as in the first 150 ms of the 200-ms mixing time. The acquisition and processing parameters are identical to those of the TOCSY-HMQC spectrum.

^{13}C–^{15}N–^{1}H Triple-Resonance Spectra. HNCA, HNCO (Kay *et al.*, 1990a; Ikura *et al.*, 1990), and HN(CO)CA (Bax & Ikura, 1991) spectra were recorded in H_2O. In all three triple-resonance experiments, the carrier frequencies were set as follows: ^{1}H, 4.71 ppm; ^{15}N, 120.00 ppm; $^{13}C\alpha$, 56.00 ppm; ^{13}CO, 176.00 ppm. The relaxation delay and the ^{1}H–^{15}N INEPT delay were set to 0.9 s and 2.25 ms, respectively. ^{15}N decoupling during acquisition was achieved by asynchronous GARP modulation of a 1.25-kHz RF field. For the HNCA experiment, the data matrix was 32 (complex) × 64 (complex) × 1024 (real) points in the t_1, t_2, and t_3 dimensions, respectively. A total of 32 scans were collected per (t_1, t_2) point with acquisition times of 17.5 ms (^{15}N, t_1), 15.1 ms ($^{13}C\alpha$, t_2) and 63.5 ms (NH, t_3). GARP modulation of the 575 Hz RF field was used to decouple the carbonyl carbons (^{13}CO) during the t_2 and 22-ms fixed delay periods immediately preceding and

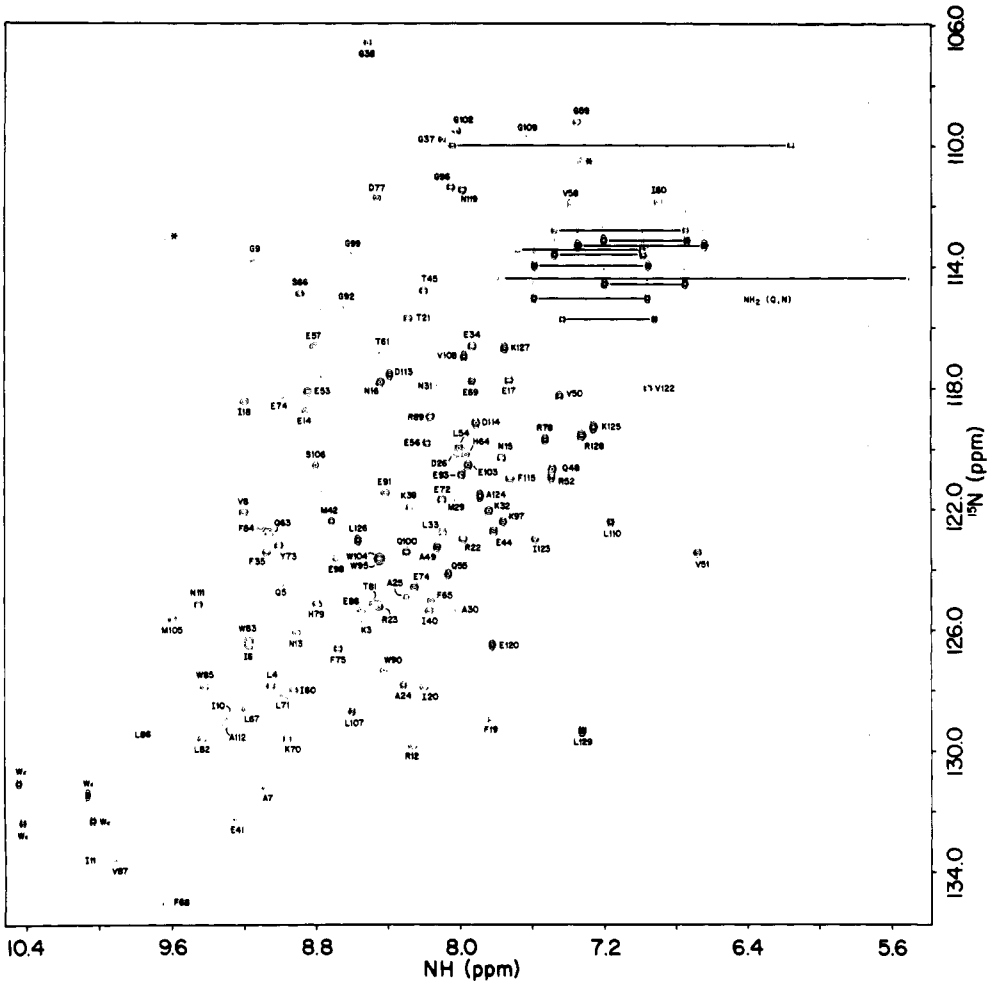

FIGURE 1: Two-dimensional ^{1}H–^{15}N HSQC spectrum of MutT labeled uniformly with ^{15}N recorded at 600-MHz ^{1}H frequency. Correlation peaks are labeled according to residue type and sequence number. The peaks connected by lines correspond to Gln and Asn side chain NH_2 groups. Peaks labeled W_ϵ correspond to ^{15}Nϵ–H correlations of Trp residues. Peaks labeled with an asterisk are folded over from side chain $N_\epsilon H$ of Arg residues.

following t_2. The total experimental time was 80 h. The final data matrix contains 64 × 256 × 1024 real points with digital resolutions of 0.47, 0.11, and 0.007 ppm/pt in the f_1, f_2, and f_3 dimensions, respectively. For the HNCO experiment (Kay *et al.*, 1990a), the data matrix was 32 (complex) × 64 (complex) × 1024 (real) points in the t_1, t_2, and t_3 dimensions, respectively. Acquisition times were 17.5 ms (^{15}N, t_1), 35.3 ms (^{13}CO, t_2) and 63.5 ms (NH, t_3). A total of 16 scans were acquired per (t_1, t_2) point with a total experimental time of 40 h. The final data matrix was 64 × 256 × 1024 real points with digital resolutions of 0.47, 0.05, and 0.007 ppm/pt in the f_1, f_2, and f_3 dimensions, respectively. For the HN(CO)CA experiment (Bax & Ikura, 1991), the data matrix was 16 (complex) × 32 (complex) × 1024 (real) points in the t_1, t_2, and t_3 dimensions, respectively. Acquisition times were 8.77 ms (^{15}N, t_1), 7.55 ms (^{13}Cα, t_2) and 63.5 ms (NH, t_3). A total of 64 scans were acquired per (t_1, t_2) point with a total experimental time of 38 h. Processing parameters were identical to those in the HNCA experiment.

HCACO and HCA(CO)N spectra (Kay *et al.*, 1990a) were recorded in D_2O. No presaturation was used in either experiment. The carrier frequencies were set as follows: ^{1}H, 4.71 ppm; ^{13}Cα, 56.00 ppm; ^{15}N, 120.00 ppm. The carbonyl pulses were generated by an off-resonance DANTE sequence as described (Kay *et al.*, 1990a). Recycle, ^{13}Cα–H INEPT, and ^{13}Cα–^{13}CO refocusing delays in both experiments were set to 0.8 s, 1.5 ms, and 3.0 ms, respectively. ^{13}Cα decoupling during the acquisition time was achieved by random GARP modulation of the 2.25- kHz RF field. For the HCACO experiment, the data matrix was 32 (complex) × 64 (complex) × 1024 (real) points in the t_1, t_2, and t_3 dimensions, respectively. A total of eight scans were signal averaged per (t_1, t_2) point using an eight-step phase cycle with acquisition times of 6.4 ms (^{13}Cα, t_1), 28.3 ms (^{13}CO, t_2) and 72.3 ms (Hα, t_3). The total experimental time was 22 h. ^{15}N decoupling during the t_2 period was achieved by GARP modulation of the 1.25-kHz RF field. The final data matrix was 64 × 256 × 1024 real points with digital resolutions of 0.52, 0.06, and 0.006 ppm/pt in the f_1, f_2, and f_3 dimensions, respectively. For the HCA(CO)N experiment, the data matrix was 32 (complex) × 32 (complex) × 1024 (real) points in the t_1, t_2, and t_3 dimensions, respectively. A total of 32 scans were signal averaged per (t_1, t_2) point using a 32-step phase cycle with acquisition times of 6.4 ms (^{13}Cα, t_1), 15.5 ms (^{15}N, t_2) and 72.3 ms (Hα, t_3). The total experimental time was 44 h. The final zero-filled data matrix was 64 × 256 × 1024 real points with digital resolutions of 0.52, 0.13, and 0.006 ppm/pt in the f_1, f_2, and f_3 dimensions, respectively.

RESULTS AND DISCUSSION

General Assignment Strategy

The ^{1}H–^{15}N HSQC spectrum of MutT protein (Figure 1) shows 140 correlation peaks, of which 22 are identified as resulting from the side chain NH_2 groups of six Asn and five Gln residues (peaks connected by horizontal lines). Four other

Table I: Backbone Resonance Assignments of MutT at 32 °C and pH 7.4[a]

residue	^{15}N	H_N	^{13}CO	$^{13}C\alpha$	$H\alpha$	residue	^{15}N	H_N	^{13}CO	$^{13}C\alpha$	$H\alpha$
M1			177.5	55.46	3.62	S66	114.89	8.89	173.46	56.97	5.04
K2			175.18	56.19	4.44	L67	128.61	9.20	175.46	56.56	3.64
K3	125.70	8.53	175.90	55.60	5.29	F68	135.00	9.63	174.79	59.57	4.57
L4	127.85	9.05	175.75	54.36	4.81	E69	117.77	7.94	172.44	55.21	4.40
Q5	124.50	9.00	174.28	54.93	5.77	K70	129.60	8.97	174.89	55.14	5.23
I6	126.52	9.18	173.98	59.56	4.61	L71	128.14	8.98	174.87	53.10	4.77
A7	131.20	9.08	175.01	50.43	5.47	E72	121.69	8.11	175.24	54.44	5.23
V8	122.10	9.20	175.43	60.65	4.77	Y73	123.21	9.01	173.21	57.32	4.51
G9	113.79	9.14	170.69	45.01	3.09, 4.89	E74	124.58	8.26	174.53	55.66	4.62
I10	128.92	9.29	173.79	60.08	4.08	F75	126.62	8.68	174.7	55.94	5.06
I11	133.27	10.05	173.70	61.19	4.32	P76	134.2		176.94	65.87	4.34
R12	129.82	8.27	175.93	54.07	5.87	D77	111.72	8.46	176.35	52.96	4.72
N13	126.07	8.91	177.10	51.10	5.26	R78	119.69	7.53	172.12	55.67	5.19
E14	118.74	8.86	176.56	58.13	4.27	H79	125.12	8.79	174.60	55.89	5.25
N15	120.31	7.78	174.84	52.19	4.91	I80	128.00	8.93	174.54	59.75	5.23
N16	117.80	8.44	173.85	54.47	4.33	T81	125.11	8.48	173.13	62.05	5.05
E17	117.73	7.73	175.32	54.07	5.06	L82	129.60	9.44	174.35	54.24	5.02
I18	118.43	9.20	175.75	57.34	6.11	W83	126.31	9.18	174.38	55.98	4.57
F19	128.94	7.83	174.03	58.68	4.15	F84	122.55	9.11	174.75	56.9	5.05
I20	127.87	8.20	174.68	57.75	4.62	W85	127.85	9.42	174.46	55.92	5.20
T21	115.69	8.29	172.65	59.70	4.63	L86	129.09	9.77	175.83	54.60	5.03
R22	122.96	7.99	175.81	54.67	4.61	V87	133.58	9.90	174.28	62.17	4.67
R23	125.20	8.45	176.08	55.37	4.30	E88	125.31	8.55	176.62	56.97	5.29
A24	127.82	8.32	177.70	52.49	4.23	R89	118.93	8.17	175.55	55.73	4.46
A25	124.89	8.30	177.4	52.85	4.21	W90	127.32	8.43	172.27	57.21	5.15
D26	120.13	8.02	176.18	53.56	4.57	E91	121.44	8.42	176.69	55.06	4.72
A27	124.65	8.02	178.0	53.30	4.13	G92	115.32	8.63	172.53	43.87	3.91
H28			175.8	56.99	4.50	E93	120.85	7.99	173.9	52.37	5.19
M29	121.72	8.02	175.2	55.12	4.35	P94	137.5		175.60	62.59	4.21
A30	125.32	8.02	177.3	52.75	4.20	W95	123.60	8.43	175.58	57.15	4.84
N31	117.75	8.12	174.37	53.5	4.51	G96	111.37	8.05	172.59	44.74	4.04, 4.19
K32	122.05	7.84	175.41	55.40	4.34	K97	122.41	7.77	176.94	55.64	4.06
L33	122.78	8.10	175.10	53.56	4.48	E98	123.64	8.69	177.5	58.05	4.13
E34	116.62	7.93	176.03	54.74	4.65	G99	113.50	8.60	174.01	45.3	3.73, 4.19
F35	123.40	9.08	174.4	55.61	5.15	Q100	123.40	8.30	172.4	53.49	4.80
P36	136.7		175.10	62.64	4.63	P101	139.2		176.90	63.45	4.41
G37	109.79	8.10	172.27	45.93	3.88, 4.67	G102	109.50	8.02	172.30	44.65	3.19, 4.44
G38	106.59	8.52	171.78	45.64	4.12	E103	120.53	7.96	173.33	54.50	4.42
K39	121.95	8.29	175.44	57.25	4.44	W104	123.60	8.45	177.24	56.32	4.98
I40	125.33	8.17	176.96	61.54	3.85	M105	125.61	9.60	175.29	54.61	4.92
E41	132.23	9.24	177.08	55.61	4.37	S106	120.54	8.81	176.01	58.81	4.48
M42	122.39	8.72	177.8	57.64	4.23	L107	128.69	8.60	178.93	58.13	4.05
G43			173.68	45.4	3.72, 4.14	V108	116.96	7.98	177.4	64.80	3.85
E44	122.79	7.82	176.36	54.81	4.65	G109	109.65	7.63	174.68	44.87	3.77, 4.15
T45	114.79	8.20	174.1	59.5	4.63	L110	122.42	7.17	175.98	55.64	4.23
P46	136.2		177.39	65.97	3.78	N111	125.14	9.46	176.3	51.51	4.93
E47	118.28	8.97	178.13	59.68	3.03	A112	129.1	9.31	179.07	55.77	3.98
Q48	120.65	7.48	179.58	58.40	3.78	D113	117.55	8.39	176.73	56.19	4.54
A49	123.24	8.13	178.98	55.06	4.03	D114	119.13	7.92	174.42	55.53	4.61
V50	118.25	7.45	175.95	65.28	3.50	F115	120.97	7.73	173.4	55.67	5.19
V51	123.42	6.68	176.47	67.05	3.07	P116	136.7		177.3	62.16	4.64
R52	120.94	7.50	179.12	59.48	4.06	P117			178.7	65.54	4.49
E53	118.11	8.85	179.04	58.93	4.03	A118			178.59	54.04	3.97
L54	119.97	8.01	179.75	58.52	3.96	N119	111.44	7.99	176.05	53.54	4.51
Q55	124.13	8.07	179.36	59.98	4.02	E120	126.48	7.82	174.5	63.19	3.84
E56	119.81	8.19	178.80	59.08	4.04	P121	135.2		180.07	65.70	4.29
E57	116.62	8.82	177.80	58.18	4.32	V122	118.01	6.95	177.00	65.53	3.43
V58	111.93	7.38	175.21	60.78	4.65	I123	122.95	7.59	177.34	64.10	3.46
G59	109.19	7.35	175.20	46.12	3.86, 4.15	A124	121.56	7.89	180.40	55.30	3.96
I60	111.84	6.90	173.65	58.19	5.12	K125	119.28	7.26	179.97	59.86	3.92
T61	116.80	8.43	174.2	59.05	4.79	L126	123.01	8.57	178.66	57.72	4.08
P62	138.7		174.13	63.07	4.34	K127	116.68	7.76	176.90	58.47	4.03
Q63	122.78	9.06	175.35	56.39	4.70	R128	119.54	7.33	175.91	55.98	4.44
H64	120.19	7.97	173.95	55.84	5.04	L129	129.33	7.33	182.6	57.66	4.09
F65	125.02	8.17	173.14	55.18	5.94						

[a] ^1H and ^{13}C chemical shifts (ppm) are relative to 3-(trimethylsilyl)propionate-*2,2,3,3-d*$_4$. ^{15}N chemical shifts (ppm) are relative to liquid NH_3 (see text for details).

peaks were also identified as NϵH correlations from four of the five Trp residues. The remaining peaks are due to the 114 out of the 120 expected backbone amide resonances, taking into account that MutT contains nine Pro residues. Five of the six missing backbone HN and ^{15}N resonances correspond to Met-1, Lys-2, His-28, Gly-43, and Ala-118, which remain unassigned. The last signal, Ala-27 is visible at lower contour levels (Table I). Overall, the HSQC spectrum shows excellent chemical shift dispersion in both dimensions, and the NH correlation peaks derived from the spectrum were used as starting points for resonance assignments.

In order to link the amide proton signal of each amino acid residue in the MutT protein to its corresponding aliphatic proton signals, a 3D TOCSY-HMQC spectrum was recorded

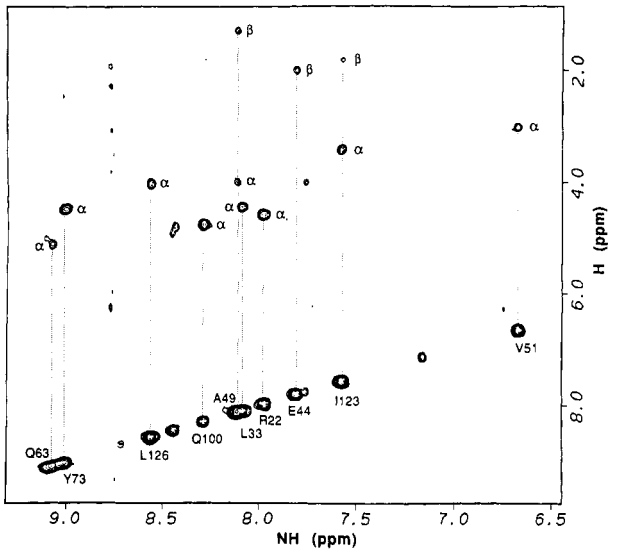

FIGURE 2: Selected $^1H(f_2)$–$^1HN(f_3)$ plane at the most crowded ^{15}N-(f_1 = 122.9 ppm) chemical shift of the 3D-TOCSY-HMQC spectrum. The extent of Hα and side chain correlations for MutT are shown. The diagonal peaks are labeled with residue type and sequence number. The unlabeled peaks appear stronger in adjacent planes.

with a mixing time of 34 ms. Figure 2, which contains a 2D TOCSY plane taken through the most crowded ^{15}N chemical shift, exemplifies the excellent resolution and the extent of side chain correlations typically seen in the 3D TOCSY-HMQC spectra. In all, we matched approximately 93% of the ^{15}N–H correlations, identified in the HSQC spectrum with corresponding Hα signals. The remaining NH–Hα connectivities were not seen, due either to fast exchanging amide resonances (M29, A30, G99, G109, and A112) or to saturation of Hα resonances along with solvent presaturation (V58, V87, and E91). In addition, we have identified at least one Hβ signal for 40% of the NH pairs. However, because of the short proton T_2 relaxation times characteristic of proteins of this size, side chain correlations to protons beyond Hβ were not observed. Furthermore, an attempt to correlate side chain resonances with the assigned Hα resonances by a 2D TOCSY/HOHAHA spectrum (Braunschweiler & Ernst, 1983; Bax & Davis, 1985) using unlabeled MutT in D_2O at two mixing times (31 and 47 ms) (data not shown) also had limited success owing to severe overlap of the aliphatic proton resonances. This difficulty of obtaining residue type assignments in MutT precluded the general strategy of sequential assignments based on ^{15}N edited spectra which has been successfully employed for several medium-size proteins (Marion *et al.*, 1989; Driscoll *et al.*, 1990; Clubb *et al.*, 1991; Stockman *et al.*, 1992).

An alternative approach has therefore been used in which the sequential connectivities are based solely on scalar coupling between different heteronuclei in the polypeptide backbone (Ikura *et al.*, 1990a,b). The heteronuclear correlations, specifically NH_i–$^{15}N_i$–$^{13}CO_{i-1}$, NH_i–$^{15}N_i$–$^{13}C\alpha_{i,i-1}$, and NH_i–$^{15}N_i$–$H\alpha_i$, were established from HNCO, HNCA, and ^1H–^{15}N TOCSY-HMQC spectra recorded in H_2O, while $C\alpha H_i$–$^{13}C\alpha_i$–$^{13}CO_i$ and $C\alpha H_i$–$^{13}C\alpha_i$–$^{15}N_{i+1}$ correlations were obtained from HCACO and HCA(CO)N spectra recorded in D_2O. The interresidue correlations thus provide three independent sequential connectivities, thereby uniquely defining each peptide linkage.

Assignment of Resonances

Experiments in H_2O. The HNCA experiment (Kay *et al.*, 1990a) links each N–H correlation with Cα resonances via $^1J_{NC\alpha}$ (~11 Hz) and $^2J_{NC\alpha}$ (~7 Hz) couplings. In general, the $^2J_{NC\alpha}$ correlation to Cα of the preceding residue shows a lower peak intensity than that of the intraresidue correlation, reflecting the difference in the magnitude of couplings involved. For example, in Figure 3, an $NH(f_3)$–$C\alpha(f_2)$ plane taken through the ^{15}N chemical shift at 122.9 ppm of the 3D HNCA spectrum of MutT shows connectivity typical of that seen in the entire experiment. In all, 90 NH pairs gave strong $C\alpha_i$, and weak $C\alpha_{i-1}$ correlations. The NH pair of I10 shows the reverse intensity pattern where the correlation to Cα of G9 was stronger than that to its own Cα. In addition, 18 NH pairs yielded single correlation peaks due either to overlap of $C\alpha_i$ and $C\alpha_{i-1}$ resonances (11 residues) or to the absence of $C\alpha_{i-1}$ correlations due largely to fast NH exchange (seven residues). Five NH pairs (K3, M29, A30, G99, and A112) did not show correlations to either Cα resonances, presumably due to fast exchanging amide protons.

To resolve ambiguities in assigning $C\alpha_{i-1}$ correlations, the 3D HN(CO)CA experiment (Bax & Ikura, 1991) was recorded on MutT. This experiment, which correlates HN, ^{15}N, and $C\alpha_{i-1}$ signals via the intervening carbonyl carbon has several advantages over the HNCA experiment. Due to its unique magnetization pathway, it simplifies NH–Cα correlation spectra providing unambiguous $C\alpha_{i-1}$ assignments (Figure 3). This experiment provided 105 NH_i–$C\alpha_{i-1}$ sequential correlations in MutT, including residues D26, N31, and I60 where $C\alpha_{i-1}$ correlations were not seen in the previous (HNCA) experiment, indicating a slight sensitivity advantage of the HN(CO)CA over the HNCA experiment.

The HNCO experiment (Kay *et al.*, 1990a), which links each intraresidue NH pair to the carbonyl resonance of the preceding residue, provided a second sequential connectivity. The HNCO spectrum of MutT (Figure 3) shows excellent resolution and sensitivity yielding 109 NH_i–CO_{i-1} sequential connectivities. In all, only six weak correlations observed in the ^1H–^{15}N HSQC spectrum (M20, A30, G99, G109, and A112) failed to show any connectivities in the HNCA, HN(CO)CA, and HNCO experiments. The data from these three experiments, together with the Hα assignments from the 3D TOCSY-HMQC experiment, yielded a subset of backbone assignments, namely, $C\alpha_{i-1}$, CO_{i-1}, N_i, NH_i, $C\alpha_i$, and $H\alpha_i$ for each backbone NH pair. The next step in the assignment process was to obtain the remaining subset of resonances, $C\alpha_i$, $H\alpha_i$, CO_i, and N_{i+1}, by means of HCACO and HCA(CO)N spectra recorded in D_2O.

Experiments in D_2O. The HCACO experiment (Kay *et al.*, 1990a) links intraresidue Hα, $^{13}C\alpha$, and ^{13}CO resonances through the $^{13}C\alpha$–Hα (140 Hz) and $^{13}C\alpha$–^{13}CO (55 Hz) couplings. Figure 4 shows a $CO(f_2)$–$H\alpha(f_3)$ plane taken through the most crowded $^{13}C\alpha$ chemical shift (55.5 ppm) of the HCACO spectrum of MutT. The excellent sensitivity of this experiment enabled us to distinguish even the correlation peaks arising from Hα protons which are very close to the residual HDO signal. The spectrum shows 127 out of the 129 expected Hα, $^{13}C\alpha$, ^{13}CO correlations, while the missing correlations are due to exact overlap of resonances in all three dimensions for the residue pairs E74/D114 and A25/A30.

Finally, in the HCA(CO)N experiment (Kay *et al.*, 1990a) each of the $^{13}C\alpha$–Hα pairs is correlated with the ^{15}N signal of the following residue, providing a third independent sequential connectivity. The $^{15}N(f_2)$–$H\alpha(f_3)$ plane (Figure 4) taken through the same $^{13}C\alpha$ chemical shift as in HCACO (Figure 4) shows corresponding N_{i+1} connectivities. The overlapping HCACO correlation from E74 and D114 now shows two N_{i+1} correlations resulting from well separated

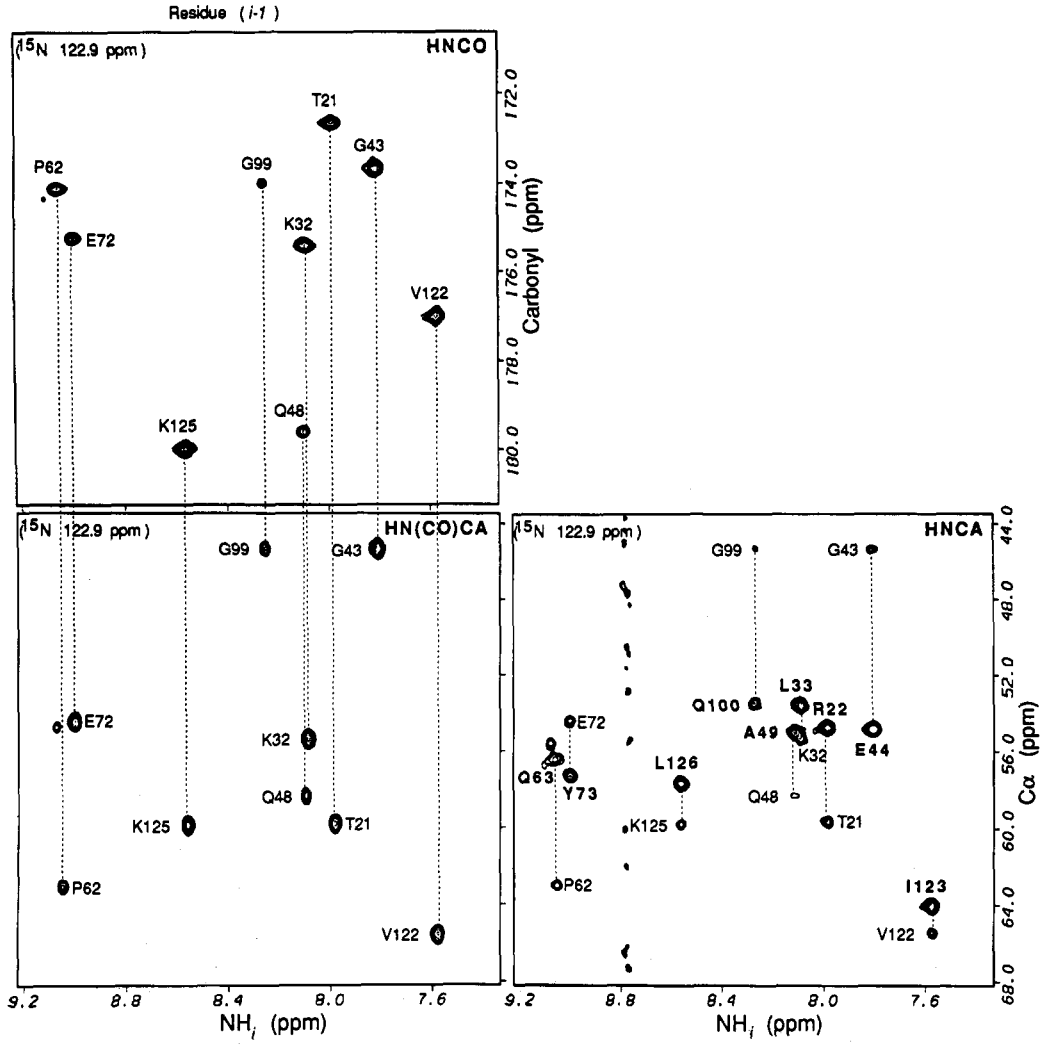

FIGURE 3: Selected $^{13}C\alpha/^{13}CO(f_2)$–$^{1}HN(f_3)$ planes at the most crowded $^{15}N(f_1)$ chemical shift of the HNCA, HN(CO)CA, and HNCO spectra of MutT. The intraresidue $^{1}HN_i$–$^{15}N_i$–$^{13}C\alpha_i$ correlations in the HNCA spectrum are labeled in bold letters with residue assignments. Interresidue correlations are labeled with the identity of the preceding residue.

^{15}N chemical shifts of the residues following them, namely, E75 and F115. We observed 124 of the 128 expected Hα, $^{13}C\alpha$, and $^{15}N_{i+1}$ correlations in the spectrum (correlation peaks are absent for C-termini). The remaining correlations (E44, Q63, V87, and P116) could not be identified because their Hα resonances fell under the strong residual HDO signal in the spectrum.

Matching of Resonances in H_2O and D_2O. The sequential assignment procedure, which involves matching the two subsets of backbone assignments obtained for the protein in H_2O and D_2O, critically depends on the reproducibility of ^{15}N and ^{13}C chemical shifts derived from the various triple-resonance experiments. The reproducibility of $^{15}N_i$ and $^{15}N_{i+1}$ chemical shifts derived respectively from the HSQC (in H_2O) and the HCA(CO)N (in D_2O) (after correcting for the deuterium isotope shift by adding 0.70 ppm) was ~0.15 ppm or better for 90% of the cases and ~0.30 ppm or better for all of the ^{15}N resonances with the exception of a few N_{i+1} signals (K39, F84, G96, and G99). In these cases, the centers of the Hα–N_{i+1} correlation peaks were shifted due to partial cancellation with the antiphase component of the signals arising from adjacent tiers (Kay *et al.*, 1990). The ^{13}CO chemical shifts derived from the HNCO and HCACO experiments show excellent reproducibility of ~0.05 ppm for all of the observable linkages. However, matching the $C\alpha_i$ chemical shifts derived from HNCA with those obtained from D_2O experiments [HCACO and HCA(CO)N] had several ambiguities, mainly due to the low digital resolution of the Cα dimension (0.52 ppm/pt) and substantial overlap present in the Hα/Cα region.

MutT contains 12 overlapping $^{13}C\alpha$–Hα pairs which have $^{13}C\alpha$ chemical shift differences of 0.25 ppm or less, as well as $^{1}H\alpha$ chemical shift differences of 0.03 ppm or less. In addition, an increase of uncertainties in the Cα dimension gives rise to a significant increase in the number of possible connectivities in the sequential assignment procedure from which the correct one must be selected. We have selected two approaches to overcome this problem. The first is to record the 2D constant time ^{1}H–^{13}C HSQC spectrum in D_2O using a uniformly ^{13}C-labeled sample (Vuister & Bax, 1992), where the Cα chemical shifts can be measured accurately due to high digital resolution, as well as suppression of ^{13}C–^{13}C couplings. Due to hardware limitations in the spectrometer, we used a slightly larger ^{13}C pulse length than optimal ($\pi/2$ = 55 μs). However, positioning the carrier at the midpoint of the Cα region yielded an optimum Hα–$^{13}C\alpha$ correlation spectrum (Figure 5).

The chemical shifts of the Cβ resonances of Ala residues are at the upfield extreme (~17 ppm) resulting in their incomplete inversion by the 180° ^{13}C pulse during the constant time period. As a result, Ala Cα/Hα cross peaks appear as split peaks in the ^{13}C dimension, while all of the other Hα/Cα correlation peaks appear as sharp singlets free of any ^{13}C–^{13}C

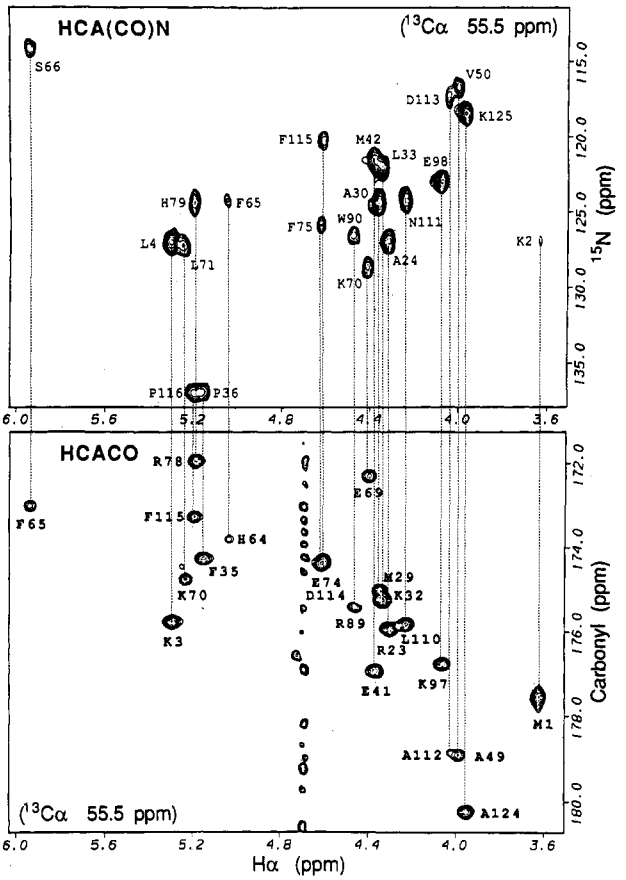

FIGURE 4: Selected $^{15}N/^{13}CO(f_2)$–$^{1}H\alpha(f_3)$ planes at the most crowded $^{13}C\alpha(f_1)$ chemical shift of the HCACO and HCA(CO)N spectra of MutT. The intraresidue $H\alpha_i$–$^{13}C\alpha_i$–$^{13}CO_i$ correlations in the HCACO spectrum are labeled in bold letters. Interresidue correlations in the HCA(CO)N spectrum are labeled with the identity of the following residue.

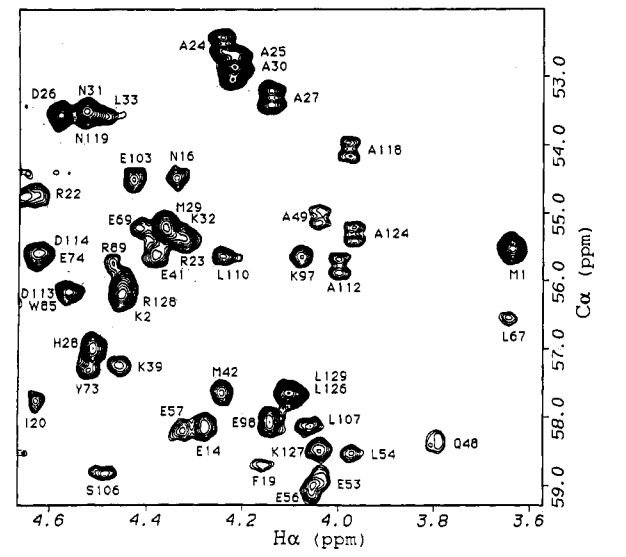

FIGURE 5: Region of the $^{1}H\alpha$–$^{13}C\alpha$ constant time HSQC spectrum of MutT uniformly labeled with ^{13}C and ^{15}N in D_2O, recorded at 600-MHz ^{1}H frequency. Correlation peaks, labeled with residue type and sequence number, show splitting of Ala Cα signals as discussed in text.

couplings (Figure 5). Such unique splitting of the Cα resonances of Ala residues is of substantive diagnostic value, since Ala cannot easily be uniquely labeled by metabolic incorporation. Figure 5 shows an expanded region of the Hα/Cα correlations in the CT-HSQC spectrum with appropriate Cα assignments. Accurate Cα assignments, together with

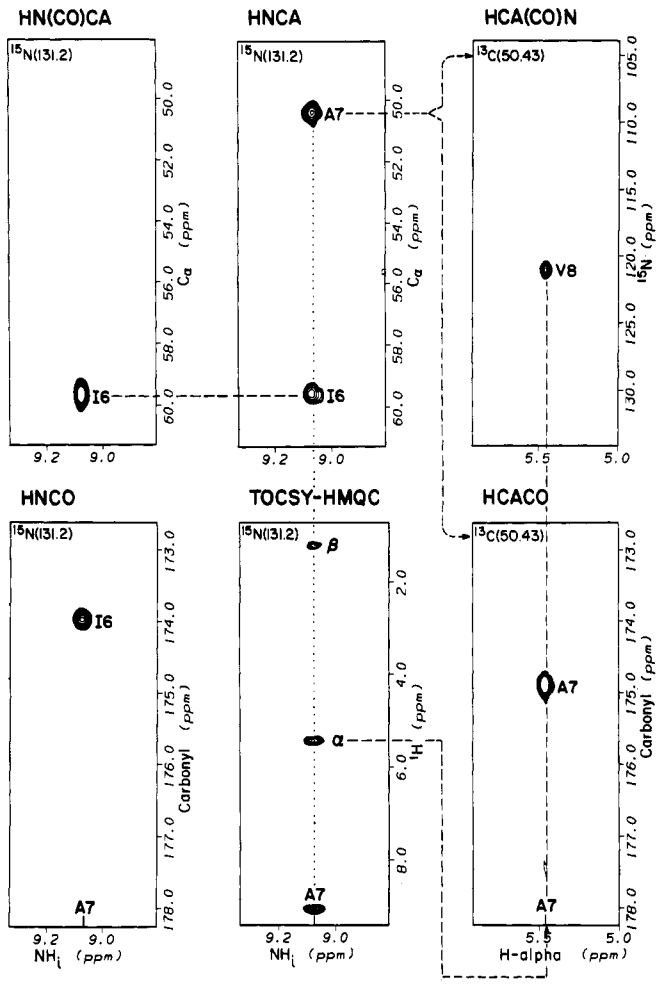

FIGURE 6: Typical example of scalar connectivities observed for a given residue in various 3D spectra of MutT illustrated for the residue A7. Sequential connectivities to the $^{13}C\alpha$ and ^{13}CO of the preceding residue (I6) and to the ^{15}N of the following residue (V8) are shown.

the identification of eight of the nine Ala Cα resonances, in particular, were found to be very useful in the sequential assignment procedure discussed below. By the same mechanism, the Cβ resonances of Thr residues retain coupling to Cγ.

The second approach to $^{13}C\alpha$-Hα assignments involves the use of 3D data sets to measure Cα chemical shifts accurately. Since it is not practical to construct large 3D data matrices with better digital resolution in all dimensions, we find that reprocessing the 1D slices (f_1) through the correlation peaks in CO/N(f_2)–Hα(f_3) planes yields Cα chemical shift values matching those of constant time HSQC data. This involves Hilbert transformation to generate complex data from real points in a 3D spectrum, inverse Fourier transformation, zero-filling to 512 points, and Fourier transformation to obtain the spectrum with high digital resolution. Since Cα line shapes in HCACO and HCA(CO)N spectra appear as resolution enhanced due to the unresolved antiphase Cα–CO couplings, the center of the peaks can be measured accurately. We find this process to be very helpful in measuring the Cα chemical shifts of overlapping Hα/Cα pairs where at least one spectrum yielded a well separated CO or N signal (f_2) correlated with an Hα in f_3. For all of the linkages, the Cα chemical shifts obtained in D_2O closely match those of the HNCA data obtained in H_2O to better than ±0.11 ppm, the Cα digital resolution of the HNCA and HN(CO)CA spectra.

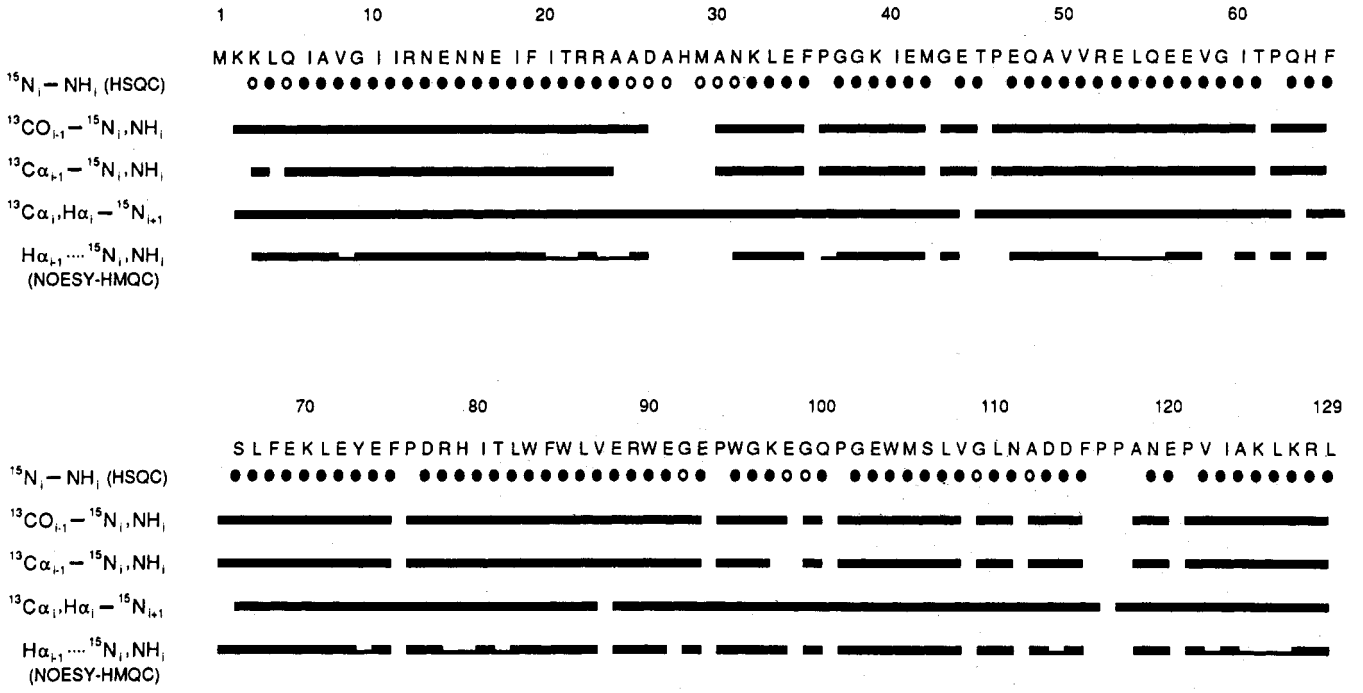

FIGURE 7: Summary of sequential connectivities observed in triple-resonance experiments of MutT. The sequential scalar connectivities ($^{13}CO_{i-1}$–$^{15}N_i$, NH_i), ($^{13}C\alpha_{i-1}$–$^{15}N_i$, NH_i), and ($H\alpha_i$, $^{13}C\alpha_i$–$^{15}N_{i+1}$) are from the HNCO, HNCA and/or HN(CO)CA, and HCA(CO)N spectra, respectively. For comparison, the sequential NOE connectivities ($H\alpha_{i-1}$–$^{15}N_i$, NH_i) from the ^1H–^{15}N NOESY-HMQC spectrum are included. The thin bars in the NOE connectivities represent ambiguities resulting from resonance overlaps. Intensities (●, strong-medium; o, weak-very weak) of the correlation peaks observed in the ^1H–^{15}N HSQC spectrum ($^{15}N_i$–NH_i) are also included for comparison.

Sequential Assignments

At this point we had obtained three types of sequential connectivities ($^{13}C\alpha_{i-1}$–$^{15}N_i$–NH_i and $^{13}CO_{i-1}$–$^{15}N_i$–NH_i in H_2O, and $^{13}C\alpha_i$–$H\alpha_i$–$^{15}N_{i+1}$ in D_2O) for most of the resonances. Figure 6 illustrates the data obtained from various triple-resonance experiments for a given residue (Ala-7) clearly showing its sequential connectivities to I6 (Cα, CO) and V8 (N).

The sequential assignments were started by selecting the N–H pair of a known residue type, i.e., either Gly, where NH pairs having characteristic chemical shifts at $C\alpha_i$ (45.0 ± 5.0 ppm) and N_i (110.0 ± 5.0 ppm), or Lys, where NH correlations for seven of the eight Lys residues in the protein were identified by ^1H–^{15}N HSQC spectra of MutT samples selectively labeled with [α-^{15}N]Lys (data not shown). This combination immediately produced two unique sites (G37–G38–K39 and G96–K97) as well as several starting points which were well dispersed within the polypeptide chain. Once the NH pair was selected, the matching subset of data (in D_2O) can be assigned to the preceding and following residues on the basis of $C\alpha_{i-1}$, CO_{i-1}, and N_{i+1} values. The next step involved finding a subset of H_2O data (N, NH, $C\alpha_i$, Hα, $C\alpha_{i-1}$, CO_{i-1}) for the i–1 residue or selecting an appropriate data set in D_2O (Cα, Hα, CO_i, N_{i+1}) for the residue itself by matching the Hα, Cα values. If a unique match was obtained, the assignment procedure was continued in both directions from the starting point.

The sequential linkages were interrupted at each of the nine prolines, because these residues lack amide protons, or at residues where amide protons were not seen in the ^1H–^{15}N HSQC spectrum (M1, K2, H28, G43, and A118), presumably due to fast exchange with water protons. In these cases sequential assignments must rely on linkages from the following residues with $C\alpha_{i-1}$ and CO_{i-1} connectivities. However, the interruptions due to Pro can easily be distinguished on the basis of the characteristic downfield shift of their amide nitrogens (139.2–134.2 ppm), which were seen only in HCA(CO)N spectra, as well as the downfield shift of the Cα resonances (66.0–62.0 ppm) where an appropriate subset of data in H_2O could not be obtained.

Once a stretch of residues was sequentially assigned, the relative positions of the known residues (i.e., Gly, Lys, Ala, and Pro) enabled us to position the stretch uniquely into the known polypeptide sequence (Akiyama *et al.*, 1987). In a similar manner, short to medium stretches of amino acid sequences were pieced together to yield nearly complete sequential assignments. In total, 105 $^{13}C\alpha_{i-1}$, 109 $^{13}CO_{i-1}$, and 123 $^{15}N_{i+1}$ connectivities were established and used to sequentially assign the 129 residues of MutT. These connectivities together with sequential NOE connectivities which were included for comparison are summarized in Figure 7. The complete backbone assignments are given in Table I.

CONCLUSIONS

The backbone NH, N, Cα, Hα, and ^{13}CO resonances of the MutT enzyme have been almost completely assigned by a combination of 2D and 3D heteronuclear methods. Two 2D and seven 3D heteronuclear experiments and the selective α-^{15}N-labeling of only one residue (Lys) were necessary. This relative economy of effort resulted from the unique identification of Ala Cα resonances in a constant time ^1H–^{13}C-HSQC experiment based on their residual coupling to Cβ resonances, the accurate measurement of Cα chemical shifts in D_2O in HCACO and HCA(CO)N experiments, and the accurate matching of these shifts with those obtained in H_2O in HNCA and HN(CO)CA experiments. The backbone assignments (Table I) have been essential in interpreting interproton NOEs in the elucidation of the solution secondary structure of this protein, as described in the accompanying paper (Weber *et al.*, 1993), since the solution structure of a protein of this size (129 residues) cannot readily be solved by proton NMR alone (Kay *et al.*, 1990a; Ikura *et al.*, 1990b).

Moreover, the sequence specifically assigned Cα and Hα chemical shifts (Table I) provide an independent approach to identifying secondary structural elements, such as helices and strands, and to define their extent (Spera & Bax, 1991; Wishart *et al.*, 1992). Thus, helices are indicated from residues E47 to V58 and from residues A118 to K127 on the basis of continuous stretches of downfield-shifted Cα resonances and upfield-shifted Hα resonances from their respective random coil values. Similarly, β-strands are less clearly suggested from K3 to N13, I18 to R23, K70 to E74, H79 to V87, and G102 to M105 by stretches of upfield-shifted Cα resonances and downfield-shifted Hα resonances. These secondary structural components are confirmed and their end points are refined by analysis of NOE data (Weber *et al.*, 1993).

The intensities of heteronuclear correlations also contain structural and dynamic information. From Figure 5, it may be seen that certain Cα–Hα correlations are intense (e.g., residues 1, 24–31), while others are relatively weak (e.g., the Ala residues 49, 112, 118, and 124). As pointed out by Vuister and Bax (1992) the strong signals probably reflect regions of high mobility, resulting in relatively long ^{13}C T_2 values. These residues also show weak or absent ^{1}H–^{15}N correlations (Figure 1), consistent with solvent exposure. As will be shown in the following paper (Weber *et al.*, 1993), residues 24–31 are located in an alanine-rich flexible loop.

ACKNOWLEDGMENT

We are grateful to David Shortle, Rolf Tschudin, Lewis Kay, and Jeffrey Pelton for valuable discussions and to Peggy Ford for secretarial assistance.

REFERENCES

Abeygunawardana, C., Weber, D. J., Frick, D. N., Bessman, M. J., & Mildvan, A. S. (1993) *Abstracts of the ENC Conference*, St. Louis, MO, March 14–18, 1993, poster 43, p 94.

Akiyama, M., Horiuchi, T., & Sekiguchi, M. (1987) *Mol. Gen. Genet. 206*, 9–16.

Bax, A., & Ikura, M. (1991) *J. Biomol. NMR 1*, 99–104.

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

Bax, A., Clore, G. M., Driscoll, P. C., Gronenborn, A. M., Ikura, M., & Kay, L. E. (1990a) *J. Magn. Reson. 87*, 620–627.

Bax, A., Ikura, M., Kay, L. E., Torchia, D. A., & Tschudin, R. (1990b) *J. Magn. Reson. 86*, 304–318.

Bhatnagar, S. K., & Bessman, M. J. (1988) *J. Biol. Chem. 263*, 8953–8957.

Bhatnagar, S. K., Bullions, L. C., Lew, G., & Bessman, M. J. (1990) *J. Bacteriol. 172*, 2802–2803.

Bhatnagar, S. K., Bullions, L. C., & Bessman, M. J. (1991) *J. Biol. Chem. 266*, 9050–9054.

Bodenhausen, G., & Ruben, D. J. (1980) *Chem. Phys. Lett. 69*, 185–188.

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

Bullions, L. C. (1993) Ph.D. Thesis, p 1–6, The Johns Hopkins University, Baltimore, MD.

Clubb, R. T., Thanabal, V., Osborne, C., & Wagner, G. (1991) *Biochemistry 30*, 7718–7730.

Cox, E. C. (1973) *Genetics 73 (Suppl.)*, 62–80.

Driscoll, P. C., Clore, G. M., Marion, D., Wingfield, P. T., & Gronenborn, A. M. (1990) *Biochemistry 29*, 3542–3556.

Frick, D. N., Weber, D. J., Gillespie, J. R., Bessman, M. J., & Mildvan, A. S. (1993) *J. Biol. Chem.* (in press).

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

Ikura, M., Kay, L. E., Tschudin, R., & Bax, A. (1990a) *J. Magn. Reson. 86.*, 204–209.

Ikura, M., Kay, L. E., Krinles, M., & Bax, A. (1991) *Biochemistry 30*, 5498–5504.

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

Levy, G. C., & Lichter, R. L. (1979) *Nitrogen-15 Nuclear Magnetic Resonance Spectroscopy*, John Wiley & Sons, New York.

Maki, H., & Sekiguchi, M. (1992) *Nature 355*, 273–275.

Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

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

Marion, D., Ikura, M., Tschudin, R., & Bax, A. (1989a) *J. Magn. Reson. 85*, 393–399.

Marion, D., Ikura, M., & Bax, A. (1989b) *J. Magn. Reson. 84*, 425–430.

Marion, D., Driscoll, P. C., Kay, L. E., Wingfield, P. T., Bax, A., Gronenborn, A. M., & Clore, G. M. (1989c) *Biochemistry 28*, 6150–6156.

Messerle, B. A., Wider, G., Otting, G., Weber, C., & Wüthrich, K. (1989) *J. Magn. Reson. 85*, 608–613.

Mo, J.-Y., Maki, H., & Sekiguchi, M. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 11021–11025.

Pelton, J. G., Torchia, D. A., Meadow, N. D., Wong, C.-Y., & Roseman, S. (1991) *Biochemistry 30*, 10043–10057.

Powers, R., Clore, G. M., Bax, A., Garrett, D. S., Stahl, S. J., Wingfield, P. T., & Gronenborn, A. M. (1991) *J. Mol. Biol. 221*, 1081–1090.

Powers, R., Garrett, D. S., March, C. J., Frieden, E. A., Gronenborn, A. M., & Clore, G. M. (1992) *Biochemistry 31*, 4334–4346.

Shaka, A. J., Barker, P. B., & Freeman, R. (1985) *J. Magn. Reson. 64*, 547–552.

Shaka, A. J., Lee, C. J., & Pines, A. (1988) *J. Magn. Reson. 77*, 274–298.

Spera, S., & Bax, A. (1991) *J. Am. Chem. Soc. 113*, 5490–5492.

Stockman, B. J., Nirmala, N. R., Wagner, G., Delcamp, T. J., DeYarman, M. T., & Freisheim, J. H. (1992) *Biochemistry 31*, 218–229.

Treffers, H. P., Spinelli, V., & Belser, N. O. (1954) *Proc. Natl. Acad. Sci. U.S.A. 40*, 1064–1071.

Vuister, G. W., & Bax, A. (1992) *J. Magn. Reson. 98*, 428–435.

Weber, D. J., Bhatnagar, S. K., Bullions, L. C., Bessman, M., & Mildvan, A. S. (1992a) *J. Biol. Chem. 267*, 16939–16942.

Weber, D. J., Gittis, A. G., Mullen, G. P., Abeygunawardana, C., Lattman, E. E., & Mildvan, A. S. (1992b) *Proteins: Struct., Funct., Genet. 13*, 275–287.

Weber, D. J., Abeygunawardana, C., Bessman, M. J., & Mildvan, A. S. (1993) *Biochemistry* (following paper in this issue).

Wishart, D. S., Sykes, B. D., & Richards, F. M. (1992) *Biochemistry 31*, 1647–1651.

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

Yanofsky, C., Cox, E. C., & Horn, V. (1966) *Proc. Natl. Acad. Sci. U.S.A. 55*, 274–281.

Zuiderweg, E. R. P., & Fesik, S. W. (1989) *Biochemistry 28*, 2387–2391.