Solution Structure of the MutT Enzyme, a Nucleoside Triphosphate Pyrophosphohydrolase[†,‡]

Chitrananda Abeygunawardana,[§] David J. Weber,[§,||] Apostolos G. Gittis,[⊥] David N. Frick,[#] Jian Lin,[§] Anne-Frances Miller,[○] Maurice J. Bessman,[#] and Albert S. Mildvan*[,§]

Department of Biological Chemistry and Department of Biophysics and Biophysical Chemistry, The Johns Hopkins School of Medicine, 725 North Wolfe Street, Baltimore, Maryland 21205-2185, and Departments of Biology and Chemistry, The Johns Hopkins University, Charles and 34th Streets, Baltimore, Maryland 21218

Received June 27, 1995; Revised Manuscript Received September 6, 1995[⊗]

ABSTRACT: The MutT enzyme (129 residues) catalyzes the hydrolysis of normal and mutagenic nucleoside triphosphates, such as 8-oxo-dGTP, by substitution at the rarely attacked β-P, to yield NMP and pyrophosphate. Previous heteronuclear NMR studies of MutT have shown the secondary structure to consist of a five-stranded mixed β-sheet connected by the loop I−α-helix I−loop II motif, by two tight turns, and by loop III, and terminated by loop IV−α-helix II [Abeygunawardana *et al.* (1993) *Biochemistry 32,* 13071−13080; Weber *et al.* (1993) *Biochemistry 32,* 13081−13087). Complete side-chain assignments of ^{1}H and ^{13}C resonances have now been made by 3D C(CO)NH and HCCH-TOCSY experiments. A total of 1461 interproton proximities (11 per residue), obtained by 3D ^{15}N-resolved NOESY-HSQC and 3D ^{13}C-resolved NOESY-HSQC spectra, including 372 long-range NOEs, as well as 65 dihedral angle (ϕ) restraints and 34 backbone hydrogen bond restraints were used to determine the tertiary structure of MutT by distance geometry, simulated annealing, and energy minimization with the program X-PLOR. The structure is globular and compact with the parallel portion of the β-sheet sandwiched between the two α-helices, forming an $\alpha + \beta$ fold. The essential divalent cation has previously been shown to bind near residues Gly-37, Gly-38, Lys-39, and Glu-57, and nucleotides have been shown to bind near residues Leu-54 and Val-58 by NMR relaxation methods [Frick *et al.* (1995) *Biochemistry 34,* 5577−5586]. The tertiary structure of MutT shows these residues to be near each other along the loop I−helix I region of the enzyme. A cluster of five glutamate residues (41, 53, 56, 57, and 98) form a patch of strongly negative electrostatic potential likely constituting the metal binding site. This site is contiguous with a deep cleft between β-strands A, C, and D and loop I which may contribute to the nucleotide binding site. This location of the active site is consistent with mutagenesis studies and with sequence homologies among MutT-like pyrophosphohydrolases.

The MutT enzyme, a pyrophosphohydrolase of 129 residues, catalyzes the unusual hydrolysis of nucleoside and deoxynucleoside triphosphates (NTP) by nucleophilic substitution at the rarely attacked β-phosphorus, yielding pyrophosphate and a nucleotide (NMP) as products (Bhatnagar *et al.*, 1991; Weber *et al.*, 1992). Like other enzymes which catalyze substitution at the electron-rich β-phosphorus, this enzyme requires two divalent cations for activity (Frick *et al.*, 1994). The biological role of this enzyme is to prevent mutations by preferentially hydrolyzing nucleotides which are misincorporated during DNA replication, thus "sanitizing" the nucleotide pool (Bhatnagar *et al.*, 1991, and references therein). It is one of a class of enzymes which specifically prevents AT → CG transversions, thereby decreasing the mutation rate in bacteria by several orders of magnitude (Kamath & Yanofsky, 1993; Mejean *et al.*, 1994). Maki and Sekiguchi (1992) have proposed that the biological substrate of the MutT enzyme is 8-oxo-dGTP, an oxidatively modified nucleotide which can mispair with template adenine during DNA synthesis *in vitro* (Cheng *et al.*, 1991), and have characterized a similar enzyme from a human cell line (Mo *et al.*, 1992; Sakumi *et al.*, 1993).

We have previously used heteronuclear multidimensional NMR methods to determine the sequence-specific assignments of the backbone Hα, Cα, N, NH, and CO resonances and the solution secondary structure of the MutT enzyme from *Escherichia coli* (Abeygunawardana *et al.*, 1993; Weber *et al.*, 1993). We have shown by nuclear Overhauser effects that the substrate analog Mg^{2+}AMPCPP[1] and the reaction product dGMP bind in high-anti conformations near the loop I−helix I−loop II motif (Frick *et al.*, 1995). The present paper completes the sequence-specific assignments of the side-chain ^{13}C and ^{1}H resonances and elucidates the tertiary

[†] This work was supported in part by National Institutes of Health Grants DK28616 (to A.S.M.), GM18649 (to M.J.B.), and GM36358 (which supports A.G.G.).

[‡] A complete listing of the distance restraints derived from the NOE data has been deposited at the Brookhaven Protein Data Bank (Chemistry Department, Brookhaven National Laboratory, Upton, NJ) together with the atomic coordinates of the family of 15 acceptable structures (file name 1MUT).

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

[§] Department of Biological Chemistry, 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.

[⊥] Department of Biophysics and Biophysical Chemistry, The Johns Hopkins School of Medicine.

[#] Department of Biology, The Johns Hopkins University.

[○] Department of Chemistry, The Johns Hopkins University.

[⊗] Abstract published in *Advance ACS Abstracts,* November 1, 1995.

0006-2960/95/0434-14997$09.00/0

structure of the MutT enzyme. The residues affected by substrate binding are shown to be near each other, thus helping to locate the active site. Preliminary abstracts of this work have been published (Abeygunawardana *et al.*, 1995a,b).

MATERIALS AND METHODS

Protein Sample Preparation. Uniformly isotopically enriched compounds including Tris-d_{11} (99% ^{2}H), $^{15}NH_4Cl$ (99% ^{15}N), D_2O (99.996% ^{2}H), and glucose (99% ^{13}C and 30% ^{13}C) were purchased from Cambridge Isotope Laboratories (Woburn, MA). The MutT protein labeled with ^{13}C and/or ^{15}N was isolated and purified as previously described (Abeygunawardana *et al.*, 1993).

General Nuclear Magnetic Resonance Methods. Unless otherwise noted, solution conditions for NMR studies were 1.5 mM labeled MutT, 0.34 mM NaN_3, 0.1 mM EDTA, 6 mM Tris-HCl-d_{11}, and 20 mM NaCl in 0.55 mL of H_2O/D_2O (90:10) at pH 7.5. The NMR data were collected with either a Bruker AM 600 spectrometer, a Varian Unityplus 500, or a Varian Unityplus 600 spectrometer and transferred to a Silicon Graphics Personal IRIS 4D/35 workstation for processing and analysis with the FELIX software package (Biosym Technologies, Inc.). Unless otherwise noted, multidimensional data sets were collected using the States–TPPI method (Marion *et al.*, 1989) in all of the indirect dimensions, with relaxation delays of 1 s. These acquired time domain data points were extended by one-third of the original size by the forward linear prediction routine in FELIX, before further processing to yield the final frequency domain data.

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 TSP at 32 °C. To ensure identical sample temperatures between different spectrometers, the temperature settings were adjusted to yield a chemical shift separation of 4.706 ppm between the HOD and TSP lines in a sample of TSP in D_2O. 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 24.93 ppm downfield from liquid NH_3 (Levy & Lichter, 1979).

A two-dimensional ^{1}H–^{15}N HMQC-J spectrum (Kay & Bax, 1990) was acquired with uniformly ^{15}N-enriched MutT. A total of 1152 t_1 values were recorded using TPPI (Marion & Wüthrich, 1983) with 32 scans per t_1 value. Acquisition times were 127 ms (t_2) and 316 ms (t_1), respectively. The data were processed with 40° (t_2) and 60° (t_1) shifted sine bell window functions and zero-filled to obtain a final data matrix of 4K × 2K real points. The digital resolution in the ^{1}H and ^{15}N dimensions were 2 and 0.90 Hz/point, respectively.

Experiments for Side-Chain Assignments. 2D ^{1}H–^{13}C CT-HSQC spectra (Vuister & Bax, 1992) were acquired with uniformly ^{13}C/^{15}N-enriched MutT in D_2O. The CT-HSQC spectrum optimized for aliphatic resonances was acquired with a 1.5 ms INEPT transfer delay and a 53.2 ms constant time period. Similarly, the spectrum optimized for aromatic resonances was acquired with a 1.6 ms INEPT transfer delay and an 18.2 ms constant time period. All carbon pulses were applied with the 10.7 kHz RF field centered at 42 ppm for the aliphatic spectrum and at 125 ppm for the aromatic spectrum. ^{13}C decoupling during the acquisition time (64 ms) was achieved by GARP-1 decoupling using a 4 kHz RF field. Using TPPI, a total of 860 t_1 values with 32 scans per FID and 160 t_1 values with 16 scans per FID were acquired for the aliphatic and aromatic CT-HSQC spectra, respectively.

The residue-specific ^{1}H/^{13}C assignments of aromatic side chains were made from two-dimensional HMQC (Bax *et al.*, 1983), HMBC (Bax & Summers, 1986), and HMQC-TOCSY spectra (Gronenborn *et al.*, 1989) using partially ^{13}C-labeled MutT (U 30 atom % ^{13}C) in D_2O. HMQC and HMQC-TOCSY spectra were recorded with the ^{13}C carrier frequency positioned at 125 ppm whereas in the HMBC spectrum the ^{13}C carrier frequency was at 133 ppm. Acquisition times and number of complex points for each data set were as follows: HMQC, 28.0 ms (^{13}C, t_1, 128 points), 127 ms (^{1}H, t_2, 1024 points), 64 scans per complex point; HMBC, 15.4 ms (^{13}C, t_1, 128 points), 254 ms (^{1}H, t_2, 2048 points), 320 scans per complex point; HMQC-TOCSY with ^{13}C decoupling during t_2, 28.0 ms (^{13}C, t_1, 128 points), 127 ms (^{1}H, t_2, 1024 points), 384 scans per complex point; HMQC-TOCSY without ^{13}C decoupling during t_2, 21.1 ms (^{13}C, t_1, 96 points), 254 ms (^{1}H, t_2, 2048 points), 320 scans per complex point. Both HMQC-TOCSY data sets were acquired with 10.9 ms DIPSI-2 mixing periods using an 8.0 kHz ^{1}H spin-lock field (Shaka *et al.*, 1988). In all of these experiments, the ^{1}H carrier was set at the residual HOD signal.

The residue-specific ^{1}H/^{13}C assignments of the aliphatic side chains were made from three-dimensional C(CO)NH (Grzesiek *et al.*, 1993) and HCCH-TOCSY spectra (Bax *et al.*, 1990) using 1.5 mM ^{13}C/^{15}N-labeled MutT in H_2O. Experiments were carried out at a 500 MHz ^{1}H field using a Varian Unityplus spectrometer equipped with three RF channels and z gradient capabilities. The 3D C(CO)NH experiment was recorded with enhanced sensitivity using gradient coherence selection (Kay *et al.*, 1992), suppressing unwanted signals with pulsed field gradients (Bax & Pochapsky, 1992; Muhandiram & Kay, 1994). The data set was acquired with spectral widths of 7625, 1650, and 4000 Hz in f_1 (^{13}C), f_2 (^{15}N), and f_3 (^{1}H), respectively, and with 50 complex points in t_1, 32 complex points in t_2, 256 complex points in t_3, and a total of 128 scans per hypercomplex t_1, t_2 pair. The positions of the carrier frequencies during frequency labeling were as follows: 43.00 ppm (^{13}C, t_1), 120.5 ppm (^{15}N, t_2), and 8.71 ppm (^{1}H, t_3). ^{13}C isotropic mixing (13 ms at a field strength of 8.3 kHz) was accomplished by two cycles of DIPSI-3 (Shaka *et al.*, 1988). The final data matrix consisted of 256 × 128 × 512 real points in the f_1, f_2, and f_3 dimensions, respectively.

A 3D HCCH-TOCSY experiment was recorded using the sequence described by Kay *et al.* (1993) with sweep widths

[1] Abbreviations: AMPCPP, adenosine 5′-(α,β-methylene)triphosphate; C(CO)NH, side-chain carbon to nitrogen (via carbonyl) correlation; CT, constant time; DGSA, distance geometry-simulated annealing; DIPSI, decoupling in the presence of scalar interactions; FID, free induction decay; GARP, globally optimized alternating-phase rectangular pulses; HMBC, heteronuclear multiple bond correlation; HMQC, heteronuclear multiple-quantum correlation; HSQC, heteronuclear single-quantum correlation; INEPT, insensitive nuclear enhancement by polarization transfer; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; 2D, two dimensional; 3D, three dimensional; RF, radiofrequency; ROESY, rotating frame Overhauser effect spectroscopy; TOCSY, total correlation spectroscopy; TPPI, time-proportional phase incrementation; TSP, sodium 3-(trimethylsilyl)propionate-*2,2,3,3-d*$_4$.

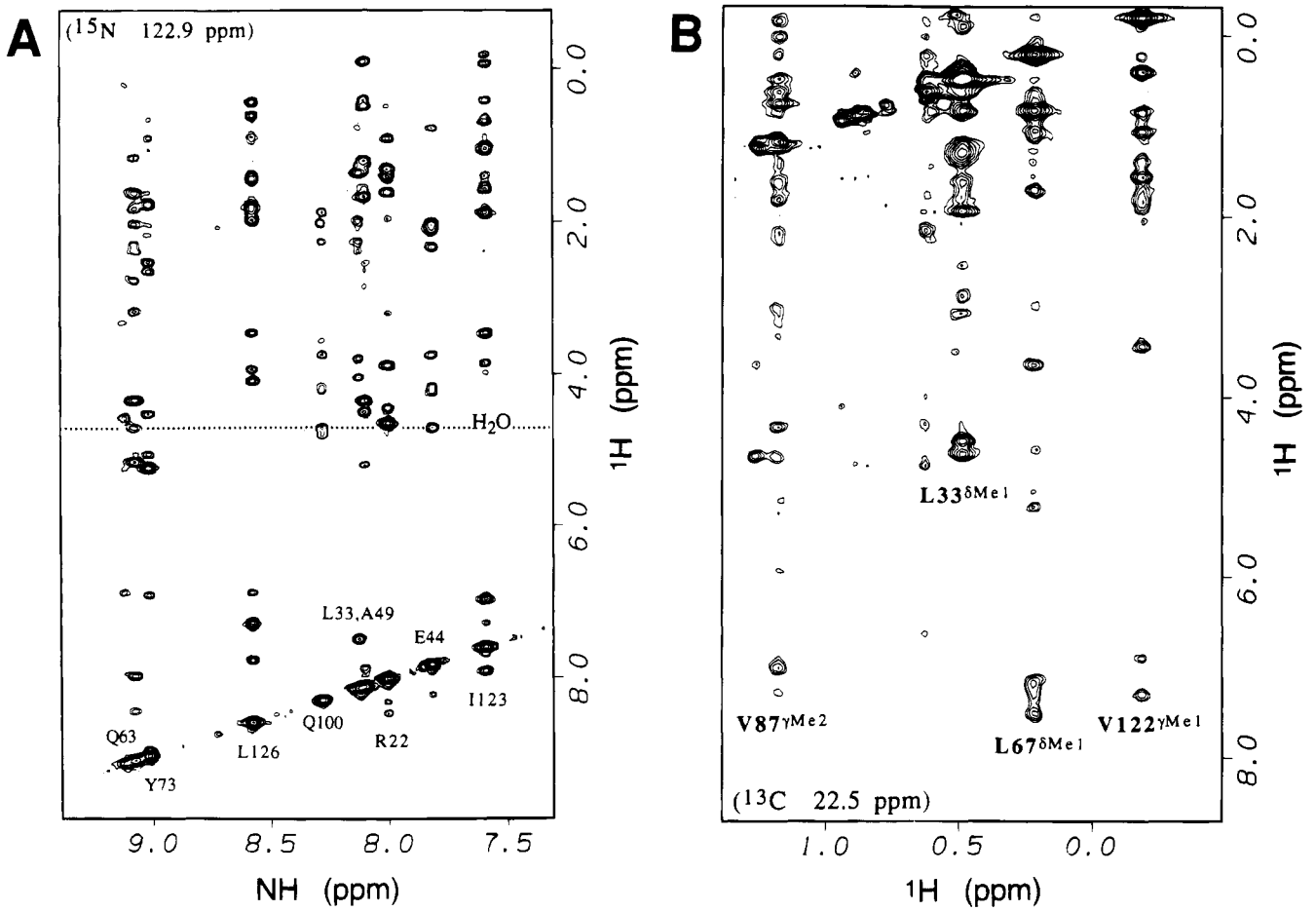

FIGURE 1: NOESY spectra of MutT. (A) Selected ^{1}H−^{1}H (^{15}N) plane at the most crowded ^{15}N chemical shift (122.90 ppm) in the 600 MHz 3D ^{1}H−^{15}N NOESY-HSQC spectrum of MutT (0.75 mM) in H_2O. The spectrum was recorded with a mixing time of 150 ms using a pulse sequence optimized to avoid water saturation (Mori *et al.*, 1995). (B) Expanded region of a selected ^{1}H−^{1}H(^{13}C) plane at the ^{13}C chemical shift of 22.5 ppm in the 500 MHz 3D ^{1}H−^{13}C NOESY-HSQC spectrum of MutT (1.5 mM) in D_2O. The spectrum was recorded with a mixing time of 150 ms. The methyl groups whose ^{1}H resonances appear on the diagonal are labeled.

and numbers of complex points in each dimension as follows: 3500 Hz (^{1}H, t_1, 96 points), 3000 Hz (^{13}C, t_2, 32 points), and 8000 Hz (^{1}H, t_3, 448 points). A total of 64 scans were collected for each hypercomplex t_1, t_2 pair. The positions of the carrier frequencies during frequency labeling were 3.00 ppm (^{1}H), 43 ppm (^{13}C), and 4.71 ppm (^{1}H) during t_1, t_2, and t_3, respectively. Three cycles of DIPSI-3 mixing (19.5 ms, 8.3 kHz field) were used in the experiment. The final data matrix was 256 × 128 × 512 real points for the f_1 (^{1}H), f_2 (^{13}C), and f_3 (^{13}C) dimensions, respectively.

Nuclear Overhauser Effect Experiments. 2D NOESY experiments with a mixing time of 150 ms (Kumar *et al.*, 1980) and 2D ROESY experiments with a spin-lock time of 70 ms (Kessler *et al.*, 1987) were carried out with unlabeled MutT (1.5 mM) in ^2H_2O containing the other components as described above. A 3D ^{1}H−^{13}C NOESY-HSQC experiment was recorded at 500 MHz with a mixing time of 150 ms using the sequence of Muhandiram *et al.* (1993) on ^{13}C/^{15}N-labeled MutT in D_2O. The data set was acquired with spectral widths of 5000, 3000, and 8000 Hz in f_1 (^{1}H), f_2 (^{13}C), and f_3 (^{1}H), respectively, and with 96, 32, and 512 complex points, respectively, in the t_1, t_2, and t_3 dimensions. A total of 64 transients were acquired for each hypercomplex t_1, t_2 pair with ^{1}H and ^{13}C carriers positioned at 4.71 and 43 ppm, respectively. The final data matrix was 512 × 128 × 512 real points for the f_1 (^{1}H), f_2 (^{13}C), and f_3 (^{1}H) dimensions, respectively.

A 3D ^{1}H−^{15}N NOESY-HSQC experiment was recorded on an 0.75 mM ^{15}N-labeled MutT solution containing the other components as described above, with a 150 ms mixing time using a sequence in which the HSQC detection scheme was optimized to avoid water saturation (Mori *et al.*, 1995). The data were obtained at 600 MHz with spectral widths of 6960, 1824, and 8200 Hz in f_1 (^{1}H), f_2 (^{15}N), and f_3 (^{1}H), respectively, and with 128, 36, and 512 complex points, respectively, in the t_1, t_2, and t_3 dimensions. A total of 64 transients were acquired for each hypercomplex t_1, t_2 pair. The final data matrix was 512 × 128 × 512 real points for the f_1 (^{1}H), f_2 (^{15}N), and f_3 (^{1}HN) dimensions, respectively

Structural Calculations. The NOE cross-peak intensities were classified as strong, medium, weak, and very weak and assigned distance constraints of 1.8−2.8, 1.8−3.2, 1.8−5.0, and 2.5−5.0 Å, respectively (Clore *et al.*, 1986). Pseudo-atom corrections were added to the upper limit for constraints involving methyl protons (1.0 Å), nonstereospecifically assigned methylene protons (1.0 Å), degenerate methyl protons of Leu and Val residues (1.7 Å), and δ and ϵ protons of Tyr and Phe residues (2.3 Å) (Wüthrich *et al.*, 1983; Wüthrich, 1986; Clore *et al.*, 1987). An additional 0.5 Å was added for NOEs involving methyl protons in ^{13}C NOESY-HSQC data to account for the higher apparent intensities. Only structurally useful intraresidue NOEs were included in the interproton distance restraints. Thus, NOEs between geminal protons or between vicinal methylene

protons were not included. Very weak distance constraints were included only between NH protons which showed reciprocal NOEs with similar intensities, thereby minimizing possible artifacts due to spin diffusion. $^3J_{NH\alpha}$ values were interpreted and incorporated as torsional angle (ϕ) restraints where $\phi = -60° \pm 20°$ for $^3J_{HN\alpha} < 6.0$ Hz and $\phi = -120° \pm 40°$ for $^3J_{HN\alpha} \geq 7.5$ Hz.

The structures were calculated using the hybrid distance geometry−simulated annealing (DGSA) protocol (Nilges *et al.*, 1988) in the program X-PLOR 3.1 (Brünger, 1992) on a Convex C220 computer. A total of 1461 NOE-derived distance restraints were applied with a square-well potential and a force constant of 50 kcal·mol^{-1}·Å^{-2}. A total of 65 ϕ dihedral angle restraints from $^3J_{HN\alpha}$ coupling constant measurements were also applied with a force constant of 200 kcal·mol^{-1}·rad^{-2}. Using these restraints, a total of 60 structures were generated by the DGSA protocol. These 60 structures were then subjected to two rounds of refinement which included, as additional distance restraints, a total of 34 backbone hydrogen bonds (NH−O) based on the secondary structural analysis (Figure 1A, Weber *et al.*, 1993) and deuterium exchange data (Weber *et al.*, 1993), as well as on the lack of H_2O exchange peaks in 3D NOESY-HSQC data.[2] By these criteria, 13 hydrogen bonds in MutT occur in α-helices and the remaining 21 occur in the β-sheet network. These refinements yielded an ensemble of 15 structures which satisfied the criteria of no NOE violations >0.5 Å and no dihedral violations >5°.[2,3] Hence, these 15 structures were used to define the tertiary structure of MutT.

Table 1: NMR Restraints and Structural Statistics for MutT

input restraints		
long-range ($\|i-j\| > 5$) NOEs		372
medium-range ($1 < \|i-j\| \leq 5$) NOEs		186
sequential ($\|i-j\| = 1$) NOEs		509
intraresidue NOEs		394
total NOEs		1461
NOEs/residue		11.3
dihedral angles (ϕ)		65
backbone hydrogen bonds		34
RMSD from idealized geometry	**best structure**	**average of 15 structures**
bonds (Å)	0.0042	0.0042 ± 0.001
angles (deg)	0.653	0.681 ± 0.018
impropers (deg)	0.359	0.368 ± 0.018
E_{NOE} (kcal/mol)	154	158 ± 8
E_{VDW} (kcal/mol)[a]	−169	−142 ± 25
RMSD of ensemble of 15 structures (Å)	**pairwise RMSD**	**RMSD from mean**
backbone, secondary[b]	1.26 ± 0.23	0.86 ± 0.22
backbone, all[c]	2.20 ± 0.37	1.51 ± 0.39
all heavy atoms[c]	3.10 ± 0.38	2.12 ± 0.55

[a] Calculated with the CHARMM parameters. [b] The 15 acceptable structures were aligned on the backbone heavy atoms of the helices and strands shown in Figure 3. [c] The 15 acceptable structures were aligned on the backbone heavy atoms of all residues.

RESULTS

Side-Chain Proton and Carbon Assignments of MutT. Figure S1 (in supporting information) shows a plane [f_3 (^{1}HN) / f_1 (^{13}C)] through the most crowded f_2 (^{15}N) chemical shift (122.9 ppm) in a 3D C(CO)NH experiment, where well-resolved side-chain ^{13}C resonances of the preceding residues are correlated with NH proton resonances. The ^{13}C resonances were initially assigned on the basis of residue type (Grzesiek & Bax, 1993) which, in turn, confirmed our earlier sequence-specific backbone assignments (Abeygunawardana *et al.*, 1993). Side-chain ^{1}H resonances were assigned, and ^{13}C resonances were confirmed by 3D HCCH-TOCSY data (Figure S2 in supporting information). These assignments are in full agreement with limited correlations seen in ^{1}H−^{15}N TOCSY-HMQC spectra of ^{15}N-labeled MutT (Abeygunawardana *et al.*, 1993). Taken together, these two 3D data sets yielded sequence-specific assignments of side-chain proton resonances and their attached ^{13}C resonances for all aliphatic amino acid residues except the Hβ/Cβ correlations of Glu-34, -47, -69, -103, and -120 and Hγ/Cγ of Glu-69. However, with the exception of Glu-69, the β-proton resonances of these Glu residues were assigned by 3D ^{15}N NOESY-HSQC spectra. Sequence-specific side-chain resonance assignments of Hα, Hβ, and Cβ were also obtained for all of the aromatic residues, by these experiments.

The aromatic side-chain proton resonances were initially identified from the cross-peak patterns in the 2D NOESY (150 ms) and 2D ROESY (70 ms) spectra and from the coupling patterns in the 2D TOCSY spectrum (Wüthrich, 1986). These assignments were confirmed and extended to aromatic carbon resonances from the direct correlations seen in the 2D ^{1}H−^{13}C HMQC and CT-HSQC spectra as well as from multiple bond and relay correlations seen in the 2D ^{1}H−^{13}C HMBC and HMQC-TOCSY spectra. Accurate assignments of the degenerate and overlapping proton resonances, Hε, Hζ of Phe-19, Hε, Hζ of Phe-84, and Hδ, Hε of Phe-75 were obtained from the relay correlations in the ^{13}C-coupled HMQC-TOCSY spectrum. The N$^{\epsilon 1}$ and NH$^{\epsilon 1}$ resonances of Trp-85 were not detected, probably due to exchange broadening.[4]

Aromatic side-chain spin systems were then sequence specifically assigned by NOEs between Hδ and Hβ resonances and/or Hα resonances in 2D NOESY spectra in D_2O. Except for histidine residues, these assignments were confirmed by NOEs between backbone NH and Hδ resonances of the aromatic ring. His-28 with a fast exchanging backbone NH proton showed a long-range connectivity of its Hα to its Cγ resonance in HMBC spectra, confirming the assignment. His-79, on the other hand, showed a weak NOE from its Hβ to its Hε resonance and a very weak NOE to its Hδ resonance. Similar confirming patterns could not be found for His-64. The backbone ^{1}H, ^{13}C, and ^{15}N chemical shift assignments of the MutT protein have previously been published (Abeygunawardana *et al.*, 1993). The complete sequence-specific ^{1}H, ^{13}C, and ^{15}N assignments of the MutT protein are given as supporting information.

[2] Refinements omitting the backbone hydrogen bonds yielded a comparable set of tertiary structures, however, lacking several experimentally detected hydrogen bonds.

[3] In preliminary refinements, only 37 dihedral angle restraints were used, involving those residues found in standard secondary structural elements such as α-helices or β-strands.

[4] Exchange of the indole ring of Trp-85 is suggested by several additional observations. The ^{1}H−^{13}C correlation peaks of the indole ring were weak in HMQC spectra, and some were absent in HSQC spectra, where exchange broadening can occur during the INEPT transfer delay (Yamazaki *et al.*, 1995). An exchange cross peak was found in a 2D ROESY spectrum for H$^{\zeta 2}$ of Trp-85 at 7.12 and 7.26 ppm, suggesting an exchange rate constant of ~530 s^{-1}. The tertiary structure of MutT shows the indole ring of Trp-85 to be solvent exposed and not sterically hindered, which would allow ring motion.

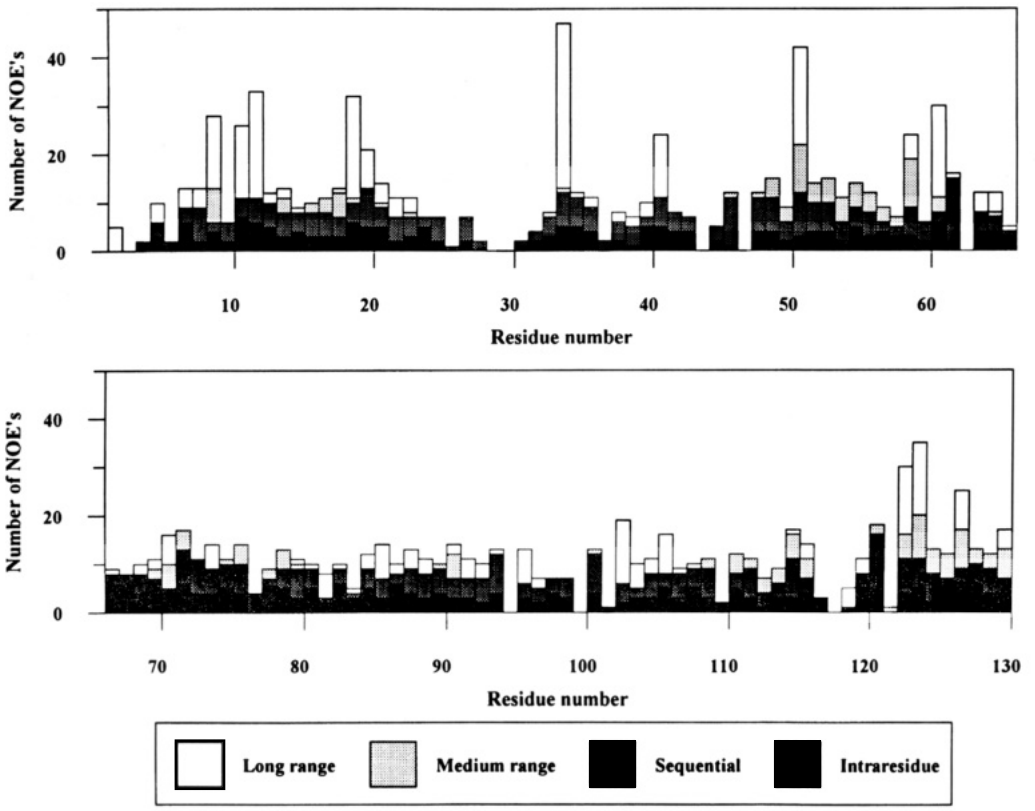

FIGURE 2: Distribution of NOE connectivities involving each residue of the MutT enzyme used to compute the solution structure. Long- and medium-range NOEs are defined in Table 1.

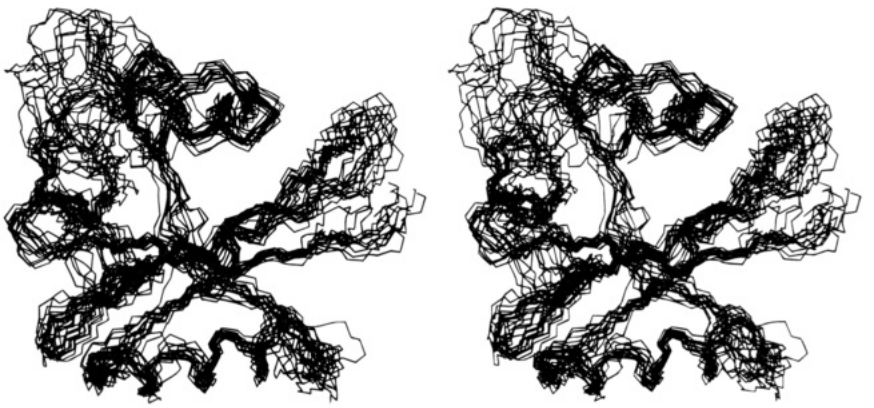

FIGURE 3: Superposition of the backbone (N, Cα, CO) atoms of secondary structural regions of 15 acceptable structures of the MutT enzyme. Table 1 gives the statistics of this set of structures.

NOEs Involving Side-Chain Protons and Hydrophobic Core of MutT. Once the resonance assignments were complete, a large number of NOEs from 3D $^{1}H-^{15}N$ NOESY-HSQC spectra (Figure 1A) were interpreted on the basis of the known secondary structure and topology of MutT (Weber *et al.*, 1993; Frick *et al.*, 1995). Figure 1A exemplifies the resolution and quality of the water flipback 3D $^{1}H-^{15}N$ NOESY-HSQC spectra. The plane was again taken through the most crowded ^{15}N chemical shift (122.9 ppm). Such NOEs, between backbone amide and side-chain proton resonances, account for a majority of the distance restraints. In addition to these, NOEs involving side-chain to side-chain proton resonances were obtained from 3D $^{1}H-^{13}C$ NOESY-HSQC spectra as well as from 2D NOESY spectra (150 ms) in D_2O.

Figure 1B shows an expanded region of an f_1 [^{1}H]–f_3 [$^{1}H(^{13}C)$] plane of the 3D $^{1}H-^{13}C$ NOESY HSQC spectrum where NOEs from several methyl groups are indicated. To remove ambiguities in NOE assignments, all NOE peaks used as distance restraints were checked for reciprocal NOE peaks. For NOEs involving aromatic protons, reciprocal peaks were checked in 2D NOESY spectra. However, due to severe spectral overlap in the 2D spectra, this process was limited to NOEs between aromatic proton resonances and resolved upfield methyl resonances.

The NOE data clearly indicate two distinct clusterings of hydrophobic side chains in the molecule. One cluster includes side chains of residues from helix II, the C-terminal helix (Leu-126, Ile-123, and Val-122), loop II (Leu-71 and Phe-68), β-strand D (Leu-86 and Phe-84), β-strand A (Ile-10), loop IV (Leu-110 and Phe-115), and loop I (Phe-35 and Leu-33). These NOEs positioned the C-terminal helix (II) on one side of the β-sheet consisting of strands A, D, and C. The other, much smaller cluster of hydrophobic side chains includes Val-50, Val-51, and Leu-54 from helix I, Phe-65 and Leu-67 from loop II, Val-87 from β-strand D,

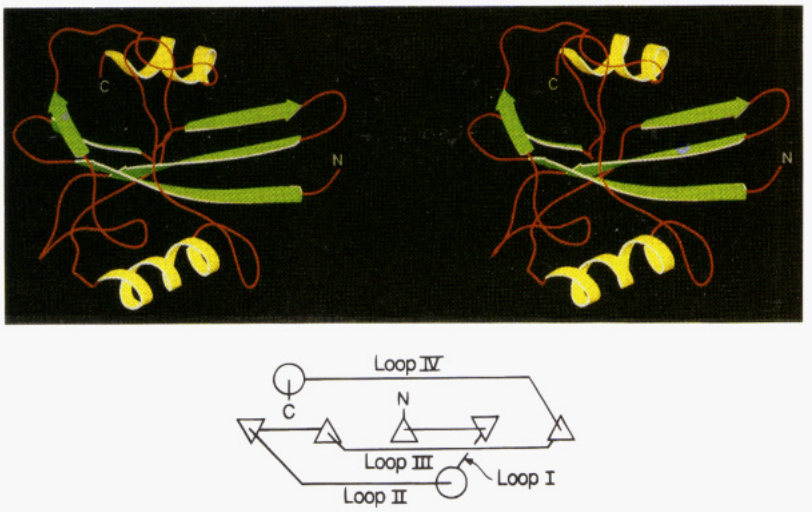

FIGURE 4: Best-fit structure of the MutT enzyme. (A, upper) Stereo pair ribbon diagram drawn with the program SETOR (Evans, 1993). The α-helices are in yellow, the β-strands are in green, and the loops and turns are in red. The view is into the β-sheet network. The potential energy values for this structure are as follows: NOE distance restraint violations, 154 kcal/mol using the force constant 50 kcal mol^{-1} $Å^{-2}$; van der Waals interactions, −169 kcal/mol using the CHARMM parameters for the Lennard-Jones potential. Deviations from ideality (RMSD) are as follows: bond lengths, 0.004 Å; bond angles, 0.65°; impropers (deviations from planarity of aromatic rings and peptide bonds), 0.36°. (B, lower) Tertiary structural diagram after Levitt and Chothia (1976) and Orengo *et al.* (1993). Helices are shown as circles, and β-strands are shown as either upright triangles which proceed toward the reader or as inverted triangles which proceed away from the reader.

and Trp-95 from loop III. These data positioned loop III near the C-terminal end of helix I on the opposite side of the β-sheet from helix II.

Tertiary Structure Calculation and Refinement. A total of 1461 distance restraints (11 per residue) were derived from the NOE data including 372 long-range, 186 medium-range, 509 sequential, and 394 intraresidue restraints (Table 1) which were well distributed over the entire enzyme (Figure 2). A total of 65 $^3J_{HN\alpha}$ values found to be in the ranges $J <$ 6 Hz or $J >$ 7.5 Hz were used for backbone dihedral angle ϕ restraints. On the basis of the criteria given in Materials and Methods, 34 backbone hydrogen bonds, 13 in the α-helices and 21 in the β-sheet network, were used as additional distance restraints in the refinement.

Description of the Tertiary Structure. Figure 3 superimposes 15 computed backbone structures which showed no NOE violations exceeding 0.5 Å and no dihedral violations exceeding 5°, Figure 4A shows the best-fit structure with the lowest van der Waals energy (Table 1), and Figure 4B diagrams the tertiary structure of MutT. The Ramachandran diagram of the best-fit structure (Figure 5) shows no residues outside the allowed range of $\phi\psi$ values. Despite the fact that 49% of the residues of MutT are in long loops (Figure 4), the structure is well defined, as reflected in Figure 3 and in the low RMSD values (Table 1). The overall stereochemical quality of the set of structures of Figure 3, assessed by the PROCHECK suite of programs (Laskowski *et al.*, 1993), was found to correspond to that of a 2.5 Å, well-refined X-ray structure with an R factor <20%.

The three-dimensional course of MutT shows it to form an α + β fold (Orengo & Thornton, 1993) consisting of a five-stranded mixed β-sheet, three long strands of which (two parallel, A and D, and one antiparallel, C) are sandwiched between two α-helices and protrude from the sandwich (Figure 4). The other two short antiparallel β strands, B and E, are connected to the remainder of the structure by three long loops, I, III, and IV, and by a nonclassical turn. The dihedral angle between the two α-helices, when viewed

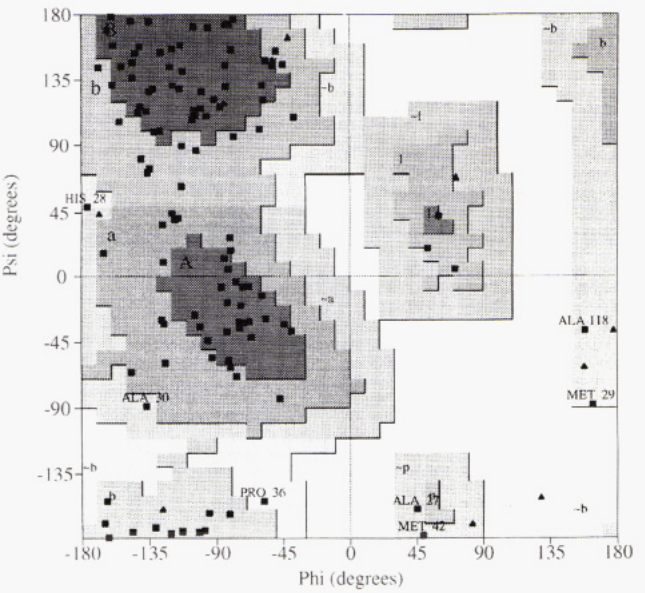

FIGURE 5: Ramachandran ($\phi\psi$) diagram of the best-fit structure of the MutT enzyme. The named residues are those in generously allowed regions, and the triangles indicate Gly residues.

through the β-sheet is ~90°. The positions of residues in the β-sheet and in the two α-helices are better defined than those in the loops (Figure 3), as reflected in the greater RMSD values of the complete structures which include the loops (Table 1). The greater uncertainties in the positions of portions of loops I and IV (Figure 3) are due to fewer NOE restraints in these regions (Figure 2). Increased flexibility of these regions was previously suggested by narrow $^{13}C\alpha$ resonances and weak or absent NH resonances (Abeygunawardana *et al.*, 1993). Flexibility, especially in loop I, may be necessary for substrate binding (see below).

As described above, MutT has two hydrophobic cores, one on each side of the β-sheet. Of the 53 hydrophobic residues of MutT, 44 (83%) are found to be buried, 7 (13%) are partially exposed, and 2, Trp-104 in loop III and Val-108 in

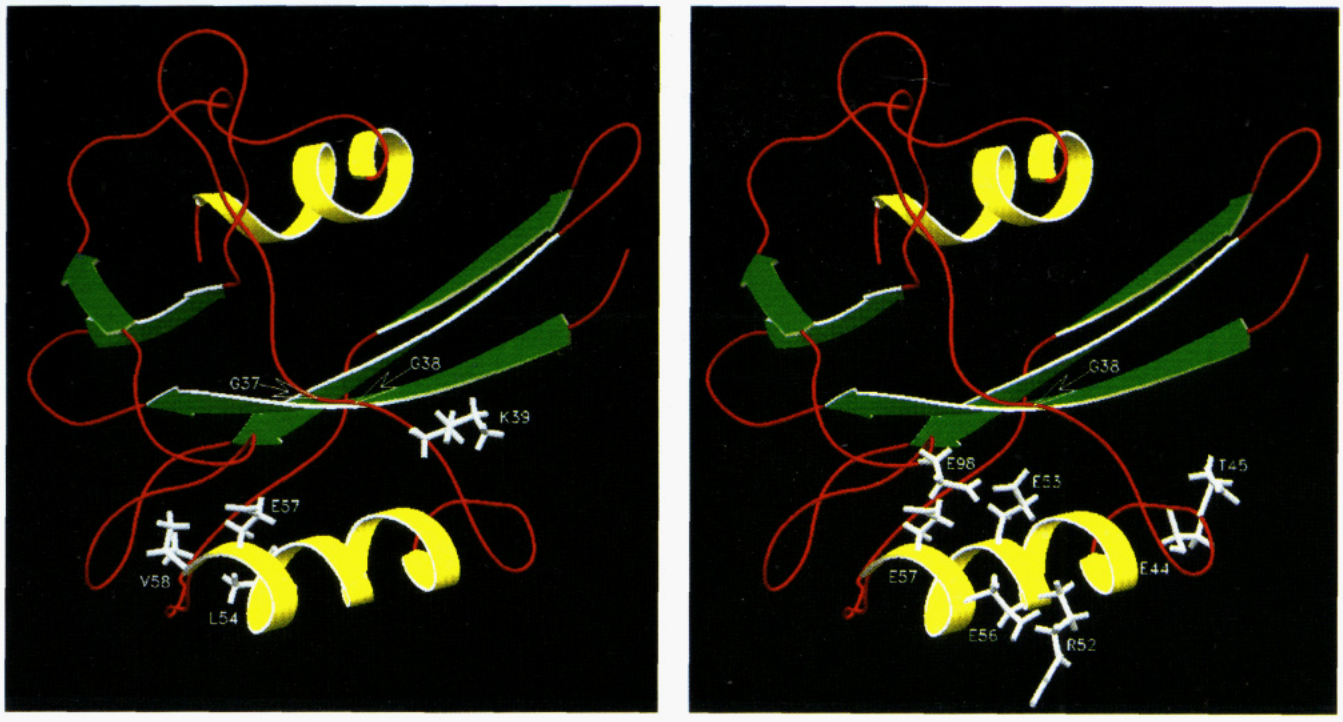

FIGURE 6: Ribbon diagram of the best-fit structure of the MutT enzyme showing (A, left) residues near the active site, as detected by paramagnetic effects of $MnCl_2$ on $^1H-^{15}N$ HSQC spectra and by intermolecular NOEs to bound Mg^{2+}AMPCPP or bound dGMP (Frick *et al.*, 1995), and (B, right) residues conserved among related nucleoside triphosphate pyrophosphohydrolases. In (A), the side-chain conformations of K39, L54, E57, and V58 are defined by a total of 8, 14, 6, and 23 NOEs, respectively. In (B), the side-chain conformations of E44, T45, R52, E53, E56, E57, and E98 are defined by a total of 4, 6, 9, 7, 4, 6, and 2 NOEs, respectively. Hence the conformations of E44, E56, and E98 are not well-defined.

helix II, are exposed. Of the 37 charged residues of MutT, 36 are exposed. The exception is Arg-22, which is buried and forms a monodentate salt bridge to a carboxylate oxygen of Glu-34, which is exposed.

All 9 prolines of MutT were previously found to be in the trans conformation (Weber *et al.*, 1993). Seven of these prolines are in loops, resulting in kinks or changes in chain direction. Thus, Pro-46 at the end of loop I forms a half-turn leading into helix I, and the consecutive prolines 116 and 117 in loop IV form the expanded apex of a wide turn leading into helix II. Of the two remaining prolines, Pro-76 is the second residue in a type I β-turn, a preferred location for proline (Chou & Fasman, 1978). Pro-121 is the third residue of helix II, a location where this residue is found at a frequency comparable to its average frequency at all locations (Rose & Wolfenden, 1993). While the $\phi\psi$ angles of Pro-121 are appropriate for a helical residue ($-67°$, $-30°$), the i to $i+4$ hydrogen bond to Lys-125 is not made, the O to NH distance being 3.9 Å. No other distortions are found in helix II.

Location of the Active Site. The active site of the MutT enzyme binds both an essential divalent cation activator and the M^{2+}-NTP substrate. Previous NMR titrations, monitoring $^1H-^{15}N$ HSQC spectra, detected paramagnetic broadenings by Mn^{2+}AMPCPP of the NH cross peaks assigned to Gly-37, Gly-38, Lys-39, and Glu-57. Transferred NOEs from proton resonances of Leu-54 and Val-58 to those of bound Mg^{2+}AMPCPP and the bound product dGMP were tentatively assigned (Frick *et al.*, 1995). These residues are grouped near each other in the loop I–helix I motif (Figure 6A). Similarly, residues conserved among the entire class of MutT-like pyrophosphohydrolases, Gly-38, Glu-44, Thr-45, Arg-52, Glu-53, Glu-56, Glu-57, and Glu-98 (Mejean *et al.*, 1994), are also mainly clustered along the loop I–helix I motif, implicating this region in catalysis (Figure 6B). With the exceptions of Glu-44, -56, and -98 in Figure 6B, the side-chain conformations shown in Figure 6 are reasonably well defined. Accordingly, mutations in the loop I–helix I region such as E34K (Mejean *et al.*, 1994), K39Q, R52Q (Frick *et al.*, 1995), and E57Q[5] profoundly decrease catalytic activity by 10^2-10^5-fold.

A surface potential model of the MutT protein (Figure 7) reveals a depressed patch of strongly negative electrostatic potential just above helix I resulting from the proximity of five Glu residues, Glu-41, -53, -56, -57, and -98, which probably constitutes the binding site for one or both of the essential divalent cations. This site is contiguous with a deep cleft defined by β-strands A, C, and D (Figure 7, upper right) and loop I (Figure 7, upper left) which may contribute to the nucleotide binding site.

DISCUSSION

The tertiary structure of the MutT enzyme, consisting of a five-stranded mixed β-sheet sandwiched between two α-helices, connected by long loops (Figure 4), places it in the structural category of $\alpha + \beta$ fold proteins (Orengo & Thornton, 1993). The $\alpha + \beta$ fold of MutT resembles, but is not identical to, certain other members of this class of proteins such as the actin binding domain of severin (Schnuchel *et al.*, 1995) and the carbohydrate recognition domain of a mannose binding protein (Weis *et al.*, 1991), both of which bind Ca^{2+}, the phosphotyrosine recognition

[5] J. Lin, D. N. Frick, M. J. Bessman, and A. S. Mildvan, unpublished observations, 1995.

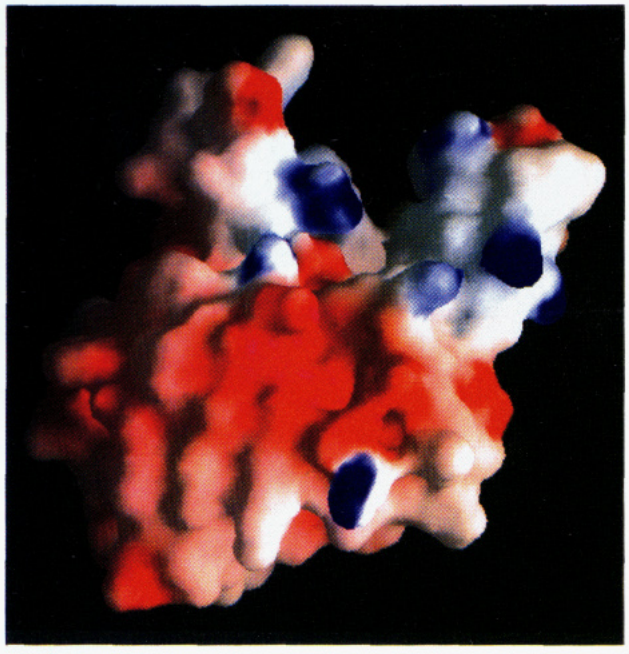

FIGURE 7: Protein surface potential model of the MutT enzyme viewed along the putative active site cleft, drawn with the program GRASP (Nicholls *et al.*, 1991). Regions of strong negative (red) and positive (blue) electrostatic potential are shown. The region of strong negative potential contains five Glu residues (41, 53, 56, 57, and 98) and probably constitutes the metal binding site. This site is contiguous with a deep cleft between β-strands A, C, and D (upper right) and loop I (upper left) which may contribute to the nucleotide binding site.

SH2 domain (Waksman *et al.*, 1992), and the mechanistically related enzyme inorganic pyrophosphatase (Kankare *et al.*, 1994; Teplyakov *et al.*, 1994) although none of these proteins show significant sequence homology to MutT (Mejean *et al.*, 1994).

Sequence homology is however found among MutT-like pyrophosphohydrolases from several bacteria, from a human cell line (Mo *et al.*, 1992; Sakumi *et al.*, 1993), and with the Orf 17 protein from *E. coli* which catalyzes a similar dATP pyrophosphohydrolase reaction (Mejean *et al.*, 1994; O'Handley *et al.*, 1995). Figure 6B shows these conserved residues to be located near each other in the solution tertiary structure, mainly in the loop I–helix I region of the enzyme. Similarly, distance-dependent NMR studies implicate residues in this region as being near the essential divalent cations, the enzyme-bound nucleotide substrate analog Mg^{2+}AMPCPP, and the bound reaction product dGMP (Figure 6A) (Frick *et al.*, 1995). Moreover, mutations of residues in loop I (E34K, K39Q) and helix I (R52Q, $E57Q^5$) profoundly decrease enzymatic activity without major alterations in protein structure, suggesting that these residues may participate directly in catalysis (Frick *et al.*, 1995; Mejean *et al.*, 1995). The contribution of Glu-57 to the negatively charged cluster of residues (Figure 7), consisting of Glu-41, -53, -56, -57, and -98, suggests that Glu-57 may be a ligand for one of the two divalent cations essential for the pyrophosphohydrolase reaction catalyzed by MutT. The precise roles of these residues in the mechanism of the MutT pyrophosphohydrolase reaction are currently under investigation. It is of interest that this cluster of Glu residues is contiguous with a deep cleft (Figure 7) which may contribute to the nucleotide binding site. Toward this end, we have recently determined the conformations of two enzyme-bound nucleotides (Frick *et al.*, 1995) and are in the process of docking them into the NMR structure.

ACKNOWLEDGMENT

We are grateful to Lewis E. Kay for providing pulse sequences and valuable discussions and to Peggy Ford for expert secretarial assistance.

SUPPORTING INFORMATION AVAILABLE

One table containing the ^1H, ^{13}C, and ^{15}N chemical shifts of the MutT protein from *E. coli* and two figures, a 2D $^{13}C-^1HN$ plane through the most crowded ^{15}N chemical shift in a sensitivity-enhanced 3D C(CO)NH spectrum of MutT in H_2O at 500 MHz with gradient coherence selection (Figure S1) and a selected $^1H-^1H(^{13}C)$ TOCSY plane of a 3D HCCH-TOCSY spectrum of MutT in H_2O at 500 MHz (Figure S2) (11 pages). Ordering information is given on any current masthead page.

REFERENCES

Abeygunawardana, C., Weber, D. J., Frick, D. N., Bessman, M. J., & Mildvan, A. S. (1993) *Biochemistry 32*, 13071–13080.

Abeygunawardana, C., Weber, D. J., Gittis, A. G., Frick, D. N., Miller, A.-F., Bessman, M. J., & Mildvan, A. S. (1995a) *J. Cell. Biochem., Abstr. Suppl. 21B*, Abstract No. D2-300, p 54.

Abeygunawardana, C., Weber, D. J., Gittis, A. G., Frick, D. N., Miller, A.-F., Bessman, M. J., Lin, J., & Mildvan, A. S. (1995b) *FASEB J. 9*, A1466.

Bax, A., & Summers, M. F. (1986) *J. Am. Chem. Soc. 108*, 2093–2094.

Bax, A., & Pochapsky, S. S. (1992) *J. Magn. Reson. 99*, 638–643.

Bax, A., Griffey, R. H., & Howkins, B. L. (1983) *J. Magn. Reson. 55*, 301–315.

Bax, A., Clore, G. M., & Gronenborn, A. M. (1990) *J. Magn. Reson. 88*, 425–431.

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

Brünger, A. T. (1992) *X-PLOR Version 3.1, A system for x-ray crystallography and NMR*, Yale University, New Haven, CT.

Cheng, K. C., Cahill, D. S., Kasai, H., Nishimura, S., & Loeb, L. A. (1991) *J. Biol. Chem. 267*, 166–172.

Chou, P. Y., & Fasman, G. D. (1978) *Adv. Enzymol. Relat. Areas Mol. Biol. 47*, 45–148.

Clore, G. M., Nilges, M., Sukumaran, D. K., Brünger, A. T., Karplus, M., & Gronenborn, A. M. (1986) *EMBO J. 5*, 2729–2735.

Clore, G. M., Gronenborn, A. M., Nilges, M., & Ryan, C. A. (1987) *Biochemistry 26*, 8012–8013.

Evans, S. V. (1993) *J. Mol. Graphics 11*, 134–138.

Frick, D. N., Weber, D. J., Gillespie, J. R., Bessman, M. J., & Mildvan, A. S. (1994) *J. Biol. Chem. 269*, 1794–1803.

Frick, D. N., Weber, D. J., Abeygunawardana, C., Gittis, A. G., Bessman, M. J., & Mildvan, A. S. (1995) *Biochemistry 34*, 5577–5586.

Gronenborn, A. M., Bax, A., Wingfield, P. T., & Clore, G. M. (1989) *FEBS Lett. 243*, 93–98.

Grzesiek, S., Anglister, J., & Bax, A. (1993) *J. Magn. Reson. 101B*, 114–119.

Kamath, A. V., & Yanofsky, C. (1993) *Gene 134*, 99–102.

Kankare, J., Neal, G. S., Salminen, T., Glumhoff, T., Cooperman, B. S., Lahti, R., & Goldman, A. (1994) *Protein Eng. 7*, 823–830.

Kay, L. E., & Bax, A. (1990) *J. Magn. Reson. 86*, 110–126.

Kay, L. E., Keifer, P., & Saarinen, T. (1992) *J. Am. Chem. Soc. 114*, 10663–10665.

Kay, L. E., Xu, G.-Y., Singer, A. U., Muhandiram, D. R., & Forman-Kay, J. D. (1993) *J. Magn. Reson. 101B*, 333–337.

Kessler, H., Griesinger, C., Kerssebaum, R., Wagner, K., & Ernst, R. R. (1987) *J. Am. Chem. Soc. 109*, 607–609.

Kumar, A., Wagner, G., Ernst, R. R., & Wüthrich, K. (1980) *Biochem. Biophys. Res. Commun. 96*, 1156−1163.
Laskowski, R. A., McArthur, M. W., Moss, D. S., & Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
Levitt, M., & Chothia, C. (1976) *Nature 261*, 552−558.
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.
Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967−974.
Marion, D., Ikura, M., Tschudin, R., & Bax, A. (1989) *J. Magn. Reson. 85*, 393−399.
Mejean, V., Salles, C., Bullions, L. C., Bessman, M. J., & Claverys, J.-P. (1994) *Mol. Microbiol. 11*, 323−330.
Mo, J. Y., Maki, H., & Sekiguchi, M. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 11021−11025.
Mori, S., Abeygunawardana, C., Johnson, M. O., & van Zijl, P. C. M. (1995) *J. Magn. Reson. B108*, 94−98.
Muhandiram, D. R., & Kay, L. E. (1994) *J. Magn. Reson. 103B*, 203−216.
Muhandiram, D. R., Farrow, N. A., Xu, G.-Y., Smallcombe, S. H., & Kay, L. E. (1993) *J. Magn. Reson. 102B*, 317−321.
Nicholls, A., Sharp, K. A., & Honig, B. H. (1991) *Proteins: Struct., Funct., Genet. 11*, 281−296.
Nilges, M., Clore, G. M., & Gronenborn, A. M. (1988) *FEBS Lett. 229*, 317−324.
O'Handley, S. F., Frick, D. N., Bullions, L. C., Mildvan, A. S., & Bessman, M. J. (1995) *FASEB J.*, A1299 (Abstract 255).
Orengo, C. A., & Thornton, J. M. (1993) *Structure 1*, 105−120.
Orengo, C. A., Flores, T. P., Taylor, W. R., & Thornton, J. M. (1993) *Protein Eng. 6*, 485−500.
Rose, G. D., & Wolfenden, R. (1993) *Annu. Rev. Biophys. Biomol. Struct. 22*, 381−415.
Sakumi, K., Furuichi, M., Tsuzuki, T., Kakuma, T., Kawahata, S., Maki, H., & Sekiguchi, M. (1993) *J. Biol. Chem. 268*, 23524−23530.
Schnuchel, A., Willschek, R., Eichinger, L., Schleicher, M., & Holak, T. A. (1995) *J. Mol. Biol. 247*, 21−27.
Shaka, A. J., Lee, C. J., & Pines, A. (1988) *J. Magn. Reson. 77*, 274−293.
Teplyakov, A., Obmolova, G., Wilson, K. S., Ishii, K., Kaji, H., Samejima, T., & Kuranova, I. (1994) *Protein Sci. 3*, 1098−1107.
Vuister, G. W., & Bax, A. (1992) *J. Magn. Reson. 98*, 428−435.
Waksman, G., Kominos, D., Robertson, S. C., Pant, N., Baltimore, D., Birge, R. B., Cowburn, D., Hanafusa, H., Mayer, B. J., Overduin, M., Resh, M. D., Rios, C. B., Silverman, L., & Kuriyan, J. (1992) *Nature 358*, 646−653.
Weber, D. J., Bhatnagar, S. K., Bullions, L. C., Bessman, M. J., & Mildvan, A. S. (1992) *J. Biol. Chem. 267*, 16939−16942.
Weber, D. J., Abeygunawardana, C., Bessman, M. J., & Mildvan, A. S. (1993) *Biochemistry 32*, 13081−13088.
Weis, W. I., Kahn, R., Fourme, R., Drickamer, K., & Hendricksen, W. A. (1991) *Science 254*, 1608−1615.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, pp 166−168, John Wiley & Sons, New York, NY.
Wüthrich, K., Billeter, M., & Braun, W. (1983) *J. Mol. Biol. 169*, 949−961.
Yamazaki, T., Pascal, S. M., Singer, A. U., Forman-Kay, J. D., & Kay, L. E. (1995) *J. Am. Chem. Soc. 117*, 3556−3564.

BI951449L