BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

Biochemistry **2010,** *49,* 5705–5713 5705
DOI: 10.1021/bi100171x

Crystal Structure of the Human *Hsmar1*-Derived Transposase Domain in the DNA Repair Enzyme Metnase[†,‡]

Kristie D. Goodwin,[§] Hongzhen He,[§] Tsuyoshi Imasaki,[§] Suk-Hee Lee,[§] and Millie M. Georgiadis*[,§,∥]

[§]*Department of Biochemistry and Molecular Biology, Indiana University School of Medicine, Indianapolis, Indiana 46202, and* [∥]*Department of Chemistry and Chemical Biology, Purdue School of Science, Indiana University-Purdue University, Indianapolis, Indiana 46202*

Received February 4, 2010; Revised Manuscript Received May 28, 2010

ABSTRACT: Although the human genome is littered with sequences derived from the *Hsmar1* transposon, the only intact *Hsmar1* transposase gene exists within a chimeric SET-transposase fusion protein referred to as Metnase or SETMAR. Metnase retains many of the transposase activities including terminal inverted repeat (TIR) specific DNA-binding activity, DNA cleavage activity, albeit uncoupled from TIR-specific binding, and the ability to form a synaptic complex. However, Metnase has evolved as a DNA repair protein that is specifically involved in nonhomologous end joining. Here, we present two crystal structures of the transposase catalytic domain of Metnase revealing a dimeric enzyme with unusual active site plasticity that may be involved in modulating metal binding. We show through characterization of a dimerization mutant, F460K, that the dimeric form of the enzyme is required for its DNA cleavage, DNA-binding, and nonhomologous end joining activities. Of significance is the conservation of F460 along with residues that we propose may be involved in the modulation of metal binding in both the predicted ancestral *Hsmar1* transposase sequence as well as in the modern enzyme. The Metnase transposase has been remarkably conserved through evolution; however, there is a clustering of substitutions located in alpha helices 4 and 5 within the putative DNA-binding site, consistent with loss of transposition specific DNA cleavage activity and acquisition of DNA repair specific cleavage activity.

While transposase activity has played an important evolutionary role accounting for half of the present organization of the human genome (*1*), little is known about the role of transposases in humans. In lower organisms, transposases mediate DNA movement by excising defined segments of DNA and then reinserting them at other locations in the genome, a process which can be repeated multiple times for a given segment (*2*−*4*). The *Hsmar1* transposon, a class II transposable element, is an ancient element within the human genome introduced at least 50 million years ago in ancestral primates (*5*). To date, only one example of an intact copy of the *Hsmar1* transposase domain has been identified within the human genome (*5*). This "functional" *Hsmar1* transposase domain exists within a chimeric fusion protein, Metnase (also known as SETMAR), which resulted from insertion of the *Hsmar1* transposon downstream of a SET gene (suppressor of variegation 3-9, enhancer-of-zeste, trithorax)[1] encoding a histone methyltransferase, ultimately fusing the SET and transposase domains (*5*−*7*).

Metnase is widely expressed in human tissues and in vitro possesses histone 3 methylation activity at lysine residues 4 and 36 (*8*) associated with chromatin opening (*9*−*12*). It promotes nonhomologous DNA end-joining (NHEJ) repair and mediates genomic integration of foreign DNA (*8*, *13*−*15*). Upon DNA damage, Metnase colocalizes with other double-strand break (DSB) repair factors and has been shown to directly interact with Pso4 (*13*), a human homologue of the 55-kDa protein encoded by the PSO4/PRP19 gene in *Saccharomyces cerevisiae* that has pleiotropic functions in DNA recombination and error-prone repair (*16*). Metnase-mediated stimulation of NHEJ repair in vivo requires both histone methyltransferase and transposase-associated activities (*8*). Metnase is also involved in chromosomal decatenation and its checkpoint pathway (*17*).

Although Metnase cannot perform transposition, it has been shown to retain a number of activities associated with transposases including 5′-terminal inverted repeats (TIR)-specific DNA binding (*6*), DNA looping activity, 5′-end processing activity, and promotion of integration at a TA dinucleotide target site (*18*, *19*). However, Metnase's DNA cleavage activity, unlike other functionally active transposases, is not coupled to its TIR-specific DNA binding (*19*). To our knowledge, there are no other DNA repair proteins in which the DNA cleavage and histone lysine methyltransferase activities reside within the same protein. Our goal in these studies is to determine how the Metnase transposase differs from other transposases as a first step toward understanding how it has acquired DNA repair activities. The crystal structures of the Metnase transposase provide a molecular basis for analyzing the importance of dimerization, the architecture of the catalytic site, and the potential importance of evolutionary amino acid substitutions as compared to the predicted ancestral sequence.

[†]This work was supported by a fellowship from the American Cancer Society (to K.D.G.) and Human Frontier Science Program Long-Term Fellowship (to T.I.).

[‡]Coordinates have been deposited with the PDB, entries 3K9J and 3K9K.

*To whom correspondence should be addressed. Telephone: (317) 278-8486; fax: (317) 274-4686; e-mail: mgeorgia@iupui.edu.

[1]Abbreviations: SET, suppressor of variegation 3-9, enhancer-of-zeste, trithorax; rmsd, root-mean-square deviation; HEPES, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; PEG, polyethylene glycol; EDTA, ethylenediaminetetraacetic acid; Tris, tris(hydroxymethyl)-aminomethane.

Published on Web 06/03/2010

EXPERIMENTAL PROCEDURES

Metnase Transposase Expression and Purification. The transposase domain of Metnase, residues 329−671, and the transposase catalytic domain, residues 433−671 (433) were subcloned and purified as described previously (*19*). A triple mutant including A649W, N566R, and G569R (433 triple mutant) was designed to improve packing in the crystal lattice based on comparison with the Mos1 structure as described in Supporting Information and purified as described (*19*). An additional mutant, L573M, was created for SeMet phasing of Metnase by site-directed mutagenesis of the triple mutant. The protein was expressed in B834 cells, grown in media replacing Met with SeMet (*20*), and purified as described for 433 triple mutant. The yield was 1−2 mg/L of culture. Purified protein was diluted in 50 mM HEPES (pH 7.0), 0.25 M NaCl and concentrated to a final concentration of 0.9 mM for use in crystallization experiments. The dimerization mutant, F460K, was created by site-directed mutagenesis of 433. The protein was expressed and purified as described for the triple mutant with a yield of 17 mg/L. See Supporting Information for further details including primers used in site-directed mutagenesis.

Crystallization, Data Collection, Structure Determination and Refinement. Crystals of 433 triple mutant ($40 \times 40 \times 300\ \mu m^3$) were grown by hanging drop vapor diffusion in 0.1 M NH_4OAc, 10 mM $CaCl_2$, 50 mM Tris-HCl pH 8.0, 10% PEG 4000 and cryosoaked in well solution containing 20% glycerol. Data sets were collected at beamline 19-ID at the Advanced Photon Source (APS) to 2.55 Å (see Table 1). Crystals of the L573M protein were microseeded with 433 triple mutant crystals, grown, and cryosoaked under the same conditions. Data sets of the SeMet substituted L573M crystals ($40 \times 40 \times 300\ \mu m^3$) were collected at APS 19-ID to 2.8 Å. See Supporting Information for further information regarding experimental phasing. Patterson searching and phasing calculations were performed in HKL3000 (*21*) using SHELXD/E (*22*, *23*), MLPHARE(*24*), and SOLVE/RESOLVE (*25*) followed by low resolution model building in RESOLVE and crystallographic refinement. The L573M dimer model was used initially and later the refined structure of the B subunit from the high resolution structure described below (Native 2 in Table 1) as the search model in molecular replacement to phase the native 433 triple mutant structure (Native 1 in Table 1). Coot (*26*) and O (*27*) were used for model building in all of the structures, and positional and individual isotropic B-factor refinement was done using CNS (*28*) and PHENIX (*29*).

The 433 triple mutant protein was crystallized with $d(GACTTGTAATATTACAAGTC)_2$ in a 1:1.2 ratio by microseeding with the 433 triple mutant crystals in 0.1 M NH_4OAc, 10 mM $CaCl_2$, 50 mM Tris-HCl pH 8.0, 16% PEG 4000. The oligonucleotide (TriLink BioTechnologies, San Diego, CA) was purified using standard reverse phase HPLC methods. Crystals were cryosoaked in well solution containing 20% glycerol. Data were collected for the triple mutant 433/oligonucleotide crystal ($30 \times 30 \times 200\ \mu m^3$) at APS 23-ID to 1.9 Å (Table 1), and the structure was solved by molecular replacement using the refined structure of the 433 triple mutant monomer as the search model. Following model building and refinement as described above, simulated annealing omit maps were calculated for the active site regions of both A and B subunits using PHENIX. Coordinates have been deposited with the PDB, 3K9J and 3K9K, for the 1.9 and 2.55 Å structures, respectively.

Gel Filtration Analysis. The oligomeric states of purified 433, the 433 triple mutant, and the 433 F460K were analyzed by gel filtration chromatography buffered in 50 mM HEPES (pH 7.0) and 0.5 M NaCl. The column was calibrated with low molecular weight standards (Pharmacia), and a calibration curve was used to calculate the molecular weights.

Cells, Enzymes, Oligonucleotides, and Antibodies. Human embryonic kidney (HEK)-293T cells used to generate stable expression of Flag-tagged wt-Metnase or the mutant (F460K) were previously described (*8*, *13*). Oligonucleotides were obtained from Integrated DNA Technologies (Coralville, IA). An anti-Metnase antiserum (polyclonal) was generated from rabbits using two peptides representing amino acids 483−495 and 659−671 (*8*). An anti-FLAG antibody was obtained from Sigma (St. Louis, MO). In addition, [γ-^{32}P]-ATP (3000 Ci/mmol) was from Perkin-Elmer and Analytical Science (Boston, MA), DE81 filters were from Whatman Bio System (Maidstone, England), and Bradford reagents and protein molecular weight markers were purchased from Bio-Rad (Hercules, CA).

Table 1: Summary of Crystallographic, Phasing and Refinement Statistics[a]

	SeMet (L573M)	Native 1	Native 2
Cell Parameters			
cell constants (Å)	$a = 82.6$	$a = 83.0$	$a = 79.2$
	$b = 82.6$	$b = 83.0$	$b = 45.4$
	$c = 288.7$	$c = 289.9$	$c = 90.7$
(deg)	$\alpha = 90$	$\alpha = 90$	$\alpha = 90$
	$\beta = 90$	$\beta = 90$	$\beta = 113.8$
	$\gamma = 120$	$\gamma = 120$	$\gamma = 90$
space group	$P6_122$	$P6_122$	$P2_1$
Data Collection[b]			
beamline	19-ID 03/2008	19-ID 11/2007	23-ID 07/2008
wavelength (Å)	0.979260	1.07203	1.10546
maximum resolution (Å)	2.8	2.55	1.9
reflections (unique)	15324	20367	46691
(total)	242795	225395	172427
completeness (%)	99.7 (100)	99.8 (99.6)	94.2 (71.1)
R_{sym} (%)[c]	8.4 (69.4)	5.8 (72.6)	6.0 (42.5)
I/σ	37.4 (3.2)	39.2 (2.6)	20.3 (1.9)
Phasing			
fom (initial)	0.33		
fom (NCS/solv flat)	0.70		
Refinement			
resolution range (Å)	N/A	50−2.55	50−1.9
atoms			
protein		3510	3395
water		103	266
average B-factor (Å^2)[#]			
A		71.3	41.6
B		64.4	37.1
waters		60.4	45.6
R_{value} (%)[d]		23.1	20.5
R_{free} (%)		29.1	23.4
rmsd bond length (Å)		0.007	0.007
bond angle (°)		1.03	1.05

[a]SeMet, Native 1, and Native 2 refer to crystals obtained of the triple mutant for the transposase catalytic domain. The Native 2 crystals were grown in the presence of DNA as described in the Experimental Procedures. [b]Data in () are for highest resolution shell. [c]$R_{sym} = \Sigma\Sigma_i|I_i - \langle I\rangle|/\Sigma\langle I\rangle$ where I is the integrated intensity of a reflection. [d]$R_{value} = \Sigma_{hkl}||F_{obs} - kF_{calc}|/\Sigma_{hkl}|F_{obs}|$. 5% of all reflections were omitted from refinement; R_{free} is the same statistic calculated for these reflections. 86% of residues and for the native 1 structure and 92.3% for native 2 fall in the most favored region of the Ramachandran plot. No residues in either structure are in the disallowed region.

Generation of HEK-293T Cells Overexpressing Either wt-Metnase or F460K. HEK-293T cells overexpressing control vector, Flag-tagged wt-Metnase, or F460K Metnase were generated by a stable transfection of HEK-293T cells with either pcDNA3.1-V5/His (Invitrogen) for the control or pcDNA3.1-wt-Metnase or -F460K for overexpression (*19*). Transfectants were selected in 800 μg/mL G418 for 14−21 days, and individual colonies were expanded and analyzed for Metnase expression by Western blot using a polyclonal antibody specific for Metnase (*8*).

Purification of Flag-Metnase and Flag-F460K Metnase. HEK-293T cells (1.6×10^8) stably expressing Flag-tagged wt-Metnase or F460K were suspended in 20 mL of cell lysis buffer (50 mM Tris-HCl pH 7.5, 1 mM EDTA, 10% glycerol, 5 mM DTT, 1.0% Nonidet-P40, and protease inhibitor cocktails containing 0.2 M NaCl), and centrifuged (100000*g*) for 60 min. The supernatant was collected and incubated at 4 °C for 60 min with anti-FLAG M2 affinity gel (Sigma) that had been pre-equilibrated with TEGDN buffer (50 mM Tris-HCl pH 7.5, 1 mM EDTA, 10% glycerol, 5 mM DTT, 0.02% Nonidet-P40, and protease inhibitor cocktails) containing 0.2 M NaCl. The beads were washed three times with TEGDN−2.0 M NaCl buffer prior to elution of the protein with TEGDN−0.2 M NaCl containing FLAG peptide (500 μg/mL). The eluant was diluted with 10 vol of TEGDN buffer, and loaded onto a heparin-Sepharose 6 Fast Flow column (Amersham Biosciences) pre-equilibrated with TEGDN buffer. After the column was washed, the protein was fractionated using a linear gradient (0−2.0 M NaCl) of TEGDN buffer. The eluted protein was dialyzed against TEGDN buffer containing 50 mM NaCl and stored at −80 °C.

Electrophoretic Mobility Shift Assay (EMSA) of Protein−DNA Interaction. Duplex DNA containing the 19-mer core of the 5′-terminal inverted repeat (TIR) (*19*) was labeled with [γ-^{32}P]-ATP (3000 Ci/mmol) and T4 polynucleotide kinase (Roche Molecular Biochemical) according to the manufacturer's instructions. Indicated amounts of purified wt-Metnase or the mutant (F460K) were incubated with 200 fmol of 5′-^{32}P-labeled DNA at room temperature for 15 min in a reaction mixture containing 50 mM Tris-HCl (pH 7.5), 5 mM DTT, bovine serum albumin (0.2 μg/μL), 5% (w/v) glycerol, and 50 mM NaCl. The protein(s)−DNA complex was analyzed on a 5% polyacrylamide gel (acrylamide: bis-acrylamide, 43.2:0.8) in 0.5× TBE. The gel was dried and exposed to X-ray films (Kodak).

DNA Cleavage Assay. Metnase's DNA cleavage assay was previously described (*19*). Briefly, the reaction mixtures (20 μL) containing 200 fmol of 5′-^{32}P-ssDNA (50-mer) were incubated with indicated amounts of wt-Metnase or the mutant in the presence of 1 mM $MgCl_2$. After incubation at 37 °C for 60 min, reaction mixtures were analyzed by 12% denaturing polyacrylamide gel for DNA cleavage.

DNA End Joining Coupled to Genomic Integration Analysis. Chromosomal integration was analyzed by the ability of cells to pass foreign nonhomologous DNA containing a selectable marker to progeny as described previously (*8*). pRNA/U6.Hygro plasmid linearized with *Asp*718I was transfected into HEK-293T cells stably expressing mock (pFlag2 vector), pFlag2-wtMetnase, or pFlag2-F460K using a calcium phosphate transfection method as previously described (*8*). Forty-eight hours after transfection, varying numbers of cells were plated into 100 mm dishes and 24 h later selective marker (hygromycin: 0.15 mg/mL) was added. Cells were incubated for 14 days in the presence of hygromycin, washed twice with PBS, stained with 0.17% methylene blue in methanol, and colonies defined as greater than 50 cells were counted. All experiments were performed in triplicate.

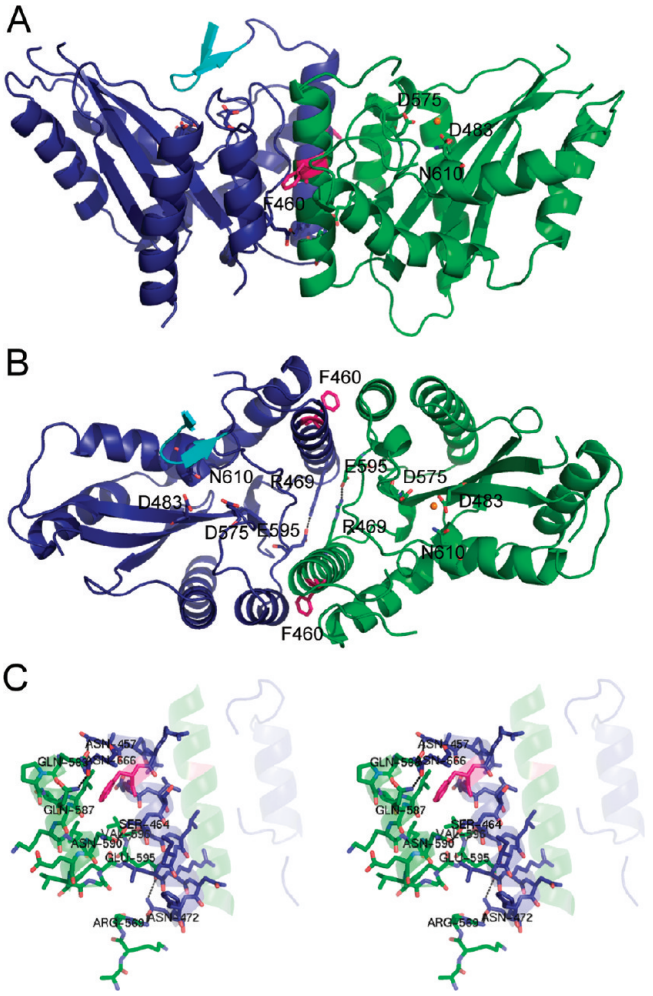

FIGURE 1: Metnase transposase domain forms a unique dimer. (A) Cartoon rendering of Metnase dimer with A subunit in green (Ca^{2+}, orange sphere) and B subunit in navy. The N-terminal region unique to the B subunit is highlighted in cyan. Residues of interest, shown as sticks, include catalytic site residues D483, D575, and N610, along with dimer interface residues R469, E595, and F460 (red). (B) View of the Metnase dimer rotated by approximately 90°. (C) Intramolecular hydrogen bonding at dimer interface with bonds shown as black dashed lines. Subunit A in green sticks and B in navy sticks with F460 in red.

RESULTS AND DISCUSSION

Crystallization and Structure Determination. Metnase is predicted to include at least one helix-turn-helix motif and a catalytic domain, similar to Mos1, a related transposase (*30*). Thus, we initially prepared proteins including residues 329−671 and 433−671, the former including the helix-turn-helix motif and the latter corresponding to the catalytic domain of the transposase. The N-terminus of the shorter construct was identified by limited proteolysis and N-terminal sequencing. As neither of these proteins crystallized, a triple mutant (433 triple mutant), A649W, N566R, and G569R, of the catalytic domain was engineered to facilitate crystal packing interactions. Of these substitutions, A649W forms critical packing interactions within the Metnase lattices, albeit different interactions than those observed for the equivalent residue in Mos1 (see Supporting Information and Figure S1). The crystal structure of the catalytic

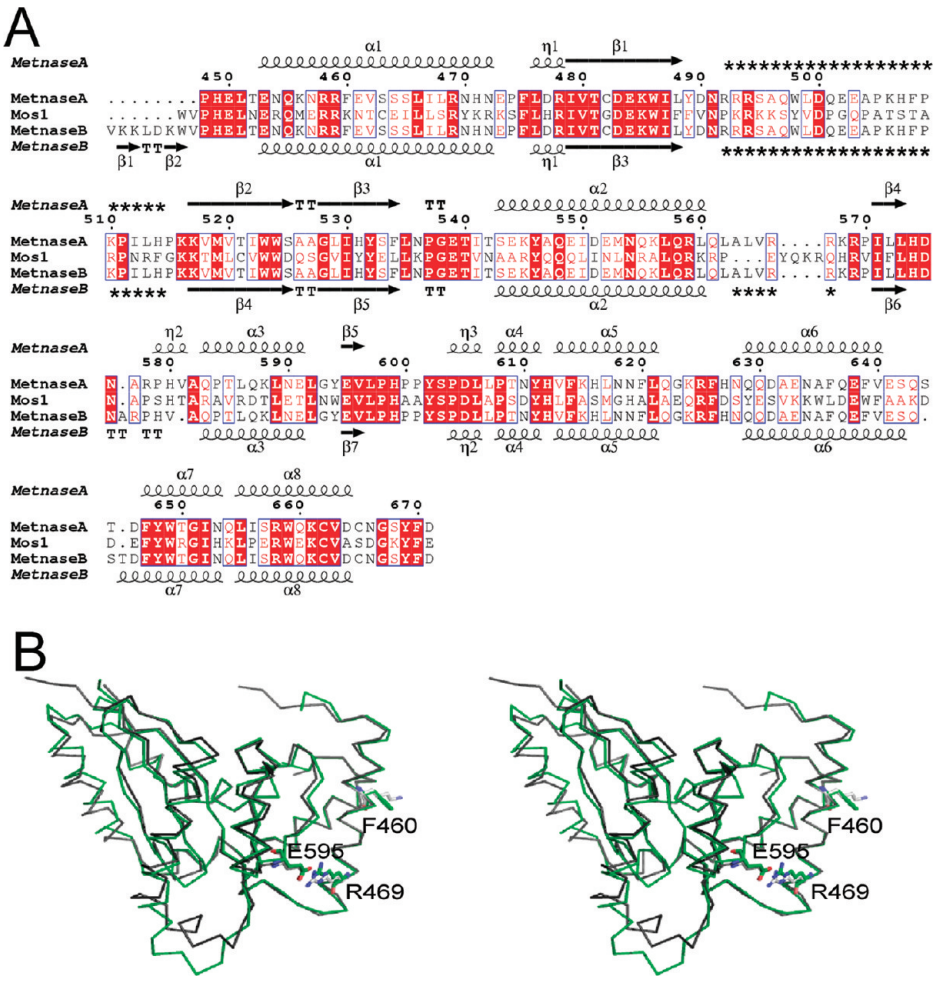

FIGURE 2: Comparison of Metnase and Mos1. (A) Structure-based sequence alignment of Metnase with Mos1. Conserved regions are shaded in red, and secondary structural elements are shown above the sequences. Asterisks denote disordered regions that are not modeled in the structure. Both subunits of Metnase are shown to highlight structural differences in the metal-bound A subunit vs the nonmetal bound B subunit. (B) Trace renderings are shown of the A subunit of Metnase (green) superimposed on Mos1 (gray) as a stereodiagram with stick renderings of residues R469, E595, and F460, and equivalent residues in Mos1, R142, E269, and K133.

domain (433 triple mutant) was determined in an hexagonal lattice using SeMet phasing (see Experimental Procedures and Table 1). Subsequently, a 1.9 Å structure was determined for 433 triple mutant crystallized with a 20-mer duplex DNA in a monoclinic lattice (see Table 1). Although the DNA is not ordered in the lattice, there is a large hole in the lattice not observed in the lattice for the protein alone (see Supporting Information and Figure S2) despite the fact that both crystal forms were obtained from nearly identical precipitant solutions containing 10 mM $CaCl_2$ as described in the Experimental Procedures. The structural results described below refer to the 1.9 Å crystal structure unless otherwise stated.

Metnase Forms a Unique Dimer. In contrast to other transposases including Mos1, the catalytic domain of the Metnase transposase forms a dimer. In both crystal lattices obtained for the transposase domain of Metnase, there are two transposase domains that form a dimer within the asymmetric unit, the unique repeating unit of the crystal (Figure 1, Figure S2). The two molecules comprising the Metnase dimer (referred to as A and B) differ somewhat from one another (rmsd of 0.7 Å for 191 Cα atoms) with the structural model for B including residues 440−491, 515−562, and 568−671 and that of A, residues 448−491 and 515−671. The dimer interface involves packing of the N-terminal helix, α1 (453−472), and the residues C-terminal of helix α8 (663−667) in one molecule with α3 (583−593) in the other molecule (Figures 1C and 2). This interface buries a surface area of approximately 2600 $Å^2$ forming 11 direct and 10 water-mediated hydrogen bonding interactions between residues in the A molecule with those in the B molecule (Supporting Information, Table S1). Most of the hydrogen-bonding interactions involve side chain atoms of polar residues. The dimer interface is not precisely symmetric with only three pairs of symmetric interactions present in both possible pairs of interactions between A and B molecules.

To identify residues critical for the formation of the observed dimer interface, we compared the structures of the catalytic domains of Metnase and Mos1, which dimerizes through a helix-turn-helix motif, rather than the catalytic domain (*30*). Although at the sequence level, the catalytic domain of the Metnase transposase is 38% identical to that of Mos1 (Figure 2A), the proposed dimer of Mos1 catalytic domain is unrelated to that formed by Metnase in that it dimerizes through an entirely different interface (Figure S3, Supporting Information). Nonetheless, the overall fold of the monomeric Mos1 catalytic domain is similar to that of a single subunit of the Metnase dimer (Figure 2B) with an rmsd of 0.9 Å for superpositioning of 191 α-carbon atoms, and many residues involved in the dimer interface of Metnase are conserved including an ion-pair,

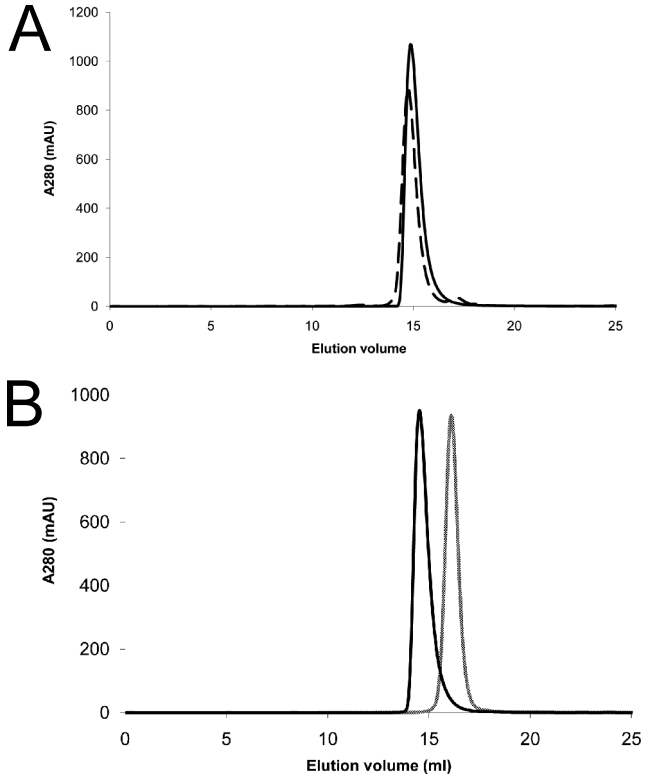

FIGURE 3: F460 is required for dimerization of the Metnase transposase domain. Gel filtration chromatography analysis of (A) wt-Metnase (residues 433−671) shown in a black solid line vs triple mutant (A649W, N566R, G569R) shown as a dashed line and (B) F460K shown as a gray line vs the triple mutant shown as a black line. Both wt-Metnase with a molecular weight of 63 kDa and the triple mutant with a molecular weight 55 kDa eluted as dimeric species, while the F460K mutant eluted with a molecular weight of 31 kDa representing a monomer in solution.

R469−E595. One residue that stood out in the comparative analysis was F460 in Metnase for which the equivalent residue in Mos1 is K133. Given the hydrophobic nature of F460, we proposed an important role in dimerization for this residue.

Consistent with the crystallographic results, the transposase domain of Metnase, both the wild-type 433−671 (433) as well as 433 triple mutant, forms a dimer in solution as characterized by gel filtration chromatography (Figure 3). To determine whether F460 is required for dimerization of the transposase domain in Metnase, we introduced a Lys substitution at this position as found in Mos1. The 433 F460K was subjected to gel filtration chromatographic analysis and found to elute as a monomer rather than a dimer (Figure 3). Thus, Phe 460 appears to be required for dimerization of the transposase domain.

Functional Implications for the Metnase Dimer. To understand the functional implications of the dimeric Metnase structure, F460K full-length Metnase was expressed, purified, and tested for TIR-specific DNA-binding, DNA cleavage activity, and DNA end joining activity (See Experimental Procedures). One of the transposase activities retained by Metnase is TIR-specific DNA-binding activity. Metnase specifically binds TIR sequences derived from *mariner* transposable elements (*6*), although these elements have been altered by deletions and insertions to the extent that they are no longer functional for transposition in the human genome. F460K Metnase does not bind the TIR sequence as analyzed by EMSA (Figure 4A). Thus, dimerization of the Metnase transposase domain appears to be required for TIR-binding activity.

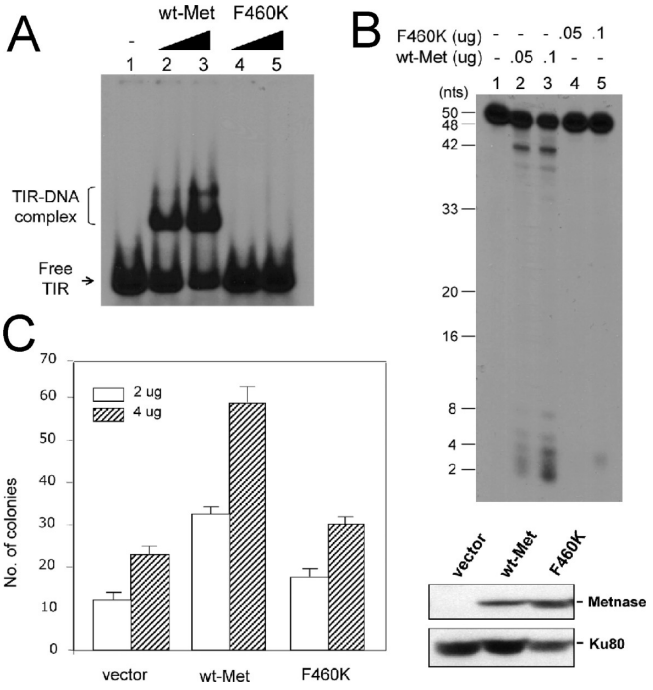

FIGURE 4: The F460K mutant fails to support Metnase's function in interaction with TIR DNA, ssDNA cleavage, and in vivo DNA end joining. (A) Gel shift assay of F460K with TIR DNA. The ^{32}P-TIR DNA (200 fmol) containing 19-mer of TIR core was incubated with 0.15 or 0.3 μg of either wt-Metnase (lanes 2−3) or F460K (lanes 4−5) and examined for protein−DNA complex formation using 5% PAGE. Positions of free TIR DNA and the Metnase−TIR complexes are indicated on the left. (B) DNA cleavage activity assay of wt-Metnase and the F460K mutant. Reaction mixtures (20 μL) containing 50 fmol of 5′-^{32}P-labeled 50-mer of ssDNA were incubated with 0 (lane 1), 0.05 μg (lanes 2 and 4) and 0.1 μg (lanes 3 and 5) of wt-Metnase (lanes 2−3) or F460K (lanes 4−5) in the presence of 2 mM $MgCl_2$. After incubation for 120 min at 37 °C, cleavage products were analyzed by 12% PAGE + 8 M urea. DNA size markers are indicated on the left. (C) Colony formation assay representing DNA end joining in the presence of wt-Metnase or the F460K mutant. Left panel, HEK-293T cells stably transfected with pFlag2 (control), pFLAG2-wt-Metnase (wt-Met), or pFLAG2-F460K (F460K) were transfected with two concentrations (2 and 4 μg) of linearized pRNA/U6-Hygro, and the number of hygromycin-resistant colonies was a measure of genomic integration. Values are averages (±SEM) of five separate experiments ($P < 0.01$). Right panels, Western blot analysis of stably transfected HEK-293T cell extracts using an anti-Flag antibody (top panel). Expression of Ku80 was included as an internal control (bottom panel).

In previous reports, Metnase retains a more robust 5′ end processing cleavage activity than 3′ end processing activity (*18*). As the end-processing activity involving cleavage of ssDNA overhangs may be more relevant in its DNA repair associated activities, we tested both Metnase and the F460K mutant for ssDNA cleavage activity. In comparison to the wild-type Metnase, F460K Metnase has significantly reduced ssDNA cleavage activity as shown in Figure 4B. Thus, we conclude that the Metnase dimer is also necessary for its DNA cleavage activity.

Since F460K Metnase was defective in dimer formation and also lacked DNA cleavage activity, we examined whether this mutant supports Metnase's function in DNA end joining. For this, we compared wt-Metnase with the F460K mutant for stimulation of genomic integration of exogenous DNA by measuring the assimilation and passage to progeny of a selective marker. Stable overexpression of wt-Metnase in HEK-293T cells lacking endogenous Metnase did not affect plating efficiency (*8*, *13*), but it

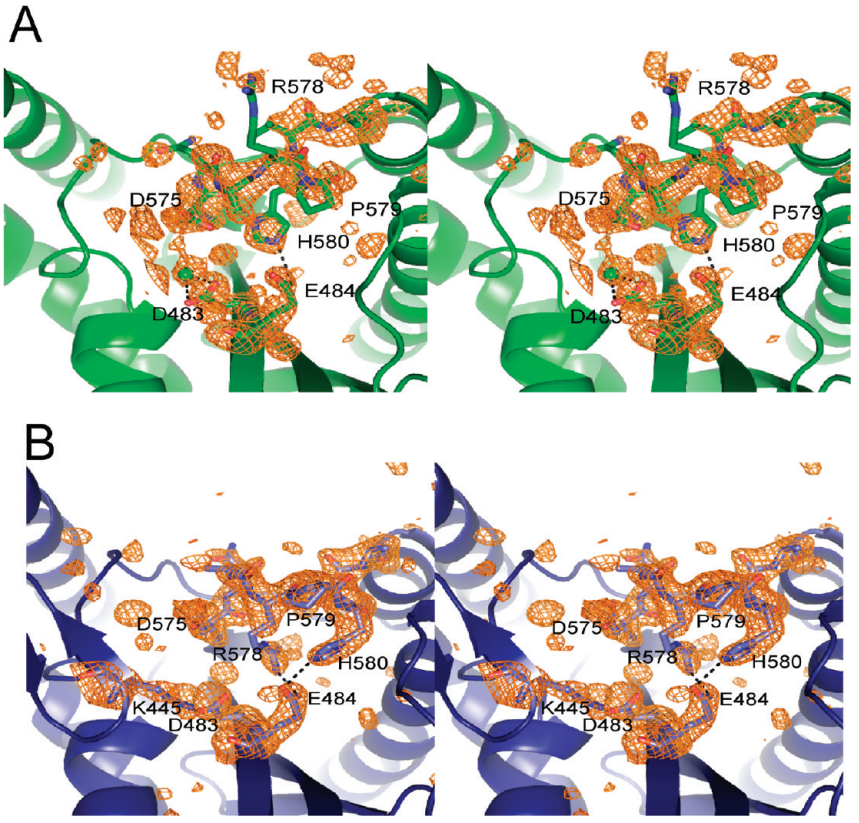

FIGURE 5: Active site plasticity in Metnase transposase domain. (A) Subunit A is shown in stereo as a green cartoon rendering with catalytic site and loop residues in sticks and Ca^{2+}as a green sphere with H-bonds represented as black dashed lines. A simulated annealing omit map ($F_o - F_c$) in which D483, D484, loop residues 575−582, and Ca^{2+} of the A subunit were omitted is shown superimposed as an orange mesh contoured at 2.5σ. (B) No metal is bound in the B active site shown in stereo as a navy cartoon rendering with K445 and active site residues and shifted loop residues shown as sticks. H-bonds are shown as black dashed lines. A simulated annealing omit map ($F_o - F_c$) in which K445, D483, D484, and loop residues 575−582 of the B subunit were omitted is shown superimposed as an orange mesh contoured at 2.5σ. Residues 575−582 within the loop adopt a different conformation in the presence of bound metal corresponding to a translation by one full residue. Consequently, R578 points out of the active site in the presence of metal and into the site in its absence. Similarly, K445 is bound in the active site in the absence of metal and is disordered in its presence.

increased genomic integration of a plasmid DNA by 3-fold, whereas overexpression of the F460K mutant had very little effect on genomic integration (Figure 4C). Together, these data suggest that Metnase dimerization is not only essential for DNA binding and DNA cleavage activity but is also necessary for its function in DNA end joining in vivo.

The Catalytic Site of Metnase Transposase. One possible mechanism by which Metnase may have acquired DNA repair activities is a change in the architecture of the catalytic site of the enzyme. To examine this possibility, we compared the catalytic sites in the dimeric Metnase transposase structures to each other and to those in Mos1 and Tn5 transposases, both active transposases (*30*, *31*). Mos1 is related both at the sequence and structural level to the Metnase transposase (Figure 1), while Tn5 is a larger and structurally distinct enzyme with no sequence similarity to Metnase (Supporting Information and Figure S1). The Metnase transposase domain includes a DDN catalytic motif (D483, D575, and N610) contrasting with the DDD motif (D156, D249, and D284) found in Mos1. And yet, despite the D to N substitution, Metnase has been reported to retain many of the activities associated with active transposases (*6*, *18*, *19*).

In one of the two molecules (the B molecule) within the dimer, the N-terminal region forms a beta hairpin structure placing Lys 445 in the active site of this molecule (B molecule, shown in blue in Figure 5B) within hydrogen bonding distance of Asp 483, one of the important catalytic residues. No metal ion is bound to this active site. In contrast, Ca^{2+} is bound to the active site of the other molecule comprising the dimer (the A molecule, Figure 5A) coordinated to both carboxylate oxygen atoms of D483 and three water molecules, one of which is also coordinated to D575. The N-terminal region (440−447) of the A molecule is disordered. Remarkably, a loop within the active site of Metnase (residues 575−582) adopts two different conformations, and the relative positions of the residues within this loop are translated by one full residue in the metal bound vs the nonmetal bound molecules as shown in Figure 5. Thus, Arg 578 is located within the catalytic site hydrogen-bonded to Glu 484 in the nonmetal bound conformation (B molecule) and flipped out of the catalytic site in the metal bound conformation (A molecule). Similarly, the position of His 580 is different in each of the two different conformations in our structure. Interestingly, the two different conformations of His 580 are each hydrogen-bonded to Glu 484. His 580 in the A molecule with Ca^{2+} bound is also hydrogen bonded to the carbonyl oxygen of Asp 575, which potentially impacts the conformation of the loop.

Although further experimentation will be required to verify a functional role for the interactions of K445 and R578 observed in the active site, we propose that a potential role for Lys 445 and Arg 578 interactions in the catalytic site might be to neutralize the negative charge arising from the catalytic residues D483 and D575 and neighboring E484 in the absence of bound metal. The observation of two different conformations within the transposase

active site in the presence and absence of bound metal is potentially a unique feature of Metnase. As there are no crystal packing interactions in the vicinity of the active sites of either molecule as shown in Supporting Information, Figure S2, we suggest that the conformations observed in our crystal are not directly impacted by crystal packing interactions. Modeling of DNA bound to Metnase based on comparison with a Tn5 transposase−DNA complex structure (PDB id 1MUS) suggests that within our lattice the active site of the A molecule would be accessible to DNA while that in the B molecule is precluded by steric clash (See Supporting Information and Figure S2). We speculate that K445 and R578 may be effectively titrated from the active site of the A molecule through interactions with DNA allowing Ca^{2+} to bind to the active site in our 1.9 Å structure.

In our 2.55 Å crystal structure (hexagonal lattice form), crystallized without DNA, there is no metal bound in either molecule comprising the dimer, and each adopts the nonmetal bound conformation including the N-terminal beta hairpin observed in the B molecule of the monoclinic lattice (Figure S4, Supporting Information). The primary difference between the hexagonal and monoclinic crystal forms of Metnase that we obtained is the inclusion of duplex DNA in the latter; the protein used for each was identical and the crystallization conditions are nearly identical. In a third structural example of the catalytic domain of the Metnase transposase (PDB id 3F2K) including residues 446−671, each molecule of the dimer has Mg^{2+} bound in the active site (Supporting Information and Figure S4). This structure reveals only the metal-bound conformation of each molecule in the dimeric enzyme. Ca^{2+} and Mg^{2+} are bound similarly in Metnase (Figure S4, Supporting Information).

In a comparative analysis of the catalytic site of Metnase with those of Mos1 and Tn5 that retain DDD or DDE active site motifs (*30*, *31*), respectively, two residues D483 and N610 in Metnase, D156 and D284 in Mos1 (PDB id 2F2K), and D97 and E324 in the Tn5-DNA complex structure (PDB id 1MUS) adopt very similar conformations (Figure 6). Metal coordination by D483 and corresponding residues in Mos1 and Tn5 is also similar. The remaining catalytic triad residue, D575, equivalent structurally to D249 in Mos1 and D188 in Tn5 does not directly coordinate the Ca^{2+} in our structure but does coordinate Mg^{2+} in the 3F2K (PDB id) structure (Figure S4, Supporting Information), whereas N610 in the 3F2K structure adopts a slightly different conformation than that observed in our structure and in Mos1 and Tn5 structures for equivalent residues. However, N610 does not affect the ability of D483 or D575 to coordinate a metal ion in the catalytic site but would likely preclude binding of a second Mg^{2+} ion requiring coordination by N610 and D483. Of potential mechanistic importance is the conservation of the active site His residue, His 254 in Mos1 and His 580 in Metnase, along with its interaction with Glu 157 in Mos1 and Glu 484 in Metnase.

In further considering the interactions of K445 and R578 in the active site of Metnase, we note that Mos1 lacks an Arg residue at the position equivalent to 578 and has a Pro instead, P252, and therefore would not be expected to modulate metal binding in a manner similar to that proposed for Metnase. In conclusion, the overall architecture of the Metnase active site is similar to that observed in other active transposases retaining the positions and conformations of catalytic residues. Thus, the DNA repair functions acquired by Metnase likely use a similarly organized catalytic site.

Implications for the Evolution of Metnase. Metnase has a number of features distinct from other transposases, most

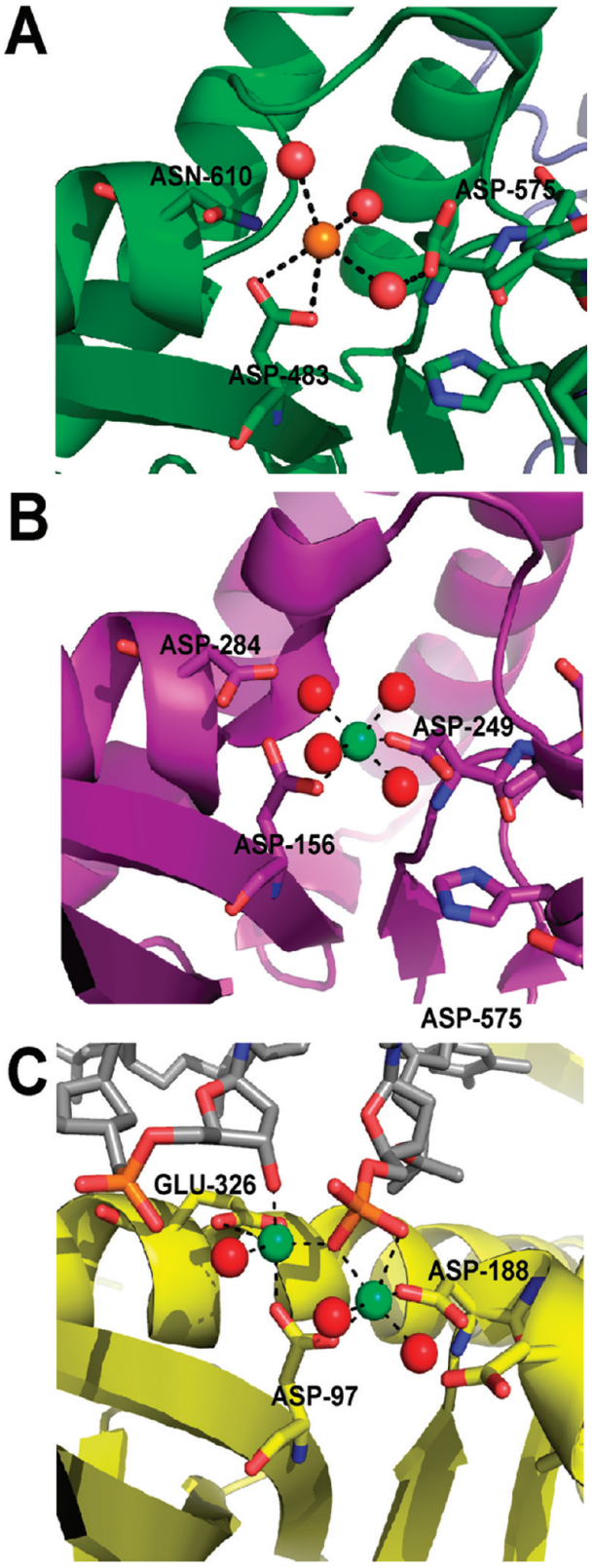

FIGURE 6: Comparison of Metnase active site with Mos1 and Tn5 transposases. (A) Subunit A, green cartoon, is shown with active site residues (D483, D575, N610) in sticks. Ca^{2+} is an orange sphere with coordinated waters in red and metal coordination represented as black dashed lines. (B) Mos1 active site is a purple cartoon with D156, D249, and D284 shown in sticks. Mg^{2+} is shown as green sphere with metal coordination in black dashed lines and waters in red. (C) Tn5 active site is in yellow with bound DNA in gray. D97, D188, E326 are shown as yellow sticks, Mg^{2+} shown in green, waters in red with metal coordination represented as black dashed lines.

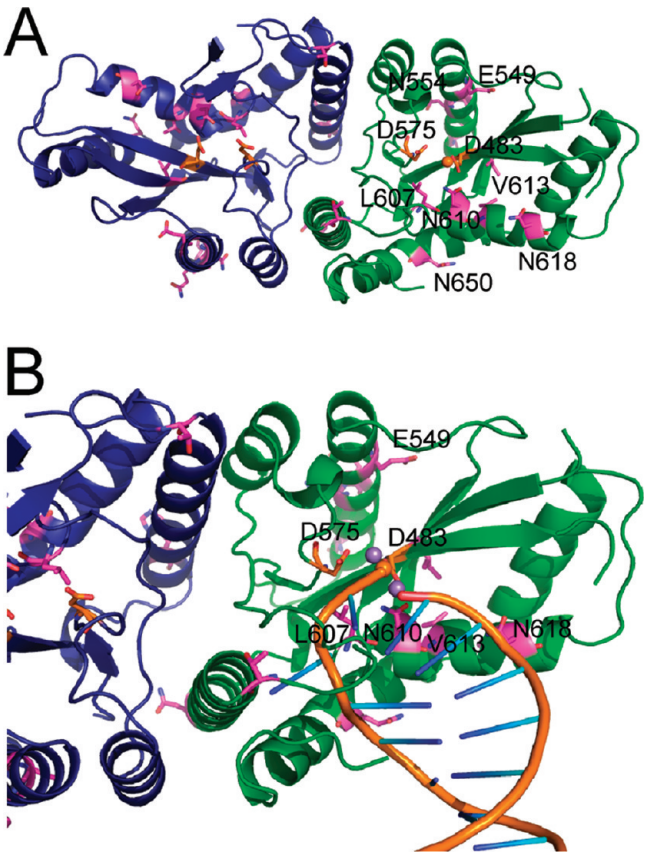

FIGURE 7: Predicted sequence differences between the ancestral and modern Metnase transposase domain. (A) Highlighted in magenta stick renderings are residues in Metnase that differ from the predicted ancestral sequence. The transposase dimer is represented as a ribbon rendering with the A subunit in green and B in navy. Active site residues D483 and D575 are shown as orange stick renderings and the active site Ca^{2+} associated with A as a yellow sphere. (B) Of particular interest are residues that fall within a predicted DNA-binding site for Metnase based on comparison with Tn5. The paired-end complex DNA found in the Tn5 complex is shown as a cartoon superimposed with the A molecule of Metnase. Tn5-DNA-associated metal ions are shown as purple spheres.

notably Mos1, its closest relative for which structural information is available. These features include the unique dimer interface, structural plasticity within the active site demonstrated in metal-bound vs apo active sites of the Metnase dimer, and the presence of a DDN vs a DDD or DDE motif within the active site. As Metnase arose from a gene fusion event placing an *Hsmar1* gene in frame and C-terminal to a SET domain, it is of interest to compare the Metnase transposase to its predicted ancestral sequence (*5*).

The transposase domain within Metnase has been reported to be under selective evolutionary pressure exhibiting only 2.4% divergence with the predicted ancestral sequence, which was generated using three genomic and 15 cDNA sequences (*5*). Residues including F460, required for dimerization, and K445, R578, and H580, involved in modulation of metal binding, are proposed to have important functional roles based on our structural analysis and are all predicted to be in the ancestral *Hsmar1* sequence. Thus, we conclude that the ancestral *Hsmar1* transposase formed a dimeric structure similar to Metnase and that it was able to modulate binding of metal to its active site as proposed for Metnase. The transposase domain of Metnase represents a unique member of a transposase family differing significantly from the related Mos1 and other transposases in requiring dimerization of the catalytic domain for function.

As compared with the predicted ancestral sequence, there are only 17 substitutions in Metnase, 4 of which N610, V613, N618, and N650 fall within α4-α5. As shown in Figure 7, these two consecutive helical elements are implicated in DNA-binding based on comparison to the Tn5-transposon DNA end complex structure (PDB id 1MUS). Included in these substitutions is D610N, with the potential consequence of altered ability to bind a second Mg^{2+} in the active site. By comparison with the Tn5 transposase, N610 is equivalent to E324, which is involved in binding metal along with D97 (equivalent to D483), while a second metal binding site is provided by D97 and D188 (equivalent to D575). The Metnase transposase may have lost the 3′ end processing activity in order to prevent transposition within its host. Alternatively, this function may not have been required for its new functions within the chimeric protein and therefore was not under selective evolutionary pressure. The predicted ancestral sequence also includes D618 located in α5, which like residue 610 is an Asn in Metnase. Thus, acquisition of new DNA repair functions by Metnase may in part result from substitutions within the DNA-binding site of the catalytic domain of the ancestral transposase.

ACKNOWLEDGMENT

Results shown in this report are derived from work performed at Argonne National Laboratory, Advanced Photon Source, at the Structural Biology Center beamline 19ID and GM/CA beamline 23ID. Argonne is operated by U Chicago Argonne, LLC, for the U.S. Department of Energy, Office of Biological and Environmental Research under contract DE-AC02-06CH11357. We thank Marianne Cuff, Stephan Ginell, and Andrzej Joachimiak at SBC and Craig Ogata at GM/CA for beamline support. We thank Sue Lee and Brian Beck for preparation of some of the reagents used in these studies, Samy Meroueh for assistance with homology modeling, Tom Hurley for helpful discussions, and Sarah Delaplane for technical assistance.

SUPPORTING INFORMATION AVAILABLE

A description of protein engineering efforts and site-directed mutagenesis associated with our studies. There is also a discussion of the differences between our Metnase structures and the PDB id 3F2K structure. Figures include protein−protein interactions in Metnase lattices compared with those of Mos1, crystal packing diagrams with modeled DNA, a comparison of the proposed Mos1 dimer with that of Metnase, and a comparison of our two structures of Metnase with that of the PDB id 3F2K structure. Hydrogen bonding interactions involved in the dimer interface (Table S1). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., and Zody, M. C.; et al. (2001) Initial sequencing and analysis of the human genome. *Nature 409*, 860–921.
2. Ivics, Z., Kaufman, C. D., Zayed, H., Miskey, C., and Walisko, O.; et al. (2004) The Sleeping Beauty transposable element: evolution, regulation and genetic applications. *Curr. Issues Mol. Biol. 6*, 43–55.
3. Miskey, C., Izsvak, Z., Kawakami, K., and Ivics, Z. (2005) DNA transposons in vertebrate functional genomics. *Cell. Mol. Life Sci. 62*, 629–641.
4. Brillet, B., Bigot, Y., and Auge-Gouillou, C. (2007) Assembly of the Tc1 and mariner transposition initiation complexes depends on the origins of their transposase DNA binding domains. *Genetica 130*, 105–120.

5. Robertson, H. M., and Zumpano, K. L. (1997) Molecular evolution of an ancient mariner transposon, Hsmar1, in the human genome. *Gene 205*, 203–217.
6. Cordaux, R., Udit, S., Batzer, M. A., and Feschotte, C. (2006) Birth of a chimeric primate gene by capture of the transposase gene from a mobile element. *Proc. Natl. Acad. Sci. U. S. A. 103*, 8101–8106.
7. Jordan, I. K. (2006) Evolutionary tinkering with transposable elements. *Proc. Natl. Acad. Sci. U. S. A. 103*, 7941–7942.
8. Lee, S. H., Oshige, M., Durant, S. T., Rasila, K. K., and Williamson, E. A.; et al. (2005) The SET domain protein Metnase mediates foreign DNA integration and links integration to nonhomologous end-joining repair. *Proc. Natl. Acad. Sci. U. S. A. 102*, 18075–18080.
9. Krogan, N. J., Kim, M., Tong, A., Golshani, A., and Cagney, G.; et al. (2003) Methylation of histone H3 by Set2 in Saccharomyces cerevisiae is linked to transcriptional elongation by RNA polymerase II. *Mol. Cell. Biol. 23*, 4207–4218.
10. Xiao, T., Hall, H., Kizer, K. O., Shibata, Y., and Hall, M. C.; et al. (2003) Phosphorylation of RNA polymerase II CTD regulates H3 methylation in yeast. *Genes Dev. 17*, 654–663.
11. Zhang, X., Yang, Z., Khan, S. I., Horton, J. R., and Tamaru, H.; et al. (2003) Structural basis for the product specificity of histone lysine methyltransferases. *Mol. Cell 12*, 177–185.
12. Hamamoto, R., Furukawa, Y., Morita, M., Iimura, Y., and Silva, F. P.; et al. (2004) SMYD3 encodes a histone methyltransferase involved in the proliferation of cancer cells. *Nat. Cell Biol. 6*, 731–740.
13. Beck, B. D., Park, S. J., Lee, Y. J., Roman, Y., and Hromas, R. A.; et al. (2008) Human Pso4 is a metnase (SETMAR)-binding partner that regulates metnase function in DNA repair. *J. Biol. Chem. 283*, 9023–9030.
14. Hromas, R., Wray, J., Lee, S. H., Martinez, L., and Farrington, J.; et al. (2008) The human set and transposase domain protein Metnase interacts with DNA Ligase IV and enhances the efficiency and accuracy of non-homologous end-joining. *DNA Repair* (*Amst*) *7*, 1927–1937.
15. Williamson, E. A., Farrington, J., Martinez, L., Ness, S., and O'Rourke, J. (2008) Expression levels of the human DNA repair protein metnase influence lentiviral genomic integration. *Biochimie 90*, 1422–1426.
16. Grey, M., Dusterhoft, A., Henriques, J. A., and Brendel, M. (1996) Allelism of PSO4 and PRP19 links pre-mRNA processing with recombination and error-prone DNA repair in Saccharomyces cerevisiae. *Nucleic Acids Res. 24*, 4009–4014.
17. Williamson, E. A., Rasila, K. K., Corwin, L. K., Wray, J., and Beck, B. D.; et al. (2008) The SET and transposase domain protein Metnase enhances chromosome decatenation: regulation by automethylation. *Nucleic Acids Res. 36*, 5822–5831.
18. Liu, D., Bischerour, J., Siddique, A., Buisine, N., and Bigot, Y.; et al. (2007) The human SETMAR protein preserves most of the activities of the ancestral Hsmar1 transposase. *Mol. Cell. Biol. 27*, 1125–1132.
19. Roman, Y., Oshige, M., Lee, Y. J., Goodwin, K., and Georgiadis, M. M.; et al. (2007) Biochemical characterization of a SET and transposase fusion protein, Metnase: its DNA binding and DNA cleavage activity. *Biochemistry 46*, 11369–11376.
20. Hendrickson, W. A., Horton, J. R., and LeMaster, D. M. (1990) Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. *EMBO J. 9*, 1665–1672.
21. Minor, W., Cymborowski, M., Otwinowski, Z., and Chruszcz, M. (2006) HKL-3000: the integration of data reduction and structure solution--from diffraction images to an initial model in minutes. *Acta Crystallogr. D Biol. Crystallogr. 62*, 859–866.
22. Schneider, T. R., and Sheldrick, G. M. (2002) Substructure solution with SHELXD. *Acta Crystallogr. D Biol. Crystallogr. 58*, 1772–1779.
23. Uson, I., Stevenson, C. E., Lawson, D. M., and Sheldrick, G. M. (2007) Structure determination of the O-methyltransferase NovP using the "free lunch algorithm" as implemented in SHELXE. *Acta Crystallogr. D Biol. Crystallogr. 63*, 1069–1074.
24. CCP4 (1994) The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. *50*, 760−763.
25. Terwilliger, T. (2004) SOLVE and RESOLVE: automated structure solution, density modification and model building. *J. Synchrotron Radiat. 11*, 49–52.
26. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr. 60*, 2126–2132.
27. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A 47* (Pt 2), 110–119.
28. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., and Gros, P.; et al. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr. 54*, 905–921.
29. Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., and McCoy, A. J.; et al. (2002) PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D Biol. Crystallogr. 58*, 1948–1954.
30. Richardson, J. M., Dawson, A., O'Hagan, N., Taylor, P., and Finnegan, D. J.; et al. (2006) Mechanism of Mos1 transposition: insights from structural analysis. *EMBO J. 25*, 1324–1334.
31. Davies, D. R., Goryshin, I. Y., Reznikoff, W. S., and Rayment, I. (2000) Three-dimensional structure of the Tn5 synaptic complex transposition intermediate. *Science 289*, 77–85.