Published online October 5, 2004

Nucleic Acids Research, 2004, Vol. 32, No. 17 ***5291–5302***
doi:10.1093/nar/gkh859

Mutational analysis of the damage-recognition and catalytic mechanism of human SMUG1 DNA glycosylase

Mayumi Matsubara, Tamon Tanaka, Hiroaki Terato, Eiji Ohmae, Shunsuke Izumi, Katsuo Katayanagi and Hiroshi Ide*

Department of Mathematical and Life Sciences, Graduate School of Science, Hiroshima University, Higashi-Hiroshima 739-8526, Japan

Received June 14, 2004; Revised August 20, 2004; Accepted September 12, 2004

ABSTRACT

Single-strand selective monofunctional uracil-DNA glycosylase (SMUG1), previously thought to be a backup enzyme for uracil-DNA glycosylase, has recently been shown to excise 5-hydroxyuracil (hoU), 5-hydroxymethyluracil (hmU) and 5-formyluracil (fU) bearing an oxidized group at ring C5 as well as an uracil. In the present study, we used site-directed mutagenesis to construct a series of mutants of human SMUG1 (hSMUG1), and tested their activity for uracil, hoU, hmU, fU and other bases to elucidate the catalytic and damage-recognition mechanism of hSMUG1. The functional analysis of the mutants, together with the homology modeling of the hSMUG1 structure based on that determined recently for *Xenopus laevis* SMUG1, revealed the crucial residues for the rupture of the N-glycosidic bond (Asn85 and His239), discrimination of pyrimidine rings through π–π stacking to the base (Phe98) and specific hydrogen bonds to the Watson–Crick face of the base (Asn163) and exquisite recognition of the C5 substituent through water-bridged (uracil) or direct (hoU, hmU and fU) hydrogen bonds (Gly87–Met91). Integration of the present results and the structural data elucidates how hSMUG1 accepts uracil, hoU, hmU and fU as substrates, but not other oxidized pyrimidines such as 5-hydroxycytosine, 5-formylcytosine and thymine glycol, and intact pyrimidines such as thymine and cytosine.

INTRODUCTION

Reactive oxygen species generated by normal aerobic metabolism, ionizing radiation and exogenous redox-active chemicals induce strand breaks and a plethora of oxidative base lesions (1–3). Many of the oxidative base lesions induce relatively minor and localized distortions in the DNA helix, but they arrest DNA synthesis or direct incorporation of non-cognate nucleotides during DNA replication, thereby resulting in cell death or mutations (4). The major cellular mechanism for coping with such genotoxic effects is the base excision repair pathway (5–7). In addition, the nucleotide incision repair pathway has been recently proposed as an alternative mechanism for repairing oxidatively damaged DNA (8–11). The base excision repair pathway is conserved from bacteria to humans, and is initiated by DNA glycosylases. In the first step of this pathway, DNA glycosylases remove the structurally altered base from the DNA backbone by the associated N-glycosylase activity. The resulting apurinic/apyrimidinic (AP) site is then incised by the AP lyase activity of DNA glycosylases via the β-elimination (or β,δ-elimination) of the 3′ phosphodiester bond. Further processing of this repair intermediate by the sequential action of AP endonuclease (Endo), DNA polymerase and ligase restores DNA.

It has been long known that in *Escherichia coli* oxidative pyrimidine lesions such as thymine glycols (Tg) are excised by Endo III and Endo VIII, whereas oxidative purine lesions such as 7,8-dihydro-8-oxoguanine and formamidopyrimidine are excised by formamidopyrimidine-DNA glycosylase (12,13). To date, the mammalian homologs of Endo III (NTH1) and Endo VIII (NEIL1, NEIL2 and NEIL3) and a functional homolog of formamidopyrimidine-DNA glycosylase (OGG1) have been identified, and their functions have been studied extensively (6,7,14,15). However, these *E.coli* and mammalian enzymes exhibit at most marginal activity toward a subset of oxidative thymine lesions including 5-hydroxymethyluracil (hmU) and 5-formyluracil (fU) (16–18). Mismatched fU is a substrate of nucleotide excision repair (19), but its physiological significance remains to be established. hmU resulting from the oxidation of thymine neither arrests DNA synthesis nor directs nucleotide misincorporation (20,21), and hence is innocuous. Conversely, the generation of hmU by the deamination of 5-methylcytosine in CpG sites and subsequent oxidation of the methyl group (or vice versa) (22) induce mutations since a non-cognate nucleotide, dAMP, is inserted opposite hmU during DNA replication. Despite suffering no damage in the Watson–Crick face, fU promotes misincorporation of dGMP and hence is mutagenic (23–25). The underlying mechanism involves the mispairing of the ionized form of template fU and incoming dGTP in an anti-wobble structure (25,26). Thus, hmU and fU need to be removed from DNA to avoid mutations.

*To whom correspondence should be addressed. Tel/Fax: +81 82 424 7457; Email: ideh@hiroshima-u.ac.jp

The hmU-excising activity in mammalian cells was demonstrated 20 years ago (27,28) and has recently been attributed to single-strand selective monofunctional uracil-DNA glycosylase (SMUG1) (29). SMUG1 was originally identified as uracil-DNA glycosylase (UDG) exhibiting a preference for single-stranded DNA (30), but later it was shown to be active for double-stranded DNA. Analyses of the residual activity for uracil (U) in UDG-knockout mice suggest a backup function of SMUG1 for UDG (31,32). The fU-excising activity in *E.coli* and mammals has also been identified recently as AlkA (3-methyladenine-DNA glycosylase II) and SMUG1, respectively (16,17,24,33). Interestingly, the mammalian functional homolog of AlkA (MPG) exhibits no activity for fU. Detailed investigations of the substrate specificity of human and rat SMUG1 have revealed that together with U, SMUG1 excises 5-hydroxyuracil (hoU), hmU and fU, bearing an oxidized group at ring C5 (17,33). The damage-excising activity for these three lesions in HeLa cell extracts was neutralized almost completely by the antiserum of SMUG1. Thus, SMUG1 is a new member of DNA glycosylases involved in the repair of oxidative base lesions (7). Interestingly, SMUG1 does not excise cytosine lesions bearing an oxidized group at ring C5 (5-hydroxycytosine (hoC) and 5-formylcytosine), C5–C6 saturated thymine (Tg and 5,6-dihdrothymine), and intact thymine and cytosine, indicating an exquisite damage-recognition mechanism (33).

The crystal structures of *Xenopus laevis* SMUG1 (xSMUG1) complexed with DNA have been recently reported (34). xSMUG1 was found to be bound to the end of the oligonucleotide substrate, and hence the complex is not a productive one. However, the structure indicates that xSMUG1 is a member of the UDG structural family and has chimeric active site residues that are topologically equivalent to those of UNG and MUG/TDG enzymes. In addition, the xSMUG1 crystals soaked in free U and hmU bases suggest hydrogen bonds that may be pertinent to the recognition of the C5 substituent for U and hmU. However, no functional analysis of amino acid residues involved in catalysis and damage recognition has been performed. In this study, we prepared a series of mutant human SMUG1 (hSMUG1) proteins by site-directed mutagenesis and analyzed their activity toward various oxidative base lesions to uncover the catalytic and exquisite damage-recognition mechanism. Here, we report the functionally important residues of hSMUG1 for catalysis, discrimination of pyrimidine bases and recognition of the oxidized group at ring C5. These data are further discussed in light of the homology modeling of the hSMUG1 protein.

MATERIALS AND METHODS

Oligonucleotide substrates

The oligonucleotides containing U (19U), hoU (25HOU), hoC (25HOC), hmU (25HMU), fU (25FU) and Tg (19TG) were prepared as previously reported (17,18) and are listed in Table 1. The oligonucleotides containing the base lesions were 5′ end labeled with [γ-^{32}P]ATP (110 TBq/mmol, Amersham Biosciences) and T4 polynucleotide kinase (New England BioLabs), and purified by a Sep-Pak cartridge (Waters). They were annealed to appropriate complementary strands and used as double-stranded substrates, which are expressed as the combination of an oligonucleotide containing the lesion and the base opposite it (e.g. 19HOU/G) throughout this paper.

Site-directed mutagenesis

The construction of the expression plasmid pGEX-hSMUG1 carrying the wild-type hSMUG1-coding sequence at the BamHI/XhoI site in pGEX4T-1 (Amersham Biosciences) has been reported elsewhere (33). *In vitro* mutagenesis of hSMUG1 was performed using pGEX-hSMUG1 as a template, mutagenesis oligonucleotide primers (Supplementary Material Table S1) and the QuikChange site-directed mutagenesis kit (Stratagene) according to the manufacturer's instructions. The constructs were verified by sequencing the entire hSMUG1-coding region of isolated plasmids.

Expression and purification of wild-type and mutant hSMUG1 proteins

The wild-type and mutant hSMUG1 proteins were expressed as N-terminal glutathione *S*-transferase (GST) fusion proteins in *E.coli* BL21(DE3) harboring the wild-type or mutant plasmids as reported previously (33). The GST-tagged hSMUG1 proteins were affinity purified on a glutathione Sepharose 4B column (Amersham Biosciences) and digested with thrombin (Amersham Biosciences). The hSMUG1 proteins without a GST tag were purified on an SP Sepharose FF column (Amersham Biosciences) and stored in 50 mM Tris–HCl (pH 8.0), 150 mM NaCl, 0.2 mM DTT, 0.005% Triton X-100 and 50% glycerol at −20°C. The wild-type and mutant hSMUG1 proteins contained two extra N-terminal amino acids (Gly-Ser) from the GST linker. Purified hSMUG1 proteins were physically homogeneous and showed a single band of the expected size (30 kDa) in SDS–PAGE analysis (Figure S1). The protein concentration was determined with a protein assay kit (Bio-Rad) using BSA as a standard.

Circular dichroism spectroscopy

The hSMUG1 proteins in stock buffer were extensively dialyzed against 25 mM Tris–HCl (pH 7.5), 0.2 mM EDTA, 50 mM NaCl, 0.1 mM DTT and 10% glycerol at 4°C to change the buffer composition, and the protein concentration was adjusted to A_{280} = 0.3–0.4 using the same buffer. Far-ultraviolet (UV) circular dichroism (CD) measurements were performed on a Jasco J-720W spectropolarimeter (calibrated with a 0.003% aqueous solution of D-camphor-10-sulfonate) at 20°C using a quartz cuvette with a 1 mm path length as reported previously (35,36). Each sample was scanned 16 times, and noise reduction was applied. The spectrum of the baseline buffer was subtracted from sample spectra to obtain molar ellipticities.

Enzyme activity assays

The 5′-labeled oligonucleotide substrates (1 nM, Table 1) were incubated with the wild-type or mutant hSMUG1 at 37°C in 25 mM Tris–HCl (pH 7.5), 0.2 mM EDTA, 2 mM DTT, 0.1 mg/ml BSA and NaCl (total volume 10 µl). The concentration of NaCl was varied to ensure optimal activity: 20 and 50 mM for the substrates containing damage:A and damage:G pairs, respectively (33). The amount of protein (0.1–200 ng) and the incubation time (3–90 min) were varied depending on the enzyme activity such that the yield of products was linearly

Table 1. Oligonucleotide substrates used in this study[a]

Substrate	Damage (X)	Sequence (5′→3′)	Paired base
19U	U	ACAGACGCCAXCAACCAGG	A, G
25HOU	hoU	GAAACACTACTATCAXGGAAGAGAG	A, G
25HOC	hoC	GAAACACTACTATCAXGGAAGAGAG	A, G, C, T
25HMU	hmU	CATCGATAGCATCCGXCACAGGCAG	A, G
25FU	fU	CATCGATAGCATCCGXCACAGGCAG	A, G
25C	C	GAAACACTACTATCAXGGAAGAGAG	A, G, C, T
19TG	Tg	ACAGACGCCAXCAACCCAGG	A

[a]Only the strands containing base lesions are shown. The complementary strands had the same length as the lesion-containing strand, and contained one of the bases indicated under paired base.

related to the amount of protein and incubation time (for detailed conditions, see tables and Figure S2). The reaction was terminated by the addition of NaOH (final concentration 0.1 M). The sample was heated at 70°C for 5 min to cleave AP sites and then neutralized with 1 M acetic acid. The products were separated by 16% denaturing PAGE, and the radioactivity in the gel was analyzed on a Fuji BAS2000 phosphorimaging analyzer. The kinetic parameters of the enzyme reactions were determined in a similar manner except that the substrate concentration was varied up to 200 nM depending on the enzyme. The appropriate range of the substrate concentration for the parameter analysis was determined from the preliminary experiments. The parameters were evaluated from the plots of substrate concentration versus initial velocity using a hyperbolic curve-fitting program.

Gel electrophoretic mobility shift assays

The oligonucleotide substrates containing U (19U/A and 19U/G) or hmU (hmU/A and hmU/G) (all 1 nM) were incubated with 100 ng of wild-type and mutant hSMUG1 in 25 mM Tris–HCl (pH 7.5), 0.2 mM EDTA, 2 mM DTT, 0.1 mg/ml BSA and 50 mM NaCl (total volume 5 μl) on ice for 30 min. After incubation, the sample was supplemented with 50% (w/v) sucrose. The sample was loaded onto a 6% non-denaturing polyacrylamide gel and electrophoresed in 40 mM Tris-acetate (pH 8.0) and 1 mM EDTA at an ice-cold temperature for 2 h. Free and bound DNA species were quantified on a Fuji BAS2000.

Molecular modeling

The three-dimensional (3D) structures of wild-type and mutant hSMUG1 were modeled from the crystal structures of xSMUG1 (Protein Data Bank accession codes: 1OE4, 1OE5 and 1OE6) (34) at the SWISS-MODEL Internet server (http://www.expasy.org/swissmod/) (37,38). The structures were displayed and analyzed by the Swiss-Pdb Viewer (http://www.expasy.org/spdbv/mainpage.html) (39).

RESULTS

Design of hSMUG1 mutants

hSMUG1 is a member of the UDG superfamily, which has been classified into several families: UNG (orthologs of *E.coli* Ung, family 1), MUG/TDG (orthologs of *E.coli* Mug, family 2), SMUG1 (family 3) and other UDGs in thermophiles and prokaryotes (family 4) (40,41). The superposition of the 3D structures of UNG and MUG/TDG enzymes shows that the active site of UDG enzymes comprises two moderately conserved polypeptides (motifs 1 and 2 in Figure 1A). In UNG enzymes (42–45), Asp in motif 1 and His in motif 2 are conserved, and they are indispensable for the unusually efficient hydrolysis of the N-glycosidic bond. The hydrogen bond from Asn between motifs 1 and 2 to the O4 carbonyl and N3H imino of U discriminates against the N4 amino and N3 *tert*-amino of cytosine. The π–π stacking interaction of Phe at the C-terminus of motif 1 stabilizes flipped-out U. Finally, Tyr in motif 1 confers a steric barrier against the methyl group of thymine, hence preventing productive binding. In MUG/TDG enzymes (46,47), Tyr in motif 1 is replaced by Gly, which confers little steric barrier, hence accepting thymine as a substrate. Asp and His involved in catalysis are also replaced by two Asn residues (MUG), or Asn and Met (TDG). In the crystal structure of *E.coli* Mug, a water molecule is positioned near deoxyribose-C1′ via hydrogen bonds to the side-chain amide and main-chain carbonyl of Asn in motif 1. Comparison of these features of UNG and MUG/TDG with those of hSMUG1 indicates that hSMUG1 has a hybrid active site structure of UNG and MUG/TDG (30): Phe98 (motif 1), Asn163 (between motifs 1 and 2) and His239 (motif 2) are conserved between hSMUG1 and UNG, whereas Asn85 and Gly87 (both in motif 1) are conserved between hSMUG1 and MUG/TDG (Figure 1A). In light of the putative functional conservation of the critical residues among the UDG superfamily, it is likely that Asn85 and His239 of hSMUG1 catalyze the rupture of the N-glycosidic bond, Asn163 and Phe98 discriminate pyrimidine bases and Gly87 or proximal residues recognize the C5 substituent (Figure 1B). Thus, the hSMUG1 proteins carrying mutations at these residues were prepared by site-directed mutagenesis in the present study (indicated by arrows in Figure 1A).

CD spectra of hSMUG1 proteins

The secondary structures of wild-type and mutant hSMUG1 proteins were analyzed by far-UV CD spectroscopy. The wild-type hSMUG1 exhibited a large negative ellipticity band centered at 208 nm that was indicative of a significant content of α-helical structure, together with a shoulder around 220 nm (Figure 2A). The CD spectrum of the N85A mutant essentially superposed that of wild-type hSMUG1 (Figure 2A); this was also the case for N163D, G87A/S/F and F89G/A/S mutants (data not shown). The H239L and H239N mutants exhibited spectra similar to that of the wild type, but the negative ellipticity of the shoulder around 220 nm was slightly decreased (Figure 2B). A similar decrease was observed for F98H/L, G90A and M91A mutants (data not shown). The CD spectra were analyzed for secondary structural elements by the SELCON3 program with 29 basal proteins (48). The wild-type hSMUG1 and mutants with superposable spectra (N85A etc.) contained 50% α-helix, 10% β-sheet, 18% β-turn and 22% random structures. The mutants that exhibited slightly different spectra (H239L etc.) contained 49–56% α-helix, 9–11% β-sheet, 17% β-turn and 19–25% random structures. Considering the estimation errors of the analysis program (5–10%), the present CD data suggest that the mutant

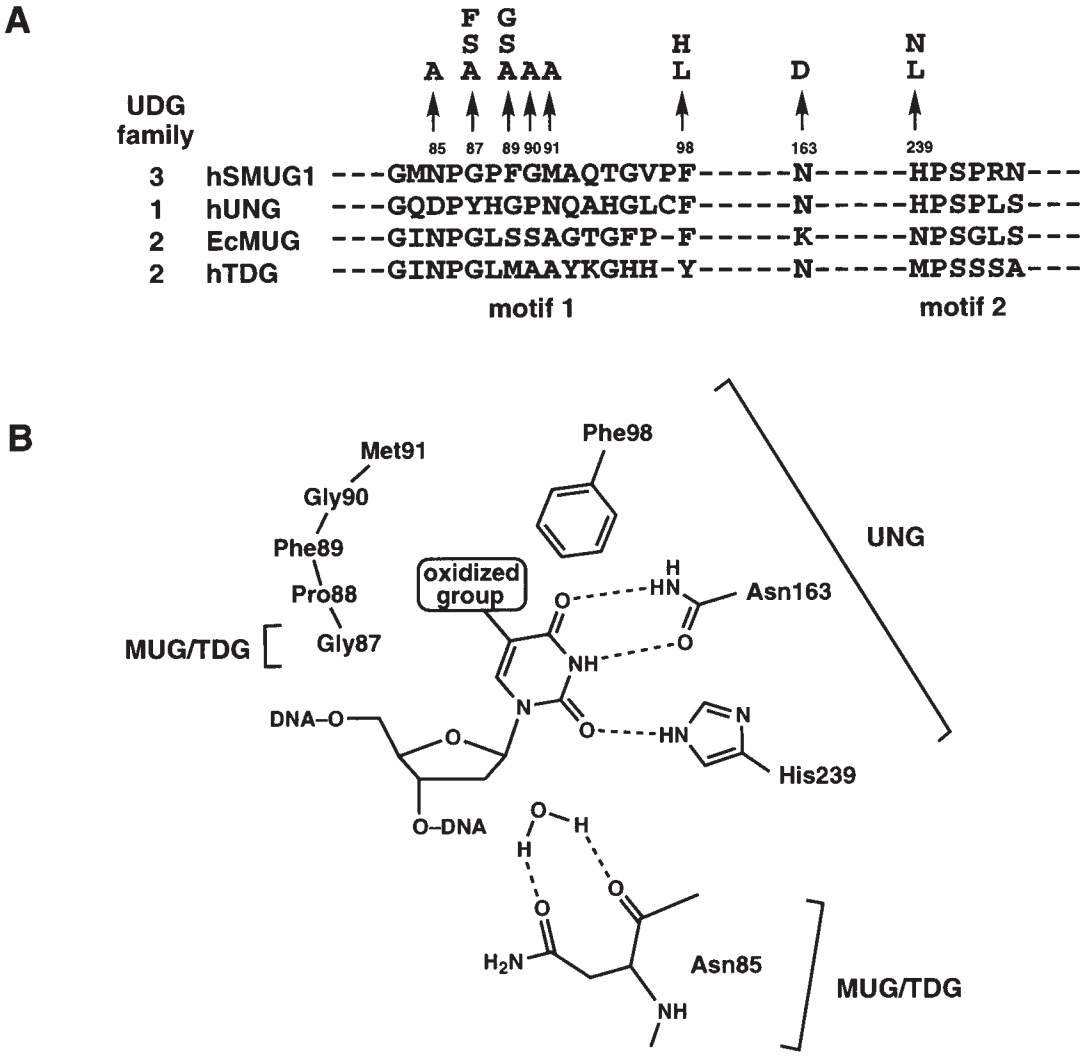

Figure 1. Conserved motifs of the UDG superfamily and expected interactions in the active site pocket of hSMUG1. (**A**) Amino acid sequences of motifs 1 and 2 and functionally important Asn that comprise the active site pocket. The position from the N-terminus of mutated residues and substituted amino acids in this study are shown above the sequence of hSMUG1. EcMUG indicates *E.coli* Mug. (**B**) Interactions of the functionally important residues in the active site pocket of hSMUG1, which is a hybrid of UNG and MUG/TDG enzymes (indicated by brackets). The hydrogen bonds of Asn85, Asn163 and His239 to the substrate base and a water molecule near deoxyribose-C1′ are indicated by broken lines. Phe98 stacks on the base through the π–π stacking interaction, and the oxidized group at ring C5 likely interacts with vicinal Gly87–Met91.

hSMUG1 proteins undergo no gross alterations of secondary structures and backbone conformations as compared with wild-type hSMUG1. We also found that there were no systematic correlations between the observed spectral changes and the activity of the mutants.

Activity of hSMUG1 mutants for U, hoU, hmU and fU

Duplex oligonucleotide substrates containing U:G, hoU:G, hmU:A (or hmU:G) and fU:A (or fU:G) pairs were incubated with wild-type and mutant hSMUG1, and nicked products were quantified by PAGE analysis after alkaline treatment. The amount of enzyme and the incubation time were varied depending on the activity of enzyme (for more details see Table 2). The damage-excising activity per nanogram of protein was determined from the initial slope of the plot of the product yield against the incubation time and normalized to that of wild-type hSMUG1 for each substrate. Table 2 summarizes these data and demonstrates the functional importance of mutated amino acids in a semi-quantitative manner (for quantitative analysis, see the next section).

The mutation of the residues involved in the hydrolysis of the N-glycosidic bond (N85A and H239L) markedly impaired the damage-excising activity of hSMUG1 (Table 2). The extent of reduction was similar for U:G, hoU:G, hmU:A and fU:A. The activities exhibited by N85A and H239L mutants were two and four orders of magnitude lower than that of wild-type hSMUG1, indicating that the role of His239 in the rupture of the N-glycosidic bond is more crucial than that of Asn85. The *E.coli* Mug has Asn at the position of His239 (Figure 1A). However, the H239N mutant carrying this mutation had significantly lower activity (by three orders of magnitude) than the wild type, though the H239N mutant showed a slightly higher activity than the H239L

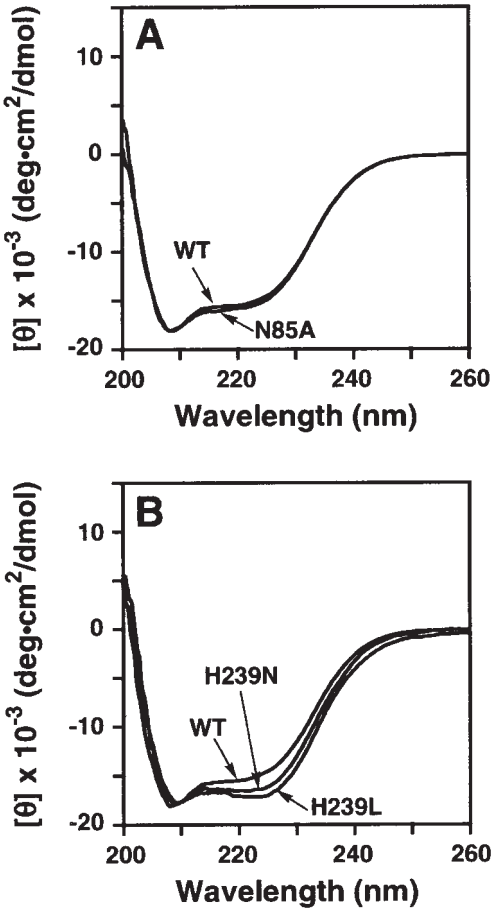

Figure 2. CD spectra of the wild-type and mutant hSMUG1 proteins. (**A**) CD spectrum of the N85A mutant superposed on that of wild-type hSMUG1 (WT). (**B**) CD spectra of the H239L and H239N mutants differed slightly from that of that of WT. The spectra were measured as described under experimental procedures.

Table 2. Relative activity of hSMUG1 mutants for U, hoU, hmU and fU

Mutant and relevant function	Activity relative to wild-type hSMUG1 (%)[a]			
	U:G	hoU:G	hmU:A	fU:A
Wild type	100	100	100	100
Hydrolysis of N-glycosidic bond				
N85A	0.60	0.70	1.9	1.4
H239L	0.0053	0.0013	0.0057	0.0045
H239N	0.081	0.0099	0.086	0.055
Discrimination of pyrimidine rings				
F98L	0.30	3.6	1.6	0.50
F98H	94	90	100	85
N163D	1.6	2.6	3.1	3.9
Recognition of C5 substituents				
G87A	7.3	5.3	1.1[b]	8.9[b]

[a]The indicated substrates (1 nM) were incubated with wild-type or mutant hSMUG1, and nicked products were quantified by PAGE analysis after alkaline treatment. The amounts of enzyme and the maximum incubation time were as follows: wild type (0.1 ng, 9 min), N85A (2 ng, 30 min), H239L (100 or 200 ng, 90 min), H239N (10 or 100 ng, 90 min), F98L (1 or 10 ng, 30 min), F98H (0.1 ng, 9 min), N163D (1 or 2 ng, 30 min) and G87A (1 ng, 9 min). The excision activity per nanogram of protein was determined from the initial slope of the plot of the amount of products against incubation time, and normalized to that of wild-type hSMUG1 for each substrate.
[b]Data for hmU:G and fU:G pairs.

mutant (~10-fold). This difference might be due to the partial restoration of the hydrogen bond between the side-chain amide of substituted Asn239 to the O2 carbonyl of the base.

The substitution of Leu for Phe (F98L) involved in the discrimination of pyrimidine ring through the π–π stacking interaction (Figure 1B) resulted in damage-excising ability for U:G, hoU:G, hmU:A and fU:A being two orders of magnitude lower than that of the wild type (Table 2). However, the F98H mutant carrying an aromatic side chain (imidazole) retained activity comparable to wild-type hSMUG1. The contrasting results with the F98L and F98H mutants are indicative of the crucial role of π–π stacking interaction in base discrimination, and further elucidates a mechanism explaining how hSMUG1 discriminates against base lesions that lose aromaticity, such as Tg and 5,6-dihydrothymine (33). It is noted that neither F98L nor F98H mutants with distinct base-stacking capacities excised Tg even when 19TG/A was incubated with copious enzymes (data not shown). The substitution of Asp for Asn (N163D) that also participates in base discrimination (i.e. U versus C) impaired the damage-excising activity for U:G, hoU:G, hmU:A and fU:A. In UNG enzymes, the hydrogen-bonding specificity of the side-chain amide of the conserved Asn (equivalent to Asn163 of hSMUG1, Figure 1B) is compatible with the Watson–Crick face of U but not with cytosine, and hence discriminates against cytosine. Consistent with this interaction, the Asn-to-Asp mutation confers a weak cytosine-excising activity on human UNG (49). Thus, we examined whether this is also the case for hSMUG1. The N163D mutant of hSMUG1 exhibited no appreciable cytosine-excising activity for a matched C:G pair and mismatched C:A, C:T and C:C pairs even when the substrates (25C/G etc.) were incubated with copious N163D mutant (100 ng) (Figure 3A). Similar results were obtained with the wild-type and N85A mutant of hSMUG1 used as controls (Figure 3A). In contrast, the N163D mutant but not the wild-type or the N85A mutant showed a clear hoC-excising activity, particularly for mismatched hoC (hoC:A, hoC:T and hoC:C) (Figure 3B). The amount of N163D mutant required to detect the activity (i.e. 100 ng) indicates that the hoC-excising activity of the N163D mutant is weak but still significant. A possible mechanism underlying the distinct activities of the N163D mutant for C and hoC is mentioned in Discussion.

The UNG enzymes have Tyr at the position of Gly87 of hSMUG1 that packs against the C5 substituent, thereby rejecting thymine but accept U, whereas MUG/TDG enzymes have Gly at this position and accept both thymine and U. Although the SMUG1 family has Gly (Gly87) as with MUG/TDG enzymes, it accepts U derivatives bearing a hydrophilic group at ring C5 (hoU, hmU and fU) but not a hydrophobic group (thymine). On the basis of these data, we have previously suggested hydrophilic interactions between the C5 substituent and the active site residue in hSMUG1 (33). In the present study, the residues lying in the vicinity of the C5 substituent (GPFGM in Figure 1B) were changed systematically, and the activities of the mutants were assayed with oligonucleotide substrates containing U:G, hoU:G, hmU:G and fU:G (all in a mismatched base pair). The activities normalized to that of wild-type hSMUG1 are shown in Figure 4. Note that both wild-type and mutant hSMUG1 exhibited a slightly higher activity for mismatched substrates (i.e. opposite base = G) than for matched substrates (i.e. opposite base = A),

but both types of substrates gave essentially the same results with respect to the activity relative to wild-type hSMUG1. The substitution of Gly87 by Ala or Ser (G87A and G87S) impaired the damage-excising activity for U bearing a small hydrogen atom at ring C5 as well as hoU, hmU and fU, indicating a discriminating interaction between the small side chains of Ala and Ser with the bases. hmU with the most bulky C5 substituent was the poorest substrate of the G87A and G87S mutants. The discrimination was further evident with the G87F mutant with a bulky side chain. The mutation of Phe89 (F89G, F89A and F89S) and Met91 (M91A) did not severely hamper the damage-excising activity for U, hoU, hmU and fU, suggesting that the side chains of Phe89 and Met91 have no contact with the C5 substituent. This proves to be the case by the structural analysis of hSMUG1 (see Figure 7A and D). Interestingly, the G90A mutant lost the U-excising activity almost completely and retained a weak but significant activity for hoU, hmU and fU, suggesting distinct interaction modes of hSMUG1 toward U and other bases (hoU, hmU and fU).

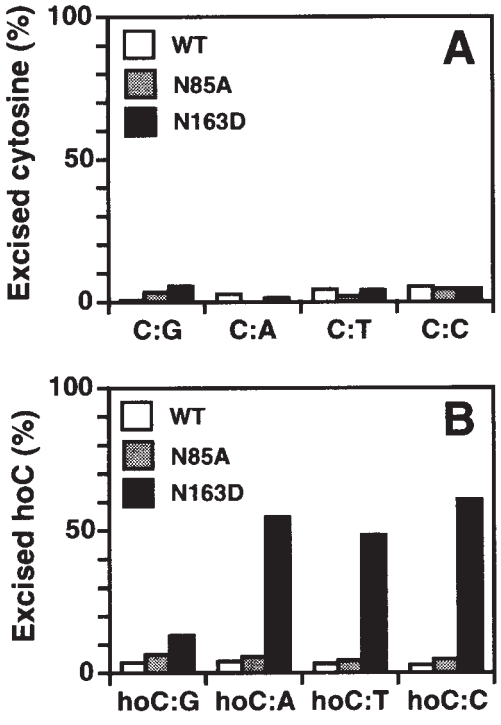

Figure 3. Cytosine- and hoC-excising activity of the N163D mutant of hSMUG1. The N163D mutant and control enzymes (wild-type and N85A mutant of hSMUG1) (all 100 ng) were incubated with 25C/N containing cytosine or 25HOC/N containing hoC (N = A, G, C, T) (1 nM) at 37°C for 30 min. AP sites resulting from the excision of cytosine or hoC were cleaved by alkaline treatment, and nicked products were quantitated by denaturing PAGE. The percentages of excised (**A**) cytosine and (**B**) hoC from the substrate are plotted: white bar, wild-type (WT); dotted bar, N85A; black bar, N163D.

Kinetic parameters of hSMUG1 mutants

To obtain further insight into the catalytic and damage-recognition mechanism, the enzymatic parameters of hSMUG1 mutants were determined using 19U/G as a substrate (Table 3). The relative catalytic efficiency of hSMUG1 mutants (k_{cat}/K_m in Table 3) correlates well with the activity data for U in Table 2. The N85A, H239L and H239N mutants carrying mutations in catalytic residues exhibited significant reductions in U-excising activity and showed the simultaneous decrease in k_{cat} and affinity for the substrate (i.e. increase in K_m). As expected from the nature of mutations, the decrease in k_{cat} was greater than that in affinity in these mutants. However, H239L and H239N mutants exhibited not only a large decrease in k_{cat} (380- and 120-fold relative to the wild type, respectively) but also a fairly large decrease in affinity (34- and 42-fold, respectively) as compared with the N85A mutant (3.4-fold). When 25HMU/G containing an hmU:G pair was used as a substrate for the catalytic mutants (N85A and H269L), similar parameter changes were observed (Table S2). The order of affinity for the substrates containing U and hmU determined by gel electrophoretic mobility shift assays (EMSAs) (wild-type $\geqslant$ N85A > N163D > H239L, Figures 5 and S3) further confirmed a large affinity decrease in the H239L mutant compared with the N85A mutant. Thus, the hydrogen bond of His239 to the O2 carbonyl of the base (Figure 1B) has two crucial roles: (i) adjustment of the base

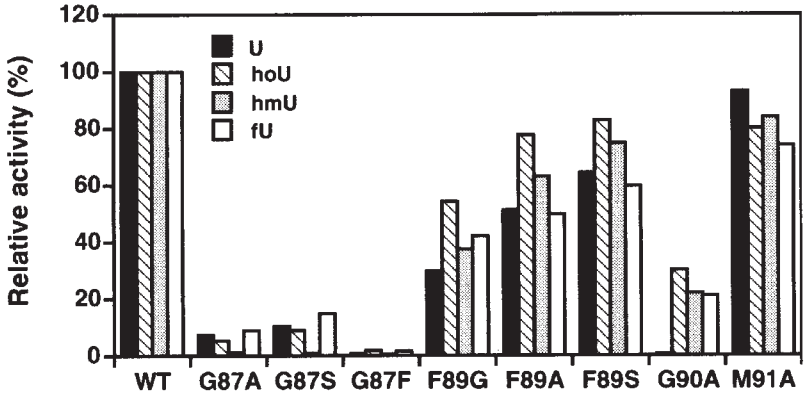

Figure 4. Activity of hSMUG1 mutants with substitutions pertinent to the recognition of the C5 substituent of bases. The hSMUG1 mutants with amino acid substitutions at Gly87 (G87A, G87S and G87F), Phe89 (F89G, F89A and F89S), Gly90 (G90A) and Met91 (M91A) were incubated with 25U/G, 25HOU/G, 25HMU/G and 25FU/G (all 1 nM) containing U, hoU, hmU and fU, respectively, at 37°C for up to 9 min. The amount of enzymes used for assays were 1 ng (G87 mutants), 0.1 ng (F89 mutants and M91A) and 0.5–10 ng (G90A). The resulting AP sites were cleaved by alkaline treatment and nicked products were quantitated by denaturing PAGE. The amount of excised bases per minute per nanogram protein for individual substrates is normalized to that for wild-type hSMUG1 (WT). The substrates are uracil (black bar), hoU (hatched bar), hmU (dotted bar) and fU (open bar).

Table 3. Kinetic parameters of hSMUG1 mutants for excision of U:G[a]

Mutant and relevant function	k_{cat} (min^{-1})	K_m (nM)	k_{cat}/K_m ($min^{-1}nM^{-1}$)	Relative k_{cat}/K_m (%)
Wild type	0.84	2.2	0.38	100
Hydrolysis of N-glycosidic bond				
N85A	0.032	7.4	0.0043	1.1
H239L	0.0022	74	0.000029	0.0076
H239N	0.0070	92	0.000076	0.020
Discrimination of pyrimidine rings				
F98L	0.0053	24	0.00022	0.058
F98H	10	30	0.33	87
N163D	0.13	15	0.0088	2.3
Recognition of C5 substituents				
G87A	2.4	41	0.058	15
F89A	8.2	19	0.43	110
G90A	0.039	20	0.0020	0.52
M91A	6.6	29	0.23	60

[a]Wild-type and mutant hSMUG1s were incubated with varying concentrations of 19U/G at 37°C. After incubation, the sample was treated with 0.1 M NaOH, and nicked products were quantified by PAGE analysis. The concentration of the substrate, the amount of enzyme and incubation time were adjusted to appropriate ranges depending on the enzyme activity such that the yield of products was linearly related to the amount of protein and incubation time. The maximum substrate concentration and incubation time, and the amount of enzyme were as follows: wild type (50 nM, 5 min, 0.1 ng), N85A (75 nM, 10 min, 1 ng), H239L (200 nM, 30 min, 100 ng), H239N (200 nM, 10 min, 50 ng), F98L (100 nM, 10 min, 20 ng), F98H (75 nM, 3 min, 0.1 ng), N163D (75 nM, 10 min, 1 ng), G87A (100 nM, 3 min, 1 ng), F89A (75 nM, 3 min, 0.1 ng), G90A (75 nM, 3 min, 10 ng) and M91A (75 nM, 3 min, 0.1 ng).

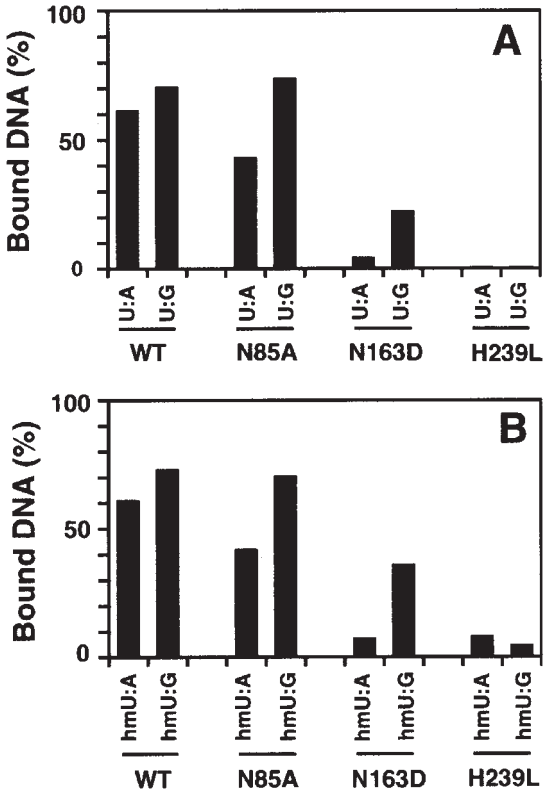

Figure 5. Affinity of wild-type and mutant hSMUG1 for substrates containing U and hmU. The oligonucleotide substrates containing U:A, U:G, hmU:A and hmU:G pairs (all 1 nM) were incubated with 100 ng of the wild type (WT) or the N85A, N163D and H239L mutants of hSMUG1 on ice for 30 min. The fraction of bond DNA was analyzed by gel EMSAs and is plotted against the substrate. (**A**) Substrates containing U:A and U:G pairs. (**B**) Substrates containing hmU:A and hmU:G pairs.

orientation in the pre-transition state to increase affinity and (ii) stabilization of the enolate intermediate in the transition state to increase k_{cat}. Interestingly, the F98L mutant also exhibited a large decrease in k_{cat} (160-fold) together with a moderate decrease in affinity (11-fold), suggesting that proper adjustment of the base orientation by Phe98 through the π–π stacking interaction is crucial not only for base binding but also for efficient catalysis. In contrast, the F98H mutant, presumably retaining the π–π stacking capacity, exhibited 12-fold higher k_{cat} than the wild type, but at the same time exhibited a 14-fold reduction of affinity, implying some as yet unknown alterations of the reaction mechanism. The opposite parameter changes of the F98H mutant resulted in activity comparable to wild-type hSMUG1. Similar results were obtained for the F89A and M91A mutants with their side chains not directly interacting with the base (see Figure 7A and D). The results obtained for the F98H, F89A and M91A mutants indicate that their apparent activity is comparable to that of wild-type hSMUG1, but it involves a slightly modified reaction mechanism. From the kinetic viewpoint, wild-type hSMUG1 with a high damage affinity is better suited than the F98H, F89A and M91A mutants when searching for rarely occurring damage in the genome. This may also explain why hSMUG1 variants such as these have not evolved.

Homology modeling of hSMUG1 structures

xSMUG1 is a member of the UDG structural family and has a chimeric active site of UNG and MUG/TDG enzymes (34) as shown in Figure 1B. We performed homology modeling of wild-type and mutant hSMUG1 by taking advantage of a high sequence homology between xSMUG1 and hSMUG1. The sequence used for modeling had 66% identity (residues 26–269 of hSMUG1 and 37–280 of xSMUG1), allowing fairly accurate prediction of the structure of hSMUG1 (37,38). The residues in motifs 1 and 2 and Asn163 (Figure 1A) are conserved between hSMUG1 and xSMUG1. The backbone of wild-type hSMUG1 virtually superposed that of xSMUG1 with a root-mean-square deviation of 0.37 Å (Figure 6A). Moreover, the position and orientation of the active site residues were well conserved between the two proteins (Figure 6B). The backbone structures of wild-type hSMUG1 and the G87A and G90A mutants were indistinguishable (data not shown). The xSMUG1 crystals soaked in free U and hmU bases have hydrogen bonds that may be pertinent to the recognition of the C5 substituent for U and hmU (Figure 6C) (34). The O4 carbonyl of U bound in the active site pocket makes water-bridged hydrogen bonds to the main-chain NHs of Gly98 and Met102 (equivalent to Gly87 and Met91 of hSMUG1), whereas the C5 hydroxymethyl of hmU displaces the water molecule and directly forms similar hydrogen bonds. If thymine is placed at the same position as hmU, the C5 methyl also displaces the water molecule but cannot make any hydrogen bonds to the main chain due to its hydrophobicity, thereby preventing productive binding of thymine. Although one must take precautions about whether the same interaction schemes are applicable to the base constrained by the DNA backbone,

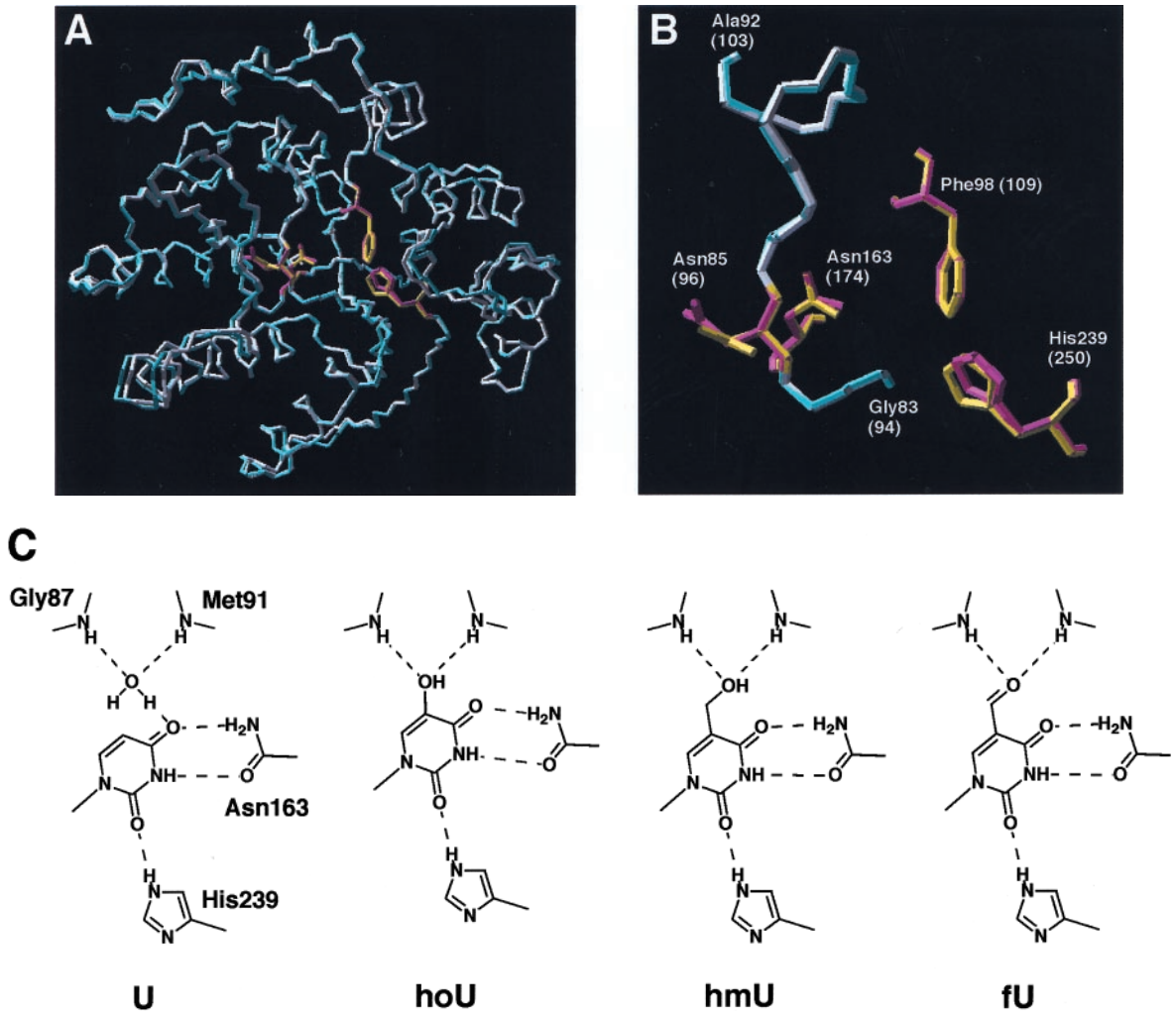

Figure 6. Comparison of the structures of hSMUG1 and xSMUG1. (**A**) The backbone structure of hSMUG1 (light blue, built by homology modeling) superposing that of xSMUG1 determined by X-ray structure analysis (white, Protein Data Bank accession code 1OE4). The side chains of the active site residues are also shown for hSMUG1 (magenta) and xSMUG1 (yellow). (**B**) A magnified view of the active site residues shown in (A). The residue labeling is based on the sequence of hSMUG1, and that for xSMUG1 is indicated in parentheses. (**C**) Hydrogen bonds (broken lines) of the hSMUG1 active site residues to the substrate base. The hydrogen bonds to Gly87 and Met91 are to the main-chain NHs of these residues.

interactions revealed for free bases provide testable models for how hSMUG1 recognizes U, hoU, hmU and fU in DNA via specific water-bridged (U) or direct (hoU, hmU and fU) hydrogen bonds to the backbone, as shown in Figure 6C. The superposed structures of hSMUG1 and the xSMUG1 complex indicate that the C5 hydroxyl of hoU displaces the bridging water molecule due to the repulsive Van der Waals interaction. Accordingly, we tentatively assume that hoU with a short C5 arm moves slightly toward the backbone of Gly87 and Met91 to optimize the hydrogen bonds to the backbone (Figure 6C). We assess the consistency between the structural data of the modeled hSMUG1 proteins and the results of the functional analysis of the mutants in Discussion.

DISCUSSION

In the present study, we used site-directed mutagenesis to identify the crucial residues of hSMUG1 involved in the rupture of the N-glycosidic bond (Asn85 and His239), discrimination of pyrimidine rings (Phe98 and Asn163) and recognition of the C5 substituent (Gly87–Met91). The substitution of Asn85 by Ala (N85A) and that of His239 by Leu (H239L) significantly hampered the damage-excising activity (Tables 2 and 3) as reported for MUG/TDG and UNG enzymes (43,50). Two distinct mechanisms have been proposed for the hydrolysis of the N-glycosidic bond by UNG enzymes. One mechanism proposes nucleophilic attack on deoxyribose-C1′ by a water molecule activated by the conserved Asp, with the conserved His coordinating the O2 carbonyl of U stabilizing the negative charge that develops on U as the reaction proceeds (associative SN2 cleavage mechanism) (42). The other mechanism proposes the dissociative cleavage of the glycosidic bond to relieve the conformational strain of U and deoxyribose in the flipped-out nucleotide (dissociative SN1-like mechanism) (45,51). The resulting U enolate anion is stabilized by the conserved His, and the deoxyribose oxocarbenium cation reacts with a water molecule positioned near C1′ by the

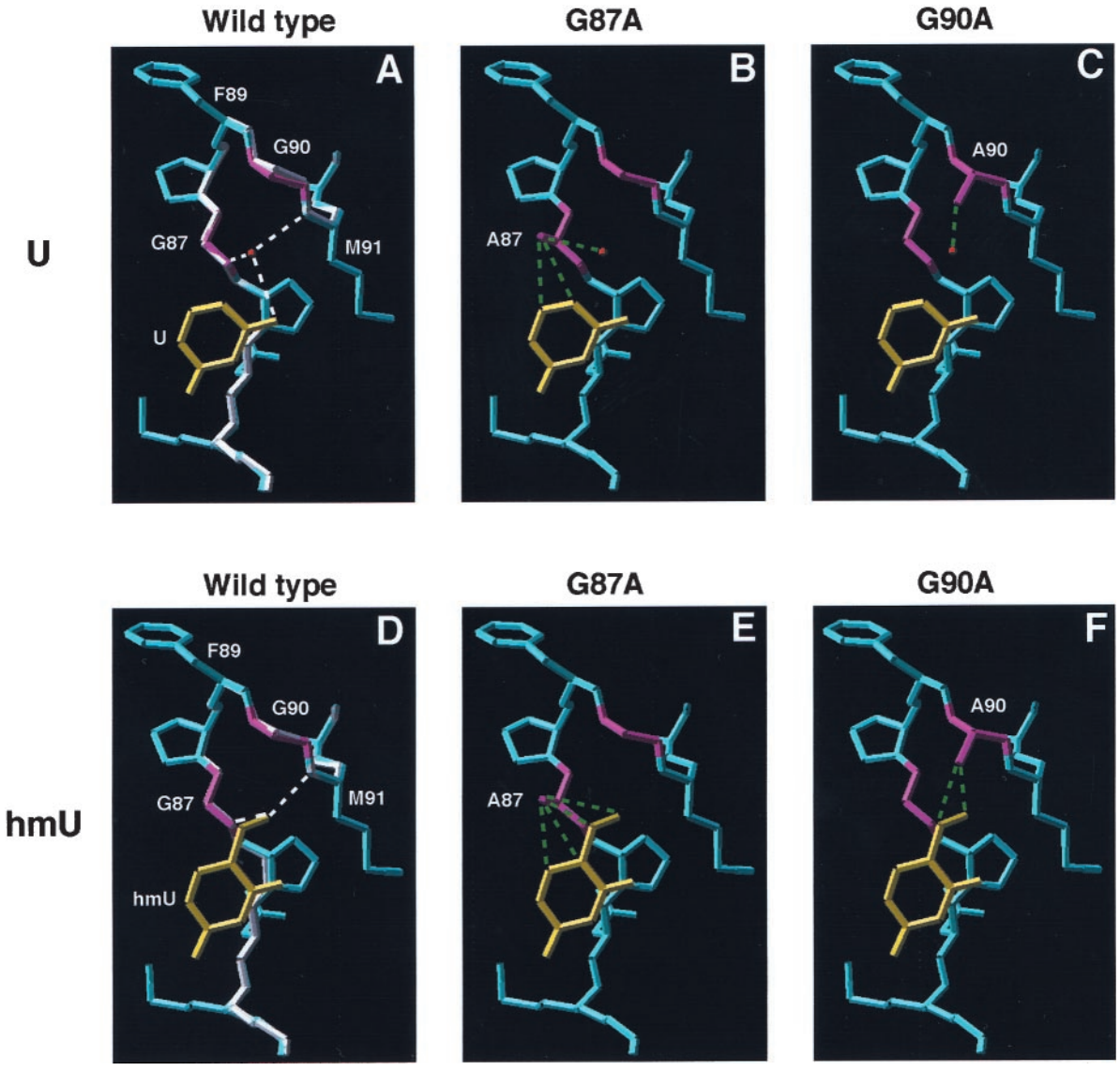

Figure 7. Discriminating interactions of hSMUG1 G87A and G90A mutants for U and hmU. (**A**) The structure of the Gly83–Ala92 region of hSMUG1 (light blue, built by modeling) superposing that of the xSMUG1–DNA–U complex (white, Protein Data Bank accession code 1OE5). The sequence of this region is conserved between the two proteins. U is shown in yellow, and the backbone oxygen of the proteins and DNA bound to the periphery of the xSMUG1 active site pocket are omitted for clarity. White broken lines show the hydrogen bond network of the O4 carbonyl of U to the main-chain NHs of Gly87 and Met91 (magenta) via bridging water (red dot). (**B**) and (**C**) Similar to (A), but with the structure of the G87A (B) or G90A (C) mutant superposing U of the xSMUG1 complex. The structure of xSMUG1 is omitted for clarity, and green broken lines indicate the repulsive Van der Waals contacts of the methyl group of substituted Ala87 and Ala90 with U and bridging water. (**D**) Similar to (A), but with the structure of hSMUG1 (light blue) superposed on that of the xSMUG1–DNA–hmU complex (white, Protein Data Bank accession code 1OE6). hmU is shown in yellow. White broken lines show the hydrogen bonds of the C5 hydroxymethyl oxygen to the main-chain NHs. (**E** and **F**) Similar to (D), but with the structure of the G87A (E) or G90A (F) mutant superposed on hmU of the xSMUG1 complex. Green broken lines indicate the repulsive Van der Waals contacts of the methyl group of substituted Ala87 and Ala90 (magenta) with hmU.

conserved Asp. Like MUG/TDG enzymes, hSMUG1 has Asn (Asn85) instead of Asp, which is conserved in UNG enzymes (Figure 1A). Thus, the Asn85 of hSMUG1 can either activate a water molecule for nucleophilic attack (associative SN2 mechanism) or position it near C1′ for the reaction with the deoxyribose oxocarbenium cation (dissociative SN1-like mechanism). As mentioned in Results, the His239 of hSMUG1 likely adjusts the orientation of the base (U, hoU, hmU or fU) and stabilizes the enolate intermediate by delocalizing the developing negative charge in both mechanisms. A new UDG family (family 4) has been recently identified that lacks polar residues for the activation or positioning of a water molecule in motif 1 (52). The present study shows that the mutation of His239 (H239L) results in a much greater reduction in the damage-excising activity of hSMUG1 than does the mutation of Asn85 (N85A) (Tables 2 and 3). These data together suggest that hSMUG1 employs the dissociative SN1-like mechanism to rupture the N-glycosidic bond.

hSMUG1 is unique among DNA glycosylases in recognizing the oxidized C5 substituent of U derivatives. We have previously suggested hydrophilic interactions between the C5 substituent and the active site pocket of hSMUG1 as

key interactions for damage recognition (33). This mechanism has been supported by the structure of xSMUG1 complexed with free U or hmU (Figure 6C) (34), though the bases are not constrained by the DNA backbone. To fill a gap between the data with free bases and those with bases constrained by actual DNA, we systematically analyzed the effect of mutations that may be pertinent to the recognition of the C5 substituent (Figures 1B and 4). The mutants of Gly87 and Gly90 exhibited reductions in damage-excising ability, whereas those of Phe89 and Met91 were only moderate to negligible (Figure 4). Figure 7 shows the structures, built by homology modeling, of Gly83–Ala92 residues of the wild-type and G87A and G90A mutants of hSMUG1. The backbone structure of wild-type hSMUG1 almost superposed those of the xSMUG1–U (Figure 7A) and xSMUG1–hmU (Figure 7D) complexes. The Gly83–Ala92 residues of hSMUG1 (light blue) are nearly indistinguishable from the corresponding ones of xSMUG1 (Gly94–Ala103, white). Thus, like xSMUG1, the O4 carbonyl of U can make water-bridged hydrogen bonds to the main-chain NHs of Gly87 and Met91 of hSMUG1 (Figure 7A), and the C5 hydroxymethyl oxygen of hmU can make direct hydrogen bonds to the same backbone (Figure 7D). When the G87A mutant superposes the xSMUG1 complex, the side-chain methyl of Ala87 makes repulsive Van der Waals contacts with the ring C5 (3.76 Å) and C6 (3.65 Å) of U and the oxygen atom (3.36 Å) of bridging water (Figure 7B), where the numbers in parentheses are the heavy-atom distance to the methyl carbon of Ala87. Similarly, the Ala87-methyl also makes repulsive contacts with the ring C5 (3.58 Å) and C6 (3.45 Å) of hmU and the CH_2 (3.15 Å) and O (3.51 Å) of the C5 hydroxymethyl (Figure 7E). In view of the structural similarity, similar discriminating interactions are also expected for fU and hoU. It is possible that these repulsive interactions displace the substrate base (and the bridging water molecule for U) and hamper the formation of hydrogen bonds crucial for the direct (hoU, hmU and fU) or indirect (U) recognition of the C5 substituent (Figure 4).

Analysis of the structure of the G90A mutant superposed on the xSMUG1 complexes also reveals repulsive Van der Waals contacts of the side-chain methyl of Ala90 with the oxygen atom (2.28 Å) of bridging water for U (Figure 7C) and the CH_2 (2.83 Å) and O (2.70 Å) of C5 hydroxymethyl of hmU (Figure 7F), where the numbers in parentheses are the heavy-atom distance to the methyl carbon of Ala90. The displacement of bridging water due to the close contact with the Ala90-methyl accounts well for the observed severe impairment of the U-excising capacity of the G90A mutant (Figure 4). Conversely, the same mutant excised hoU, hmU and fU with a low but significant efficiency (Figure 4). The distinct distances of the Ala90-methyl to the bridging water (~2.3 Å) and the C5 substituent (~2.8 Å) correlates with the observed activity difference (U < hmU, hoU and fU). However, the C5 substituents of hmU, hoU and fU are too close to the Ala90-methyl to partially relieve the repulsive interaction, implying a limitation of analyses based on the crystal structure with the diffused free base. It is likely that the C5 substituents of hoU, hmU and fU lie slightly distant from the modeled Ala90-methyl when they are actually constrained by the DNA backbone. It is also not clear from the present model as to why the displacement of the bridging water for U in the G90A and G87A mutants leads to opposite effects on k_{cat} (22-fold decrease with G90A and 3-fold increase with G87A relative to the wild type), while affinity reduces in both mutants (Table 3).

The N163D mutant but not wild-type hSMUG1 excised hoC with a low but significant efficiency (Figure 3B). The C5 hydroxyl of hoC can form hydrogen bonds to the backbone like hoU (Figure 8A and B), but the hydrogen-bonding specificity of the side-chain amide of Asn163 of wild-type hSMUG1 is incompatible with the Watson–Crick face of hoC (Figure 8B). The present modeling data show that the orientation of the side-chain amide of Asn163 is locked by the hydrogen bond to the main-chain NH of Met84. This conformational restriction is crucial since otherwise a side-chain rotamer with the flipped amide group can make hydrogen bonds to the Watson–Crick face of hoC. Moreover, the amide NH_2 in the locked conformation makes a repulsive contact with the N4 amino of hoC (Figure 8B). The same unfavorable interaction is expected for 5-formylcytosine,

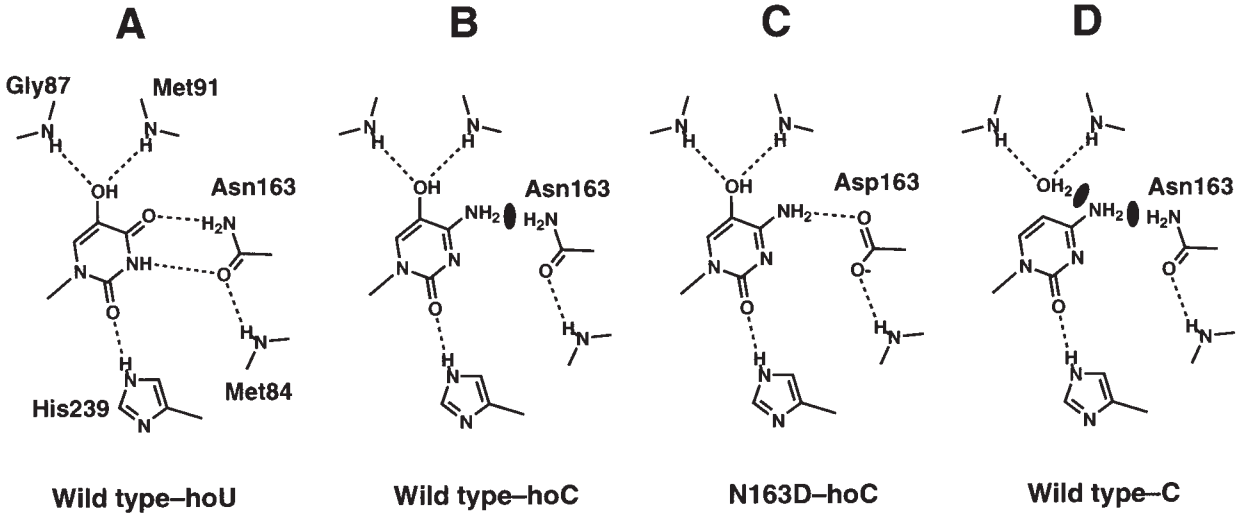

Figure 8. Interactions of the active site residues of the wild-type and N163D mutant of hSMUG1 with hoC. Interactions of wild-type hSMUG1 with (**A**) hoU, (**B**) hoC, and (**D**) cytosine and (**C**) those of the N163D mutant with hoC. Broken lines indicate hydrogen bonds, and filled ellipsoids show repulsive contacts.

which is not a substrate of hSMUG1 (33). Presumably, the N163D mutant accommodates hoC through the hydrogen bonds to the C5 hydroxyl and partially restores one to the N4 amino (Figure 8C), and excises hoC with a low but detectable efficiency. Conversely, the N4 amino of cytosine has a repulsive Van der Waals contact with the water molecule bound to the backbone of the wild-type (Figure 8D) and N163D mutant of hSMUG1, and is unable to make any stable hydrogen bonds to the backbone. Thus, both the wild-type and N163D mutant of hSMUG1 can discriminate against cytosine (Figure 3B). The N163D mutant prefers mismatched hoC to matched hoC as a substrate (Figure 3B), as does wild-type hSMUG1 for U, hoU, hmU and fU (33,53) (this study). The molecular basis of the preference has not been fully established, but is possibly related to the flip-out mechanism (54).

In conclusion, we have identified the crucial residues of hSMUG1, a new member of DNA glycosylases for oxidative damage, involved in damage-recognition and catalysis. Determination of the 3D structure of the productive hSMUG1–DNA complex will provide further insight into the underlying mechanisms of action.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at NAR Online.

ACKNOWLEDGEMENTS

We thank Koichi Matsuo for the analyses of CD spectra and Yoshihiko Ohyama for technical assistance. This work was supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology, Japan (H.I.), and in part by the National Project on Protein Structural and Functional Analyses from the Ministry of Education, Culture, Sports, Science and Technology, Japan (K.K.).

REFERENCES

1. Lindahl,T. (1993) Instability and decay of the primary structure of DNA. *Nature*, **362**, 709–715.
2. Friedberg,E.C., Walker,G.C. and Siede,W. (1995) *DNA Repair and Mutagenesis*. American Society for Microbiology, Washington, DC.
3. Friedberg,E.C. (2003) DNA damage and repair. *Nature*, **421**, 436–440.
4. Kunkel,T.A. and Bebenek,K. (2000) DNA replication fidelity. *Annu. Rev. Biochem.*, **69**, 497–529.
5. Scharer,O.D. and Jiricny,J. (2001) Recent progress in the biology, chemistry and structural biology of DNA glycosylases. *Bioessays*, **23**, 270–281.
6. Wallace,S.S. (2002) Biological consequences of free radical-damaged DNA bases. *Free Radic. Biol. Med.*, **33**, 1–14.
7. Ide,H. and Kotera,M. (2004) Human DNA glycosylases involved in the repair of oxidatively damaged DNA. *Biol. Pharm. Bull.*, **27**, 480–485.
8. Ischenko,A.A. and Saparbaev,M.K. (2002) Alternative nucleotide incision repair pathway for oxidative DNA damage. *Nature*, **415**, 183–187.
9. Ishchenko,A.A., Sanz,G., Privezentzev,C.V., Maksimenko,A.V. and Saparbaev,M. (2003) Characterisation of new substrate specificities of *Escherichia coli* and *Saccharomyces cerevisiae* AP endonucleases. *Nucleic Acids Res.*, **31**, 6344–6353.
10. Gros,L., Ishchenko,A.A., Ide,H., Elder,R.H. and Saparbaev,M.K. (2004) The major human AP endonuclease (Ape1) is involved in the nucleotide incision repair pathway. *Nucleic Acids Res.*, **32**, 73–81.
11. Ide,H., Tedzuka,K., Shimzu,H., Kimura,Y., Purmal,A.A., Wallace,S.S. and Kow,Y.W. (1994) Alpha-deoxyadenosine, a major anoxic radiolysis product of adenine in DNA, is a substrate for *Escherichia coli* endonuclease IV. *Biochemistry*, **33**, 7842–7847.
12. Wallace,S.S. (1998) Enzymatic processing of radiation-induced free radical damage in DNA. *Radiat. Res.*, **150**, S60–S79.
13. Zharkov,D.O., Shoham,G. and Grollman,A.P. (2003) Structural characterization of the Fpg family of DNA glycosylases. *DNA Repair*, **2**, 839–862.
14. Wallace,S.S., Bandaru,V., Kathe,S.D. and Bond,J.P. (2003) The enigma of endonuclease VIII. *DNA Repair*, **2**, 441–453.
15. Hazra,T.K., Izumi,T., Kow,Y.W. and Mitra,S. (2003) The discovery of a new family of mammalian enzymes for repair of oxidatively damaged DNA, and its physiological implications. *Carcinogenesis*, **24**, 155–157.
16. Masaoka,A., Terato,H., Kobayashi,M., Honsho,A., Ohyama,Y. and Ide,H. (1999) Enzymatic repair of 5-formyluracil. I. Excision of 5-formyluracil site-specifically incorporated into oligonucleotide substrates by AlkA protein (*Escherichia coli* 3-methyladenine DNA glycosylase II). *J. Biol. Chem.*, **274**, 25136–25143.
17. Matsubara,M., Masaoka,A., Tanaka,T., Miyano,T., Kato,N., Terato,H., Ohyama,Y., Iwai,S. and Ide,H. (2003) Mammalian 5-formyluracil-DNA glycosylase. 1. Identification and characterization of a novel activity that releases 5-formyluracil from DNA. *Biochemistry*, **42**, 4993–5002.
18. Katafuchi,A., Nakano,T., Masaoka,A., Terato,H., Iwai,S., Hanaoka,F. and Ide,H. (2004) Differential specificity of human and *Escherichia coli* endonuclease III and VIII homologues for oxidative base lesions. *J. Biol. Chem.*, **279**, 14464–14471.
19. Kino,K., Shimizu,Y., Sugasawa,K., Sugiyama,H. and Hanaoka,F. (2004) Nucleotide excision repair of 5-formyluracil *in vitro* is enhanced by the presence of mismatched bases. *Biochemistry*, **43**, 2682–2687.
20. Levy,D.D. and Teebor,G.W. (1991) Site directed substitution of 5-hydroxymethyluracil for thymine in replicating phi X-174am3 DNA via synthesis of 5-hydroxymethyl-2′-deoxyuridine-5′-triphosphate. *Nucleic Acids Res.*, **19**, 3337–3343.
21. Zhang,Q.M., Sugiyama,H., Miyabe,I., Matsuda,S., Kino,K., Saito,I. and Yonei,S. (1999) Replication *in vitro* and cleavage by restriction endonuclease of 5-formyluracil- and 5-hydroxymethyluracil-containing oligonucleotides. *Int. J. Radiat. Biol.*, **75**, 59–65.
22. Cannon-Carlson,S.V., Gokhale,H. and Teebor,G.W. (1989) Purification and characterization of 5-hydroxymethyluracil-DNA glycosylase from calf thymus. Its possible role in the maintenance of methylated cytosine residues. *J. Biol. Chem.*, **264**, 13306–13312.
23. Yoshida,M., Makino,K., Morita,H., Terato,H., Ohyama,Y. and Ide,H. (1997) Substrate and mispairing properties of 5-formyl-2′-deoxyuridine 5′-triphosphate assessed by *in vitro* DNA polymerase reactions. *Nucleic Acids Res.*, **25**, 1570–1577.
24. Terato,H., Masaoka,A., Kobayashi,M., Fukushima,S., Ohyama,Y., Yoshida,M. and Ide,H. (1999) Enzymatic repair of 5-formyluracil. II. Mismatch formation between 5-formyluracil and guanine during DNA replication and its recognition by two proteins involved in base excision repair (AlkA) and mismatch repair (MutS). *J. Biol. Chem.*, **274**, 25144–25150.
25. Masaoka,A., Terato,H., Kobayashi,M., Ohyama,Y. and Ide,H. (2001) Oxidation of thymine to 5-formyluracil in DNA promotes misincorporation of dGMP and subsequent elongation of a mismatched primer terminus by DNA polymerase. *J. Biol. Chem.*, **276**, 16501–16510.
26. Tsunoda,M., Sakaue,T., Naito,S., Sunami,T., Karino,N., Ueno,Y., Matsuda,A. and Takenaka,A. (2001) X-ray analyses of DNA dodecamers containing 2′-deoxy-5-formyluridine. *Nucleic Acids Res. Suppl.*, **1**, 279–280.
27. Hollstein,M.C., Brooks,P., Linn,S. and Ames,B.N. (1984) Hydroxymethyluracil DNA glycosylase in mammalian cells. *Proc. Natl Acad. Sci. USA*, **81**, 4003–4007.
28. Boorstein,R.J., Levy,D.D. and Teebor,G.W. (1987) 5-Hydroxymethyluracil-DNA glycosylase activity may be a differentiated mammalian function. *Mutat. Res.*, **183**, 257–263.
29. Boorstein,R.J., Cummings,A.,Jr, Marenstein,D.R., Chan,M.K., Ma,Y., Neubert,T.A., Brown,S.M. and Teebor,G.W. (2001) Definitive identification of mammalian 5-hydroxymethyluracil DNA N-glycosylase activity as SMUG1. *J. Biol. Chem.*, **276**, 41991–41997.
30. Haushalter,K.A., Todd Stukenberg,M.W., Kirschner,M.W. and Verdine,G.L. (1999) Identification of a new uracil-DNA glycosylase family by expression cloning using synthetic inhibitors. *Curr. Biol.*, **9**, 174–185.

31. Nilsen,H., Rosewell,I., Robins,P., Skjelbred,C.F., Andersen,S., Slupphaug,G., Daly,G., Krokan,H.E., Lindahl,T. and Barnes,D.E. (2000) Uracil-DNA glycosylase (UNG)-deficient mice reveal a primary role of the enzyme during DNA replication. *Mol. Cell*, **5**, 1059–1065.
32. Nilsen,H., Haushalter,K.A., Robins,P., Barnes,D.E., Verdine,G.L. and Lindahl,T. (2001) Excision of deaminated cytosine from the vertebrate genome: role of the SMUG1 uracil-DNA glycosylase. *EMBO J.*, **20**, 4278–4286.
33. Masaoka,A., Matsubara,M., Hasegawa,R., Tanaka,T., Kurisu,S., Terato,H., Ohyama,Y., Karino,N., Matsuda,A. and Ide,H. (2003) Mammalian 5-formyluracil-DNA glycosylase. 2. Role of SMUG1 uracil-DNA glycosylase in repair of 5-formyluracil and other oxidized and deaminated base lesions. *Biochemistry*, **42**, 5003–5012.
34. Wibley,J.E., Waters,T.R., Haushalter,K., Verdine,G.L. and Pearl,L.H. (2003) Structure and specificity of the vertebrate anti-mutator uracil-DNA glycosylase SMUG1. *Mol. Cell*, **11**, 1647–1659.
35. Ohmae,E., Iriyama,K., Ichihara,S. and Gekko,K. (1996) Effects of point mutations at the flexible loop glycine-67 of *Escherichia coli* dihydrofolate reductase on its stability and function. *J. Biochem.*, **119**, 703–710.
36. Ohmae,E., Sasaki,Y. and Gekko,K. (2001) Effects of five-tryptophan mutations on structure, stability and function of *Escherichia coli* dihydrofolate reductase. *J. Biochem.*, **130**, 439–447.
37. Schwede,T., Kopp,J., Guex,N. and Peitsch,M.C. (2003) SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Res.*, **31**, 3381–3385.
38. Kopp,J. and Schwede,T. (2004) The SWISS-MODEL Repository of annotated three-dimensional protein structure homology models. *Nucleic Acids Res.*, **32**, D230–D234.
39. Guex,N. and Peitsch,M.C. (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*, **18**, 2714–2723.
40. Pearl,L.H. (2000) Structure and function in the uracil-DNA glycosylase superfamily. *Mutat. Res.*, **460**, 165–181.
41. Aravind,L. and Koonin,E.V. (2000) The alpha/beta fold uracil DNA glycosylases: a common origin with diverse fates. *Genome Biol.*, **1**, research0007.1–research0007.8.
42. Savva,R., McAuley-Hecht,K., Brown,T. and Pearl,L. (1995) The structural basis of specific base-excision repair by uracil-DNA glycosylase. *Nature*, **373**, 487–493.
43. Mol,C.D., Arvai,A.S., Slupphaug,G., Kavli,B., Alseth,I., Krokan,H.E. and Tainer,J.A. (1995) Crystal structure and mutational analysis of human uracil-DNA glycosylase: structural basis for specificity and catalysis. *Cell*, **80**, 869–878.
44. Slupphaug,G., Mol,C.D., Kavli,B., Arvai,A.S., Krokan,H.E. and Tainer,J.A. (1996) A nucleotide-flipping mechanism from the structure of human uracil-DNA glycosylase bound to DNA. *Nature*, **384**, 87–92.
45. Parikh,S.S., Walcher,G., Jones,G.D., Slupphaug,G., Krokan,H.E., Blackburn,G.M. and Tainer,J.A. (2000) Uracil-DNA glycosylase-DNA substrate and product structures: conformational strain promotes catalytic efficiency by coupled stereoelectronic effects. *Proc. Natl Acad. Sci. USA*, **97**, 5083–5088.
46. Barrett,T.E., Savva,R., Panayotou,G., Barlow,T., Brown,T., Jiricny,J. and Pearl,L.H. (1998) Crystal structure of a G:T/U mismatch-specific DNA glycosylase: mismatch recognition by complementary-strand interactions. *Cell*, **92**, 117–129.
47. Barrett,T.E., Scharer,O.D., Savva,R., Brown,T., Jiricny,J., Verdine,G.L. and Pearl,L.H. (1999) Crystal structure of a thwarted mismatch glycosylase DNA repair complex. *EMBO J.*, **18**, 6599–6609.
48. Sreerama,N. and Woody,R.W. (2000) Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal. Biochem.*, **287**, 252–260.
49. Kavli,B., Slupphaug,G., Mol,C.D., Arvai,A.S., Peterson,S.B., Tainer,J.A. and Krokan,H.E. (1996) Excision of cytosine and thymine from DNA by mutants of human uracil-DNA glycosylase. *EMBO J.*, **15**, 3442–3447.
50. Hardeland,U., Bentele,M., Jiricny,J. and Schar,P. (2000) Separating substrate recognition from base hydrolysis in human thymine DNA glycosylase by mutational analysis. *J. Biol. Chem.*, **275**, 33449–33456.
51. Werner,R.M. and Stivers,J.T. (2000) Kinetic isotope effect studies of the reaction catalyzed by uracil DNA glycosylase: evidence for an oxocarbenium ion-uracil anion intermediate. *Biochemistry*, **39**, 14054–14064.
52. Sartori,A.A., Fitz-Gibbon,S., Yang,H., Miller,J.H. and Jiricny,J. (2002) A novel uracil-DNA glycosylase with broad substrate specificity and an unusual active site. *EMBO J.*, **21**, 3182–3191.
53. Kavli,B., Sundheim,O., Akbari,M., Otterlei,M., Nilsen,H., Skorpen,F., Aas,P.A., Hagen,L., Krokan,H.E. and Slupphaug,G. (2002) hUNG2 is the major repair enzyme for removal of uracil from U:A matches, U:G mismatches, and U in single-stranded DNA, with hSMUG1 as a broad specificity backup. *J. Biol. Chem.*, **277**, 39926–39936.
54. Asagoshi,K., Odawara,H., Nakano,H., Miyano,T., Terato,H., Ohyama,Y., Seki,S. and Ide,H. (2000) Comparison of substrate specificities of *Escherichia coli* endonuclease III and its mouse homologue (mNTH1) using defined oligonucleotide substrates. *Biochemistry*, **39**, 11389–11398.