Definitive Identification of Mammalian 5-Hydroxymethyluracil DNA N-Glycosylase Activity as SMUG1*

Received for publication, July 23, 2001, and in revised form, August 23, 2001
Published, JBC Papers in Press, August 28, 2001, DOI 10.1074/jbc.M106953200

Robert J. Boorstein‡, Archie Cummings, Jr.‡, Dina R. Marenstein‡, Michael K. Chan‡, Yuliang Ma§, Thomas A. Neubert§, Stuart M. Brown¶, and George W. Teebor∥

From the ‡Department of Pathology and Kaplan Comprehensive Cancer Center, the §Department of Pharmacology and Skirball Institute of Biomolecular Medicine, and the ¶Department of Cell Biology, New York University School of Medicine, New York, New York 10016

© 2001 by The American Society for Biochemistry and Molecular Biology, Inc.
Printed in U.S.A.

5-Hydroxymethyluracil (5hmUra)1 is formed in DNA exposed to ionizing radiation both in solution and in cells. Its formation has been posited to be caused by the indirect action of ionizing radiation under aerobic conditions via the hydroxyl radical-mediated abstraction of a hydrogen radical from the 5-methyl-moiety of thymine followed by oxidation via molecular oxygen (1). 5hmUra has also been identified in human urine and in the DNA of white cells, and the amount has been found to vary under differing conditions of oxidative stress (2–4).

5-Hydroxymethyluracil (5hmUra)1 is formed in DNA by the action of 5-hydroxymethyluracil N-glycosylase (HMUDG) that released 5-hydroxymethyluracil (5hmUra) from the DNA of Bacillus subtilis phage SPO1. Analysis of the most purified fraction by SDS-polyacrylamide gel electrophoresis revealed a multiplicity of protein species making it impossible to identify HMUDG by inspection. Therefore, we retracted the enzyme after SDS-polyacrylamide gel electrophoresis and assayed slices of the gel for DNA N-glycosylase activity directed against 5hmUra. Maximum enzymatic activity was identified between molecular mass markers 30 and 34 kDa. Protein was extracted from gel slices and subjected to tryptic digestion and analysis by mass spectrometry. Analysis revealed the presence of 11 peptides that were homologous or identical to the sequence of the recently characterized human single-stranded monofunctional uracil DNA N-glycosylase (hSMUG1). The cDNA of hSMUG1 was isolated and expressed as a recombinant glutathione S-transferase fusion protein that was shown to release 5hmUra with 20× the specific activity of the most purified bovine fraction. We conclude that hSMUG1 and HMUDG are the same protein.

Human autoantibodies directed against 5-hydroxymethyl-2'-deoxyuridine in DNA have been identified, and the titer has been correlated with relative risk for certain types of cancer (5).

Despite identification as a radiation product nearly 20 years ago, the biological consequences of the formation of 5hmUra from thymine in DNA remain uncertain. 5hmUra is a normal constituent of SPO1 Bacillus subtilis phage DNA (6), functioning as a coding partner for adenine in place of thymine, and is present in large amounts in the DNA of dinoflagellates (7). Nonetheless, 5hmUra is a substrate for a DNA N-glycosylase activity, 5hmUra DNA N-glycosylase (HMUDG), which releases 5hmUra from the DNA backbone (8, 9). Compared with most other DNA N-glycosylases, HMUDG activity is unusual in that no comparable activity has been detected in bacteria (10) or yeast (9).

We undertook a study of the phylogenetic distribution of HMUDG activity (11), and our results indicate that the phylogenetic distribution of HMUDG was linked to the use by the organism of 5-methylcytosine (5meCyt) as a mediator of gene expression. We suggested that 5hmUra could be formed in DNA via two pathways. One pathway was the oxidation of the 5-methyl group of thymine yielding 5hmUra directly. Alternatively, 5hmUra may arise via oxidation of the 5-methyl group of 5meCyt yielding 5hmCyt followed by deamination of 5hmCyt to 5hmUra (12). The propensity of 5hmCyt residues to deaminate to 5hmUra has long ago been observed in the DNA of the Escherichia coli T-even phages. These phages utilize glucosylated 5hmCyt as a normal constituent of DNA and exhibit a very high rate of C to T transitions upon heating (13). Thus, in cellular DNA the oxidation and subsequent deamination of 5hmCyt, if left unrepaired, may result in a C to T transition in CpG islands.

We previously reported the partial purification of HMUDG from calf thymus and demonstrated that HMUDG activity was distinct from the major uracil DNA N-glycosylase (UDG) activity of the cell (12). However, we have been unable to achieve a degree of purification sufficient to identify the HMUDG protein species with certainty from examination of an SDS-PAGE analysis of our most purified calf thymus fraction. In this report we describe the methods by which we were able to unequivocally demonstrate that HMUDG is the recently characterized single-stranded monofunctional uracil DNA N-glycosylase (SMUG1) (14).

EXPERIMENTAL PROCEDURES

Purification of HMUDG from Calf Thymus—Fresh calf thymus obtained from local abattoirs, generally weighing between 1.2 and 2 kg, was trimmed of connective tissue. 600-g portions were homogenized in a Waring Blender together with 3 volumes of Buffer A (25 mM HEPES pH 7.4, 7.5 mM EDTA, 1 mM dithiothreitol, and 100 μM of protease inhibitor mixture (Sigma) per 1.0 liter of buffer). The total volume of
homogenate from a typical preparation was between 5 and 7 liters. To this homogenate 5 mM NaCl was added to a final concentration of 320 mM NaCl. The resulting gelatinous precipitate of nucleoprotein was removed by spooling onto a 10-ml glass pipette. The remaining liquid was filtered through cheesecloth followed by clearing via centrifugation in the GS-3 rotor of a Sorvall RC-5B centrifuge at 9000 rpm for 20 min to yield Fraction I (15).

Fraction I was precipitated with solid (NH₄)₂SO₄ to a final concentration of 70%, the pH continually being adjusted with NH₄OH to maintain pH 7.5. After centrifugation, the pellets were washed with 100% (NH₄)₂SO₄ and stored at 4°C in aliquots representing the pellet material from 500 ml of the original 70% (NH₄)₂SO₄ suspension, yielding Fraction II.

Fraction II was dialyzed overnight against Buffer A and applied to a 100-ml bed cation exchange resin in an XK 26/20 column (SP Fastflow, Amersham Pharmacia Biotech). Aliquots of Fraction II applied to the column averaged 3–4 g of protein/sample, which was half the maximal binding capacity of the resin. The column was washed with Buffer A until absorbance of the eluent at 280 nm was at a minimum. Protein was eluted with a 1-liter gradient of Buffer A increasing in NaCl concentration from 0 to 1 M at a flow rate of 5 ml/min. Fractions were collected as 10-ml aliquots and were assayed for HMUDG activity as described below. Fractions containing HMUDG activity were pooled as Fraction III. Fraction III was adjusted to 70% (NH₄)₂SO₄ by the addition of solid (NH₄)₂SO₄ and then pelleted by centrifugation to make Fraction IV, which was stored in aliquots at 4°C.

Aliquots of Fraction IV, dissolved in Buffer A in 400 mM NaCl, were applied to a Hiloass chromatography Superdex-75 preparative column that had been calibrated with the gel filtration low molecular weight calibration kit (Amersham Pharmacia Biotech). The column was eluted at a flow rate of 2.5 ml/min using 400 mM NaCl Buffer A, and fractions were collected as 5-ml aliquots. Fifty microliters of each aliquot were assayed for HMUDG activity, and aliquots containing HMUDG activity were pooled as Fraction V.

Electrophoresis was performed according to the protocol of Nadlermann et al. (21). The slices were homogenized (using a microsample Teflon pestle (Scientware) attached to a hand drill) until no large pieces of gel remained. The suspension was then vigorously shaken overnight at 37°C. The gel suspension was centrifuged at 14,000 rpm for 15 min at 4°C in the GS-3 rotor of a Sorvall RC-5B centrifuge. The supernatant was transferred to a 2-ml Eppendorf tube, and 1.6 ml of 4 M guanidine HCl, 1 M dithiothreitol, and 0.1 mg/ml acetylated BSA). The following renaturation procedure was adopted from that of Nadlermann et al. (21). The slices were homogenized (using a microsample Teflon pestle (Scientware) attached to a hand drill) until no large pieces of gel remained. The suspension was then vigorously shaken overnight at 37°C. The gel suspension was centrifuged at 14,000 rpm for 15 min at 4°C in the GS-3 rotor of a Sorvall RC-5B centrifuge. The supernatant was transferred to a 2-ml Eppendorf tube, and 1.6 ml of 4 M guanidine HCl, 1 M dithiothreitol, and 0.1 mg/ml acetylated BSA). The following renaturation procedure was adopted from that of Nadlermann et al. (21). The slices were homogenized (using a microsample Teflon pestle (Scientware) attached to a hand drill) until no large pieces of gel remained. The suspension was then vigorously shaken overnight at 37°C. The gel suspension was centrifuged at 14,000 rpm for 15 min at 4°C in the GS-3 rotor of a Sorvall RC-5B centrifuge. The supernatant was transferred to a 2-ml Eppendorf tube, and 1.6 ml of 4 M guanidine HCl, 1 M dithiothreitol, and 0.1 mg/ml acetylated BSA). The following renaturation procedure was adopted from that of Nadlermann et al. (21). The slices were homogenized (using a microsample Teflon pestle (Scientware) attached to a hand drill) until no large pieces of gel remained. The suspension was then vigorously shaken overnight at 37°C. The gel suspension was centrifuged at 14,000 rpm for 15 min at 4°C in the GS-3 rotor of a Sorvall RC-5B centrifuge. The supernatant was transferred to a 2-ml Eppendorf tube, and 1.6 ml of 4 M guanidine HCl, 1 M dithiothreitol, and 0.1 mg/ml acetylated BSA). The following renaturation procedure was adopted from that of Nadlermann et al. (21). The slices were homogenized (using a microsample Teflon pestle (Scientware) attached to a hand drill) until no large pieces of gel remained. The suspension was then vigorously shaken overnight at 37°C. The gel suspension was centrifuged at 14,000 rpm for 15 min at 4°C in the GS-3 rotor of a Sorvall RC-5B centrifuge. The supernatant was transferred to a 2-ml Eppendorf tube, and 1.6 ml of 4 M guanidine HCl, 1 M dithiothreitol, and 0.1 mg/ml acetylated BSA). The following renaturation procedure was adopted from that of Nadlermann et al. (21). The slices were homogenized (using a microsample Teflon pestle (Scientware) attached to a hand drill) until no large pieces of gel remained. The suspension was then vigorously shaken overnight at 37°C. The gel suspension was centrifuged at 14,000 rpm for 15 min at 4°C in the GS-3 rotor of a Sorvall RC-5B centrifuge. The supernatant was transferred to a 2-ml Eppendorf tube, and 1.6 ml of 4 M guanidine HCl, 1 M dithiothreitol, and 0.1 mg/ml acetylated BSA). The following renaturation procedure was adopted from that of Nadlermann et al. (21). The slices were homogenized (using a microsample Teflon pestle (Scientware) attached to a hand drill) until no large pieces of gel remained. The suspension was then vigorously shaken overnight at 37°C. The gel suspension was centrifuged at 14,000 rpm for 15 min at 4°C in the GS-3 rotor of a Sorvall RC-5B centrifuge. The supernatant was transferred to a 2-ml Eppendorf tube, and 1.6 ml of 4 M guanidine HCl, 1 M dithiothreitol, and 0.1 mg/ml acetylated BSA). The following renaturation procedure was adopted from that of Nadlermann et al. (21). The slices were homogenized (using a microsample Teflon pestle (Scientware) attached to a hand drill) until no large pieces of gel remained. The suspension was then vigorously shaken overnight at 37°C. The gel suspension was centrifuged at 14,000 rpm for 15 min at 4°C in the GS-3 rotor of a Sorvall RC-5B centrifuge. The supernatant was transferred to a 2-ml Eppendorf tube, and 1.6 ml of 4 M guanidine HCl, 1 M dithiothreitol, and 0.1 mg/ml acetylated BSA). The following renaturation procedure was adopted from that of Nadlermann et al. (21). The slices were homogenized (using a microsample Teflon pestle (Scientware) attached to a hand drill) until no large pieces of gel remained. The suspension was then vigorously shaken overnight at 37°C. The gel suspension was centrifuged at 14,000 rpm for 15 min at 4°C in the GS-3 rotor of a Sorvall RC-5B centrifuge. The supernatant was transferred to a 2-ml Eppendorf tube, and 1.6 ml of 4 M guanidine HCl, 1 M dithiothreitol, and 0.1 mg/ml acetylated BSA). The following renaturation procedure was adopted from that of Nadlermann et al. (21). The slices were homogenized (using a microsample Teflon pestle (Scientware) attached to a hand drill) until no large pieces of gel remained. The suspension was then vigorously shaken overnight at 37°C. The gel suspension was centrifuged at 14,000 rpm for 15 min at 4°C in the GS-3 rotor of a Sorvall RC-5B centrifuge. The supernatant was transferred to a 2-ml Eppendorf tube, and 1.6 ml of 4 M guanidine HCl, 1 M dithiothreitol, and 0.1 mg/ml acetylated BSA). The following renaturation procedure was adopted from that of Nadlermann et al. (21). The slices were homogenized (using a microsample Teflon pestle (Scientware) attached to a hand drill) until no large pieces of gel remained. The suspension was then vigorously shaken overnight at 37°C. The gel suspension was centrifuged at 14,000 rpm for 15 min at 4°C in the GS-3 rotor of a Sorvall RC-5B centrifuge. The supernatant was transferred to a 2-ml Eppendorf tube, and 1.6 ml of 4 M guanidine HCl, 1 M dithiothreitol, and 0.1 mg/ml acetylated BSA).
washed with 50% methanol and water to remove the remaining SDS and acetic acid. The gel slices were cut into 1 mm² pieces, reduced, alkylated using iodoacetamide, and digested in-gel with trypsin (17). The peptides were extracted, dried under vacuum, and redissolved in 20 μl of 5% acetic acid. Samples were desalted and concentrated using Zip-Tip™ C18 micro-columns (Millipore). Peptides were eluted with 2 μl of 70% acetonitrile. The final solution was adjusted to contain 50% acetonitrile, 1% acetic acid and then loaded into a gold-coated medium nanoelectrospray needle (Protana).

MS and MS/MS spectra were acquired on a quadrupole time-of-flight tandem mass spectrometer equipped with a Z-spray nano-ESI ion source (Micromass, Manchester, UK). The instrument was calibrated in MS/MS mode using 1 pmol/μl [Glu¹]-Fibrinopeptide B (Sigma) in 50% methanol/1% acetic acid. The capillary voltage used was 1400 V, and the cone voltage was set to 40 V. Tryptic peptide precursor ions (singly, doubly, or triply charged) were selected by the quadrupole mass filter (MS1) and induced to fragment by collision with Argon, and the resulting product ions were analyzed by the time-of-flight mass analyzer (MS2). The collision energy was varied between 15 and 45 V depending on the mass and charge-state of the peptides. The recorded MS/MS spectra were processed using MassLynx MaxEnt 3 (Micromass).

The charged state deconvoluted MS/MS spectra were directly used to search the public data base using the Mascot search program (Matrix Science, UK). For peptides with no exact matches in the public data bases, the MS/MS spectra were sequenced de novo either manually or with the aid of the PepSeq program (Micromass). The amino acid sequences obtained in this way were searched against the NCBI non-redundant protein data base using the BLAST search program (18).

**Expression and Purification of hSMUG1**—The hSMUG1 coding sequence was polymerase chain reaction-amplified as a BamHI-XhoI fragment from a hSMUG1 cDNA clone (ATCC 294606) purchased from ATCC. The polymerase chain reaction product was cloned into...
pGEX6P-1 (Amersham Pharmacia Biotech) so as to express the protein as an N-terminal GST fusion protein for purification on a glutathione column. Expression of the recombinant protein was induced in *E. coli* BL21-DE3 cells by the addition of isopropyl-β-D-thiogalactopyranoside to a final concentration of 0.5 mM. A 1-liter bacterial culture at an *A*$_{600}$ of 0.6 was induced at 30°C for 5 h and harvested by centrifugation at 4°C at 3200 × *g* in a Beckman 6KR centrifuge. The recombinant protein was purified from the induced cell pellet using glutathione-Sepharose 4B affinity medium (Amersham Pharmacia Biotech) per the manufacturer’s protocols for batch purification of fusion proteins. Following purification, the glutathione S-transferase moiety was removed according to the manufacturer’s instructions using Prescission Protease, leaving five amino acid residues remaining on the N terminus of the recombinant hSMUG1.

Assay for Uracil DNA N-Glycosylase Activity by Cleavage of Ura-Containing 2'-Deoxyribosel Oligonucleotides—2'-Deoxyribosel oligonucleotides were synthesized in the core facility of the NYU School of Medicine, Department of Cell Biology. All 2'-deoxyribosel oligonucleo-

### TABLE II

| Peptide sequence$^a$ | Precursor ions | Molecular mass$^b$ |
|----------------------|----------------|------------------|
|                      | *m/z*          | Observed         | Calculated       |
| FWGFFR               | 430.22$^{2+}$  | 858.42           | 858.42           |
| QVQLLGVR             | 456.80$^{2+}$  | 911.58           | 911.56           |
| RPVLGECQPQSEVSAR     | 618.99$^{2+}$  | 1281.75          | 1281.75          |
| NLEGLLLPLIis        | 1282.78$^{2+}$ | 1281.75          | 1281.75          |
| NLCGQFEVVFr          | 683.85$^{2+}$  | 1390.70          | 1390.71          |
| SLAEgFLqEELR         | 696.36$^{2+}$  | 1390.70          | 1390.71          |
| ALAsLMPeVQEGILLHPSPR| 715.40$^{2+}$  | 2143.18          | 2143.15          |
| EQLLGlvCDAALCR       | 752.89$^{2+}$  | 1503.76          | 1503.71          |
| QFSEIvGLoYNPVEYAEPHR| 925.46$^{2+}$  | 2773.36          | 2773.39          |
| (qVlFLGMPMPGMAQTGVFPGEVSvVR| 978.84$^{2+}$| 2933.49          | 2933.49          |
|                      | 1467.79$^{2+}$ | 2933.56          | 2933.49          |
|                      | 1374.74$^{2+}$ | 2747.46          |                 |

$^a$ C represents carboxamidomethylcysteine. Lowercase letters indicate the changed amino acid residues compared to the sequence of hSMUG1 in the data base. The letters in brackets indicate the order could not be determined.

$^b$ All values are reported as monoisotopic mass.
tides were deblocked, deprotected, and purified by 20% denaturing PAGE. A 26-mer uracil-containing 2'-deoxyribose oligonucleotide of the sequence d(CGCGAAACGCCTAGUGATTGGTAGGG) was labeled at the 5'-end using [γ-32P]ATP (PerkinElmer Life Sciences) and T4 polynucleotide kinase (Life Technologies, Inc.). For assays involving DNA duplexes, the 5'-end-labeled 26-mer uracil-containing 2'-deoxyribose oligonucleotide was annealed to the corresponding complementary strand (containing either an A or G at position 12) in a 1:2 molar ratio.

Assays were performed at 37 °C in buffer containing 50 mM HEPES, pH 7.5, 20 mM NaCl, 0.1 mg/ml BSA, 1 mM EDTA, and 1 mM dithiothreitol. Enzyme, protein, and substrate were diluted to working conditions in assay buffer and equilibrated at 37 °C. Assays contained 2 μM 5'-end-labeled single-stranded or duplex 26-mer uracil-containing 2'-deoxyribose oligonucleotide and 20 nM hSMUG1 with or without 0.3 units of uracil DNA N-glycosylase inhibitor (New England Biolabs). The concentration of hSMUG1 was determined using the Bradford assay with albumin as standard and was assigned a molecular weight of 30,142.0 to the fusion protein. The control assays for effectiveness of the uracil DNA N-glycosylase inhibitor contained 0.3 units of uracil DNA N-glycosylase (Life Technologies, Inc.). The enzyme solutions were incubated at 37 °C for 1 min prior to the addition of substrate. The final volume was 10 μl. The reactions were incubated at 37 °C for 60 min and were terminated by snap-freezing in ethanol and dry ice and treated with 0.5 M putrescine, pH 8.0 (19). The treated reaction mixtures were then heated at 95 °C for 5 min followed by the addition of 15 μl of loading dye (95% deionized formamide, 10 mM EDTA, 0.05% bromphenol blue, and 0.05% xylene cyanol). Samples were then heated at 95 °C

Fig. 3. Alignment of peptides from tryptic digests of purified bovine HMUDG with the sequence of hSMUG1. The amino acid sequences were obtained by mass spectroscopic analysis.

Fig. 4. Expression of the cDNA of hSMUG1. The cDNA was isolated, and hSMUG1 was expressed as described under “Experimental Procedures.” An aliquot was analyzed by SDS-PAGE.

Fig. 5. Comparison of the release of 5hmUra from SPO1 DNA by hSMUG1 and purified bovine HMUDG. Indicated amounts of purified recombinant hSMUG1 (black bars) and Fraction VIII (gray bars) of bovine HMUDG were analyzed for the ability to release [3H]5hmUra from [3H]SPO1 DNA.
The activity of renatured protein corresponded to a single Coo-
tered. The greatest amount of HMUDG activity was present in
into fractions (as indicated in Fig. 1). The HMUDG
renatured proteins were assayed for HMUDG activ-
stered with uracil opposite A or G, as shown. The
shown oligonucleotide in the ab-
ence of enzyme. The right two lanes (8 and 9) show the activity of recombinant
UDG in the presence or absence of UGI.

Renaturation of HMUDG
—

increase in specific enzyme activity was somewhere between
5-fold from that of the crude homo-
activity of HMUDG by 3–
the addition of NaCl. This step actually increased the specific
out of the gelatinous nucleoprotein precipitate resulting from
homogenate, so the final degree of purification (as measured by an
increase in specific enzyme activity) was somewhere between
1500- and 2500-fold.

Results

Purification of HMUDG from Calf Thymus—A mammalian
protein with HMUDG activity was purified from fresh calf
thymus on the basis of the ability to excise \(^{3}H\)5hmUra from
\[^{3}H\]SP01 DNA, as shown in Table I. Calf thymus was used as
the source of enzyme because large amounts of fresh tissue are
available, and our previous studies revealed that the specific
HMUDG activity of thymus was higher than any other tissue
except brain (12). The purification scheme was similar to that
used previously to purify human endonuclease III homologue 1
(15), with the addition of an (NH\(_{4}\))\(_{2}\)SO\(_{4}\) concentration step.

Experimental Procedures

The recorded MS/MS spectra were used for a data base search
(Table II), were found to match or nearly
match sequences from hSMUG1. The alignment of these bovine
peptide ions in Fig. 1
—

Alignment with hSMUG1 Amino Acid Sequence

FIG. 6. Activity of hSMUG1 against uracil containing single- and double-
stranded DNA containing matched and mismatched uracil. Recombinant
hSMUG1 was assayed for the ability to cleave uracil containing 2'-deoxyribose o-
ligonucleotides in the presence or absence of UGI (lanes 2–7). The oligonucleotides
were either single- or double-stranded with uracil opposite A or G, as shown. The
left lane shows oligonucleotide in the absence of enzyme. The right two lanes (8
and 9) show the activity of recombinant UD
observed in Fig. 2A were significantly weaker or absent in the
nano-ESI spectra of other gel slices (data not shown), indicat-
ing that these peptides correlated with HMUDG activity.
Therefore, these peptides were further fragmented by collision-
induced dissociation to generate sequences or sequence tags.
The recorded MS/MS spectra were used for a database search
(see “Experimental Procedures”). To illustrate this identifica-
tion process, the MS/MS spectrum of the doubly charged pep-
tides at \(m/z\) 696.36 is shown in Fig. 2B. The complete \(y\) ion
series (\(y_1–y_{14}\)) could be assigned, allowing the determination of
the full peptide sequence as SLAEGFLQELR (low collision
energy MS/MS analysis as used here cannot distinguish be-
tween the isobaric Ile and Leu amino acid residues).

Direct data base searching with the uninterpreted MS/MS
spectrum did not result in a positive protein identification.
However, searching the NCBI protein data base using BLAST
revealed that the bovine sequence nearly matched a peptide
sequence from human SMUG1, a recently identified member of a
family of uracil DNA N-glycosylases. Two changed amino
acids were Ser\( \rightarrow \)Gly(G) and Glu\( \rightarrow \)Ghs(Q). The probability that
this identification was correct was high because a total of 13
peptide ions in Fig. 1.A, corresponding to 11 peptide sequences
obtained by MS/MS (Table II), were found to match or nearly
match sequences from hSMUG1. The alignment of these bovine
sequences with the hSMUG1 sequence is shown in Fig. 3.

Isolation of the cDNA of hSMUG1 and Expression of the
Protein—The human SMUG1 cDNA was isolated and ex-
pressed as a GST fusion protein, which was purified and the
GST moiety removed, leaving 5 amino acids on the N terminus
of the recombinant hSMUG1. When analyzed by PAGE, the
purified recombinant protein was visualized as a single band
(Fig. 4) with an apparent molecular mass slightly larger than
that of the region of the gel where the maximal activity of
purified bovine HMUDG was recovered (Fig. 3).

Release of 5hmUra by hSMUG1 Compared with That of
Purified Bovine HMUDG (Fraction VIII)—Purified expressed
hSMUG1 was assayed for its ability to release \(^{3}H\)5hmUra from
\[^{3}H\]SP01. At the concentrations assayed, there was a
linear increase of release of 5hmUra with the increase in
hSMUG1 concentration (Fig. 5). The identity of the released
material was confirmed by HPLC. The specific activity of pu-
rified recombinant hSMUG1 was \(20\times\) greater than the most
purified fraction (Fraction VIII) of bovine HMUDG.

Confirmation of Uracil DNA N-Glycosylase Activity of Re-
combinant hSMUG1—Recombinant hSMUG1 was analyzed for
its ability to remove Ura from single- and double-stranded
2’-deoxyribose oligonucleotides containing a single Ura residue

for 5 min, and products were separated by 20% PAGE in 7 \(\times\) urea and
Tris borate/EDTA (TBE). Products were analyzed via phosphorimaging
using a Molecular Imager FX System with Quantity One Software
(Bio-Rad) (20).

Determination of Bovine HMUDG Amino Acid Sequence and
Alignment with hSMUG1 Amino Acid Sequence—The nano-
ESI mass spectrum (\(m/z\) range of 300–1600) of peptides from
slice 2 that showed the highest HMUDG activity is shown in
Fig. 2A. In this complex spectrum, more than 150 singly, dou-
bly, or triply charged tryptic peptides were observed, indicating
that a mixture of proteins was present in this gel slice. In the
MS survey, we found that the signals of a number of peptides

U
I
G
I
- 
- 
- 
- 
- 
- 
- 
- 
substrate 
ssU 
ssU 
ssU 
U:A 
U:A 
U:G 
U:G 
U:A 
U:A

Downloaded from http://www.jbc.org/ by guest on July 25, 2018
(Fig. 6). The recombinant protein showed maximal activity in a cleavage assay against single-stranded DNA. hSMUG1 also displayed activity against double-stranded DNA containing a single Ura residue. As reported by Haushalter et al. (14), hSMUG1 activity was greater against mismatches (U:G) than against matches (U:A).

The activity of the recombinant hSMUG1 was distinct from UDG (in this case recombinant bacterial UDG). UDG had greater activity against double-stranded DNA containing U opposite A (Fig. 6, lane 7). Furthermore, UDG activity was suppressed by UGI in contrast to hSMUG1 activity, which was not (Fig. 6, lane 8).

**DISCUSSION**

SMUG1 was identified by binding to synthetic uracil DNA N'-glycosylase inhibitors and has been definitively shown to be distinct from the primary uracil DNA N'-glycosylase TDG which preferentially repairs U:G and T:G mismatches (14). In this study we have shown that a bovine protein with HMUDG activity, as defined previously by Hollstein et al. (8) and this group (9, 12), has peptide sequences homologous to those of hSMUG1, whereas recombinant hSMUG1 is shown here for the first time to have HMUDG activity.

Previous observations of HMUDG activity in cells and tissues can now be reevaluated in light of the determination of the identity of HMUDG as SMUG1. The HMUDG activity was first searched for in bacteria, where it could not be found (10), and it was subsequently shown to be absent in yeast but present in more complex multicellular organisms (11). This unique phylogenetic distribution has been confirmed for SMUG1 (14) in contrast to UDG, which is highly conserved throughout phylogeny. The uracil DNA N'-glycosylase activity of SMUG1 is not inhibited by the phage UGI. Our most purified fractions of bovine HMUDG also displayed uracil DNA N'-glycosylase activity that was not inhibited by UGI (data not shown), further confirming the identity of HMUDG as SMUG1. Additional support for the identity of HMUDG as SMUG1 comes from the finding that SMUG1 ESTs are present in a wide variety of tissues (14, 22), consistent with our previous observation that HMUDG activity is present in all tissues, and is most abundant in brain and thymus (11).

The biological role of HMUDG/SMUG1 is still uncertain. Haushalter and co-workers (14, 22) speculated that SMUG1 might play a role in the repair of Ura residues resulting from deamination of Cyt residues in DNA that is transiently single-stranded during transcription, recombination, or replication. Recent work by the same group has suggested that SMUG1 functions as an antimutator, with preference for repair of Ura opposite guanine (22).

We previously proposed that HMUDG might repair 5hmUra residues arising from oxidation and deamination of 5meCyt residues in DNA that, if un repaired, would be mutagenic. Such a hypothesis is supported by the common phylogenetic distributions of the use of 5meCyt as a regulatory element in transcription, HMUDG enzyme activity, and SMUG1 EST sequences. Recent evidence further supports this hypothesis. It was shown that 5hmUra, as part of 5hmUra:G mismatches, was repaired to a significantly greater extent than 5hmUra:A in identical sequences using partially purified HMUDG from Hela cells (23).

Regardless of which of the two putative functions is of greater significance to the maintenance of genomic integrity by HMUDG/SMUG1, it should be noted that a mutant Syrian hamster cell line (V79) deficient in HMUDG activity, which we developed, is viable and grows normally (24). Although it seems evident that HMUDG/SMUG1 protects cells from mutations arising from conversion of 5meCyt to 5hmUra and deamination of Cyt to Ura, the possibility remains that HMUDG/SMUG1 is also involved in a sequence of controlled oxidation of 5meCyt followed by deamination and removal of 5hmUra as a mechanism of changing the methylation status of specific regions of DNA.

Full understanding of the biological significance of HMUDG/SMUG1 will require an explanation of both its unique in vitro substrate specificities and its phylogenetic distribution.

**Acknowledgment—**We thank Dr. Lawrence C. Sowers for valuable discussions and remarkable insight.

**REFERENCES**

1. Cerutti, P. A. (1976) in Photochemistry and Photobiology of Nucleic Acids (Wang, S. Y., ed) pp. 375–401, Academic Press, New York
2. Faure, H., Mousseau, M., Cadet, J., Guimier, C., Tripier, M., Hida, H., and Favier, A. (1998) Free Radic. Res. 28, 377–382
3. Pourcelot, S., Faure, H., Fivizzani, P., Ducros, V., Tripier, M., Hee, J., Cadet, J., and Favier, A. (1999) Free Radic. Res. 30, 173–180
4. Djuric, Z., Ilijin, L. K., Reading, B. A., Boomer, A., Valeriote, F. A., and Martino, S. (1991) Natl. Cancer Inst. 83, 766–769
5. Frenkel, K., Karkozyka, J., Glassman, T., Dubin, N., Tenani, P., Taieb, E., Moorey, L. A., and Kato, I. (1998) Cancer Epidemiol. Biomark. Prev. 7, 49–57
6. Kallen, R. G., Simon, M., and Marmur, J. (1962) J. Biol. Chem. 245, 248–250
7. Herzog, M., Seyer, M. O., and Denay de Marcillac, G. (1992) J. Cell Biol. 105, 151–155
8. Hollstein, M. C., Brooks, P. L., and Ames, B. N. (1984) Proc. Natl. Acad. Sci. U.S.A. 81, 4003–4007
9. Boorstein, R. J., Levy, D. D., and Teever, G. W. (1987) Mutat. Res. 183, 257–263
10. Friedman, E. C., Ganesan, A. K., and Minton, K. (1975) J. Biol. Chem. 250, 13306–13312
11. Balts, R. H., Bingham, P. M., and Drake, J. W. (1976) Proc. Natl. Acad. Sci. U. S. A. 73, 1289–1273
12. Haushalter, K. A., Todd Stukenberg, M. W., Kirschner, M. W., and Verdine, G. L. (1999) J. Biol. Chem. 274, 4527–4533
13. Shevchenko, A., Wilm, M., Mann, O., and Mann, M. (1999) Anal. Chem. 71, 858–868
14. Alschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Nucleic Acids Res. 25, 3389–3402
15. Zharkov, D. O., Rosenquist, T. A., Gerchman, S. E., and Grollman, A. P. (2000) EMBO J. 19, 4278–4286
16. Boorstein, R. J., Chung, L. N., and Teever, G. W. (1985) J. Biol. Chem. 260, 13306–13312
17. Ball, R. P., and Teever, G. W. (1996) Biochemistry 35, 2503–2511
18. Frenkel, K., Cummings, A., Solomon, J., Cadet, J., Steinberg, J. J., and Teever, G. W. (1998) Biochemistry 27, 4527–4533
19. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Nucleic Acids Res. 25, 3389–3402
20. Zharkov, D. O., Rosenquist, T. A., Gerchman, S. E., and Grollman, A. P. (2000) J. Biol. Chem. 275, 28607–28617
21. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Nucleic Acids Res. 25, 3389–3402
22. Zharkov, D. O., Rosenquist, T. A., Gerchman, S. E., and Grollman, A. P. (2000) J. Biol. Chem. 275, 28607–28617
23. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Nucleic Acids Res. 25, 3389–3402
24. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Nucleic Acids Res. 25, 3389–3402
