

# Identification and characterization of a human DNA glycosylase for repair of modified bases in oxidatively damaged DNA

Tapas K. Hazra\*, Tadahide Izumi\*, Istvan Boldogh<sup>†</sup>, Barry Imhoff<sup>‡</sup>, Yoke W. Kow<sup>‡</sup>, Pawel Jaruga<sup>§</sup>, Miral Dizdaroglu<sup>§</sup>, and Sankar Mitra\*<sup>¶1</sup>

\*Sealy Center for Molecular Science and Department of Human Biological Chemistry and Genetics, <sup>†</sup>Department of Microbiology and Immunology, University of Texas Medical Branch, Galveston, TX 77555; <sup>‡</sup>Department of Radiation Oncology, Emory University School of Medicine, Atlanta, GA 30335; and <sup>§</sup>Chemical Science and Technology Laboratory, National Institute of Standards and Technology, Gaithersburg, MD 20899-8311

Communicated by Sankar Adhya, National Institutes of Health, Bethesda, MD, January 30, 2002 (received for review December 2, 2001)

**8-oxoguanine (8-oxoG), ring-opened purines (formamidopyrimidines or Fapys), and other oxidized DNA base lesions generated by reactive oxygen species are often mutagenic and toxic, and have been implicated in the etiology of many diseases, including cancer, and in aging. Repair of these lesions in all organisms occurs primarily via the DNA base excision repair pathway, initiated with their excision by DNA glycosylase/AP lyases, which are of two classes. One class utilizes an internal Lys residue as the active site nucleophile, and includes *Escherichia coli* Nth and both known mammalian DNA glycosylase/AP lyases, namely, OGG1 and NTH1. *E. coli* MutM and its paralog Nei, which comprise the second class, use N-terminal Pro as the active site. Here, we report the presence of two human orthologs of *E. coli* mutM nei genes in the human genome database, and characterize one of their products. Based on the substrate preference, we have named it NEH1 (Nei homolog). The 44-kDa, wild-type recombinant NEH1, purified to homogeneity from *E. coli*, excises Fapys from damaged DNA, and oxidized pyrimidines and 8-oxoG from oligodeoxynucleotides. Inactivation of the enzyme because of either deletion of N-terminal Pro or Histag fusion at the N terminus supports the role of N-terminal Pro as its active site. The tissue-specific levels of NEH1 and OGG1 mRNAs are distinct, and S phase-specific increase in NEH1 at both RNA and protein levels suggests that NEH1 is involved in replication-associated repair of oxidized bases.**

**R**eactive oxygen species (ROS), including the O<sub>2</sub><sup>-</sup> radical, H<sub>2</sub>O<sub>2</sub>, and the OH<sup>•</sup> radical, may be the most important genotoxic agents because of their ubiquitous and continuous production as by-products of respiration (1). They are also generated in mammalian cells during the inflammatory response, and have been implicated in the etiology of a variety of pathophysiologies ranging from cancer to rheumatoid arthritis, cardiovascular diseases, and Alzheimer's disease, and also in aging (2–4). Genotoxicity of ROS results from their reaction with DNA to produce a plethora of DNA base lesions as well as strand breaks (5). All of these lesions except the double-strand breaks are repaired primarily via the base excision repair (BER) pathway in both bacteria and eukaryotes (6). BER is initiated with excision of oxidatively damaged bases by several DNA glycosylase/AP lyases. These glycosylases have broad substrate ranges, commensurate with the need to excise lesions of widely different chemical structures by using a limited number of enzymes. However, these enzymes have preferred substrates. Thus 8-oxoguanine (8-oxoG), an abundant and arguably the most critical mutagenic lesion, is repaired primarily by 8-oxoG-DNA glycosylase (OGG1) in mammalian cells and by MutM/Fpg in *Escherichia coli*. Formamidopyrimidines (Fapys), derived from purines by ROS, are more abundant than 8-oxoG in normal mammalian DNA (7). Although not extensively investigated, Fapys, which are likely to inhibit replication (8) and may be mutagenic, are also excised by OGG1 and other DNA glycosylases. A number of pyrimidine lesions, including 5,6 dihydrouracil (DHU) and 5-hydroxyuracil (5-OHU), generated from cytosine by ionizing radiation and ROS, are excised primarily by *E. coli* endonuclease III (Nth) and its

mammalian ortholog NTH1, and also by *E. coli* endonuclease VIII (Nei; refs. 5, 9, and 10). DNA glycosylase/AP lyases can be divided into two classes. The two mammalian enzymes, specific for oxidized base lesions and identified so far, namely, NTH1 and OGG1, belong to the *E. coli* Nth class, based on similar structural features and reaction mechanism (11). These enzymes use an internal Lys residue as the active site nucleophile, and carry out  $\beta$  elimination (after excising the damaged base), to cleave the DNA strand, and generate 3' phospho  $\alpha,\beta$ -unsaturated aldehyde and 5' phosphate termini (11, 12). Only prokaryotic MutM/Fpg and its paralog Nei comprise the second class, and use N-terminal Pro as the active site nucleophile. These enzymes carry out  $\beta\delta$ -elimination and cleave the DNA strand at the site of lesion to generate 3' phosphate and 5' phosphate termini (13). Both classes of enzymes form covalent Schiff base intermediates with the lesion-containing DNA strand, both before and after its cleavage, and these unstable intermediates can be converted into stable "trapped complexes" by reduction with NaCNBH<sub>3</sub>. With oligo substrates, the mobility of these complexes in SDS/PAGE depends on the relative size of the enzyme.

MutM and human OGG1 (hOGG1) have no obvious structural similarity despite their common substrate preference for excising 8-oxoG from an 8-oxoG-C pair. Their activity toward Fapys is more complex. Whereas MutM cleaves both G-derived Fapy (FapyG) and A-derived Fapy (FapyA), human OGG1 (hOGG1) excises only FapyG (14–16). The teleological reason for evolution of several DNA glycosylases with overlapping substrate range is not clear. It is possible that they have distinct roles in *in vivo* repair. For example, oxidized base lesions have now been shown to be preferentially repaired in the transcribed strand in mammalian cells (17, 18). This "transcription-coupled repair" (TCR) was discovered for DNA bulky adducts, which are repaired via the nucleotide excision repair (NER) pathway in which XPG functions as an essential endonuclease (18, 19). TCR was also observed for thymine glycol (Tg) in human cells, a substrate of Nth, Nei, and human NTH1 that requires the presence of the XPG protein but not its endonuclease activity (20). Thus TCR of Tg is likely to use the base excision repair and not the NER pathway. Evidence for TCR of 8-oxoG in human cells was elegantly provided by Le Page *et al.* (21). Accumulation of 8-oxoG in DNA extracted from Ogg1 (-/-) mouse cells and tissues indicates that OGG1 is the major repair enzyme for 8-oxoG

Abbreviations: DHU, dihydrouracil; 8-oxoG, 8-oxoguanine; Fapy, formamidopyrimidine; 5-OHU, 5-hydroxyuracil; hOGG1, human 8-oxoguanine-DNA glycosylase 1; NEH, Nei homolog; RAR, replication-associated repair; TCR, transcription-coupled repair; ROS, reactive oxygen species; WT, wild type.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database [accession nos. AAH10876 (NEH1) and XL055431 (NEH2)].

<sup>¶1</sup>To whom reprint requests should be addressed at: Sealy Center for Molecular Science and Department of Human Biological Chemistry and Genetics, University of Texas Medical Branch, 6.136 Medical Research Building, Galveston, TX 77555. E-mail: samitra@utmb.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

(22, 23). However, complete absence of 8-oxoG removal from the nontranscribed strand but proficient and nearly normal removal of 8-oxoG from the transcribed sequences in Ogg1 (-/-) cells indicated TCR of 8-oxoG in mouse cells as well (24). These results are consistent with either the presence of a distinct DNA glycosylase that functions primarily in TCR, or removal of this lesion during TCR by an altogether different mechanism. Although TCR of Fapys has not been demonstrated, it is likely that Fapys, like 8-oxoG and Tg, are also subject to TCR.

We identified and partially characterized a second OGG (OGG2), in HeLa cells. It excises 8-oxoG preferentially from 8-oxoG-G or 8-oxoG-A pairs (12). In contrast, the major substrate for both OGG1 and MutM is the 8-oxoG-C pair (12, 25). We showed subsequently that *E. coli* Nei also has OGG activity and its substrate preference mirrors that of human OGG2 (26). These observations led us to propose that OGG2 preferentially excises 8-oxoG when it is misincorporated in the nascent strand opposite A (12). Similar replication-associated repair (RAR) also occurs for normal bases, e.g., A or G, when misincorporated opposite 8-oxoG. Such repair is mediated by MYH in human cells, an ortholog of *E. coli* MutY (27, 28).

Searching human genomic databases, we identified two candidate DNA glycosylases with sequence homology to MutM and Nei. In this report, we describe characterization of one of these DNA glycosylases, and discuss its potential role in repair of oxidatively damaged bases via RAR and TCR processes.

## Materials and Methods

**Human Genome Database Analysis and Molecular Cloning of NEH1.** The NCBI and Celera Genomic Databases were searched by using the BLASTP program to identify human cDNA clones with significant homology to *E. coli* MutM or Nei. Two putative clones of human genes along with their mouse homologs were identified that we tentatively named NEH1 and NEH2 (Nei Homolog). The cDNA for human NEH1 (hNEH1; accession no. AAH10876), obtained from Research Genetics (Huntsville, AL), was inserted between the *Nde*I and *Xho*I sites of pRSETB plasmid, and its identity was confirmed by sequence analysis. Predicting that N-terminal Pro is the catalytic site (after cleavage of the initiator Met residue), we constructed expression plasmids encoding the Pro 1 deletion mutant of NEH1, and its N-terminal Histag fusion by using PCR-mediated site-specific mutagenesis.

**Purification of NEH1.** Because plasmid-encoded expression of hNEH1 was low in *mutM nei E. coli* DE884, we introduced the recombinant plasmid into *E. coli* BL21 (DE3) codon plus (Stratagene) for large-scale purification of the enzyme after its induction with 0.2 mM isopropyl  $\beta$ -thiogalactoside at 16°C for 16 h. We purified wild type (WT) NEH1 from the sonicated extract of log-phase bacteria in a series of steps starting with the removal of nucleic acids by Polymin P, followed by enzyme precipitation with 60% saturation ammonium sulfate. After dissolving the precipitate in a minimum volume of buffer A (25 mM Tris-HCl, pH 7.5/10% glycerol/1 mM DTT/0.1 mM EDTA) containing 150 mM NaCl, and chromatography in a Superdex 75 column, fractions containing NEH1 were adjusted to 50 mM NaCl, and then applied to a 5 ml HiTrap Q and SP columns (Amersham Pharmacia) connected in tandem. After a wash with the starting buffer, the Q column was disconnected, and the proteins were eluted from the SP column at 500 mM NaCl with a 100-ml linear gradient of 50 mM to 1 M NaCl. The final step involved chromatography on a 1-ml Mono S column, and NEH1 was eluted without detectable impurity at 500 mM NaCl. The active fractions were frozen in liquid nitrogen and stored at -80°C. The final yield of the protein was about 5 mg per liter of culture.

**Enzymatic Assay of NEH1 by GC-MS Analysis.** The release of modified bases by NEH1 from irradiated calf thymus DNA was assayed as

described earlier (29). In short, 100  $\mu$ g of irradiated calf thymus DNA was incubated with or without the enzyme at 37°C for 30 min in 100  $\mu$ l of 50 mM phosphate buffer (pH 7.4), containing 100 mM KCl, 1 mM EDTA, and 0.1 mM DTT. Aliquots of stable isotope-labeled analogs of modified bases were added, and the samples were treated and analyzed by GC/MS as previously described (29).

**Incision Assay with Duplex Oligo Substrates.** We used the standard procedure for assaying DNA glycosylase/AP lyase activity of NEH1 with <sup>32</sup>P-labeled duplex oligo substrates. A 31-mer oligo containing either 8-oxoG or DHU at position 16 and <sup>32</sup>P-labeled at the 5' terminus (12) was annealed with the complementary strand, which contained either A, G, C, or T opposite the lesion site. In a different experiment, we used a 32-mer oligo containing 5-OHU in the sequence, 5'-TTCCAGACTGTCCTTCGTXACTTTCCTCTCAA-3', where X is 5-OHU and the complementary strand had G opposite this lesion. The <sup>32</sup>P-labeled duplex oligo (4 pmol) was incubated with 0.5 pmol NEH1 at 37°C in 20  $\mu$ l mixture containing 20 mM Tris-HCl (pH 7.5), 50 mM NaCl, 0.1 mM EDTA, and 100  $\mu$ g/ml BSA for 20 min, unless otherwise stated. After terminating the reaction with a stop buffer, the substrates and the cleaved products were separated by denaturing gel electrophoresis in 15% polyacrylamide containing 7 M urea in 90 mM Tris-borate (pH 8.3) and 2 mM EDTA and detected by PhosphorImager (Molecular Dynamics) analysis. NEH1-cleaved products were identified by comparing with those generated by *E. coli* Nth, Nei, and MutM.

**Analysis of Trapped Complexes of NEH1.** The <sup>32</sup>P-labeled duplex oligo substrate (100 fmol) was incubated with 1  $\mu$ g crude extract in 20  $\mu$ l assay buffer in the presence of 25 mM NaCNBH<sub>3</sub> at 37°C for 30 min, and the trapped complexes were separated by SDS/PAGE (12% polyacrylamide; ref. 12).

**Cell Synchronization and Analysis of Cell Cycle.** Primary human diploid MRC5 fibroblasts were synchronized by serum starvation for 120 h, and then stimulated to proliferate by adding FBS (10%) to the Eagle's minimal essential medium (EMEM) culture medium. The cells were harvested at various times; the aliquots were used for cell cycle analysis, and for preparing lysates for blot analysis of NEH1 mRNA and protein (30).

**Blot Analysis for NEH1.** Nitrocellulose membranes containing 2  $\mu$ g poly(A)<sup>+</sup> RNA from various human tissues (CLONTECH) or 50  $\mu$ g total RNA from synchronized MRC5 cells were probed with <sup>32</sup>P-labeled NEH1 cDNA or an 18S ribosomal RNA gene probe following standard protocol (31). Forty micrograms total cell lysate proteins were used for immunoblot analysis with anti-NEH1 or anti-tubulin antibody (Santa Cruz Biotechnology) by using the enhanced chemiluminescence system (Amersham Pharmacia; ref. 12).

**Other Reagents.** *E. coli* MutM, Nei and Nth, and human OGG1 and NTH1, used as references in this study, were purified to apparent homogeneity (26, 32, 33). Purified hNEH1 was used for raising polyclonal antibodies in rabbits (Alpha Diagnostic, San Antonio, TX).

## Results

**Identification of Mammalian Orthologs of *E. coli* MutM/Nei.** Fig. 1 shows comparative sequence homology of hNEH1 and hNEH2, with *E. coli* MutM and Nei as analyzed by CLUSTALW. Although the overall identity between the human NEH1 and MutM is only 14.6% (57 identical residues of 390 residues in NEH1, and 14.9% between NEH1 and Nei), key residues, in particular the N-terminal PE (L/G) PE motif, are completely conserved in NEH1, along with an internal Lys residue (Lys-56 in MutM; ref. 34). Furthermore, NEH1 shares a potential helix-2-turn-helix (H2TH) motif with MutM and



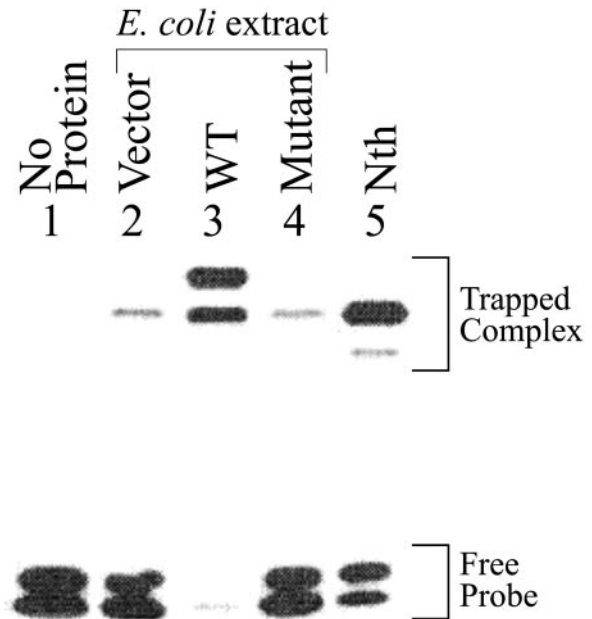
**Fig. 1.** Sequence alignment of hNEH1 and hNEH2 with *E. coli* MutM/Nei. Identical residues are underlined. Box A contains catalytic N-terminal Pro; box B, helix-2-turn-helix motif; box C, Zn-finger motif. Potential nuclear localization signal (NLS; by pSort) in NEH1 is underlined. N-terminal Met is cleaved after synthesis.

Nei (35). However, unlike in MutM, the absence of conserved cysteines suggests the absence of a Zn-finger motif in NEH1. It is thus expected that Pro-1 (after proteolytic cleavage of the initiator Met) is the active site in NEH1, which, like MutM/Nei, should carry out  $\beta\delta$  elimination. Several coding sequences for hNEH1, deposited in GenBank, show two polymorphic sites in nonconserved regions, with Ser or Asn at position 147 and Arg or Lys at position 242. It is not known whether NEH1 activity is affected by polymorphic substitution. Our clone (accession no. AAH10876) contains Asn and Arg at positions 147 and 242 respectively. In any event, we decided to investigate the DNA glycosylase activity of NEH1.

**Chromosomal Map and Predicted Size of NEH1 and NEH2 of Humans and Mice.** Table 1 summarizes the information gleaned from the genomic database on chromosomal maps of NEH1 and NEH2 in the human and mouse genomes. No information is yet available about the potential linkage of these genes to any known disease. HNEH1 and hNEH2, with predicted molecular masses of 44 kDa

**Table 1. Protein size and chromosomal location of human and mouse NEHs**

Size of human gene	Species	Protein size, kDa	Chromosomal location
NEH1	Human	43.7	15q 25
	Mouse		9a5.3
NEH2	Human	36.8	4q35
	Mouse		8b.1.2

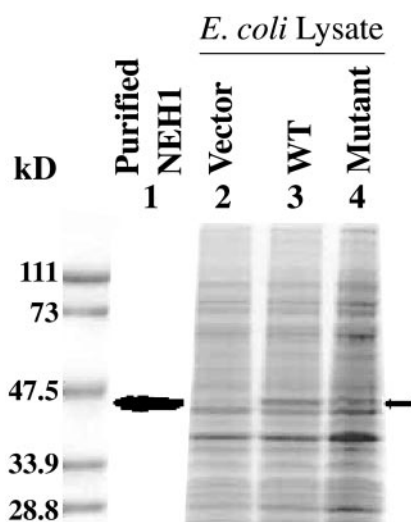


**Fig. 2.** Trapping analysis of recombinant NEH1 in extracts of *mutM nei E. coli*. Lane 1, no protein; lanes 2–4, trapping assay with lysates (1  $\mu$ g) of *E. coli* containing empty vector, WT NEH1, and Pro 1 deletion mutant, after incubation with 0.1 pmol of DHU-A-containing duplex oligo; lane 5, trapped complex of purified Nth (5 ng) used as a reference with 0.1 pmol of DHU-A. The positions of trapped complexes and free DNA are indicated.

and 36 kDa, respectively, are larger than the 30-kDa MutM or Nei, but are comparable in size to other mammalian DNA glycosylases.

**Detection of DNA Glycosylase Activity of Recombinant NEH1 by Trapping Analysis.** We tested DNA glycosylase/AP lyase activity of WT NEH1 and its Pro-1 deletion mutant in crude extracts of *mutM nei E. coli* harboring NEH1 expression plasmids. Fig. 2 shows SDS/PAGE of  $^{32}$ P-labeled trapped complexes generated by using extracts of *mutM nei E. coli* with or without expression of hNEH1. Because the mobility of such complexes reflects the size of the DNA glycosylase when the same oligo substrate is used (12), it is evident that, in control *E. coli* lacking MutM and Nei, only endogenous Nth formed a major trapped complex with the uncleaved oligo (lane 2). A minor species of smaller size was observed with the Nth standard, which corresponds to its complex with the cleaved oligo (lane 5). The WT NEH1 similarly forms two trapped complexes with uncleaved and cleaved oligos, respectively (lane 3). Coincidentally, the mobility of the NEH1 complex with the cleaved oligo was similar to that of the uncleaved oligo complex of Nth. Lane 4 shows that no NEH1-specific trapped complex was formed with the Pro-1 deletion mutant even though its level in the *E. coli* lysate was comparable to that of the WT protein as indicated by Western analysis (data not shown). Furthermore, a significant increase in strand incision activity with the DHU-A pair was observed in the sonicate of *mutM nei E. coli* expressing WT NEH1, but not the Pro-1 mutant, over low basal level (because of endogenous Nth) present in the extract of cells harboring the empty vector (data not shown). The requirement of N-terminal Pro for activity was further confirmed by the observation that recombinant NEH1 with N-terminal Histag fusion was inactive (data not shown). Taken together, these results show that recombinant NEH1 is active as a DNA glycosylase/AP lyase with the DHU oligo substrate, and that Pro-1 is likely to be its active site.

**Purification and Properties of hNEH1.** SDS/PAGE of extracts of *E. coli* codon plus harboring WT NEH1 and its Pro 1 mutant

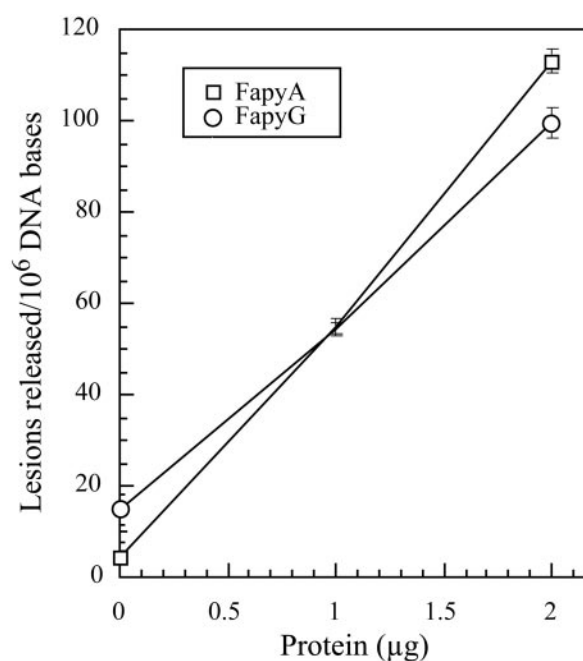


**Fig. 3.** Expression and purification of NEH1. Lysates (15  $\mu$ g) of *E. coli* expressing NEH1 were analyzed by SDS/PAGE and Coomassie blue staining. The Bio-Rad protein markers are shown on the left. Purified WT NEH1 (lane 1); lanes 2–4, lysates of *E. coli* with empty vector, WT, and Pro-1 mutant NEH1 expression plasmid, respectively. The position of NEH1 is indicated by arrow.

expression plasmids showed significant levels of expression of these proteins, most of which was present in the soluble form (Fig. 3, lanes 2–4). The WT NEH1 was then purified to apparent homogeneity by using conventional steps (Fig. 3, lane 1). The N-terminal sequence of this protein, Pro-Glu-Gly-Pro-Glu-Leu-His-Leu-Ala-Ser, matched perfectly with its predicted sequence (after cleavage of the initiator Met residue). Its molecular weight (43,633 as determined by MS analysis) was close to the predicted value of 43,582.

**DNA Glycosylase Activity of hNEH1 with Irradiated Natural Duplex DNA Substrate.** By using GC/MS, we examined the substrate range of NEH1 with irradiated DNA, which contained a wide variety of modified bases (29). Irradiated DNA used for these experiments contained FapyA, FapyG, and 8-oxoG at levels of  $508 \pm 37$ ,  $1825 \pm 180$ , and  $915 \pm 55$  lesions per  $10^6$  DNA bases, respectively. MutM/Fpg, used as the control, excised FapyA and FapyG, as well as 8-oxoG, from the same DNA, as expected (14, 16). Surprisingly, only FapyA and FapyG were found to be excised by hNEH1 (Fig. 4). The release of  $173 \pm 10$ ,  $721 \pm 5$ , and  $839 \pm 1$  of FapyA, FapyG, and 8-oxoG per  $10^6$  DNA bases, respectively, per microgram of 30-kDa MutM under the same conditions indicates that NEH1 excises FapyA and FapyG about 2- and 6-fold less efficiently than MutM on an equimolar basis.

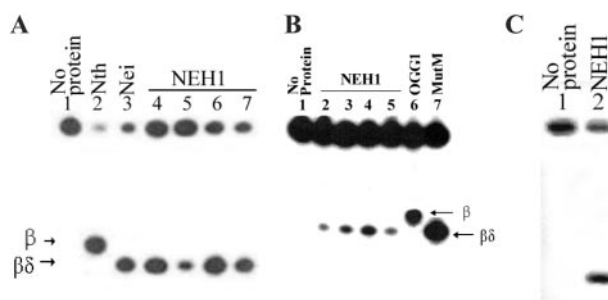
**Analysis of NEH1 Activity with Duplex Oligo Substrates.** The sensitivity of NEH1 assay with irradiated DNA substrate containing low levels of multiple lesions is much less than with oligo substrates containing specific lesions. Furthermore, the effect of complementary base on the activity of DNA glycosylases could not be investigated in natural DNA. We, therefore, examined the activity of NEH1 for 8-oxoG and DHU present in 31-mer duplex oligos that had identical sequence, except for the lesion at position 16 (12, 26). For measuring relative reaction rates with various substrates, we assayed with excess substrates in the linear kinetic range. Fig. 5 shows that, under these conditions, NEH1 incised both DHU- and 8-oxoG-containing oligos at the lesion site, although its activity was higher for the pyrimidine lesion. The effect of the opposite base on NEH1 activity was also investigated. Incision of the DHU-containing strand was similar with A, C, or G opposite the lesion, and is reduced when T is present (Fig. 5A). The effect of different



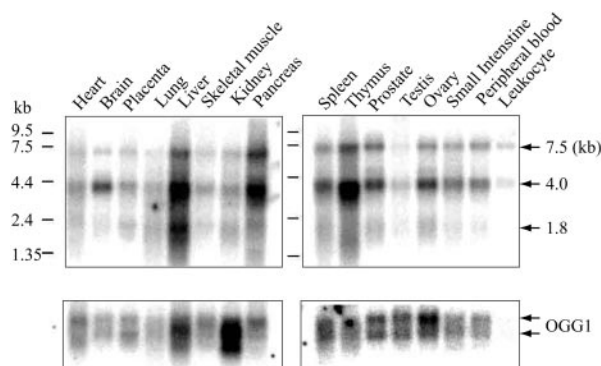
**Fig. 4.** Excision of FapyA ( $\square$ ) and FapyG ( $\circ$ ) by hNEH1 from irradiated DNA. Other details are described in *Materials and Methods*.

base pairs on hNEH1 incision activity for 8-oxoG was more pronounced (Fig. 5B). The enzyme was most active with the 8-oxoG-G pair (lane 4) followed by the 8-oxoG-T pair (lane 3) and had weak activity with 8-oxoG-C (lane 5) and 8-oxoG-A pairs (lane 2). Similar studies showed that hOGG1 was most active with the 8-oxoG-C pair, and the activity was barely detectable with the 8-oxoG-A pair (12, 36). NEH1 was also active in incising 5-OHU-strand from a duplex oligo when the lesion is paired with G (Fig. 5C). Thus, NEH1 mimics Nei more closely than MutM with respect to substrate preferences (26).

*E. coli* MutM and Nei differ from other oxidized base lesion-specific DNA glycosylase/AP lyases because they generate 3' phosphate termini after  $\beta\delta$  elimination (13, 26). A comparison of the mobility of NEH1 digestion products with those of OGG1, and *E. coli* MutM, Nei, and Nth from 8-oxoG- or DHU-containing oligo substrate of the same sequences shows that NEH1 carries out  $\beta\delta$



**Fig. 5.** Substrate specificity and AP lyase activity of purified hNEH1. (A) Purified NEH1 (0.5 pmol) was incubated with  $5'$   $^{32}$ P-labeled DHU-containing duplex oligo (4 pmol) with A, T, G, and C in the complementary strand (lanes 4–7) opposite DHU. Lanes 2 and 3, Incision activity of Nth and Nei, respectively, with DHU-A oligo. (B) Incision activity of NEH1 with 8-oxoG-A-, 8-oxoG-T-, 8-oxoG-G-, and 8-oxoG-C-containing oligos, respectively (lanes 2–5). Incision activity of 8-oxoG-C with hOGG1 and MutM (lanes 6 and 7). Positions of  $\beta$  and  $\beta\delta$ -elimination products are indicated. (C) Cleavage of 5-OHU-containing oligo (100 fmol) by NEH1. Lane 1, No enzyme; lane 2, 1 pmol NEH1.



**Fig. 6.** Tissue-specific expression of NEH1. The blots containing 2  $\mu$ g poly(A)<sup>+</sup> RNA purified from various human tissues (CLONTECH) were probed sequentially for Northern analysis with <sup>32</sup>P-labeled cDNAs of hNEH1 (Upper) and hOGG1 (Lower).

elimination like Nei and MutM, and not  $\beta$  elimination like Nth and OGG1 (Fig. 5A and B).

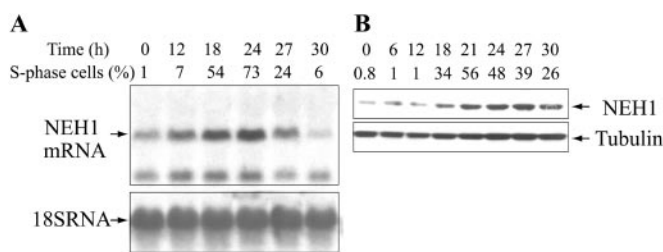
**Expression of NEH1 and OGG1 mRNAs in Human Tissues.** We compared the levels of NEH1 and OGG1 mRNAs in human tissues (Fig. 6). Three species of NEH1 mRNA ( $\approx$ 7.5, 4.0, and 1.8 kb) were observed. Because the hNEH1 gene is 8.1 kb long, it is likely that the largest RNA species corresponds to its primary transcript, and that the 4.0-kb species is a processing intermediate whereas the 1.8-kb species corresponds to mature NEH1 mRNA. We do not know the reason why hNEH1 mRNA was processed poorly in various tissue samples, although 7.5- and 4-kb precursor RNAs were not detected in the total RNA of *in vitro* cultured cells (see below). In any case, we observed the highest expression of the NEH1 gene in the liver, pancreas, and thymus; moderate expression in the brain, spleen, prostate, and ovary; and a relatively low level in the testis and leukocytes (Fig. 6).

When the same membranes were probed for OGG1 expression, two mRNA species were detected in all of the tissues, as observed earlier (36). More importantly, the patterns of expression for OGG1 and NEH1 were different. For example, the brain has low expression of OGG1 but high expression level of NEH1. Assuming that the levels of NEH1 and OGG1 polypeptides reflect their mRNA levels, it is possible that these two enzymes serve as each other's back-up system, but that their levels may also reflect the need to use distinct types of repair mechanisms such as RAR and TCR in various tissues.

**S Phase-Specific Expression of hNEH1.** Cell cycle dependence of NEH1 expression was examined by both Northern analysis with total RNA isolated from synchronized human diploid fibroblasts and Western analysis with whole cell lysates. Fig. 7A shows that the NEH1 mRNA level reached its maximum at 24 h after release from serum starvation, paralleling the fraction of cells in the S phase. About a 4-fold increase over the level in G<sub>0</sub>/G<sub>1</sub> cells was observed, which declined at 27 h. The level of 18s ribosomal RNA, used as an internal control, did not change significantly with the cell cycle (Fig. 7A).

It is interesting to note that only 1.8-kb mRNA species were observed in MRC5 cells, along with a smaller RNA species whose origin is not clear. The absence of the 7.5- and 4-kb RNA species suggests that NEH1 mRNA is efficiently processed in these cells.

In a separate experiment, Western analysis of the level of the NEH1 polypeptide in synchronized cells showed a similar increase of the enzyme level during the S phase, reaching a maximum that was at 27 h about 6-fold higher than that in G<sub>0</sub>/G<sub>1</sub> cells (Fig. 7B). Interestingly, the protein level did not go down as rapidly as the RNA level in cells beyond the S phase. These results showing S



**Fig. 7.** S phase-specific activation of NEH1. (A) Northern analysis of NEH1 mRNA, and (B) Western analysis of NEH1 polypeptide in synchronized MRC5 cells collected at various times after serum addition. The percentage of S phase cells in these different experiments is shown.

phase-specific activation of NEH1 suggest a linkage of this enzyme to DNA replication.

## Discussion

Identification and cloning of hNEH1 and hNEH2, the first mammalian orthologs of *E. coli* Nei and MutM that carry out  $\beta\delta$  elimination, became possible because of the human genome project, for several reasons. First, the cellular abundance of these enzymes, particularly in non-S phase cells, may be too low for their detection in cell extracts on the basis of activity. Second, 8-oxoG·C pair-containing DNA or oligos, commonly used for assaying DNA glycosylases for oxidized bases, are poor substrates, at least for NEH1, whereas FapyA and FapyG, the major substrates of NEH1, are not routinely used as substrates. Although hNEH1 has significant homology with both MutM and Nei in the catalytic domain (Fig. 1), we named it NEH (Nei Homolog) based on its substrate preference. NEH1, unlike MutM/Fpg but like Nei, has very weak 8-oxoG·C-specific OGG activity, the hallmark of MutM, and stronger activity with the 8-oxoG·G pair. Furthermore, it resembles Nei, and not MutM, in its stronger excision activity for pyrimidine lesions than for 8-oxoG. Although we have tentatively named the second candidate enzyme NEH2, its definitive naming should depend on its substrate preference, once determined. Interestingly, NEH1 acts like MutM in excising both FapyA and FapyG in natural DNA, whereas hOGG1 cleaves only FapyG (14, 16). Among all of the lesions in irradiated natural DNA, only FapyA and FapyG were found to be substrates of NEH1, and this is the only human enzyme identified so far that excises FapyA. Interestingly, FapyA is repaired by *E. coli* Nth but not its human ortholog NTH1 (29, 37). It appears that multiple glycosylases have evolved to excise Fapys because of their abundance in the genomes of both normal and oxidatively stressed cells (7).

We had identified and partially purified an enzyme from HeLa cells with an apparent molecular mass of 36 kDa, which we named OGG2 (12). Both OGG2 and NEH1 prefer the 8-oxoG·G pair as a substrate. However, apart from the size difference between OGG2 and NEH1, the fact that OGG2 carried out  $\beta$  elimination, whereas NEH1 has  $\beta\delta$  elimination activity, indicates that OGG2 is not identical to NEH1. It is thus evident that the hOGG2 gene would not have been detected in the search that led to the discovery of hNEH1.

NEH1 has no activity toward 2-hydroxyadenine, 1-N<sup>6</sup> etheno-adenine, 3-N<sup>4</sup> ethenocytosine, hypoxanthine, and xanthine (data not shown). As expected, NEH1 efficiently cleaves AP site in DNA by functioning as an AP lyase (data not shown). 8-oxoG is present only as 8-oxoG·C pairs in nonreplicating DNA. Thus, our inability to observe significant excision of 8-oxoG from natural DNA is likely to be due to the weak activity of NEH1 toward an 8-oxoG·C pair. A similar lack of activity toward pyrimidine lesions in natural DNA (whereas a robust activity was observed with synthetic oligos) may be due to the limiting lesion concentration and competition among

substrates in natural DNA, whereas excess substrates were used to compare reaction rates with oligos.

The larger size of hNEH1 compared with its ortholog Nei follows the pattern of other DNA glycosylases in that the mammalian ortholog is larger than the prokaryotic enzyme. It is likely that the nonhomologous C-terminal sequence in hNEH1 is involved in noncatalytic functions of the enzyme, e.g., in interaction with other proteins and/or in nuclear translocation. The C-terminal region of NEH1 contains several putative nuclear localization signal (NLS) sequences (Fig. 1). Splice variants of hOGG1 are targeted to the nucleus or mitochondria (38). The mitochondrial target sequence (MTS) is usually present at the N terminus, which is cleaved off during transit through the mitochondrial membrane (39). The presence of the active site Pro at the start of the coding sequence suggests that hNEH1 is not localized in the mitochondria.

**Potential Involvement of NEH1 in TCR and RAR.** As pointed out in the introduction, preferential repair of 8-oxoG located in the transcribed sequences does not involve OGG1. It is possible that NEH1 (and NEH2) are responsible for TCR of 8-oxoG. Although the observed activity of NEH1 for 8-oxoG and oxidized purines is rather low, this result may not reflect the *in vivo* situation where the activity could be significantly enhanced in the presence of other proteins, e.g., XPG, which is involved in TCR (17, 40). Furthermore, although not tested, Fapys may also be subject to TCR by using NEHs in this process.

NEH1 may also be involved in replication-associated repair of oxidatively damaged bases. We proposed that RAR is necessary to prevent misincorporation of damaged bases in the nascent strand during DNA replication (12). In view of the broad substrate specificity of NEH1, at least with oligo substrates, it is possible that NEH1 preferentially excises 8-oxoG, as well as pyrimidine lesions, when incorporated in the nascent strand. We have shown earlier that both human OGG2 and *E. coli* Nei preferentially incise 8-oxoG-strand when 8-oxoG is paired with G and A (12, 26). The present study shows that NEH1 is also more active with the

8-oxoG-G pair than with the 8-oxoG-C pair. Electrophoretic gel mobility-shift assay (EMSA) showed distinct patterns of affinities of hOGG1 and hNEH1 for different 8-oxoG base pairs. OGG1 bound much more tightly to the 8-oxoG-C pair than to the 8-oxoG-A pair, whereas NEH1 preferred 8-oxoG-G and 8-oxoG-A pairs (data not shown). 8-oxoG in *syn* conformation was shown to mispair with A during *in vitro* replication (41). However, it can also pair with G as a 6-enolate-8-keto tautomer (42). Whether 8-oxoG-G pairing occurs during DNA replication *in vivo*, which would lead to GC → CG transversion mutation, is not known. Such mutations have, in fact, been found in oxidatively stressed *E. coli* and mammalian cells (43, 44). That these mutations in *E. coli* are prevented by the presence of Nei suggests that Nei in *E. coli* (45), and possibly NEH1 in mammalian cells, are responsible for repair of 8-oxoG (and other oxidized bases) in the nascent strand. We have observed that the level of hNEH1 polypeptide is highest in the S phase, with 6- to 7-fold enhancement over that in G<sub>0</sub>/G<sub>1</sub>-phase. In contrast, the level of hOGG1 is not significantly affected by the cell cycle (46). These results further support our prediction that NEH1 is involved in RAR. RAR of oxidized bases requires targeting of the DNA glycosylase to the nascent strand, which is presumably mediated by its interaction with replication-associated proteins. Identification of the interacting partners of hNEH1 should be an important topic for future studies.

We thank Eric Wilmore and Julie Lock for excellent technical assistance, and Alex Kurosky and Stephen Smith of the National Institute on Environmental Health Sciences (NIEHS) Center for protein sequence analysis, David Konkel for editing the manuscript, and Wanda Smith for expert secretarial assistance. We also thank Rabindra Roy for a gift of human NTH1. Certain commercial equipment or materials are identified in this paper to specify adequately the experimental procedure. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the materials or equipment identified are necessarily the best available for the purpose. This research was supported by U.S. Public Health Service Grants CA81063 (to S.M.), CA90860 (to Y.W.K.), and CA 84461 (to I.B.), and NIEHS Center Grant ES06676.

- Grisham, M. B. & McCord, J. M. (1986) in *Physiology of Oxygen Radicals*, eds. Taylor, A. E., Matalon, S. & Ward, P. A. (Waverly Press, Baltimore), pp. 1–18.
- Ames, B. N., Shigenaga, M. K. & Hagan, T. M. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 7915–7922.
- Gotz, M. E., Kunig, G., Riederer, P. & Youdim, M. B. (1994) *Pharmacol. Ther.* **63**, 37–122.
- Lovell, M. A., Xie, C. & Markesbery, W. R. (2000) *Brain Res.* **855**, 116–123.
- Breen, A. P. & Murphy, J. A. (1995) *Free Rad. Biol. Med.* **18**, 1033–1077.
- Krokan, H. E., Nilsen, H., Skorpen, F., Otterlei, M. & Slupphaug, G. (2000) *FEBS Lett.* **476**, 73–77.
- Jaruga, P., Speina, E., Gackowski, D., Tudek, B. & Olinski, R. (2000) *Nucleic Acids Res.* **28**, E16.
- O'Connor, T. R., Boiteux, S. & Laval, J. (1988) *Nucleic Acids Res.* **16**, 5879–5894.
- Kow, Y. W. & Wallace, S. S. (1987) *Biochemistry* **26**, 8200–8206.
- Jiang, D., Hatahet, Z., Blaisdell, J. O., Melamed, R. J. & Wallace, S. S. (1997) *J. Bacteriol.* **179**, 3773–3782.
- Nash, H. M., Bruner, S. D., Schärer, O. D., Kawate, T., Addona, T. A., Spooner, E., Lane, W. S. & Verdine, G. L. (1996) *Curr. Biol.* **6**, 968–980.
- Hazra, T. K., Izumi, T., Maidt, L., Floyd, R. A. & Mitra, S. (1998) *Nucleic Acids Res.* **26**, 5116–5122.
- Zharkov, D. O., Rieger, R. A., Iden, C. R. & Grollman, A. P. (1997) *J. Biol. Chem.* **272**, 5335–5341.
- Boiteux, S., Gajewski, E., Laval, J. & Dizdaroglu, M. (1992) *Biochemistry* **31**, 106–110.
- Dherin, C., Radicella, J. P., Dizdaroglu, M. & Boiteux, S. (1999) *Nucleic Acids Res.* **27**, 4001–4007.
- Karakaya, A., Jaruga, P., Bohr, V. A., Grollman, A. P. & Dizdaroglu, M. (1997) *Nucleic Acids Res.* **25**, 474–479.
- Leadon, S. A. (2000) *Cold Spring Harbor Symp. Quant. Biol.* **65**, 561–566.
- Tsutakawa, S. E. & Cooper, P. K. (2000) *Cold Spring Harbor Symp. Quant. Biol.* **65**, 201–215.
- Mellon, I., Bohr, V. A., Smith, C. A. & Hanawalt, P. C. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 8878–8882.
- Cooper, P. K., Nouspikel, T., Clarkson, S. G. & Leadon, S. A. (1997) *Science* **275**, 990–993.
- Le Page, F., Kwoh, E. E., Avrutskaya, A., Gentil, A., Leadon, S. A., Sarasin, A. & Cooper, P. K. (2000) *Cell* **101**, 159–171.
- Klungland, A., Rosewell, I., Hollenbach, S., Larsen, E., Daly, G., Epe, B., Seeberg, E., Lindahl, T. & Barnes, D. E. (1999) *Proc. Natl. Acad. Sci. USA* **96**, 13300–13305.
- Minowa, O., Arai, T., Hirano, M., Monden, Y., Nakai, S., Fukuda, M., Itoh, M., Takano, H., Hippo, Y., Aburatani, H., et al. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 4156–4161.
- Le Page, F., Klungland, A., Barnes, D. E., Sarasin, A. & Boiteux, S. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 8397–8402.
- Tchou, J., Kasai, H., Shibutani, S., Chung, M.-H., Laval, J., Grollman, A. P. & Nishimura, S. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 4690–4694.
- Hazra, T. K., Izumi, T., Venkataraman, R., Kow, Y. W., Dizdaroglu, M. & Mitra, S. (2000) *J. Biol. Chem.* **275**, 27762–27767.
- Zhang, Q. M., Ishikawa, N., Nakahara, T. & Yonei, S. (1998) *Nucleic Acids Res.* **26**, 4669–4675.
- Boldogh, I., Milligan, D., Lee, M. S., Bassett, H., Lloyd, R. S. & McCullough, A. K. (2001) *Nucleic Acids Res.* **29**, 2802–2809.
- Dizdaroglu, M., Bauche, C., Rodriguez, H. & Laval, J. (2000) *Biochemistry* **39**, 5586–5592.
- Bresnahan, W. A., Boldogh, I., Ma, T., Albrecht, T. & Thompson, E. A. (1996) *Cell Growth Differ.* **7**, 1283–1290.
- Ramana, C. V., Boldogh, I., Izumi, T. & Mitra, S. (1998) *Proc. Natl. Acad. Sci. USA* **95**, 5061–5066.
- Hill, J. W., Hazra, T. K., Izumi, T. & Mitra, S. (2001) *Nucleic Acids Res.* **29**, 430–438.
- Ikeda, S., Biswas, T., Roy, R., Izumi, T., Boldogh, I., Kurosky, A., Sarker, A. H., Seki, S. & Mitra, S. (1998) *J. Biol. Chem.* **273**, 21585–21593.
- Sidorkina, O. M. & Laval, J. (1998) *Nucleic Acids Res.* **26**, 5351–5357.
- Sugahara, M., Mikawa, T., Kumasaka, T., Yamamoto, M., Kato, R., Fukuyama, K., Inoue, Y. & Kuramitsu, S. (2000) *EMBO J.* **19**, 3857–3869.
- Radicella, J. P., Dherin, C., Desmaze, C., Fox, M. S. & Boiteux, S. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 8010–8015.
- Dizdaroglu, M., Karahalil, B., Senturker, S., Buckley, T. J. & Roldan-Arjona, T. (1999) *Biochemistry* **38**, 243–246.
- Nishioka, K., Ohtsubo, T., Oda, H., Fujiwara, T., Kang, D., Sugimachi, K. & Nakabeppu, Y. (1999) *Mol. Biol. Cell* **10**, 1637–1652.
- Lithgow, T. (2000) *FEBS Lett.* **476**, 22–26.
- Klungland, A., Hoss, M., Gunz, D., Constantinou, A., Clarkson, S. G., Doetsch, P. W., Bolton, P. H., Wood, R. D. & Lindahl, T. (1999) *Mol. Cell* **3**, 33–42.
- Grollman, A. P. & Moriya, M. (1993) *Trends Genet.* **9**, 246–249.
- Culp, S. J., Cho, B. P., Kadlubar, F. F. & Evans, F. E. (1989) *Chem. Res. Toxicol.* **2**, 416–422.
- McBride, T. J., Preston, B. D. & Loeb, L. A. (1991) *Biochemistry* **30**, 207–213.
- de Oliveira, R. C., Ribeiro, D. T., Nigro, R. G., Di Mascio, P. & Menck, C. F. (1992) *Nucleic Acids Res.* **20**, 4319–4323.
- Matsumoto, Y., Zhang, Q. M., Takao, M., Yasui, A. & Yonei, S. (2001) *Nucleic Acids Res.* **29**, 1975–1981.
- Dhenaut, A., Boiteux, S. & Radicella, J. P. (2000) *Mutat. Res.* **461**, 109–118.