A Back-up Glycosylase in Nth1 Knock-out Mice Is a Functional Nei (Endonucl ease VIII) Homologue*

Masashi Takao‡‡, Shin-ichiro Kanno‡, Kumiko Kobayashi‡, Qiu-Mei Zhang‡, Shuji Yonei‡, Gijbvertus T. J. van der Horst**, and Akira Yasui‡ ‡‡

From the ‡Department of Molecular Genetics, Institute of Development, Aging and Cancer, Tohoku University, Sendai 980-8575, Japan, *)Department of Chemistry and Laboratory of Radiation Biology, Graduate School of Science, Kyoto University, Kyoto 606-8502, Japan, and **Medical Genetics Centre, Department of Cell Biology and Genetics, Erasmus University, P. O. Box 1738, 3000 DR Rotterdam, The Netherlands

Thymine glycol, a potentially lethal DNA lesion produced by reactive oxygen species, can be removed by DNA glycosylase, Escherichia coli Nth (endonuclease III), or its mammalian homologue NTH1. We have found previously that mice deleted in the Nth homologue still retain at least two residual glycosylase activities for thymine glycol. We report herein that in cell extracts from the mNth1 knock-out mouse there is a third thymine glycol glycosylase activity that is encoded by one of three mammalian proteins with sequence similarity to E. coli Fpg (MutM) and Nei (endonuclease VIII). Tissue expression of this mouse Nei-like (designated as Neil1) gene is ubiquitous but much lower than that of mNth1 except in heart, spleen, and skeletal muscle. Recombinant NEIL1 can remove thymine glycol and 5-hydroxyuracil in double- and single-stranded DNA much more efficiently than 8-oxoguanine and can nick the strand by an associated (β-θ) apurinic/apyrimidinic lyase activity. In addition, the mouse NEIL1 has a unique DNA glycosylase/lyase activity toward mismatched uracil and thymine, especially in U:C and T:C mismatches. These results suggest that NEIL1 is a back-up glycosylase for NTH1 with unique substrate specificity and tissue-specific expression.

Oxidative damage in DNA is widely acknowledged to be a causative factor in cancer and aging (1, 2). Major oxidative damage includes premutagenic lesions and lesions that block replicative DNA polymerases. A typical example of a premutagenic lesion is 8-oxoguanine (8-oxoG), which can form a base pair with adenine, as well as with cytosine, resulting in G:C to T:A transversion. In contrast to 8-oxoG, a major replication-blocking lesion generated by reactive oxygen species in vivo is thymine glycol (Tg). Therefore, both 8-oxoG and Tg are often used as biomarkers for oxidative stress and aging (3, 4). Despite the importance of Tg, repair pathway of Tg in mammalian cells is not well understood.

It has been well established that the most important repair pathway for the removal of mutagenic or toxic oxidative base damage is base excision repair (BER). The marked evolutionary conservation of BER enzymes from bacteria to mammals (5, 6) suggests the importance of BER for life and has enabled the identification of various mammalian homologues through database searches using the nucleotide sequences encoding well-characterized Escherichia coli and yeast enzymes. Three mammalian genes for oxidative DNA repair glycosylases, OGG1, MYH, and NTH1, have been identified in this way. In addition, the growing amount of information provided by genome and cdNA sequencing projects has elucidated species-specific differences between enzymes involved in the conserved BER reaction. Prokaryotes employ Fpg glycosylase for repair of oxidized purines, whereas eukaryotes use OGG1, which has almost no amino acid sequence similarity with the bacterial Fpg protein. A homologue of OGG1 is also found in archaea (7), whereas Arabidopsis thaliana has both Fpg and OGG1 (8). E. coli possesses two DNA glycosylases for repair of oxidized pyrimidine, Nth and Nei, which are structurally unrelated but functional homologues. nth/nei double-deficient bacteria, but not the single mutant, show sensitivity to hydrogen peroxide and ionizing radiation (9, 10). Nth is one of the most widespread DNA glycosylases in eukaryotes, as well as prokaryotes, whereas the Nei orthologs are rarely found in bacteria and are even absent in the Saccharomyces cerevisiae, Schizosaccharomyces pombe, and Dro sophila melanogaster genomes. Although this distribution suggests that mammalian species may also lack Nei homologues, three Fpg/Nei DNA glycosylase-like sequences have recently been registered in a full-length cdNA database (NEDO human cdNA sequencing project; www. nedo.go.jp/bio/). We have characterized these Fpg Nei-Like genes and their repair activities in recombinant proteins. During preparation and submission of this manuscript, these clones have been reported (11–13) and, accordingly, we use the name NEIL (Nei-Like) for the genes in this paper.

We have recently established Nth1 knock-out mice to investigate the biological consequences of oxidative pyrimidine repair deficiency (14). Surprisingly, the NTH1-deficient mice lack overt phenotypic abnormalities. Moreover, Tg produced in the mutant mouse liver DNA by X-irradiation disappeared with time, though more slowly than in the wild-type mouse. Biochemical analysis of NTH1-deficient mouse liver extracts

* This work was supported by Grants 13680618 (to M.T.) and 13480162 and 12143201 (to A.Y.) from the Ministry of Education, Culture, Sports, Science and Technology. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the DDBJ/GenBank™/EMBL Data Bank with accession number(s) AB079068 (human NEIL1), AB079069 (mouse NEIL1/FLJ22402), AB079070 (human NEIL2/FLJ31644), and AB079071 (human NEIL3/FLJ10858).

† To whom correspondence may be addressed. Tel.: 81-22-717-8470; Fax: 81-22-717-8470; E-mail: mtakao@idac.tohoku.ac.jp.
‡ To whom correspondence may be addressed. Tel.: 81-22-717-8469; Fax: 81-22-717-8470; E-mail: ayasu@idac.tohoku.ac.jp.
§ The abbreviations used are: 8-oxoG, 8-oxoguanine; AP, apurinic/apyrimidinic; Tg, thymine glycol; BER, base excision repair; 5-OHU, 5-hydroxyuracil; Ni-NTA, nickel-nitrilotriacetic acid; HPLC, high pressure liquid chromatography; H2TH, helix-two-turn-helix; TCR, transcription-coupled repair.

This paper is available on line at http://www.jbc.org

Published, JBC Papers in Press, August 27, 2002, DOI 10.1074/jbc.M206884200

© 2002 by The American Society for Biochemistry and Molecular Biology, Inc.
showed that there are at least two Tg-DNA glycosylase back-up activities. One activity designated as TGG1 (thymine glycol glycosylase 1) is a monofunctional DNA glycosylase found mainly in the mitochondria, and the other TGG2 forms a Schiff base reaction intermediate, which is indicative of a nuclear AP lyase-associating enzyme. Because we found three human NEIL genes in the database, we analyzed the enzymatic activity of the proteins in vitro and in vivo and their relationships to residual glycosylase activities in mNth1 knock-out mice.

EXPERIMENTAL PROCEDURES

Cloning of Mammalian cDNAs—Primers for the PCR were designed based on the cDNA sequences in GenBank™ (AK026055 for FLJ22402/nH1E, AK056206 for FLJ31644, AK001720 for FLJ10858, and AK013322 for mouse ortholog of FLJ22402/mNEIL1). The cDNA for human FLJ31644, FLJ10858, or mNEIL1 was amplified from a cDNA pool reverse-transcribed from mRNA from HeLa cells, keratinocytes, or primary mouse embryonic fibroblasts. The cDNA for hNEIL1 was amplified from a commercial human testis cDNA library (Invitrogen). The cDNA was synthesized with a kit using oligo(dT) primers (Roche Molecular Biochemicals). The amplified fragment was subcloned in pSTBlue-1 (Novagen). Multiple clones obtained were sequenced extensively to exclude mutants generated by PCR. The sequence of any recombinant fragment was also verified. These sequence data were submitted to the DDBJ/GenBank™/EBI database under accession number AB079068 (human NEIL1), AB079069 (mouse NEIL1/FLJ22402), AB079070 (human NEIL2/FLJ31644), and AB079071 (human NEIL3/FLJ10858).

In Vitro Transcription and Translation—The initial PCR mentioned above was made with a 5′-primer containing Kozak’s sequence (ACC) in front of the initiation codon. The cDNA fragment was subcloned into a pSPUTK vector (Stratagene). In vitro expression was performed with a Quick TNT system (Promega) in the presence of [35S]methionine for the detection of protein production or cold methionine for activity screening. The translated products were analyzed on SDS-PAGE with a 10–20% gradient gel (Daichi Pure Chemicals).

Real-time, Quantitative PCR—Quantitative PCR was performed using a LightCycler (Roche Molecular Biochemicals) with a DNA Master SYBR Green 1 kit (Roche Molecular Biochemicals) according to the system and instructions. A mouse tissue specific cDNA panel was purchased from Clontech. Amplimers used were as follows: mouse glyceraldehyde-3-phosphate dehydrogenase gene (Gapdh) primers, 5′-CTTGGCTCTGCTGATGTCGAAT and 5′-CGATGAGCCATGGAGGGTCACACCAC; Nth1 primers, 5′-GCGACACGGGGCAGCATGAGCACCAGGAGA and 5′-GCGACTTCTAGCATGAGCAGTCCGC; Nei1 primers, 5′-CTGACGTTCAACAGAAGCGTATCCCTTTCTG and 5′-CCACGCTGTCAGCACCTCTG; and Nei2 primers, 5′-CTCAGCTAGCTGAGGCTGAGCCCCGTCG and 5′-CCACGCTGTCAGCACCTCTG. Standard curves were obtained by quantitative PCR with a luminometer unit of linearized plasmid DNA containing a single copy of the cDNA. Prior to quantification, the cDNA panel was re-equalized by PCR for Gapdh. Data were expressed relative to the expression of Gapdh. Immunofluorescence—An epitope tag (FLAG sequence) was added to mNeil1 cDNA by PCR. The tagged cDNA was subcloned in a mammalian expression vector pTargeT (Promega). COS-7 cells were transfected with the construct using FuGENE6 (Roche Molecular Biochemicals). The transiently expressed protein was detected by an anti-FLAG antibody (Sigma-Aldrich) and an Alexa Fluor 488-conjugated second antibody (Molecular Probes).

Incision Assay—The incision assay was made with a 30-mer oligonucleotide containing a single modified base indicated by X, 5′-CTCAGTCGACATCTCATACATACGTCGATGCT-3′. Oligos containing Tg, 8-oxoG, and the tetrahydrofuran AP site (AP) were generously provided by Dr. S. Iwai (Tokyo University). Oligos containing 5-hydroxycytosine (5-OH) and uracil were synthesized and purified by Japan Bio-Service. The oligo was labeled using either T4-polynucleotide kinase [32P]-ATP (5000 Ci/mmol; Amersham Biosciences) or terminal deoxynucleotidyl transferase (Invitrogen) with [a-32P]-dATP (5000 Ci/mmol; Amersham Biosciences). The incision reaction was performed in reaction buffer (20 mM sodium phosphate (pH 7.5), 50 mM NaCl, 5 mM EDTA, and 100 μg/ml bovine serum albumin) containing 1 fmol/μl of either double-stranded DNA or single-stranded DNA at 30 °C for 30 min. The reaction was stopped by addition of an equal volume of ice-cold loading dye solution. The samples were heated at 75 °C for 2 min and run on a 20% polyacrylamide gel containing 7 M urea. The gel was exposed to an imaging plate and analyzed with a BAS2000 image analyzer (Fuji photo film). The densitometric analysis of the gel image was performed with NIH Image software. For the reaction with in vitro translated gene products, the reacted DNA was hydrolyzed with alkaline as described previously (15). Briefly, the DNA was phenol/chloroform-extracted and ethanol-precipitated. The DNA was dissolved in 0.1 M NaOH and incubated at 95 °C for 5 min to convert the abasic DNA into the nicked form. In some experiments, we used T4 endonuclease V as “AP lyase” enzyme for detecting monofunctional DNA glycosylase, as well as for assays with crude extract (14). To characterize the terminal structure of the incision products, 5′-labeled or 5′-labeled DNA digests were treated with bacterial alkaline phosphatase (TaKaRa) or with T4 polynucleotide kinase in a buffer containing 20 mM sodium phosphate (pH 6.0), 10 mM MgCl2, 5 mM 2-mercaptoethanol, respectively (16).

Preparation of Recombinant NEIL1 Protein—A His6-tagged expression construct was made by subcloning mNeil1 into pET21 vector (Novagen). The E. coli BL21 codon plus (Stratagene) was transformed and cultured at 37 °C until anOD reached 0.6. After induction with 0.5 mM isopropyl-1-thio-β-D-galactopyranoside at 25 °C for 6 h, Ni-NTA (Qiagen) column chromatography was performed under native conditions according to the manufacturer’s instructions. The recombinant NEIL1 protein was eluted in a 20 mM sodium phosphate buffer (pH 8.0) containing 100 mM NaCl and 250 mM imidazole. The eluate was directly loaded onto Hitrap-SP (Amersham Biosciences), and the active fractions eluted in phosphate buffer containing 0.5–0.7 M NaCl were pooled. The purified NEIL1 protein was concentrated to 1 mg/ml and stored in 25 mM sodium phosphate buffer (pH 8.0), 50 mM NaCl, 1 mM dithiothreitol, 1 mM EDTA, and 50% glycerol at −30 °C. Because the activity gradually decreased, the quantitative and comparative experiments were made with the same purification lot within 2 days.

Trapping Assay—For DNA trapping, the recombinant NEIL1 protein or column fraction of mouse cell extracts was mixed with 100 mM NaN3BH3 in 50 mM sodium phosphate buffer (pH 7.5) containing 10 mM EDTA on ice. Labeled Tg-containing oligo (final concentration, 5 fmol/μl) was added and incubated at 30 °C for 60 min. The 50-μl reaction mix was applied on a Sephadex G-50 microcolumn (Amersham Biosciences) to remove the redundant. The enzyme-DNA complex was separated on a 10–20% SDS-polyacrylamide gel and autoradiographed. For limited proteolysis experiments, a portion of the G5 gel-filtrated sample was digested further either by trypsin or chymotrypsin.

Preparation of DNA Glycosylase Active Fractions from Nth1−/− Mice—Partial purification of Tg-DNA glycosylase activities from NTH1-deficient mouse liver nuclear extracts has been described in a previous paper (14). In brief, nuclear proteins were extracted in Buffer A (50 mM Tris-HCl (pH 7.5), 0.1% Nonidet P-40, 2 mM dithiothreitol) with 0.3 mM NaCl. The supernatant was dialyzed in Buffer B (50 mM Tris-HCl (pH 7.5), 2 mM dithiothreitol, 1 mM EDTA, 10% glycerol) with 0.1 mM NaCl and applied to heparin-agarose and eluted with Buffer B containing 0.8 mM NaCl. The fraction was dialyzed and applied to an UNO S-1 column in an HPLC system (Bio-Rad). Incision assays and trapping assays were conducted on the flow-through fraction, the fractions separated by a linear 0–0.6 M NaCl gradient, and the fraction eluted by stepwise increase of NaCl from 0.6 to 1.0 M.

RESULTS

Three Putative Human DNA Glycosylases Similar to Bacterial Fpg and Nei—A BLAST homology search of GenBank™ using the E. coli Nei sequence as query yielded three human sequences for hypothetical proteins (FLJ22402, FLJ31644, and FLJ10858) that exhibit significant similarity to E. coli Fpg and Nei (Fig. 1A). The identity and similarities were designated as follows (13): NEIL1 for FLJ22402, NEIL2 for FLJ31644, and NEIL3 for FLJ10858. The human NEIL1, NEIL2, and NEIL3 genes map on chromosomes 15, 8, and 4, respectively. The respective cDNA structure and approximate positions of introns inherited from its genomic sequence are shown in Fig. 1B.

The crystal structures of Thermus thermophilus Fpg and Nei show a significant homology (17, 18), delineating a DNA binding cleft composed of common structural elements and conserved amino acid residues. The structural analysis suggests further that a protein potentially belonging to the Fpg-Nei protein family has six conserved DNA-contacting elements in its sequence: the N-terminal active site, turn β3-β4, turn β5-β6, turn β8-β9, helix-two-turn-helix (H2TH) motif, and zinc-finger motif (numbering of the β-sheets refers to Tth-Fpg struc-
ture; see Ref. 17). Indeed, these domains are highly conserved among the known and putative prokaryotic Fpg and Nei sequences in the database.

To assess a possible DNA binding capacity of the human proteins, we searched for the DNA-contacting elements and re-aligned the NEIL sequences with Fpg or Nei (Fig. 1A). All three proteins contain a H2TH motif, representing a helix-hairpin-helix motif widespread in DNA-binding proteins including many DNA glycosylases (19). The prototype helix-hairpin-helix motif has a consensus G\textsubscript{h}G(\textsubscript{h} is a hydrophobic residue) sequence between the two helices that is found in all three proteins. Moreover, the three NEIL proteins have absolutely conserved residues (Asp, Asn, and Glu) specific for the H2TH motif of the Fpg-Nei family. The NEIL1 and NEIL2 proteins contain the conserved N-terminal active site residues, Pro-2 and Glu-3, whereas in the NEIL3 protein the Schiff base-forming Pro-2 (18, 20, 21) is replaced by valine. The zinc-finger motif is conserved in NEIL3 and to a lesser extent in NEIL2, whereas both retain the invariable Arg and Gln at the appropriate positions. Interestingly, the NEIL1 protein does not show any homology to Fpg or Nei in the C-terminal region of the H2TH motif. Nevertheless, the NEIL1 alone exhibits all the three DNA-contacting turns (\(\beta3-\beta4\), \(\beta5-\beta6\), and \(\beta8-\beta9\)) with great resemblance to Fpg or Nei. The relative position of these potential DNA-contacting elements is depicted schematically in the respective cDNA structure (Fig. 1B).

**NEIL1 Is a Functional Nei Homologue**—To establish the function of these putative glycosylases, we first cloned the cDNAs for human NEIL1, NEIL2, and NEIL3, as well as the cDNA for mouse NEIL1 ortholog, by PCR. The NEIL3 protein appears to have an additional domain in the C-terminal half. This domain shows local homology to topoisomerase III (TOP3) and a C-terminal domain of APE2 (22), a protein having an AP-endonuclease-like sequence with unknown function. We used in vitro expressed proteins to screen for enzymatic activity. As shown in Fig. 1C, we could produce full-length NEIL1 and NEIL2 proteins using the rabbit reticulocyte lysate. For NEIL3, we were unable to obtain a full-size product and thus examined a truncated protein, NEIL3 (1–289), that is encoded by exons from 1 to 6. This polypeptide lacks the C-terminal domain but covers the entire Fpg-Nei-like domain.

The translation mixtures were incubated with \(5'\)-labeled 30-mer DNA containing a single Tg at a defined position to detect the strand incision activity. The reaction was conducted by inducing a strand incision at an abasic site to follow up a monofunctional DNA glycosylase having no incision activity. For NEIL3, we were unable to obtain a full-size product and thus examined a truncated protein, NEIL3 (1–289), that is encoded by exons from 1 to 6. This polypeptide lacks the C-terminal domain but covers the entire Fpg-Nei-like domain.

The translation mixtures were incubated with \(5'\)-labeled 30-mer DNA containing a single Tg at a defined position to detect the strand incision activity. The reaction was conducted by inducing a strand incision at an abasic site to follow up a monofunctional DNA glycosylase having no incision activity. As shown in Fig. 1D, the rabbit lysate itself showed a slight background incision activity, probably because of rabbit enzymes (luciferase-expressing lane as a negative control in Fig. 1D). The reactions for NEIL2 and NEIL3 (1–289) did not yield nicked products above the background signal. On the other hand, the reactions for human NEIL1 and its mouse ortholog...
resulted in an obvious incision, suggesting that NEIL1 is a Tg-nicking DNA glycosylase.

**Characterization of Mouse NEIL1 as Tg-DNA Glycosylase and AP Lyase**—Because we anticipated that the NEIL1 gene product could be a back-up enzyme for Tg removal found in Nth1−/− mice, the mouse NEIL1 protein was characterized intensively. We produced a recombinant mouse NEIL1 tagged with His6 at its C terminus in E. coli (Fig. 2A, arrow) and recovered the protein from the soluble fraction of the whole cell extracts. Mouse NEIL1 was purified to apparent homogeneity by affinity chromatography through a Ni-NTA column and a HiTrap-SP column chromatography. This recombinant NEIL1 enzyme was used to further characterize its substrate specificity in WCE. We produced a recombinant mouse NEIL1 tagged with His6 at its C terminus in E. coli and recovered the protein from the soluble fraction of the whole cell extracts. The purity and apparent homogeneity of the observed Tg-DNA nicking activity (Fig. 2B) was derived from the recombinant NEIL1, rather than from contaminating host cell enzymes. This recombinant NEIL1 enzyme was used to further optimize the reaction conditions. The enzyme did not require a divalent cation, was resistant to EDTA up to 20 mM, and displayed comparable activity between pH 6 and 9 (data not shown).

### Substrate Specificity of NEIL1—Primary substrates for E. coli Nth and Nei are oxidized pyrimidines whereas Nei also shows 8-oxoguanine DNA glycosylase activity (25, 26). To determine the substrate specificity of mammalian NEIL1, we prepared oligonucleotide substrates containing a modified base at a defined position in the same sequence context. DNA lesions tested include two oxidized pyrimidine lesions, Tg and 5-OHU; an oxidized purine lesion, 8-oxoG; and an AP lesion. Using an amount of enzyme that cleaved about 50% of the Tg:A substrate as a reference, the nicking efficiency toward the other substrates was evaluated (Fig. 4A). NEIL1 cleaved Tg and 5-OHU to a similar extent but cleaved 8-oxoG much less efficiently. A further increase in the amount of NEIL1 added to 8-oxoG:C or 8-oxoG:A substrates invoked only marginal incision (Fig. 4C). The nicking of the AP substrate was a little more efficient than that of the Tg:A substrate (Fig. 4C). Interestingly, NEIL1 was able to cleave a single-strand substrate containing Tg, 5-OHU, and AP (Fig. 4A, lanes ss). Also, nicking at the modified base occurred with weak preference for the opposite DNA base (Fig. 4D, C = T ≈ G ≈ A). NEIL1 (with an excess amount) did not cleave an unmodified base opposite Tg and 5-OHU (Fig. 4B), which implies that the enzyme does not produce a double-strand break at the lesion. Taken together, the findings indicate that the primary substrates of NEIL1 are oxidized pyrimidines in double-stranded, as well as single-stranded, DNA in any opposite base context.

### Uracil/Thymine Mismatch DNA Glycosylase Activity—While examining the substrate specificity for NEIL1 and other NEIL proteins, we noticed that NEIL1 showed some activity toward uracil. As shown in Fig. 5, A and B, the recombinant mouse NEIL1 protein incised the uracil-containing substrate only when the uracil was mispaired (U:C > U:G ≈ U:T). The incision activity of NEIL1 to U:C substrate was much stronger than
that to 8-oxoG substrates (Fig. 4). Activity toward a U:A substrate could not be detected, whereas nicking of U:G and U:T was less efficient than that of U:C, and the plot of the nicking of U:G and U:T as a function of protein amount showed a threshold effect (Fig. 5, arrowheads). Because we showed that NEIL1 activity toward oxidized pyrimidines does not depend on the opposite base, we assumed that a major determinant for the observed nicking activity would be a mismatched state rather than the uracil base structure. Therefore, we also examined the opposite (undamaged) strand. An unmodified complementary strand (A, G, C, or T) was labeled and used either as a single-strand substrate (lane ss) or as a duplex substrate with A, G, C, or T in the complementary strand. The reaction with 1 nM substrate and 10 nM protein (including an inert fraction) resulted in about 50% incision for Tg:A substrate. This amount of protein was applied to the reactions of the other substrates. The reaction of E. coli Fpg with 8-oxoG:C substrate was included as a positive control for the 8-oxoG substrate. The band with an asterisk in the 5-OHU reaction is an impurity of the oligonucleotide. B, no incision at the opposite (undamaged) strand. An unmodified complementary strand (A, G, C, or T) was labeled individually and annealed with the Tg- or 5-OHU-containing strand. An equimolar mixture of each of the four duplexes (total 1 nM) was reacted with 100 nM NEIL1 (an amount 10-fold that used in A). C, incision activity on oxidative base lesions as a function of enzyme amount. The representative duplex substrates (Tg:A, AP:A, 8-oxoG:A, and 8-oxoG:C) were examined with various amounts of NEIL1 in separate experiments. D, incision activity on single strand Tg substrate and duplex Tg substrate mispaired with each of the four bases.

Subcellular Localization and Tissue Expression Profiles—Both mouse and human NEIL1 protein sequences show potential nuclear localization signals near the C terminus (see Fig. 8). Indeed, as expected for a DNA repair protein, FLAG-tagged NEIL1 protein localized in the nucleus when expressed in COS-7 cells (Fig. 6A).

It has been reported that mRNA for mouse NTH1 is not expressed homogeneously throughout the various tissues (27). Because NTH1 and NEIL1 target the same Tg substrate, this redundancy might also be reflected at the level of tissue-dependent expression. To address whether NeiI expression is ubiquitous or specific in tissues, we performed a quantitative PCR assay using an equalized tissue mouse cDNA panel. Fig. 6B shows the expression levels of Nth1 and NeiI compared in various tissues, relative to the mouse glyceraldehyde-3-phosphate dehydrogenase (Gapdh) used as a control. The highest expression levels of Nth1 are observed in oxygen-exposed lung, whereas the lowest Nth1 mRNA levels are detected in skeletal muscle (approximately 50 times lower than lung). For most...
tissues, expression of Neil1 is lower than that of Nth1. Exceptions are the spleen (comparable expression level to Neil1 and Nth1) and heart and skeletal muscle (Neil1 more abundantly expressed than Nth1). Neil1 may compensate the repair capacity in some tissues where Nth1 is less expressed.

Identification of the Endogenous NEIL1 Activity in Mouse Cells—Because exogenously expressed NEIL1 has a nuclear localization with Tg-DNA glycosylase activity, we considered the possibility that endogenous NEIL1 is responsible for one of the two nuclear Tg-DNA glycosylase activities (TGG1 and TGG2) detected previously in liver extracts from the Nth1 knockout mouse. The TGG1 and TGG2 activities could be separated in HPLC using an UNO S column with NaCl gradient elution. Although TGG1 was present in the flow-through fraction containing 0.1 M NaCl, TGG2 eluted in the fractions containing about 0.4 M NaCl (14). When recombinant NEIL1 protein was subjected to UNO S column chromatography, the protein eluted in the high salt fraction (data not shown). Our observation that high salt conditions inhibit NEIL1 activity (see Fig. 2C), this suggests strongly that we have missed the endogenous NEIL1 activity in our efforts to identify the Tg-glycosylase back-up activity from NTH1-deficient mouse livers. To clarify this point, we re-examined the UNO S fractions of the Nth1 knockout nuclear extracts under conditions where the NaCl concentration was kept below 0.15 M NaCl in the incision reaction. With this modification, we found two novel activities (Fig. 7A, upper panel, fractions 20 and 23) in addition to TGG1 (fraction 3) and TGG2 (fraction 16). Because recombinant NEIL1 protein (termed rNEIL1 to distinguish it from endogenous protein) also eluted in fraction 23 under the same chromatographic conditions, we next investigated whether the endogenous activity in fraction 23 is derived from the mouse NEIL1 protein. To this end, we performed a sodium cyanoborohydride (NaCNBH₃) trapping assay. This reagent reduces the Schiff base intermediate and provokes the formation of a stable cross-link between enzyme and DNA (28). As expected, the migration of the enzyme-DNA complex from fraction 23 in SDS-PAGE was identical to that of the rNEIL1-DNA complex (Fig. 7A). Moreover, when the enzyme-DNA complexes were subjected to limited proteolysis with trypsin or chymotrypsin as indicated. Protein markers that merged in the autoradiogram are shown.
smaller molecular size than the rNEIL1-DNA complex. Interestingly however, the trypsin- and chymotrypsin-derived fingerprint are indistinguishable from those of the rNEIL1-DNA complex. Therefore, we designate this enzyme as NEIL1/H9252, which would be a processed version of NEIL1 or a degradation product resulting from purification, yet retaining its activity.

The proteolytic fingerprint of cross-linked fraction 16 is clearly different from that of NEIL1, indicating that TGG2 activity does not relate to NEIL1. TGG1 was not trapped by NaCNBH3, indicating that TGG1 is a monofunctional DNA glycosylase for Tg as reported previously (14).

We traced nicking activities to the 8-oxoG:C substrate in the UNO S column fractions (Fig. 7B, middle panel). Consistent with the substrate specificity of rNEIL1, fraction 23 showed only marginal nicking activity toward the substrate. AP endonuclease activities of the UNO S fractions were also monitored using a substrate containing a synthetic AP site (fAP) that is resistant to incision by AP lyase (29). As shown in Fig. 7B (lower panel), AP endonuclease activity was detected in a wide range of fractions (0.1–0.4 M NaCl). Fortunately, fraction 23 was hardly contaminated with AP endonuclease activity, which would modify the 3’-end structure of the product nicked by AP lyase. The absence of AP endonuclease activity allowed examination of how the endogenous NEIL1 catalyzed strand incision. As shown in Fig. 7C, fraction 23 generated β- and δ-elimination products, as did rNEIL1, indicating that the endogenous NEIL1 is associated with (β-δ) AP lyase activity. From these results, it is concluded that the active NEIL1 protein is expressed and contributes to the repair of oxidized pyrimidines in Nth1 knock-out mice.

**DISCUSSION**

In the present study, we identified a mouse counterpart of the hypothetical protein FLJ22402 as a functional Nei homologue and designated it as mouse NEIL1 as proposed previously (13). In addition to NEIL1, there are two other Fpg-Nei-like proteins, NEIL2 (FLJ31644) and NEIL3 (FLJ10858) as reported previously (12). These three sequences are quite diverged, and the exon-intron composition of the corresponding genes are also different (Fig. 1B), suggesting that the three Fpg-Nei-like genes did not emerge by recent gene duplications. An intronless pseudogene for NEIL2 is found on chromosome 4, as well. Strikingly, possible counterparts of any of the three human sequences were not found in the genomes of S. cerevisiae, S. pombe, or D. melanogaster. This situation is in sharp contrast to the case of Nth, whose homologues are widespread in prokaryotes and eukaryotes. A zebrafish ortholog of NEIL1 is predicted in the EST database, but other eukaryotic candidates for a functional NEIL1 homologue are as yet undetected. Therefore, the presence of Nei homologues in human, mouse, and zebrafish is suggestive of special functional significance in vertebrates.

Functions of NEIL2 and NEIL3 proteins remain to be elucidated. In preliminary tests with in vitro translated NEIL2 and NEIL3 (1–289), DNA glycosylase activities could not be detected toward the following substrates: Tg:A (Fig. 1), 5-OHU:G, 8-oxoG:C, and 8-oxoG:A. Among the three NEIL proteins NEIL2 shows only low (and local) homology to the Fpg-Nei family proteins, whereas the DNA glycosylase domain of NEIL3 shows significant homology over the entire sequence. Within the 268 N-terminal amino acids, NEIL3 displays 22% identity and 38% similarity to E. coli Nei. However, NEIL3 lacks absolutely conserved Pro2 (Pro1 in prokaryote), which serves as a nucleophilic attack and is a residue forming the Schiff base intermediate with the resulting AP site. Because we have only tested the truncated version of NEIL3, we have not excluded the possibility that a full-length NEIL3 protein may...
have enzymatic activity. Very recently, NEIL2 was reported to have DNA glycosylase/AP lyase activity for oxidative cytosome, with highest activity for 5-hydroxuracil (13). In Fig. 8, the human and mouse NEIL1 sequences are aligned with a consensus sequence composed from the known Fpg and Nei sequences by a conserved domain search (30). This alignment highlights the strong similarity among the DNA-contacting elements predicted from the structures of Tth-Fpg and Nei. In this regard, the structure of the mammalian NEIL1 protein from the N terminus through the H2TH motif should be similar to those of the bacterial Tth-Fpg and Nei. However, the homology to the bacterial Nei disappears in exon 5–10. The lack of a zinc-finger motif means a loss of the known active site residue (i.e. Arg-252 in E. coli Nei). Therefore, the varied mammalian C-terminal region should contain an alternative active site residue and a required DNA binding element. When comparing human and mouse sequences, the homology is reduced dramatically behind exon 8 (89 to 50%). This suggests that the region between exon 5 and exon 7 contain an important structural element for the two mammalian enzymes to substitute the zinc-finger motif in bacterial proteins.

The substrate specificity of E. coli Nei has been examined with oligonucleotides containing Tg, uracil glycol, 5-OHU, 5-hydroxycytosine, dihydrouracil, dihyrothymine, and formyluracil (24, 31, 32). We showed the activity of mouse NEIL1 toward Tg and 5-OHU substrates. These data are in agreement with those reported recently (12). 5-OHU is a potentially mutagenic lesion (33), whereas Tg is known to be a strong block of replicative DNA polymerases (34), although both might be mutagenic if error-prone translesion synthesis by the UmuC family of DNA polymerases (35) takes place at the lesions. In addition to oxidative pyrimidines, E. coli Nei has some activity toward 8-oxoG (26), contributing to the suppression of G:C to T:A transversions in a fpg and mutY double mutant (25). In contrast to a recent report (11), we detected only a marginal incision activity toward the 8-oxoG substrates (Fig. 4). Therefore, the relevance of this activity to in vivo repair should be discussed with more data. NEIL1 does not show a significant preference as to the bases opposite Tg, 5-OHU, and AP lesions. This implies that the enzyme does not contribute to preventive repair as is the case of Fpg and OGG1, which incise 8-oxoG specifically from the 8-oxoG:C pair. Oxidative base lesions such as 8-oxoG (36) and Tg (37) are the target of transcription-coupled repair (TCR), although involvement of DNA glycosylase in TCR has not been established. NEIL1 is a possible candidate if the observed TCR to Tg requires a specific lesion recognition, because NEIL1 can act on a single-stranded DNA that is formed along the transcription. The ability to act on lesions in single-stranded DNA has been reported only for uracil-DNA glycosylases (38).

In contrast to the weak activity toward 8-oxoG, mouse NEIL1 possesses a more significant nicking activity to uracil or thymine mismatch. To date four DNA glycosylases for uracil repair have been described:UNG, SMUG1, TDG, and MBD4 (6, 39). UNG and SMUG1 repair U:G and U:A lesions but do not act on Tg, whereas TDG and MBD4 act on both U:G and Tg but not on U:A. The U:G and T:G mismatches are generated by deamination of cytosine and 5-methylcytosine, whereas U:A results from incorporation of dUMP opposite adenine during replication. Because NEIL1 does not act on U:A lesions, its activity and substrate recognition resemble those of TDG and MBD4. However, the incision occurs in the context of U:C and T:C rather than in U:G and T:G mismatches. Therefore, the role of mismatch incision of NEIL1 should be different from the primary function of the members of uracil-DNA glycosylase.

The physiological generation of U:C and T:C mismatches is unclear, but one can speculate that it might be the result of replication errors. These data suggest that NEIL1 may play a role in mismatch repair during replication.

In conclusion, we have identified a novel DNA glycosylase (NEIL1) in Nth1−/− mice and have characterized further DNA glycosylases (TGG1 and TGG2) found previously. Thus, the genetic of the wild-type mouse is protected from harmful pyrimidine oxidation by at least four DNA glycosylases, namely, NTH1, NEIL1, TGG1, and TGG2. Two of the glycosylases, NTH1 and TGG1, are present in mitochondria, and three of them, NTH1, NEIL1 and TGG2, are present in nuclei. It is an interesting and important question why mammals provide so many enzymes for oxidative pyrimidine damage. The simplest answer would be that the damage is too extensive to be repaired by a single enzyme. However, if so, cells might express a greater amount of one enzyme rather than more kinds. We think it more likely that the substrate range, subcellular localization, or repair mode (e.g. replication-associated repair or TCR) of these enzymes may be different, and these should be elucidated in further studies.

References—We thank Dr. S. Iwai (University of Tokyo) for proving DNA substrate containing Tg, 8-oxoG, and IAP. We thank Drs. S. McCready and S. Nakajima for editing text and figures, respectively.

Acknowledgments—
32. Zhang, Q.-M., Miyabe, I., Matsumoto Y., Kino K., Sugiyama, H., and Yonei, S. (2000) J. Biol. Chem. 275, 35471–35477
33. Kreutzer, D. A., and Essigmann, J. M. (1998) Proc. Natl. Acad. Sci. U. S. A. 95, 3578–3582
34. Evans, J., Maccabee, M., Hatahet, Z., Courcelle, J., Bockrath, R., Ide, H., and Wallace, S. (1993) Mutat. Res. 299, 147–156
35. Wang, Z. (2001) Mutat. Res. 486, 59–70
36. Le Page, F., Kwoh, E. E., Avrutskaya, A., Gentil, A., Leadon, S. A., Sarasin, A., and Cooper, P. K. (2000) Cell 101, 159–171
37. Avrutskaya, A. V., and Leadon, S. A. (2000) Methods 22, 127–134
38. Schaerer, O. D., and Jiricny J. (2001) BioEssays 23, 270–281
39. Pearl, L. H. (2000) Mutat. Res. 460, 165–181