Article No. mb982042 *J. Mol. Biol.* (1998) **282**, 761–774

JMB

Cloning and Characterization of a Mouse Homologue (mNthl1) of *Escherichia coli* Endonuclease III

Altaf H. Sarker[1], Shogo Ikeda[2], Hironobu Nakano[3], Hiroaki Terato[3] Hiroshi Ide[3], Kotoe Imai[1], Kosuke Akiyama[1], Ken Tsutsui[1], Zhang Bo[1] Kihei Kubo[4], Kazuo Yamamoto[5], Akira Yasui[6], Michihiro C. Yoshida[7] and Shuji Seki[1]*

[1]*Department of Molecular Biology, Institute of Cellular & Molecular Biology, Okayama University Medical School Okayama 700-8558, Japan*

[2]*Department of Biological Chemistry, Okayama University of Science, Okayama 700-0005, Japan*

[3]*Faculty of Science, Hiroshima University, Higashi-Hiroshima 739-8526, Japan*

[4]*Department of Radiology School of Veterinary Medicine Osaka Prefectural University Sakai 599-8531, Japan*

[5]*Biological Institute, Graduate School of Science, Tohoku University, Sendai 980-8578 Japan*

[6]*Department of Molecular Genetics, Institute of Development, Aging and Cancer, Tohoku University Sendai 980-8578, Japan*

[7]*Chromosome Research Unit Faculty of Science, Hokkaido University, Sapporo 060-0180 Japan*

**Corresponding author*

Endonuclease III (endoIII; *nth* gene product) of *Escherichia coli* is known to be a DNA repair enzyme having a relatively broad specificity for damaged pyrimidine bases of DNA. Here, we describe the cloning and characterization of the cDNA and the gene for a mouse homologue (mNthl1/mNth1) of endoIII. The cDNA was cloned from a mouse T-cell cDNA library with a probe prepared by PCR using the library and specific PCR primers synthesized based on the reported information of partial amino acid sequences of bovine NTHL1/NTH1 and of EST Data Bases. The cDNA is 1025 nucleotides long and encodes a protein consisting of 300 amino acids with a predicted molecular mass of 33.6 kDa. The amino acid sequence exhibits significant homologies to those of endoIII and its prokaryotic and eukaryotic homologues. The recombinant mNthl1 with a hexahistidine tag was overexpressed in a *nth*::cm^r *nei*::Km^r double mutant of *E. coli*, and purified to apparent homogeneity. The enzyme showed thymine glycol DNA glycosylase, urea DNA glycosylase and AP lyase activities. Northern blot analysis indicated that mNthl1 mRNA is about 1 kb and is expressed ubiquitously. A 15 kb DNA fragment containing the mNthl1 gene was cloned from a mouse genomic library and sequenced. The gene consists of six exons and five introns spanning 6.09 kb. The sequenced 5′ flanking region lacks a typical TATA box, but contains a CAAT box and putative binding sites for several transcription factors such as Ets, Sp1, AP-1 and AP-2. The mNthl1 gene was shown to lie immediately adjacent to the tuberous sclerosis 2 (Tsc2) gene in a 5′-to-5′ orientation by sequence analysis and was assigned to chromosome 17A3 by *in situ* hybridization.

Keywords: thymine glycol DNA glycosylase; urea DNA glycosylase; mouse Nthl1 gene; DNA glycosylase; AP lyase

Introduction

Oxygen radicals and other reactive oxygen species generated as a by-product of normal metabolism or by external sources such as ionizing radiation can cause extensive base damage and strand breaks in DNA. Thymine residues of DNA are known to be especially susceptible to oxidative

Abbreviations used: endoIII, endonuclease III; endoIV, endonuclease IV; PAGE, denatured polyacrylamide gel electrophoresis; Fapy DNA glycosylase, formamidopyrimidine DNA glycosylase; TG, thymine glycol; APEX, APEX nuclease.

E-mail address of the corresponding author: spseki@cc.okayama-u.ac.jp

0022–2836/98/390761–14 $30.00/0

DNA damage, and the major detectable product of this damage is thymine glycol (Cathcart *et al.*, 1984). Repair of the lesion is generally initiated by DNA glycosylase activity of such enzymes as endonuclease III (endoIII) of *Escherichia coli* which is followed by the AP lyase activity of the same enzymes (Bailly & Verly, 1987; Friedberg *et al.*, 1995). EndoIII can excise not only thymine glycol but also other damaged pyrimidine derivatives such as 5,6-dihydrothymine, 5-hydroxy dihydrothymine, 5-hydroxycytosine, 5-hydroxyuracil, uracil glycol, β-ureidobutyric acid, 5-hydroxy-5-methyl hydantoin and urea (Friedberg *et al.*, 1995; Rolden-Arjona *et al.*, 1996; Krokan *et al.*, 1997). These lesions are cytotoxic or mutagenic, unless they are repaired properly (Friedberg *et al.*, 1995; Rolden-Arjona *et al.*, 1996; Krokan *et al.*, 1997; Zou *et al.*, 1995).

Enzyme activities functionally analogous to endoIII have been identified in several prokaryotic and eukaryotic cells through using UV-irradiated DNA, chemically oxidized DNA, and γ-irradiated DNA as substrates (Friedberg *et al.*, 1995; Demple & Harrison, 1994; Bacchetti & Benne, 1975; Kim & Linn, 1989; Nes & Nissen-Meyer, 1978; Doetsch *et al.*, 1986). Recently, a bovine 31 kDa enzyme functionally homologous to endoIII was highly purified, and its partial amino acid sequences revealed homology to the predicted amino acid sequence of a *Caenorhabditis elegans* 27.8 kDa protein which, in turn, has homologies to endoIII (Hilbert *et al.*, 1996). While we were conducting the present study, cDNA cloning and characterization of a functional human homologue of endoIII were reported by two groups (Hilbert *et al.*, 1997; Aspinwall *et al.*, 1997). The gene for the human enzyme is located on chromosome 16p13.3, immediately adjacent to the tuberous sclerosis 2 (TSC2) gene in a 5′-to-5′ orientation (Hilbert *et al.*, 1997; Aspinwall *et al.*, 1997).

In the present study, we cloned cDNA for a mouse homologue (Mouse Nthl1/Nth1, abbreviated as mNthl1/mNth1†) of endoIII using a probe amplified by PCR based on the bovine homologue data reported by Hilbert *et al.* (1996) and the data of 3′-expressed sequence tags (ESTs) from *Rattus sp.* PC 12 cells (accession number H33255) and from *H. sapiens* brain tissue (accession number F04657) in Data Bases. Characterization of the recombinant mNthl1 protein expressed in *E. coli*, mNthl1 gene cloning, sequencing and chromosome mapping were also performed.

† After the submission of this paper, the gene name, nth (*E. coli* endonuclease III)-like 1 and the gene symbol, NTHL1 have been given by the HUGO/GDB Nomenclature Committee for the human homologue of *Escherichia coli* endonuclease III. Therefore, the present paper is adjusted using the new symbol, Nthl1 (or mNthl1) instead of the present Nth1 (or mNth1) for the symbol of the mouse homologue of *Escherichia coli* endonuclease III with some exceptions.

Results

Cloning and characterization of the mNthl1 cDNA

A mouse homologue (designated as mNthl1/mNth1) of *E. coli* endonuclease III (*nth* gene product) was cloned and sequenced as described in Materials and Methods. The transcription initiation site was determined by the 5′ RACE technique (Hirzmann *et al.*, 1993; Akiyama *et al.*, 1995). The 5′ RACE product carried an extra G residue at the 5′ end which is not present in the sequence of the genomic clone. The extra G residue was believed to be derived from the terminal cap G nucleotide (Hirzmann *et al.*, 1993; Akiyama *et al.*, 1995). Therefore, the following A residue was assigned as the transcription initiation site. The complete mNthl1 cDNA determined was 1025 bp long with an open reading frame of 900 bp encoding 300 amino acid residues. The cDNA terminated with a poly(A) tail, 22 nucleotides downstream from an authentic polyadenylation signal, AATAAA (Proudfoot & Brownlee, 1976). The mNthl1 protein encoded by the cDNA has a calculated molecular weight of 33,646 and a calculated p*I* of 9.97. Northern blot analysis using Mouse Multiple Tissue Northern Blot (Clontech) with the mNthl1 cDNA as a probe identified a single 1 kb transcript in all the tissues examined (Figure 1).

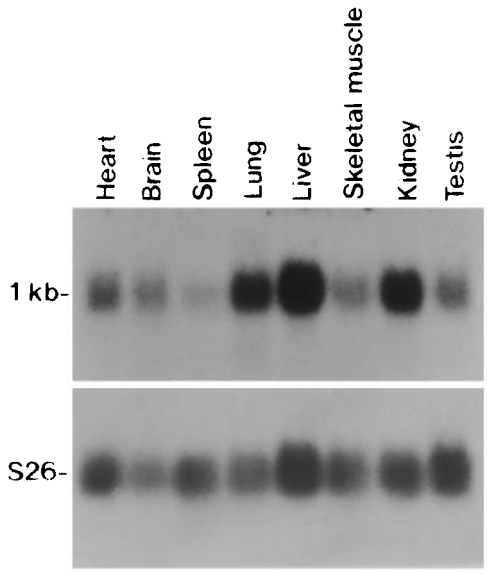

Figure 1. Expression of mNthl1 gene in mouse tissues. Polyadenylated RNA obtained from the tissues indicated was electrophoresed at 2 μg in each lane and blotted on a nylon membrane (Mouse Multiple Tissue Northern Blot). The blot was hybridized with the ^{32}P-labelled mNthl1 cDNA probe, and autoradiographed. The membrane was rehybridized with a cDNA probe for the mouse S26 ribosomal protein to correct for the amount of mRNA loaded.

Analysis of the deduced amino acid sequence of mNthl1

The deduced amino acid sequence of the mNthl1 was compared with those of endoIII and its eukaryotic homologues (Figure 2). The amino acid sequence of the C-terminal, 2/3 region of mNthl1 has a significant homology with that of almost the entire region of endoIII (26.5% identity/204 amino acids). The helix-hairpin-helix (HhH) motif, which is common in enzymes having DNA glycosylase activity and is involved in interaction with DNA, is preserved in mNthl1. The four cysteine residue motif (Cys-X_6-Cys-X_2-Cys-X_5-Cys) that constitutes the iron/sulphur (4Fe-4S) cluster is present at the C-terminal region of all of the reported endoIII homologues except in that of a putative non-Fe-S homologue (Eide *et al*., 1996) of *Saccharomyces cerevisiae*. The potential nuclear targeting sequences are also observed at positions 44/48 (RHPRR), 91/94 (RSKK) and 113/119 (PKVRRYQ) (Nakai & Kanehisa, 1992). The N-terminal 24-amino acid region of mNthl1, which is not present in the endoIII or in the other prokaryotic homologues, contains putative mitochondrial targeting sequence (rich in arginine, hydrophobic and hydroxylated residues) and potentially sensitive sequences for mitochondrial matrix protease at positions 10/13 (RSRS) and 17/20 (RIAS) (Nakai & Kanehisa, 1992; Gavel & Heijne, 1990).

Purification and functional characterization of the overexpressed mNthl1 protein

The hexahistidine-tagged mNthl1 protein was overexpressed in the *nth*$^-$ and *nei*$^-$ double mutant of *E.coli*. The overexpressed protein was purified to apparent homogeneity by a combination of phosphocellulose chromatography and affinity chromatography using a nickel chelate column (data not shown). Contamination with endoIII and endoVIII of *E. coli* was avoided by using the *nth*$^-$ and *nei*$^-$ double mutant. The purified enzyme had an apparent molecular mass of 36 kDa on SDS-PAGE, which is close to the calculated molecular mass of 35 kDa of the recombinant protein.

mNthl1-dose-dependent release of ^{3}H-labelled material was observed when the enzyme was incubated with M13 double-stranded DNA containing scattered ^{3}H-labelled thymine glycol residues ([^{3}H]Tg-DNA), but not when it was incubated with M13 double-stranded DNA containing scattered ^{3}H-labelled thymine residues ([^{3}H]T-DNA; Figure 3). The released ^{3}H-labelled material was shown to be thymine glycol by analysis using reverse phase HPLC (data not shown). This result indicates that mNthl1 exhibits thymine glycol DNA glycosylase activity.

The enzymatic properties of mNthl1 were further analysed using synthetic oligonucleotides (see Material and Methods for the sequences and abbreviations of the synthetic oligonucleotides used). First, the site-specific introduction of thymine glycol in the synthesized oligonucleotide, 19TG, was confirmed as follows: the 5′-labelled 19TG was treated at 90°C for 30 minutes with 10% piperidine. After removing the piperidine by vacuum evaporation, the sample was analysed by 16% denatured polyacrylamide gel electrophoresis (16% PAGE). The 19TG was annealed with the complementary strand, 19COM, and then incubated at 37°C for 30 minutes with *E. coli* Fapy DNA glycosylase. The product was also analysed by 16% PAGE. As shown in Figure 4A (see also Figure 8), the 19TG was cleaved completely by the treatment with piperidine and produced a single band (lane 4), although the control, 19T, was not cleaved (lane 2). The product cleaved by the piperidine-treatment was the same as that produced from the 19TG/19COM by β,δ-elimination reaction with Fapy DNA glycosylase (Figure 4A, lane 5). The results indicated that all the oligonucleotides 19TG had the thymine glycol residue at the same, expected position.

Annealed, double-stranded substrates, 19T/19COM (control), 19TG/19COM and 19UREA/19COM were treated at 37°C for 30 minutes with Fapy DNA glycosylase, endoIV, endoIII or mNthl1. The reaction products were analysed by 16% PAGE (Figures 4B and 8). Although the control 19T was not cleaved by mNthl1 (Figure 4B, lane 5), the 19TG containing a thymine glycol residue and the 19UREA containing a urea residue at the specific position were cleaved by mNthl1 (lanes 10 and 15, respectively), and the cleaved products were electrophoresed at the same position as that of the product cleaved by endoIII. Comparing the mobility of the mNthl1-cleaved product with the mobilities of the products produced by treating the 19TG/19COM and 19UREA/19COM with Fapy DNA glycosylase, endoIV and endoIII, the incision mode of mNthl1 on the oligonucleotide having the thymine glycol or urea residue is thought to be β-elimination, similar to that of endoIII (when treated with Fapy DNA glycosylase, the β,δ-elimination product is produced, lanes 7 and 12; based on a personal communication from Hatahet, Wallace *et al*.; when treated with endoIV, the product having a 3′-OH terminus is produced, lane 13; when treated with endoIII, the β-elimination product is produced, lanes 9 and 14). In addition, the conversion of the thymine glycol on the 19TG into urea by alkaline treatment at pH 12 was confirmed by differential recognition with the two enzymes, endoIII and endoIV (e.g. the 19TG was incised only by endoIII, whereas the 19UREA was incised by both enzymes, endoIII and endoIV; Figure 4B, lanes 8,9,13,14). When AP sites are incised by AP lyase activity of mNthl1, the nicked sites are thought to have 5′ termini bearing 5′-phosphonucleotide and 3′ termini bearing the α,β-unsaturated aldose 5-phosphate (3′ tags). APEX nuclease (a major mammalian AP endonuclease; also called as HAP1, APE, Ref-1; Friedberg *et al*., 1995) is thought to remove the 3′ tags by it's DNA 3′ repair diesterase activity and to restore 3′

Figure 2. Alignment of the amino acid sequence of mNthl1 with those of putatively homologous proteins. The amino acid sequence of mouse endonuclease III homologue (mNthl1) is aligned with homologous sequences (abbreviated name and accession number in brackets) from human (hNTHL1: AB001575), *C. elegans* (C. ele.: Z50874), *S. pombe* (Nth-spo: Z67961), *S. cerevisiae* (Sec: Z74785), *S. cerevisiae* non-Fe-S homologue (Sec-nFeS: L05146), and *E. coli* (Eco-nth: J02857). The alignment was plotted by the *Clustal W* (ver.1.6) program. Amino acids that are identical in at least five of the sequences are boxed. Numbers in the right column refer to the last amino acid residue in each line of the respective protein sequence. Continuous lines indicate the helix-hairpin-helix motif and the four cysteine [4Fe-4S] motif. The proposed active site lysine residue is indicated by a filled circle and the conserved aspartic acid residue is indicated by an open circle.

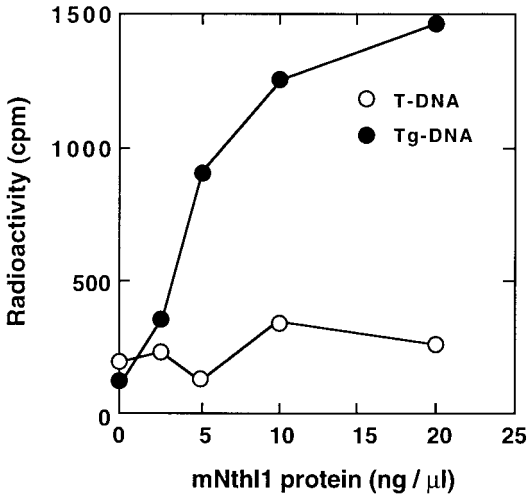

Figure 3. Enzymatic release of radioactive material (^{3}H-labelled thymine glycol) by mNthl1 protein. [^{3}H]Tg-DNA and [^{3}H]T-DNA (both 1.5×10^5 cpm, ca 0.1 μg) were incubated in 10 μl of buffer C (20 mM Hepes-KOH (pH 8.0), 50 mM KCl, 0.25 mM EDTA, 0.25 mM DTT, 0.1 mg/ml BSA) at 37°C for 30 minutes. To separate DNA and released materials, the reaction mixture was diluted with 90 μl of distilled water and loaded onto a Bio-Rad AG 1-X2 (Cl-form) column (0.5 ml). The column was washed with 1.5 ml of distilled water, and radioactivity of the eluant was counted with a liquid scintillation counter.

hydroxyl termini. Sequential treatment of the 19TG/19COM with mNthl1 and APEX nuclease (hAPEX) produced the expected product with 3′ hydroxyl terminus, the same product as that produced from the 19UREA/19COM by endoIV (Figure 5A, lanes 7 and 9, respectively).

AP lyase activity of mNthl1 was also tested using the annealed oligonucleotides, 25U/25COM and 21AP/21COM, as described in Materials and Methods. These oligonucleotides were incised by mNthl1 through the β-elimination reaction (AP lyase activity) at the abasic site, which was introduced by uracil DNA glycosylase on the 25U/25COM (Figure 5B, lane 4) or was pre-existent on the 21AP/21COM (data not shown). When the sample was treated with endoIII in place of mNthl1, an additional band appeared above the β-elimination product (Figure 5B, lane 5). The appearance of such an additional band has been reported previously, and there is some argument concerning its nature (Bailly & Verly, 1987, 1988; Purmal *et al*., 1996; Mazumder *et al*., 1991; Bailly *et al*., 1989a,b).

Southern blot analysis

The rough mNthl1 gene structure in the mouse genome was studied by Southern blot analysis of mouse liver DNA digested with restriction enzymes using the mNthl1 cDNA as a probe.

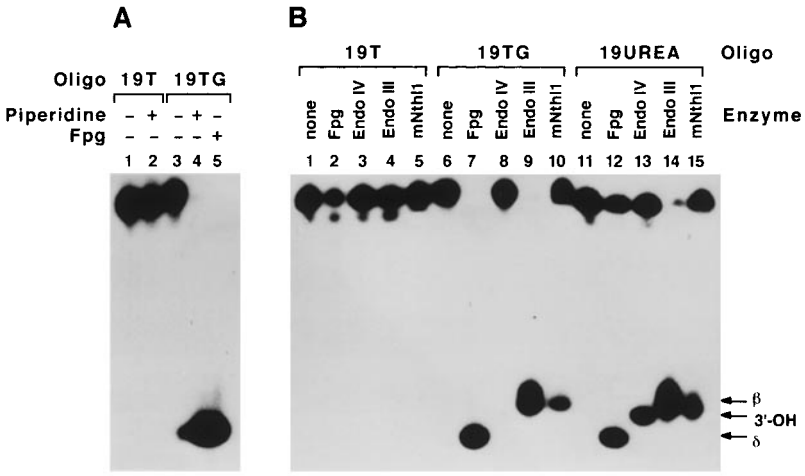

Figure 4. Characterization of oligonucleotide 19TG by treatment with piperidine and Fapy glycosylase (A), and analysis of mNthl1 action of 5′-labelled oligonucleotide substrate containing thymine glycol or urea (B). A, Piperidine treatment: 5′-^{32}P-labelled 19T and 19TG (0.5 pmol each; see the 'Preparation of DNA and oligonucleotide substrates for enzymatic analysis' section of Materials and Methods for the oligonucleotide sequences and abbreviations) were incubated at 90°C for 30 minutes in 200 μl of 10% piperidine solution. After removing piperidine under reduced pressure, samples were washed twice with distilled water, and then dissolved in the gel loading buffer (0.1% (w/v) xylene cyanol, 20 mM EDTA, 95% formamide). Fapy glycosylase-treatment: The labelled 19T or 19TG was annealed with the complementary strand, 19COM. The annealed substrate (2 nM) in 10 μl of buffer A (10 mM Tris-HCl (pH 7.5), 100 mM NaCl, 1 mM EDTA) was incubated at 37°C for 30 minutes with 100 ng of Fapy glycosylase. The reaction was stopped by adding the gel loading buffer. The products were analysed by 16% denatured polyacrylamide gel electrophoresis (16% PAGE). B, The reaction buffer used was the buffer A for endoIII and Fapy glycosylase (Fpg), buffer B (120 mM Tris-HCl (pH 7.5), 50 mM NaCl, 1 mM EDTA) for endoIV and the buffer C (described in the legend to Figure 3) for mNthl1. The annealed duplex substrate (19T/19COM, 19TG/19COM or 19UREA/19COM) at 2 nM in 10 μl of an appropriate buffer was incubated at 37°C for 30 minutes with 100 ng Fpg, 20 ng endoIV, 100 ng endoIII or 10 ng mNthl1. The reaction product was analysed by 16% PAGE. Positions β, 3′-OH and δ indicate the electrophoresed positions of the β-elimination product, 3′-OH product and β,δ-elimination product, respectively (see also Figure 8).

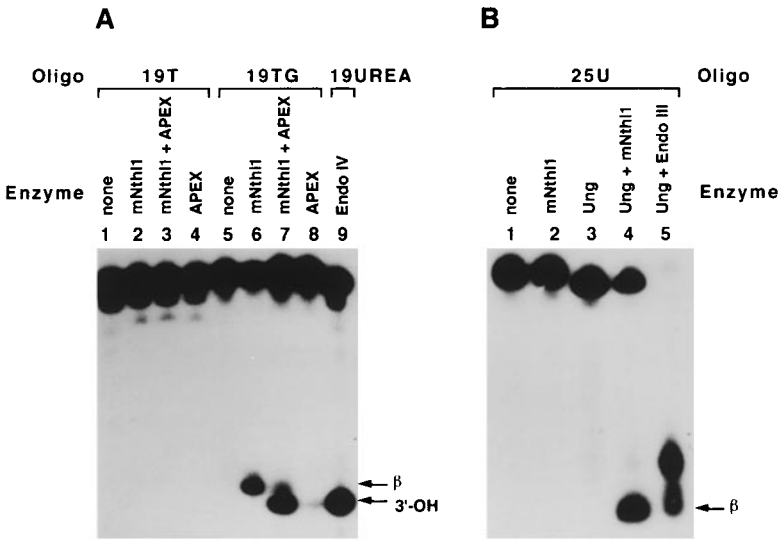

Figure 5. Removal of the 3′-tags produced by enzymatic reaction on 19TG/19COM by APEX nuclease (hAPEX) (A), and AP lyase activity of mNthl1 (B). A, The annealed duplex substrate (19T/19COM, 19TG/19COM or 19UREA/19COM) was incubated with mNthl1 or endoIV under the same conditions as described in the legend to Figure 4. After incubation, $MgCl_2$ (3 mM at the final concentration) and 10 ng hAPEX were added to the appropriate tubes, and incubation at 37°C was continued for an additional 30 minutes. Treatment with hAPEX alone was conducted at 37°C for 30 minutes in buffer D (10 mM Tris-HCl (pH 7.5), 3 mM $MgCl_2$, 6 mM 2-mercaptoethanol). B, The annealed duplex substrate containing uracil (25U/25COM) in 10 μl of buffer E (20 mM Tris-HCl (pH 8.0), 50 mM NaCl, 1 mM DTT) was treated in advance with 0.1 unit uracil *N*-glycosylase (Ung) to produce an abasic site by removing the uracil. After the reaction, 10 ng mNthl1 or 100 ng endoIII was added to the reaction mixture, and incubation was conducted at 37°C for 30 minutes. The other conditions are the same as those described in the legend to Figure 4.

The results (fragment numbers and sizes produced by these restriction enzymes) shown in Figure 6 correlate well with the restriction fragment patterns deduced from the 15 kb genomic sequence containing the mNthl1 gene described later and the mNthl1 cDNA sequence described above. These results suggest that the mNthl1 gene is present as a single copy in the haploid genome.

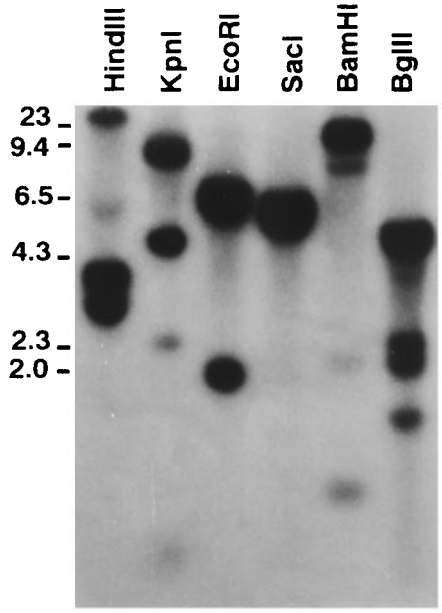

Figure 6. Southern blot analysis of the mNthl1 gene. Mouse genomic DNA was digested with the restriction enzymes indicated at the top of the Figure. Five micrograms of DNA per lane were subjected to electrophoresis on a 0.7% agarose gel, and then transferred onto a nylon membrane. The blots were hybridized with the mNthl1 cDNA excised from the clone A and labelled with ^{32}P. Numbers on the left indicate the size marker.

Cloning and sequencing of the mNthl1 gene

A mouse leukocyte genomic library in EMBL-3 bacteriophage vectors was screened with randomly labelled mNthl1 cDNA as a probe. Sixteen positive clones were isolated from 1×10^6 independent clones. Among them, two clones containing the entire region of the mNthl1 gene were analysed. An 11 kb insert in one of two clones was subcloned into the pBluescript KS(−) plasmid vector, and the recombinant was designated as mNthG-3. A 13 kb insert in the other clone was subcloned into the pUC18 plasmid vector, and the recombinant was designated as mNthG-5. The inserts of both mNthG-3 and mNthG-5 clones were sequenced. By comparing the sequence with the mNthl1 cDNA sequence, we defined the exon/intron structure of the mNthl1 gene (Figure 7). The gene consists of six exons and five introns, with a total length of 6090 bp. All exon/intron junctions follow the GT/AG rule (Breathnach & Chambon, 1981). The translation initiation and termination codons (ATG and TAA) are located in exons 1 and 6, respectively (Figure 7). The total length of all exons (1025 bp) is consistent with the length of the mNthl1 cDNA (mRNA) described above, and the nucleotide sequence of each exon is completely identical to that of the corresponding region of the mNthl1 cDNA. In order to tentatively assign a base position within the gene and in the flanking regions, the adenine residue in the translation initiation codon of the mNthl1 gene was designated as position +1 and the 5′ neighbouring residue C as position −1. With this as a starting point, positive numbers are given to positions downstream from the starting point and negative numbers to those upstream. The mNthG-5 and mNthG-3 clones contained a 5′ flanking region of

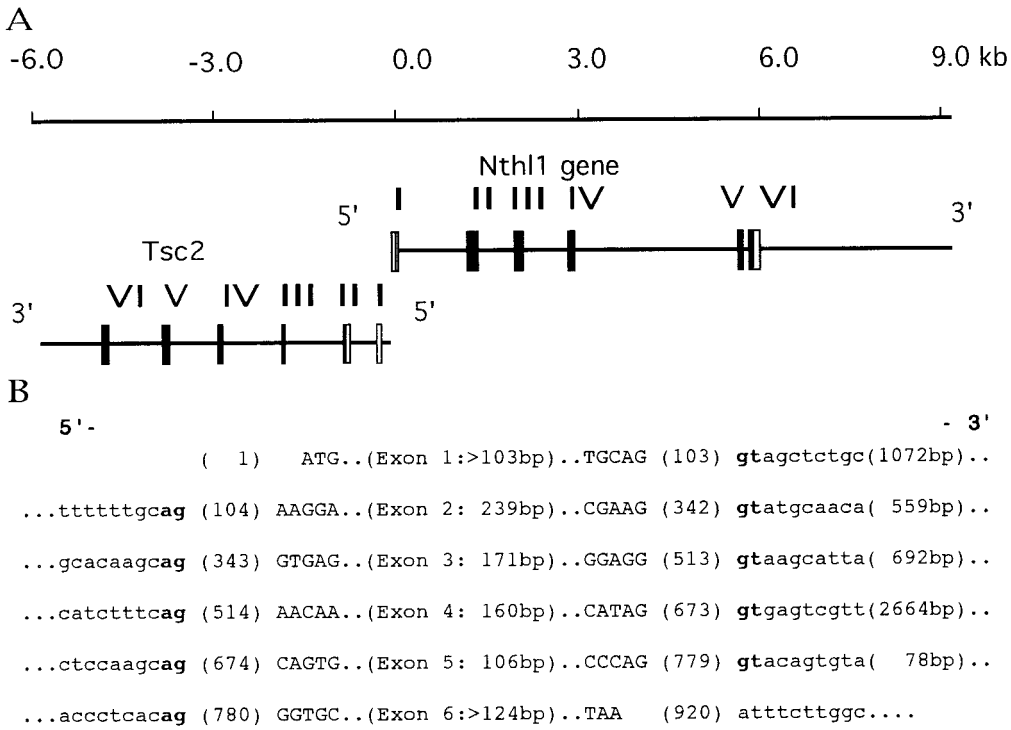

Figure 7. The structure and exon-intron organization of the mouse Nthl1 gene and its flanking regions. In order to assign tentatively a base position within the gene and in the flanking regions, the adenine residue in the translation initiation codon of the mNthl1 gene was designated as position +1 and the 5′ neighbouring residue C as position −1. With this as a starting point, positive numbers are given to positions downstream from the starting point and negative numbers to those upstream. A, From the top: scale in kilobases, a diagram of the mouse Nthl1 gene and the 5′ region of the mouse tuberous sclerosis 2 gene (Tsc2). The Nthl1 gene lies adjacent to the Tsc2 gene in a head-to-head orientation. Therefore, the sequence is complementary for the Tsc2 gene. The exons are numbered from the 5′ end of the gene. Filled and open boxes represent translated and untranslated regions, receptively. B, Nucleotide sequences of the exon-intron boundaries of the Nthl1 gene. Sequences of introns are given in lowercase, and those of exons in uppercase. On both sides of the parentheses for the number and size of each exon, the first and last five nucleotides are shown. The nucleotide numbers of the first and last residue of each exon are shown in parentheses before and after the first and last five nucleotides, respectively. The size of each intron is shown in parentheses on the right side.

5.83 kb and a 3′-flanking region of 3.32 kb, respectively. These regions were also sequenced, and were found to contain several repetitive sequences such as the GT, PuC and GAA repeats. The sequence analysis of the 5′-flanking region using the mouse tuberous sclerosis 2 (Tsc2) mRNA sequence reported by Kim *et al.* (1995) revealed that the mNthl1 gene lies immediately adjacent to and in a 5′-to-5′ orientation with the Tsc2 gene. The exon 1 (119 bp) of the mNthl1 gene and the exon 1 (61 bp) of the Tsc2 gene both lie within the 317 bp sequence between positions −214 and +103. In the 5′-flanking region (5.83 kb) of the mNthl1, exons 1 through 6 of the Tsc2 gene could be mapped in reference to the mNthl1 gene (Figure 7). The promoters of these genes may be overlapping or bidirectional. There is no typical TATA box at its characteristic position in relation to the possible transcription initiation site(s) of the mNthl1 gene. The 5′-flanking region of the mNthl1 gene contains a CAAT box at position −293/ − 296 and multiple GC boxes that are thought to be a consensus sequence for nuclear factor Spl-binding sites, located in intron 1 of the Tsc2 gene. The 5′ flanking region also contains several consensus sequences of binding sites for transcription factors such as Ets1 (CCGGAAGC; −141/ − 134), AP-1 (TGACTA; −102/ − 97) and AP-2 (CCCAGGGC; −154/ − 147).

Chromosome mapping

A total of 20 metaphases were examined. In 12 of them, twin-spot signals (total, 36 signals) specific for mNthl1 were detected on both homologous chromosomes 17. Such specific signals were distributed between 17A3 and 17B1 with the highest peak of signals (62% of all signals) over the distal region of 17A3 (data not shown). Therefore, the mouse Nthl1 gene was assigned to chromosome 17A3. The mNthl1 gene lies immediately adjacent to the tuberous sclerosis 2 (Tsc2) gene in a head-to-head orientation. The first exons of these two opposing genes are located within only 317 bp, indicating their close linkage.

Discussion

Here, we have cloned and characterized a mouse homologue (mNthl1) of *E. coli* endoIII. The mNthl1 cDNA is 1025 nucleotides long, and encodes a protein consisting of 300 amino acid residues. Recently, cloning of cDNA for the human homologue was reported by two groups (Hilbert *et al.*, 1997; Aspinwall *et al.*, 1997). Two other groups deposited human cDNA sequence data in Data Bases (Rognes *et al.*, accession number: Y09687; and Ikeda *et al.*, a member of our group (1997), accession number AB001575). cDNA sequence data for the mouse homologue was also deposited in Data Bases by two groups (ours, accession number: AB006812; and Rognes, accession number: Y09688). The amino acid sequence of mNthl1 showed 81.0% identity in 300 residues to the human homologue (hNTHL1 or hNTH1; nth (*E. coli* endonuclease III)-like 1; accession number: AB001575), 51.4% identity in 208 residues to the homologue of *Caenorhabditis elegans* (accession number: Z50874; Wilson *et al.*, 1994), 48.2% identity in 226 residues to the homologue of *Schizosaccharomyces pombe* (Roldan-Arjona *et al.*, 1996), and also exhibited significant homologies with the sequences of the other eukaryotic and prokaryotic homologues (Augeri *et al.*, 1997; Eide *et al.*, 1996; Asahara *et al.*, 1989; Piersen *et al.*, 1995; Bruand *et al.*, 1995; Figure 2). The helix-hairpin-helix motif and four-cysteine-residue motif (Cys-X_6-Cys-X_2-Cys-X_5-Cys) are well preserved in the mNthl1 protein similar to those of the other homologues. These two motifs are thought to form a pocket, and are implicated in the recognition of specific DNA damage and in DNA binding (Thayer *et al.*, 1995). The ε-amino group of the Lys120 in endoIII, which corresponds to Lys208 in mNthl1 and Lys212 in hNTHL1, is thought to be critical for the formation of enzyme-substrate intermediates (Thayer *et al.*, 1995). Asp227 in mNthl1 corresponds to Asp138 in endoIII, which is also known to be involved in catalytic functions (Thayer *et al.*, 1995). The fact that these critical motifs and residues are all preserved in the core region of these homologues suggests that the mNthl1 protein has, overall, a similar three-dimensional structure to that of endoIII, whose structure was recently established (Thayer *et al.*, 1995), and similar functions to those of endoIII.

The hexahistidine-tagged mNthl1 protein was overexpressed in *nth*$^-$, *nei*$^-$ *E. coli* mutant cells, and purified to apparent homogeneity. When tested on M13 double-stranded DNA containing thymine glycol and synthetic oligonucleotide containing thymine glycol, urea or an abasic site at a define position, the recombinant mNthl1 showed almost the same activities as those of endoIII. These results indicate that mNthl1 has thymine glycol DNA glycosylase, urea DNA glycosylase and AP lyase activities, similar to endoIII. Figure 8 shows schematically the key structure of the substrates used for the enzyme-assisted functional analysis and the results.

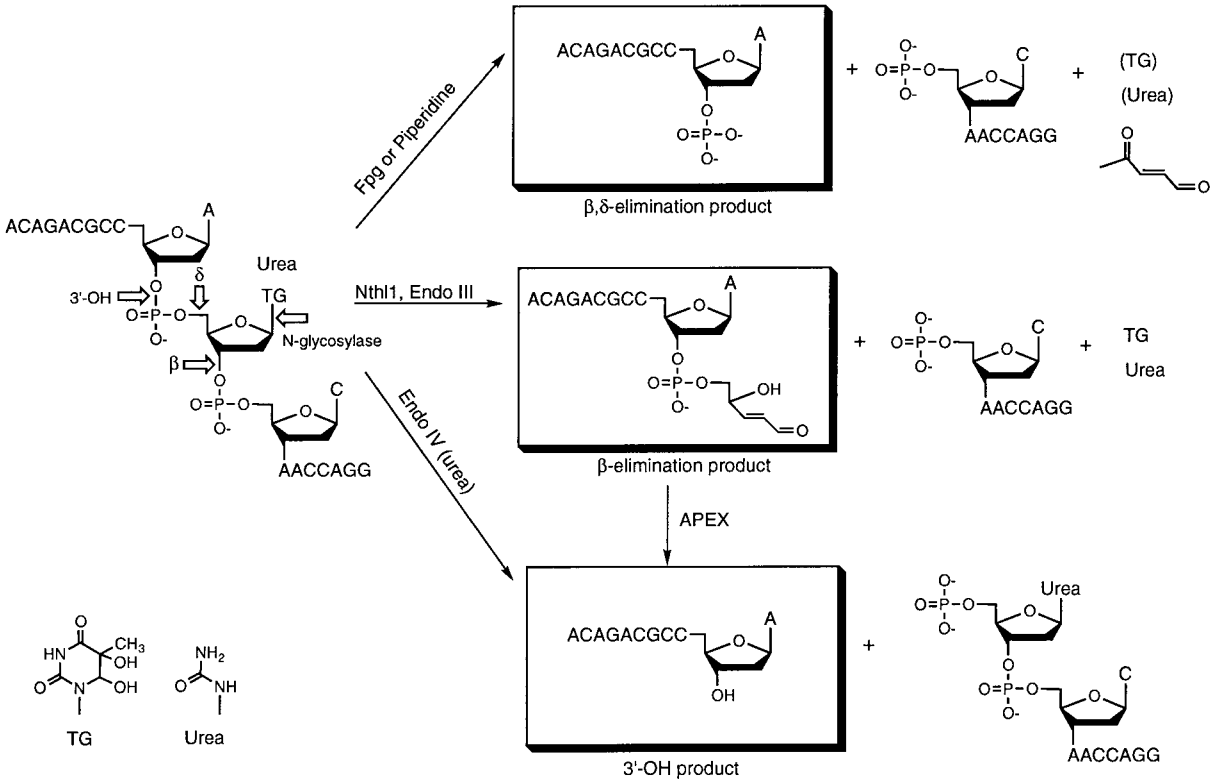

Figure 8. Schematic representation of substrates and products of mNthl1/mNth1 and the result of enzyme-assisted analysis using repair enzymes with known enzymatic properties.

The result of northern blot analysis indicate ubiquitous expression of mNthl1 mRNA, although the level of the expression is variable from tissue to tissue. Further analysis, in the form of well-controlled experiments, is necessary to better understand these tissue-to-tissue variations in expression.

As a first step toward understanding the regulation of the gene expression and to analyse its *in vivo* function, the mNthl1 gene was cloned and analysed here. The mNthl1 gene consists of six exons and five introns spanning 6.09 kb. The results of Southern blot analysis suggest a single mNthl1 gene in a haploid genome. We determined the transcription initiation site by the 5′RACE technique. After we deposited the mNthl1/mNth1 cDNA sequence in Data Bases, Rognes deposited a mouse Nth1/Nthl1 cDNA sequence in Data Bases (accession number: Y09688). Both sequences are almost identical, except that his sequence is 26 bases (gggatcgcggggctgctcggaggagt) longer at the 5′ end than ours, and has one base change in the middle region. Considering the fact that the mNthl1 gene has no typical TATA box, the above result may indicate that the transcription initiation sites of the mNthl1 gene are multiple.

Analysis of the 5′ flanking region of the mNthl1 gene revealed that the gene lies adjacent to the Tsc2 gene in the head-to-head orientation. This finding is consistent with the results reported by Aspinwall *et al*. (1997). They reported that the human NTH1/NTHL1 gene (OCTS3 gene) lies adjacent to the TSC2 gene on chromosome 16 (16p13.3) in a head-to-head orientation, although the complete sequence of the human NTHL1 gene has not yet been reported. The 5′ flanking, 5.8 kb region of the mouse Nthl1 gene reported here was analysed using the mouse Tsc2 cDNA sequence reported by Kim *et al*. (1995). The results indicated that the 5′ flanking region contains exons 1 through 6 of the Tsc2 gene. The two oppositely oriented first exons of the mouse Nthl1 and Tsc2 are located within 317 bp. The mouse Tsc2-exon 1 here corresponds to the exon 1a of rat Tsc2 which was reported recently (Kobayashi *et al*., 1997). The mouse Tsc2-exon 2 here corresponds to the mouse, rat or human tuberous sclerosis-exon 1 previously mapped (Kim *et al*., 1995; Kobayashi *et al*., 1995; Maheshwar *et al*., 1996). The difference might be due to the fact that the cDNA used for the analysis of the TSC2 gene structure was not large enough to show the present exon 1 or rat Tsc2-exon 1a, or that the transcription of TSC2 gene initiates at multiple sites. The finding of the novel leader exon at the 5′ end of Tsc2 (or TSC2) is thought to be important for studying transcriptional regulation of both Tsc2 (or TSC2) and Nthl1 (or NTHL1) genes.

We determined the position of the mNthl1 gene on chromosome 17A3 by *in situ* hybridization. It is known that the mouse Tsc2 and the polycystic kidney disease 1 (Pkd1) genes are oriented in a tail-to-tail orientation with a distance of only 63 bp between the polyadenylation signals of these genes, and that the Tsc2 and Pkd1 genes are mapped on chromosome 17 at 10.3 and 10.4 map units from the centromere, respectively (Olsson *et al*., 1995, 1996; Pilz *et al*., 1995). Based on this information and the present results, the mNthl1 gene is thought to be located proximal (centromere side) to the Tsc2 gene, opposite the Pkd1 gene on mouse chromosome 17.

Tuberous sclerosis in humans is a dominantly inherited hamartomatous disorder characterized by seizures, mental retardation, skin lesions, eye lesions and widespread development of growths (hamartoma and neoplasm; European Chromosome 16 TSC Consortium, 1993). Tuberous sclerosis is extremely variable in its manifestations and severity, and linkage studies have shown locus heterogeneity with the TSC1 gene mapped on chromosome 9q34.3 and with the TSC2 gene mapped on chromosome 16p13.3. It is also reported that the TSC2 gene functions as a tumour suppressor gene, in accordance with Knudson's hypothesis (Green *et al*., 1994; Yeung *et al*., 1994). Previous reports have shown that the size of TSC2-associated deletions on chromosome 16p13.3 are fairly large, and deletion of the TSC2 gene in tuberous sclerosis patients is frequently associated with deletion of the OCTS3 gene (The European Chromosome 16 TSC Consortium, 1993; Green *et al*., 1994). The OCTS3 gene has been designated as the NTHL1 gene here, because the gene product is a homologue of endonuclease III (nth (*E. coli* endonuclease III)-like 1). Considering the frequent occurrence of NTHL1 gene deletion in TSC patients, structural and/or functional changes in the NTHL1 gene may be involved in some in a wide variety of the signs, symptoms, and complications of tuberous sclerosis. Some mutation in the NTHL1 gene may possibly become a mutator phenotype. The nature of the relationship between gene regulation in the TSC2 and NTHL1 genes may also be important for understanding the role each gene plays in the phenotypic expression of tuberous sclerosis, considering the fact that the TSC2 gene lies immediately adjacent to the NTHL1 gene in a head-to-head orientation (Aspinwall *et al*., 1997; and our unpublished result). Our present results indicate that the orientation of the mouse Tsc2 and Nthl1 genes in the genome is the same as that of the corresponding human genes. We have determined the complete sequence of the mouse Nthl1 gene and its 5′- and 3′-flanking regions. These data will be valuable for genetic approaches to tuberous sclerosis as well as for analysis of the *in vivo* functions of the Nthl1 (or NTHL1) gene.

Materials and Methods

Enzymes and chemicals

E. coli formamidopyrimidine DNA glycosylase (Fapy DNA glycosylase) and endonuclease III were gifts from Drs Z. Hatahet and S. S. Wallace of the University of Vermont. Uracil DNA glycosylase of *E. coli* was pur-

chased from USB, and polynucleotide kinase and restriction enzymes from Toyobo Biochem., Japan. Recombinant human class II AP endonuclease (APEX nuclease, hAPEX) was purified as described previously (Seki *et al.*, 1991, 1992). Osmium tetroxide was purchased from Nacalai Tesque, Japan, [γ-^{32}P]ATP (7000 Ci/mmol) from ICN and [α-^{32}P]dCTP (3000 Ci/mmol) and [methyl-^{3}H]TTP (49 Ci/mmol) from Amersham. A plasmid containing cDNA for the mouse S26 ribosomal protein was kindly provided by Professor J. Fujita of Kyoto University, Kyoto, Japan (Kaneko *et al.*, 1995). A *nth*::Cm^{r} *nei*::Km^{r} double mutant of *E. coli* was prepared as described previously (Saito *et al.*, 1997). The other materials used were obtained or prepared as described previously (Sarker *et al.*, 1995).

Cloning of the mNthl1 cDNA

Oligonucleotide primers (rF_0-sense strand, 5′TGCACAGAATAGCCAACAGA-3′; rR_0-antisense strand, 5′-AGACCCTGGGCAGCAGGACA-3′) for PCR were prepared based on the rat 3′-EST sequence (accession number: H33255). Using a mouse T-cell cDNA library in Uni-Zap XR vector (Stratagene) and the PCR primers, 211 bp cDNA fragments of the size expected from the EST sequence were amplified with hotstart under the following PCR conditions: 1 cycle of 95°C for five minutes and 40 cycles of 94°C for one minute, 60°C for one minute, 72°C for two minutes, followed by a final cycle of seven minutes at 72°C. The PCR product of the expected size (211 bp) was subcloned into a pGEM-T vector (Promega Biotec, WI), and sequenced for conformation. The 211 bp cDNA fragments were amplified and labelled with [α-^{32}P]dCTP using a random primer DNA labelling system (Amersham). The mNthl1 cDNA was obtained by recloning using the same DNA library and the ^{32}P-labelled probe. The inserts in the Lambda ZAP-II vectors in the isolated bacteriophage clone were excised with helper phage infection and recircularized to generate a subclone in the pBluescript SK(−) phagemid vector (Short *et al.*, 1988). The nucleotide sequence of the insert of the subclone (designated as clone A) was determined by the dideoxy chain termination method (Sanger *et al.*, 1977) using either M13 universal primers or oligonucleotide primers synthesized according to the sequences determined using Ampli*Taq* FS Ready Mix (Perkin-Elmer) by an Applied Biosystem Model 373A DNA sequencer. The insert of clone A was an incomplete cDNA containing a 717 bp sequence which corresponds to the cDNA region from position 309 to the 3′ end of the full size mNthl1/mNth1 cDNA deposited in Data Bases (accession number: AB006812).

5′ RACE and construction of the full size mNthl1 cDNA

The transcription initiation site and the 5′ region of the mNthl1 cDNA were determined by the 5′ RACE (rapid amplification of the 5′ ends) technique (Hirzmann *et al.*, 1993). Briefly, poly(A)$^{+}$ RNA isolated from SR-C3H/He (mouse ascites sarcoma) cells was heated at 65°C for three minutes and cooled on ice. The reverse transcription reaction was performed at 37°C for two hours in a 10 µl reaction mixture containing 600 ng of the 65°C-treated RNA, 50 mM Tris-HCl (pH 8.3), 75 mM KCl. 3 mM $MgCl_2$, 5 mM DTT, 1 mM of each dNTP, 0.5 unit/µl RNasin (Promega Biotec), 10 unit/µl Moloney murine leukemia virus reverse transcriptase and the 1 µM oligonucleotide primer (R_0: 5′-AGATCCTGGGCAGGGCA-3′, complementary to the nucleotide sequence from position 896 to position 915 in the mNthl1/mNth1 cDNA-accession number AB006812). Excess primer molecules and substrate were removed by washing the product using a Micron-30 membrane (Amicon Corp., MA). For dG tailing, an aliquot of the reaction product (equivalent to 100 ng of mRNA) was incubated at 37°C for 20 minutes in a 20 µl reaction mixture containing 100 mM cacodylate buffer (pH 6.8), 1 mM $CoCl_2$, 0.1 mM DTT, 100 µg/ml BSA, 0.2 mM dGTP and 10 units terminal deoxynucleotidyl transferase (Promega Biotec). The reaction was stopped by heating to 75°C for ten minutes. An aliquot of the dG-tailed product was used for PCR amplification with an *Eco*RI-*Bam*HI-$(dC)_{12}$-oligonucleotide primer (5′TTCTAGAATTCGGATC$_{12}$-3), and the R_2 primer (5′-GTCTGGCTGGAGAGCATCAG-3′, complementary to the nucleotide sequence from position 386 to position 405 in the mNthl1 cDNA). The PCR product was cloned into a PGEM-T vector and sequenced. The full-length mNthl1 cDNA was constructed from the 717 bp cDNA fragment in the clone A and the 5′-RACE product. Briefly, the PGEM-T vector containing the 5′-RACE product was digested with *Eco*RI and *Nar*I, and the fragments were gel-purified. The recombinant plasmid (designated as pBlue/mNthl1) containing the full length cDNA was constructed by replacing the gel-purified 5′-RACE fragment with the *Eco*RI and *Nar*I fragment in the pBluescript SK(−) containing the 717 bp cDNA fragment.

Preparation of expression construct

The open reading frame of the pBlue/mNth1 was amplified by PCR using the following primers: F_1 (5′-CATGACGAATTCAACTCAGGGGTGCGGA-3′) and R_1 (5′-TACTAGGATCCAGAGGAAGCCCTATGGC-3′) both of which have unique sequences and *Eco*RI or *Bam*HI site (underlined). The amplified fragments were introduced into the *Eco*RI and *Bam*HI sites of a PEK318 expression vector (containing Trc and T7 promoter sites and six histidine codons close to the *Eco*RI site). The resulting construct designated as PEK/mNthl1 was introduced into the *nth*$^{-}$ (nth::Cm^{r}) and *nei*$^{-}$ (nei::Km^{r}) double mutant of *E. coli* for transformation. The plasmid, PEK/mNthl1, was isolated from the recombinant cells using the Qiagen plasmid isolation kit and the insert was sequenced to confirm the correctness of the construct.

Overexpression and purification of recombinant protein

The recombinant cells in 5 ml of the overnight culture were cultured at 37°C in 500 ml of LB broth supplemented with 200 µg ampicillin/ml, 50 µg chloramphenicol/ml, 30 µg kanamycin/ml, and 15 µg tetracycline/ml until the absorbance at 600 nm of the culture reached 0.6 (~five hours). Expression of the recombinant protein was induced by incubating the culture at 37°C for two hours with 0.5 mM isopropyl β-D-thiogalactoside (IPTG), the cells were collected by centrifugation at 4000 *g* for 20 minutes, and the pellet was stored at −70°C until use. Protein purification was conducted as described by Roldan-Arjona *et al.* (1996) with some modifications. Briefly, after thawing, the pellet was resuspended in 5 ml of a sonication buffer (50 mM Hepes-KOH (pH 8.0), 50 mM NaCl, 0.1 mM EDTA, 1 mM PMSF). The cells were disrupted by sonication

using a Branson Sonifier Model 250D, the sonicate was centrifuged at 10,000 *g* for 30 minutes, and the supernatant was transferred into a clean tube. It was then loaded on a phosphocellulose column previously equilibrated with the sonication buffer. The column was washed with five column volumes of the buffer. The adsorbed protein was eluted stepwise with the sonication buffer supplemented with 100, 200, 500, 700 and 1000 mM NaCl. An aliquot of each fraction was analysed by 12% SDS-polyacrylamide gel electrophoresis (SDS-PAGE). Fractions containing the mNthl1 protein, which was largely eluted between 200 mM and 500 mM NaCl, were pooled, and mixed with 1 ml of 50% slurry of Ni^{2+}-NTA agarose (Qiagen) previously equilibrated with the sonication buffer. The mixture was gently stirred for one hour, and the resin was packed into a Poly-Prep Chromatography Column (Bio-Rad). The column was washed twice with 10 ml of the sonication buffer and once with 5 ml of washing buffer (50 mM Hepes-KOH (pH 8.0), 10% glycerol, 100 mM NaCl, 0.1 mM EDTA) containing 40 mM imidazole. Proteins were eluted stepwise with 3 ml each of the washing buffer supplemented with 100, 200, 400 and 500 mM imidazole, and collected in fraction tubes (1 ml each). An aliquot of each fraction was analysed by 12% SDS-PAGE. Fractions (eluted between 200 and 400 mM imidazole) containing the purified, recombinant mNthl1 protein were pooled, and dialysed against a dialysis buffer (50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 50% glycerol, 30 mM NaCl) overnight. Then, the dialysed sample was distributed into Eppendorf tubes and stored at −70°C until use. An apparent homogeneous recombinant mNthl1 preparation was also obtained by direct affinity purification using the Ni^{2+}-NTA agarose column without the step of the phosphocellulose column chromatography, although the recovery was low. All purification procedures on the recombinant protein were carried out at 4°C or on ice.

Preparation of DNA and oligonucleotide substrates for enzymatic analysis

M13 double-stranded DNA containing scattered ^{3}H-labelled thymine glycol residues ([^{3}H]Tg-DNA) was prepared, essentially as described previously (Ide *et al.*, 1991). Briefly, M13 mp18 single-stranded DNA (10 μg, ca 4 pmol), primed with synthetic 20-mer (5′-CCATTCGCCATTCAGGCTGA-3′, 20 pmol), was incubated at 25°C for one hour with the Klenow fragment (132 units) of *E. coli* DNA polymerase I in a mixture (final volume, 100 μl) containing 66 mM Tris-HCl (pH 7.5), 1.5 mM 2-mercaptoethanol, 6.6 mM $MgCl_2$, and 50 μM each of dATP, dGTP, dCTP and [*methyl*-^{3}H]dTTP at 3600 cpm/pmol) to prepare [^{3}H]dTMP-labelled, double-stranded M13 DNA ([^{3}H]T-DNA). [^{3}H]T-DNA was extracted with phenol from the reaction mixture, precipitated with ethanol, and dissolved in 100 μl of distilled water. DNA was gel-filtered with a Sephadex G-25 column (3 ml) to remove the remaining substrate. [^{3}H]T-DNA thus obtained was treated with 0.1% osmium tetroxide at 70°C for 30 minutes, and [^{3}H]Tg-DNA was obtained from the reaction mixture after removing the osmium tetroxide by gel filtration as described above.

The following oligonucleotides were prepared and purified essentially as described previously (Ide *et al.*, 1991; Kao *et al.*, 1995); 5′-ACAGACGCCATCAACCAGG-3′ (abbreviated as 19T), 5′-ACAGACGCCAXCAACCAGG-3′ (X = thymine glycol; abbreviated as 19TG), 5′-ACAGACGCCAYCAACCAGG-3′ (Y = urea; abbreviated as 19UREA), 3′-TGTCTGCGGTAGTTGGTCC-5′ (abbreviated as 19COM), 5′CATCGATAGCATCCGUCACAGGCAG-3′ (U, uracil; abbreviated as 25U), and 3′-GTAGCTATCGTAGGCAGTGTCCGTC-5′ (abbreviated as 25COM). Briefly, the 19TG was prepared by specifically oxidizing the thymine of the 19T into TG with $KMnO_4$, and purified by HPLC (Kao *et al.*, 1993). The 19UREA was prepared from the 19TG by alkaline treatment at pH 12 (Ide *et al.*, 1991). The oligonucleotide containing an AP site (5′-CCTGCCCTGOGCAGCTGTGGG-3′ where O is the position of AP site; abbreviated as 21AP) and the complementary oligonucleotide (5′-CCCACAGCTGCACAGGGCAGG-3′; abbreviated as 21COM) were purchased from Trevigen Inc. Oligonucleotides with an abnormal base or an abasic site were labelled at the 5′ end with [γ-^{32}P]ATP and polynucleotide kinase, and purified as described (Ide *et al.*, 1993).

Assay of DNA glycosylase/AP lyase activity

To measure thymine glycol DNA glycosylase activity of mNthl1, [^{3}H]Tg-DNA and [^{3}H]T-DNA (both 1.5×10^5 cpm, ca 0.1 μg) were treated with mNthl1 or endoIII. The material released by DNA glycosylase activity was separated from DNA using an anion exchange column, Bio-Rad AG 1-X2 (Cl^- form). Released ^{3}H radioactivity was measured by a liquid scintillation counter. To confirm the release of thymine glycol, the released material was concentrated by a vacuum centrifugal evaporator, and then analysed by reverse phase HPLC using a Wakosil DNA C18 column (4.6 mm × 1590 mm, Wako, Japan), using authentic thymine, hydroxymethyluracil and thymine glycol (Tg) as a standard. The column was eluted with distilled water. Fractions were collected every 20 seconds and assayed for radioactivity. DNA glycosylase/AP lyase activity of mNthl1 was assayed using ^{32}P-labelled oligonucleotides. Each procedure is described in the corresponding Figure legend.

Southern blot analysis of mNthl1 genomic DNA

High molecular weight genomic DNA was isolated from mouse liver as described (Sambrook *et al.*, 1989). Five micrograms of the DNA digested with restriction enzymes were electrophoresed on a 0.7% agarose gel at 7 V for 16 hours, and transferred onto a positively charged nylon membrane (Hybond N^+; Boehringer Mannheim Biochem.) as described (Sambrook *et al.*, 1989). The mNthl1 cDNA fragment was excised from the plasmid, clone A, by treatment with *Eco*RI and *Xho*I, gel-purified, and labelled with [α-^{32}P]dCTP (3000 Ci/mmol) using the random priming method. The blotted membrane was hybridized with the ^{32}P-labelled probe. Hybridization was carried out at 68°C for 12 hours in 6 × SSC (1 × SSC is 150 mM NaCl, 15 mM sodium citrate, pH 7.0), 5 × Denhardt's solution, 0.5% SDS, denatured salmon testis DNA at 100 μg/ml and the labelled cDNA probe at 3×10^5 cpm/ml. After washing at room temperature for ten minutes in 2 × SSC and 0.1% SDS and then twice at 68°C for one hour in 1 × SSC and 0.1% SDS, the membrane was air-dried, and exposed on Kodak X-ray film for autoradiography.

Northern blot analysis

The ''Mouse Multiple Tissue Northern Blot'' (MTN) purchased from Clontech Laboratories contained polyadenylated RNA (2 μg per lane) from mouse heart, brain, spleen, lung, liver, skeletal muscle, kidney and testis.

The blot was hybridized with the same probe as used in the Southern blot analysis. Hybridization and washing were conducted according to the manufacturer's instructions. Briefly, the membrane was incubated at 68°C for one hour in the ExpressHyb™ solution (supplied by Clontech) for hybridization with the labelled mNthl1 cDNA probe (8×10^5 cpm/ml). The membrane was then washed at room temperature for 40 minutes in 2 × SSC, 0.05% SDS, washed again at 50°C for 40 minutes in 0.1 × SSC, 0.1% SDS, and processed for autoradiography.

Cloning of the mNthl1 gene

The genomic library constructed using BALB/c adult mouse leukocyte DNA and cloning vector EMBL-3 SP6/T7 (Clontech Laboratories, CA) was screened using the same DNA probe as that used for the Southern blot analysis. Among the 16 positive clones obtained, inserts of two clones (clones 3 and 5 designated as mNth1G-3 and mNth1G-5, respectively) were subcloned in plasmid pBluescript KS(−), and sequenced. The nucleotide sequence data were deposited in the GSDB, DDBJ, EMBL and NCBI nucleotide sequence data bases under the accession no. AB009371.

Chromosome mapping by fluorescence *in situ* hybridization (FISH)

Metaphase chromosome preparations were made from fibroblasts derived from primary cultures of a DBA strain mouse. *In situ* hybridization was performed using a biotin-labelled, 8 kb genomic fragment which contained the entire region of the mNthl1 gene. The biotin label was detected with an anti-biotin Fab′-alkaline phosphatase conjugate, and fluorescent signals were amplified by 3-hydroxy-*N*-2′-biphenyl-2-naphthalebcarboxamide phosphate ester (HNPP) an Fast red TR (Kondoh *et al.*, 1995). Chromosomes were counterstained with Hoechst 33258 and quinacrine mustard for precise localization of the signals by Q-banding (Yoshida *et al.*, 1975).

Acknowledgements

We thank Dr S. Mitra of The University of Texas Medical Branch for his valuable insights and encouragement, Drs Z. Hatahet and S. S. Wallace of The University of Vermont for providing DNA repair enzymes of *E. coli*, Dr J. Fujita of The Kyoto University for providing a plasmid containing cDNA of the mouse S26 ribosomal protein, and Mr T. Nakamura and Ms K. Tabuchi for their technical assistance. This work was supported in part by a Grant-in-Aid for Scientific Research from The Ministry of Education, Science, Sports and Culture of Japan, and in part by a Grant-in-Aid for Scientific Research on Priority Areas from the Ministry of Education, Science, Sports and Culture of Japan (Repair, Recombination and Mutagenesis, 08280101).

References

Akiyama, K., Nagao, K., Oshida, T., Tsutsui, K., Yoshida, M. C. & Seki, S. (1995). Cloning sequence analysis, and chromosomal assignment of the mouse Apex gene. *Genomics,* **26**, 63–69.

Asahara, H., Wistort, P. M., Bank, J. F., Bakerian, R. H. & Cunningham, R. P. (1989). Purification and characterization of *Escherichia coli* endonuclease III from the cloned nth gene. *Biochemistry,* **28**, 4444–4449.

Aspinwall, R., Rothwell, D. G., Roldan-Arjona, T., Anselmino, C., Ward, C. J., Cheadle, J. P., Sampson, J. R., Lindahl, T., Harris, P. C. & Hickson, I. D. (1997). Cloning and characterization off a functional human homolog of a functional human homolog of *Escherichia coli* endonuclease III. *Proc. Natl Acad. Sci. USA,* **94**, 109–114.

Augeri, L., Lee, Y. M., Barton, A. B. & Doestsch, P. W. (1997). Purification, characterization, gene cloning, an expression of *Saccharomyces cerevisiae* redoxyendonuclease, a homologue of *Escherichia coli* endonuclease III. *Biochemistry,* **36**, 721–729.

Bacchetti, S. & Benne, R. (1975). Purification and characterization of an endonuclease from calf thymus acting on irradiated DNA. *Biochim. Biophys. Acta,* **390**, 285–297.

Bailly, V. & Verly, W. G. (1987). *Escherichia coli* endonuclease III is not an endonuclease by a beta-elimination catalyst. *Biochem. J.* **242**, 565–572.

Bailly, V. & Verly, W. G. (1988). Importance of thiols in the repair mechanisms of DNA containing AP (apurinic or apyrimidinic) sites. *Nucl. Acids Res.* **16**, 9489–9496.

Bailly, V., Sente, B. & Verly, W. G. (1989a). Bacteriophage-T4 and Micrococcus luteus UV endonucleases are not endonucleases but beta-elimination and sometimes beta delta-elimination catalysts. *Biochem. J.* **259**, 751–759.

Bailly, V., Verly, W. G., O'Connor, T. & Laval, J. (1989b). Mechanism of DNA strand nicking at apurinic/apyrimidinic sites by *Escherichia coli* [formamidopyrimidine]DNA glycosylase. *Biochem. J.* **262**, 581–589.

Breathnach, R. & Chambon, P. (1981). Organization of expression of eukaryotic split genes coding for proteins. *Annu. Rev. Biochem.* **50**, 349–383.

Bruand, C., Sorokin, A., Serror, P. & Ehrlich, S. D. (1995). The *Bacillus subtilis dna*I gene is part of the dnaB operon. *Microbiology,* **141**, 1199–1200.

Cathcart, R., Schwiers, E., Saul, R. L. & Ames, B. N. (1984). Thymine glycol and thymidine glycol in human and rat urine: a possible assay for oxidative DNA damage. *Proc. Natl Acad. Sci. USA,* **81**, 5633–5637.

Demple, B. & Harrison, L. (1994). Repair of oxidative damage to DNA: enzymology and biology. *Annu. Rev. Biochem.* **63**, 915–948.

Doestsch, P. W., Helland, D. E. & Haseltine, W. A. (1986). Mechanism of action of mammalian DNA repair endonuclease. *Biochemistry,* **25**, 2212–2220.

Eide, L., Bjoras, M., Pirovano, M., Alseth, I., Berdal, K. G. & Seeberg, E. (1996). Base excision of oxidative purine and pyrimidine DNA damage in *Saccharomyces cerevisiae* by a DNA glycosylase with sequence similarity to endonuclease III from *Escherichia coli*. *Proc. Natl Acad. Sci, USA,* **93**, 10735–10740.

European chromosome 16 Tuberous Sclerosis consortium (1993). Identification and characterization of the Tuberous Sclerosis gene on chromosome 16. *Cell,* **75**, 1305–1315.

Friedberg, E. C., Walker, G. C. & Siede, W. (1995). *DNA Repair and Mutagenesis,* ASM Press, Washington, DC.

Gavel, Y. & Heijne, G. V. (1990). Cleavage-site motifs in mitochondrial targeting peptides. *Protein Eng.* **4**, 33–37.

Green, A. J., Smith, M. & Yates, J. R. (1994). Loss of heterozygosity on chromosome 16p13.3 in hamartomas from tuberous sclerosis patients. *Nature Genet.* **6**, 193–196.

Hilbert, T. P., Boorstein, R. J., King, H. C., Bolton, P. H., Xing, D., Cunningham, R. P. & Teebor, G. W. (1996). Purification of a mammalian homologue of *Escherichia coli* endonuclease III: Identification of a bovine pyrimidine hydrate-thymine glycol DNA-glycosylase/AP lyase by irreversible cross linking to a thymine glycol-containing oligodeoxynucleotide. *Biochemistry*, **35**, 2505–2511.

Hilbert, T. P., Chaung, W., Boorstein, R. J., Cunningham, R. P. & Teebor, G. W. (1997). Cloning and expression of the cDNA encoding the human homologue of the DNA repair enzyme, *Escherichia coli* endonuclease III. *J. Biol. Chem.* **272**, 6733–6740.

Hirzmann, J., Luo, D., Hahnen, J. & Hobom, G. (1993). Determination of messenger RNA 5′–ends by reverse transcription of the cap structure. *Nucl. Acids Res.* **21**, 3597–3598.

Ide, H., Petrullo, L. A., Hatahet, Z. & Wallace, S. S. (1991). Processing of DNA base damage by DNA polymerases. *J. Biol. Chem.* **266**, 1469–1477.

Ide, H., Okagami, M., Murayama, H., Kimura, Y. & Makino, K. (1993). Synthesis and characterization of oligonucleotides containing the alpha-anomer of deoxyadenosine to study its influence on DNA replication. *Biochem. Mol. Biol. Int.* **31**, 485–491.

Ikeda, S., Biswas, T., Roy, R., Izumi, T., Boldogh, I., Kurosky, A., Sarker, A. H., Seki, S. & Mitra, S. (1997). Cloning of human endonuclease III cDNA and characterization of the enzyme. Abstracts of NATO Advanced Study Institute, Tekirova-Antalya, Turkey, Oct. 14–24, DNA Damage and Repair, P. 121.

Kaneko, Y., Nonoguchi, K., Fukuyama, H., Takano, S., Higashitsuji, H., Nishiyama, H., Takenawa, J., Nakayama, H. & Fujita, J. (1995). Presence of alternative 5′ untranslated sequences and identification of cells expressing ctk transcripts in the brain and testis. *Oncogene,* **10**, 945–952.

Kao, J. Y., Goljer, I., Phan, T. A. & Bolton, P. H. (1993). Characterization of the effects of the thymine glycol residue on the structure, dynamics, and stability of duplex DNA by NMR. *J. Biol. Chem.* **268**, 17787–17793.

Kim, J. & Linn, S. (1989). Purification and characterization of UV endonucleases I and II from murine plasmacytoma cells. *J. Biol. Chem.* **264**, 2739–2745.

Kim, K. K., Pajak, L., Wang, H. & Field, L. J. (1995). Cloning, developmental expression, and evidence for alternative splicing of the murine tuberous sclerosis (TSC2) gene product. *Cell. Mol. Biol. Res.* **41**, 515–526.

Kobayashi, T., Nishizawa, M., Hirayama, Y., Kobayashi, E. & Hino, O. (1995). cDNA structure, alternative splicing and exon-intron organization of the predisposition tuberous sclerosis (Tsc2) gene of the Eker rat model. *Nucl. Acids Res.* **23**, 2608–2613.

Kobayashi, T., Urakami, S., Cheadle, P. J., Aspinwall, R., Harris, P., Sampson, J. R. & Hino, O. (1997). Identification of a leader exon and a core promoter for the rat tuberous sclerosis 2 (Tsc2) gene and structural comparison with the human homolog. *Mammal. Genome,* **8**, 554–558.

Kondoh, Y., Ono, T., Kagiyama, N., Fujita, S., Momiyama, M., Hori, S. H. & Yoshida, M. C. (1995). Simultaneous visualization of Q-bands and FISH signals using a novel fluorochrome. *Cytogenet. Cell Genet.* **71**, 96–98.

Krokan, H. E., Standal, R. & Slupphaug, G. (1997). DNA glycosylases in the base excision repair of DNA. *Biochem. J.* **325**, 1–16.

Maheshwar, M. M., Sandford, R., Nellist, M., Cheadle, J. P., Sgotto, B., Vaudin, M. & Sampson, J. R. (1996). Comparative analysis and genomic structure of the tuberous sclerosis 2 (TSC2) gene in human and pufferfish. *Hum. Mol. Genet.* **5**, 131–137.

Mazumder, A., Gerlt, J. A., Absalon, M. J., Stubbe, J., Cunningham, R. P., Withka, J. & Bolton, P. H. (1991). Stereochemical studies of the beta-elimination reactions at aldehydic abasic sites in DNA: endonuclease III from *Escherichia coli*, sodium hydroxide, and Lys-Trp-Lys. *Biochemistry,* **30**, 1119–1126.

Nakai, K. & Kanehisa, M. (1992). A knowledge base for predicting protein localization sites in eukaryotic cells. *Genomics,* **14**, 897–911.

Nes, I. F. & Nissen-Meyer, J. (1978). Endonuclease activities from a permanently established mouse cell line that act upon damaged DNA damaged by ultraviolet light, acid and osmium tetroxide. *Biochim. Biophys. Acta,* **520**, 111–121.

Olsson, P. G., Sutherland, H. F., Nowicka, U., Korn, B., Poustka, A. & Frischauf, A. M. (1995). The mouse homologue of the tuberin gene (TSC2) maps to a conserved synteny group between mouse chromosome 17 and human 16p13.3. *Genomics,* **25**, 339–340.

Olsson, P. G., Lohning, C., Horsley, S., Kearney, L., Harris, P. C. & Frischauf, A. (1996). The mouse homologue of the polycystic kidney disease gene (pkd1) is a single-copy gene. *Gemonics,* **34**, 233–235.

Piersen, C. E., Prince, M. A., Augustine, M. L., Dodson, M. L. & Lloyd, R. S. (1995). Purification and cloning of *Micrococcus luteus* ultraviolet endonuclease, an N-glycosylase/abasic lyase that proceeds via an imino enzyme-DNA intermediate. *J. Biol. Chem.* **270**, 23475–23484.

Pilz, A., Glenister, P., Povey, S., Lyon, M. & Abbott, C. (1995). The mouse homologue of the human tuberous sclerosis 2 (TSC2) gene maps to chromosome 17, but does not fall within the tw18 or th20 deletions. *Genomics,* **26**, 420–421.

Proudfoot, N. J. & Brownlee, G. G. (1976). 3′ Non-coding region sequences in eukaryotic messenger RNA. *Nature,* **263**, 211–214.

Purmal, A. A., Rabow, L. E., Lampman, G. W., Cunningham, R. P. & Kow, Y. W. (1996). A common mechanism of action for the N-glycosylase activity of DNA N-glycosylase/AP lyases from *E. coli* and T4. *Mutat. Res.* **364**, 193–207.

Roldan-Arjona, T., Anselmino, C. & Lindahl, T. (1996). Molecular cloning and functional analysis of a *Schizosaccharomyces pombe* homologue of *Escherichia coli* endonuclease III. *Nucl. Acids Res.* **24**, 3307–3312.

Saito, Y., Uraki, F., Nakajima, S., Asaeda, A., Ono, K., Kubo, K. & Yamamoto, K. (1997). Characterization of endonuclease III (nth) and endonuclease VIII (nei) mutants of *Escherichia coli* K-12. *J. Bacteriol.* **179**, 3783–3785.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*, 2nd edit., Cold Spring Harbor Laboratory Press, Plainview, NY.

Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl Acad. Sci. USA,* **74**, 5463–5467.

Sarker, A. H., Watanabe, S., Akiyama, K., Nakagawa, Y., Wakabayashi, H., Tan, Y. & Seki, S. (1995). Purification and characterization of an AP endonuclease/DNA 3′ repair diesterase from mouse ascites sarcoma cells. *Biochim. Biophys. Acta,* **1245**, 299–304.

Seki, S., Akiyama, K., Watanabe, S., Hatsushika, M., Ikeda, S. & Tsutsui, K. (1991). cDNA and deduced amino acid sequence of a mouse DNA repair enzyme (APEX nuclease) with significant homology to *Escherichia coli* exonuclease III. *J. Biol. Chem.* **266**, 20797–20802.

Seki, S., Hatsushika, M., Watanabe, S., Akiyama, K., Nagao, K. & Tsutsui, K. (1992). cDNA cloning, sequencing, expression and possible domain structure of human APEX nuclease homologies to *Escherichia coli* exonuclease III. *Biochim. Biophys. Acta,* **1131**, 287–299.

Short, J. M., Fernandez, J. M., Sorge, J. A. & Huse, W. D. (1988). Lambda ZAP: A bacteriophage lambda expression vector with *in vivo* excision properties. *Nucl. Acids Res.* **16**, 7583–7600.

Thayer, M. M., Ahern, H., Xing, D., Cunningham, R. P. & Tainer, J. A. (1995). Novel DNA binding motifs in the DNA repair enzyme endonuclease III crystal structure. *EMBO J.* **14**, 4108–4120.

Wilson, R., Ainscough, R., Anderson, K., Baynes, C., Berks, M., Bonfield, J., Burton, J., Connell, M., Copsey, T. & Cooper, J., *et al*. (1994). 2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans. *Nature,* **368**, 32–38.

Yeung, R. S., Xiao, G.-H., Jin, F., Lee, W.-C., Testa, J. R. & Knudson, A. G. (1994). Predisposition to renal carcinoma in the Eker rat is determined by germline mutation of the tuberous sclerosis 2 (TSC2) gene. *Proc. Natl Acad. Sci. USA,* **91**, 11413–11416.

Yoshida, M. C., Ikeuchi, T. & Sasaki, M. (1975). Differential staining of parental chromosomes in interspecific cell hybrids with a combined quinacrine and 33258 Hoechst technique. *Proc. Jpn. Acad.* **51**, 184–187.

Zuo, S., Boorstein, R. J. & Teebor, G. W. (1995). Oxidative damage to 5-methylcytosine in DNA. *Nucl. Acids Res.* **23**, 3239–3243.

Edited by J. Miller

(Received 3 February 1998; received in revised form 3 June 1998; accepted 20 June 1998)