

## *Escherichia coli* Endonuclease VIII: Cloning, Sequencing, and Overexpression of the *nei* Structural Gene and Characterization of *nei* and *nei nth* Mutants

DONGYAN JIANG,<sup>†</sup> ZAFER HATAHET, JEFFREY O. BLAISDELL, ROBERT J. MELAMEDE,  
AND SUSAN S. WALLACE\*

Department of Microbiology and Molecular Genetics, The Markey Center for Molecular Genetics,  
The University of Vermont, Burlington, Vermont 05405

Received 13 February 1997/Accepted 25 March 1997

*Escherichia coli* possesses two DNA glycosylase/apurinic lyase activities with overlapping substrate specificities, endonuclease III and endonuclease VIII, that recognize and remove oxidized pyrimidines from DNA. Endonuclease III is encoded by the *nth* gene. Endonuclease VIII has now been purified to apparent homogeneity, and the gene, *nei*, has been cloned by using reverse genetics. The gene *nei* is located at 16 min on the *E. coli* chromosome and encodes a 263-amino-acid protein which shows significant homology in the N-terminal and C-terminal regions to five bacterial Fpg proteins. A *nei* partial deletion replacement mutant was constructed, and deletion of *nei* was confirmed by genomic PCR, activity analysis, and Western blot analysis. *nth nei* double mutants were hypersensitive to ionizing radiation and hydrogen peroxide but not as sensitive as mutants devoid of base excision repair (*xth nfo*). Single *nth* mutants exhibited wild-type sensitivity to X rays, while *nei* mutants were consistently slightly more sensitive than the wild type. Double mutants lacking both endonucleases III and VIII exhibited a strong spontaneous mutator phenotype (about 20-fold) as determined by a rifampin forward mutation assay. In contrast to *nth* mutants, which showed a weak mutator phenotype, *nei* single mutants behaved as the wild type.

Free radicals are produced in cells by ionizing radiation, a variety of chemical agents, and normal oxidative metabolism. The spectrum of free radical-induced damage to DNA is broad and includes a wide variety of modifications to the purine and pyrimidine bases, sites of base loss and single-strand breaks (for reviews, see references 11, 39, and 66). The 5,6 double bond of pyrimidines is particularly vulnerable to free radical attack (66), and for DNA thymine, the products include 5,6-dihydroxy-5,6-dihydrothymine (thymine glycol [Tg]), 5-hydroxy-5,6-dihydrothymine, 6-hydroxy-5,6-dihydrothymine, and 5,6-dihydrothymine (produced under anaerobic conditions). 5-Hydroxymethyluracil (24) and a number of ring contraction and fragmentation products such as 5-hydroxy-5-methylhydantoin, methyltartronyl urea, and urea are also formed. A major product of free radical attack on cytosine is 5,6-dihydroxy-5,6-dihydrocytosine (cytosine glycol) (67), which is unstable and dehydrates to form 5-hydroxycytosine (5-OHC) or deaminates to form uracil glycol (Ug), which can also dehydrate to form 5-hydroxyuracil (5-OHU). Contraction and fragmentation products can also be found after free radical attack on DNA cytosine. Of the above-mentioned structures that retain intact ring conformation and base-pairing capabilities, only Tg has been shown to be a potentially lethal lesion. Tg is a potent block to DNA synthesis in vitro (17, 28, 31, 56), using a variety of DNA polymerases, and when present in biologically active single-stranded phage DNA molecules is a lethal lesion with an inactivation efficiency of 1 (1, 29). Tg is also lethal in duplex

ϕX DNA, where it takes about 10 to 12 Tg lesions to kill (41). Other intact pyrimidine products such as Ug, dihydrothymine, 5-OHC, and 5-OHU do not block synthesis in vitro (32, 52, 53) and, where tested, in vivo (22, 23). However, the pyrimidine contraction and fragmentation products, of which only urea has been well studied (41, 42), are presumed to be potentially lethal, but these are less prevalent.

Many of the oxidized cytosines are premutagenic lesions. 5-OHC mispairs in vitro (53) and has been shown to be mutagenic in *Escherichia coli* (23). Ug and 5-OHU pair correctly with A (52, 53); however, since these products are derived from C, they are presumably potent premutagenic lesions. In contrast, dihydrothymine is not a mutagenic lesion (22) and Tg is rarely mutagenic (29) except when present in a sequence that is readily bypassed, and in this case it pairs with G (6). Urea is a premutagenic lesion in *E. coli* under SOS conditions which favor lesion bypass (42).

Oxidative damage produced by normal metabolism appears to be responsible for a substantial fraction of endogenous DNA damage, and cells have evolved an efficient and accurate repair mechanism, base excision repair, to remove these lesions from DNA. Base excision repair is highly conserved from bacteria to humans (for reviews, see references 20 and 68). In *E. coli*, there are three known DNA glycosylases that recognize oxidized purines and pyrimidines. FAPY-DNA glycosylase (Fpg) recognizes and removes formamidopyrimidine products of adenine and guanine (16) as well as 8-oxoguanine (45, 63) and 5-hydroxypyrimidines (27) from DNA. Pyrimidine damages are recognized and removed by endonuclease III (endo III) (21, 27, 54, 62; for reviews, see references 20 and 68) and endonuclease VIII (endo VIII) (43), which share a common range of substrates. In addition to their glycosylase activities, all three of these enzymes possess a lyase activity which cleaves the DNA backbone, leaving a blocked 3' terminus in the resulting nick (references 20 and 68 and references therein). The

\* Corresponding author. Mailing address: Department of Microbiology and Molecular Genetics, The Markey Center for Molecular Genetics, The University of Vermont, Stafford Hall, Burlington, VT 05405. Phone: (802) 656-2164. Fax: (802) 656-8749. E-mail: swallace@zoo.uvm.edu.

<sup>†</sup> Present address: Department of Molecular and Cellular Toxicology, Harvard School of Public Health, Boston, MA 02115.

blocked 3' terminus is removed by the phosphodiesterase activities of either endonuclease IV (endo IV) or exonuclease III (exo III), leaving a single-base gap that is filled in by DNA polymerase I and sealed by DNA ligase (references 20 and 68 and references therein).

The genes for endo III, *nth* (18), and Fpg, *fpg* (10), have been cloned and sequenced, and mutants have been constructed. The product of *fpg* is the same as that of *mutM* (45), originally isolated as a mutator for G→T transversions. *fpg* mutants were reported to be insensitive to any damaging agents that produce free radicals (9), although recent evidence suggests that *fpg* mutants may be sensitive to hydrogen peroxide (71). *nth* mutants exhibit a weak mutator phenotype (69) but are not sensitive to any DNA-damaging agents that produce the substrates for the enzyme (18). However, when duplex  $\phi$ X174 DNA containing Tg is transfected into wild-type and *nth* mutant hosts, the survival of the phage is decreased about 2.5-fold in the mutant cells (37, 41), indicating that the enzyme functions in vivo as predicted by its substrate specificity. It was not clear why *nth* mutants exhibited wild-type sensitivity to DNA-damaging agents while at the same time they were less effective at supporting replication of damaged viral transfecting DNA. It seemed likely that either the DNA-damaging agents (X rays and H<sub>2</sub>O<sub>2</sub>) may not have produced enough lethal lesions for expression of a phenotype or an additional enzyme was present in the cell that did not act on exogenous supercoiled DNA but functioned to protect cellular DNA. It was the lack of a cellular phenotype for *nth* mutants that originally led to the examination of cell extracts for an additional pyrimidine-specific enzyme and the identification and characterization of endo VIII (43).

In this paper, we report the purification of endo VIII to apparent homogeneity and the use of the purified protein in a reverse genetic approach to clone the gene. The amino acid sequence of the protein shares significant homology, especially at the amino- and carboxy-terminal ends, to bacterial Fpg proteins. Finally, we have constructed *nei* and *nth nei* mutants and show that mutants deficient in both endo III and endo VIII are hypersensitive to killing by ionizing radiation-induced free radicals and exhibit a spontaneous mutator phenotype as measured by a forward mutation assay.

#### MATERIALS AND METHODS

***E. coli* genomic DNA library and cell strains.** The *E. coli* Kohara library was provided by Yuji Kohara, Department of Molecular Biology, School of Science, Nagoya University. *E. coli* BW434 [F<sup>-</sup> *nth::Kan*  $\Delta$ (*xth-pncA90*) *leuB6*(amber SuII) *thr-1*  $\Delta$ (*gpt-proA*)2 *hisG4*(ochre) *argE3*(ochre) *lacY1* *galK2* *ara-14* *mit-1* *xyI-5* *thi-1* *tsx-33* *rpsL31* *supE44*(amber SuII) *rac*  $\lambda$ ^-], BW853(*recD1903::mini-Tn10*), BW35 (KL16) {Hfr *KL16* *PO-45*:[*lysA*(61)-*serA*(62)]/*thi* *relA* *spoT1*} and its mutant BW402 [like BW35 but *nth::Kan*] were provided by Bernard Weiss, Department of Pathology, University of Michigan. SR2263 or NR9102[F<sup>-</sup> *lacI<sup>q</sup>* *lacI204* (hotspotless) *lacZ<sup>L8</sup>* *pro<sup>+</sup>*  $\Delta$ (*lac-pro*) *thi* *ara*] and SR2617 [like SR2263 but *nfo::Kan*  $\Delta$ (*xth-pncA*)] were provided by Neil Sargentini, Department of Microbiology and Immunology, Kirkwood College of Osteopathic Medicine. The genotype of BL21(DE3) is F<sup>-</sup> *ompT* *hdsS<sub>B</sub>* (*r<sub>B</sub>*<sup>-</sup> *m<sub>B</sub>*<sup>-</sup>) *gal* *dcm* (DE3), and the genotype of XL1-Blue is *recA1* *endA1* *gvrA96* *thi-1* *hdsR17* *supE44* *relA1* *lac* [F<sup>-</sup> *proAB* *lacI<sup>q</sup>* *Z $\Delta$ M15* *Tn10* (Tet<sup>r</sup>)]. All cultures were grown at 37°C in Luria-Bertani (LB) liquid medium (Life Technologies) (with aeration by shaking at 250 to 260 rpm) or on LB-agar plates.

**Purification of endo VIII.** The purification method used was similar to that previously reported (43), with several modifications. All purification steps were done at 4°C. Cell paste was collected from an *E. coli xth nth* double-mutant strain, BW434, and resuspended in 1× volume of lysis buffer A (10 mM HEPES [pH 7.5], 1 mM EDTA, 1 M NaCl, 14 mM 2-mercaptoethanol, 0.5 mM phenylmethylsulfonyl fluoride, 1 mM benzimidazole, 1.85  $\mu$ M leupeptin, 1.46  $\mu$ M pepstatin). The cells were lysed with a Braun MSK homogenizer. Nucleic acids were precipitated with lysis buffer A plus 30% polyethylene glycol 8000 and removed by centrifugation. The crude extract thus obtained was subjected to a series of chromatographic steps which included S-Sepharose radial flow followed by Mono S and phenyl-Superose fast protein liquid chromatography (FPLC) (43).

The pooled peak fraction from phenyl-Superose was dialyzed against buffer B

(10 mM potassium phosphate [pH 7.4], 100 mM KCl, 14 mM 2-mercaptoethanol) for 4 h and then loaded onto a hydroxylapatite FPLC column. Proteins were eluted with a linear gradient of 10 to 500 mM potassium phosphate in buffer B at a flow rate of 0.5 ml/min and collected at 1 ml per tube. Affinity chromatography was used as the final step of purification. The pooled peak fraction from the hydroxylapatite column was diluted sevenfold into buffer B and loaded onto a reduced apurinic (redAP) site-containing DNA (redAP-DNA)-cellulose affinity column (see below) with an 8-ml bed volume. Proteins were eluted with a linear gradient of 0 to 1 M NaCl in buffer C (10 mM HEPES, 1 mM EDTA, 14 mM 2-mercaptoethanol) at a flow rate of about 0.6 ml/min, and 2.5-ml fractions were collected per tube. The matrix of the affinity column consisted of redAP-DNA-cellulose prepared as described by Alberts and Herrick (2).

**Protein assays.** During purification of endo VIII, the protein concentrations of the fractions were measured according to the instructions for the Bio-Rad protein assay kit, with bovine serum albumin as a standard. To visualize the extent of purification after each chromatographic step, the protein fractions containing endo VIII activity were analyzed by sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) and silver stained (50).

**Enzymatic activity assay of endo VIII.** The enzyme activity was assayed by using double-stranded BlueScript pSK DNA containing Tg (Tg-pSK) in the alkali fluorometric assay (38). Specifically, 100 ng of double-stranded Tg-pSK was incubated with 1  $\mu$ l of eluate at 37°C for 5 min and immediately transferred to ice to stop the reaction; 2 ml of buffer D (20 mM potassium phosphate [pH 11.8], 0.5 mM EDTA, 500  $\mu$ g of ethidium bromide per liter) was added. Fluorescence of the reaction mixture was measured with a Ratio-2 System filter fluorometer (Farrand Optical Inc., Valhalla, N.Y.) both before and after boiling the sample. The number of nicks per DNA molecule was calculated from the two readings as previously described (36, 38). One unit was defined as the amount of enzyme needed to nick 1 pmol of DNA in 1 min.

**Cloning and sequencing of the *nei* gene.** Amino acid sequencing analysis of the purified protein was performed by the Microchemistry Facility of Harvard University and the Protein Chemistry Laboratory of the University of Texas Medical Branch at Galveston. The N-terminal amino acid sequence, PEGPEIRRAADN LEAAIKGKPLT?V(S)FAFPQ(LKP)Y, and two internal amino acid sequences, [D]LNAAOLDALAHALLEIP(R) (int1) and GOVDENKHHGALFR (int2) (brackets indicate probable/reasonable and parentheses indicate probable/low), were obtained. Three oligodeoxyribonucleotides, whose degenerate sequences either code for the N-terminal amino acid sequence (excluding the uncertain sequence of the last 12 amino acids) or are complementary to the coding sequence for each of the two internal amino acid sequences, were synthesized and used as primers (designated primer-N, primer-int1, and primer-int2, respectively) for PCR. *E. coli* genomic DNA, prepared as previously described (4), was used as the PCR template. The PCRs were set up with 1  $\mu$ g of *E. coli* genomic DNA, 2  $\mu$ M each primer (primer-N/primer-int1 or primer-N/primer-int2), and 200  $\mu$ M each of four deoxynucleotides in 50  $\mu$ l of PCR buffer supplemented with 1.5 mM Mg<sup>2+</sup> and 2.5 U of *Taq* polymerase (Life Technologies) and run with denaturation for 5 min at 95°C, followed by 30 cycles of 94°C for 1 min, 55°C for 1 min, and 72°C for 2 min and extension for 10 min at 72°C. The reaction mixtures were then electrophoresed on a 2% agarose gel (NuSieve; FMC Bio-products) with 0.5  $\mu$ g of ethidium bromide/ml. The band containing the PCR product of ~650 bp from the reaction with primer-N/primer-int1 was cut from the gel; the DNA was recovered by electroelution and labeled by <sup>32</sup>P with a nick translation kit (U.S. Biochemical [USB] as instructed by the supplier). The Kohara library was screened by in situ hybridization as described by Sambrook et al. (58), using the labeled DNA probe. DNA from the hybridizing clone was prepared from lambda phage particles (58). By using the restriction map provided with the Kohara library (34), the DNA clone was digested separately by restriction enzymes *Bam*HI, *Eco*RI, *Eco*RV, *Kpn*I, *Pst*I, *Bgl*I, and *Pvu*II (New England Biolabs and USB), and Southern blotting was performed (58) with the same labeled probe. The DNA fragment (~6 kb) from the *Pst*I digestion that contained the *nei* gene was purified from the agarose gel and ligated into the BlueScript phagemid pSK(-) DNA which was linearized by *Pst*I digestion, producing subclone pSKNEI.

The two strands of *nei* and its up- and downstream regions were then sequenced. The sequencing reactions were conducted with a Taq DyeDeoxy Terminator cycle sequencing kit (Applied Biosystems) as instructed by the supplier. Sequencing electrophoresis and data analysis were performed with Applied Biosystems equipment (University of Vermont Molecular Biology Facility). A sequence homology search was done through the National Center for Biotechnology Information BLAST E-mail server, and alignments were done by GES Matrix (55).

**Overexpression of the *nei* gene.** *Pfu* DNA polymerase (Stratagene) was used to PCR amplify *nei* from BW434 genomic DNA. The PCR primers contained linker sequences on their 5' ends to facilitate cloning of the gene into the pCYB1 expression vector by the ligase-free method (72). A T7 RNA polymerase promoter and a ribosome binding site were then cloned 5' to the start codon of *nei* to generate the vector E8-pCYB1-T7. pCYB1 is part of the novel IMPACT I One-Step protein purification system (New England Biolabs). *nei* was expressed as a fusion with an intein-chitin binding domain (56 kDa). E8-pCYB1-T7 was transformed into *E. coli* BL21(DE3), which was grown at 37°C in LB to mid-log phase and then induced with 0.4 mM isopropylthiogalactopyranoside for 3 h. Cells were harvested, resuspended in 0.05 volume (relative to culture volume) of

TES (50 mM Tris-HCl [pH 7.5], 0.1 mM EDTA, 0.5 M NaCl), and lysed by using glass beads and a Braun homogenizer as described earlier. Lysed cells were centrifuged for 10 min at  $4,000 \times g$  followed by 15 min at  $20,000 \times g$  at 4°C. The supernatant was then loaded onto a chitin agarose column equilibrated with TES. The column was washed with 10 bed volumes of TES followed by 3 bed volumes of intein cleavage buffer (50 mM Tris-HCl [pH 7.5], 0.1 mM EDTA, 100 mM NaCl, 30 mM dithiothreitol). The column was then closed off and left in cleavage buffer overnight at 4°C. Endo VIII was eluted from the column by using 50 mM Tris-HCl (pH 7.5)–0.1 mM EDTA–200 mM NaCl.

**Construction of *nei* single and *nei nth* double null mutants.** The construction of a *nei* single null mutant included two steps: making a *nei* deletion DNA construct carrying a chloramphenicol resistance gene (*Cm*), and introducing the construct into the *E. coli* genome. A DNA fragment containing the *nei* gene and its upstream (260 bp) and downstream (400 bp) flanking regions was made by PCR using pSKNEI as the template, purified from the agarose gel after electrophoresis, and then cloned into the pUC18 vector at the *HincII* site, producing clone pUCNEI. pUCNEI was restriction digested with *BglII* and *HpaI* (New England Biolabs), whose sites are at positions 167 and 564 in the *nei* gene, resulting in removal of 397 bp (~50%) of the coding region from the middle of *nei*. For the following blunt-end ligation, the 3'-recessive end produced by *BglII* was filled in by an elongation reaction with Klenow fragment (USB) to produce a blunt end. A DNA fragment containing the structural gene coding for chloramphenicol resistance, *Cm*, and its promoter was obtained by PCR. The linearized pUCNEI containing a deleted *nei* and the DNA fragment containing *Cm* were purified from agarose gels and ligated, producing pUCJ( $\Delta nei::Cm$ ). To introduce  $\Delta nei::Cm$  into the *E. coli* genome, linear DNA transformation was employed, using BW853(*recD::mini-Tn10*) as described previously (57). The DNA fragment of  $\Delta nei::Cm$  used for transformation was obtained by PCR amplification. The culture, transformed by linear  $\Delta nei::Cm$  DNA, was grown in liquid LB without chloramphenicol overnight at 37°C and then selected on LB-agar plates containing  $4 \mu\text{g}$  of chloramphenicol per ml for the transformants containing chromosomal  $\Delta nei::Cm$ . The final *nei* single-mutant and *nth nei* double-mutant strains, SW2-8( $\Delta nei::Cm$ ) and SW2-38(*nth::Kan*  $\Delta nei::Cm$ ), were made from BW35 (KL16) and the extant congenic *nth* mutant, BW402 (*nth::Kan*), respectively, by phage P1 transduction (46) of the chromosomal  $\Delta nei::Cm$  from the BW853 transformant described above.

**PCR analysis of mutant strains.** The *nei* null mutants were confirmed at both the DNA and protein levels. Genomic DNA preparations of presumptive *nei* mutants (*Cm*<sup>r</sup>) were analyzed with two different sets of primers as described in the legend to Fig. 5. The reactions were done by using Idaho Technologies buffers as follows: 1 hold cycle of 94°C for 10 s; 30 cycles of 94°C for 5 s, 60°C for 5 s, and 72°C for 1 min; and 1 hold cycle of 72°C for 3 min.

**Isolation of monoclonal Fabs specific for endo VIII.** Repertoire cloning and phage display (5, 30, 59) were performed by using vector pCOMB3 to isolate monoclonal Fabs that specifically bound to endo VIII. Mice were immunized with purified endo VIII, and antibodies were cloned as described by Bespalov et al. (7). RNA was reverse transcribed, and the cDNA was used for PCR of the repertoire of heavy and light antibody chains that were then randomly combined to produce a combinatorial library containing about  $10^6$  members.

The library was panned to enrich for clones that bound to endo VIII. The panning procedure, as described by Barbas and Lerner (5), is a modification of that first described by Parmley and Smith (49). Microtiter plate (Coster 3690) wells (two per panning) were coated with 50  $\mu\text{l}$  of phosphate-buffered saline (PBS; 0.14 M NaCl, 1 mM  $\text{KH}_2\text{PO}_4$ , 20 mM  $\text{Na}_2\text{HPO}_4$ , 3 mM KCl [pH 7.4]) containing 1  $\mu\text{g}$  of endo VIII overnight at 4°C. The endo VIII solution was shaken out, and the wells were blocked with 3% bovine serum albumin in PBS for 1 h. The wells were washed twice with PBS. Phage (50  $\mu\text{l}$  of  $\sim 10^{12}/\text{ml}$ ) were added to each well for binding enrichment (1 h at 37°C). The phage solution was then removed, and the wells were washed two times with deionized  $\text{H}_2\text{O}$ , once with PBS-Tween (0.5%), and an additional two times with  $\text{H}_2\text{O}$ . Then 50  $\mu\text{l}$  of elution buffer was added per well, and the plates were incubated at room temperature for 15 min. The wells were washed as described above. Five rounds of panning and growth were carried out as described previously (7).

Soluble Fabs were produced by eliminating the heavy chain-gene 3 fusion after three to four rounds of panning by *NheI/SpeI* digestion and self-ligating the library. Cells were grown, induced, and harvested as described previously (5). Soluble Fabs were released from the periplasm by three cycles of freeze-thawing.

**Activity and Western blot analysis of endo VIII in crude extracts.** Crude lysates were prepared by diluting overnight cultures of BW35 (KL16), BW402, SW2-8, and SW2-38 1 to 50 in Superbroth (Bio 101, Inc.) and grown to an optical density (OD) of 1.85. Fifty milliliters of each culture was centrifuged and resuspended in 5 ml of TE buffer (10 mM Tris, 1 mM EDTA). The cells were lysed with lysozyme (100  $\mu\text{g}/\text{ml}$ , final concentration) on ice and sonicated until the viscosity was reduced. The extracts were then centrifuged at 15,000 rpm for 45 min in an SS34 rotor (Sorvall).

For determination of enzyme activity, 5  $\mu\text{l}$  of extract was incubated with 10  $\mu\text{l}$  (350 ng) of  $\text{OsO}_4$ -treated DNA (43) containing Tg for 4 min at 37°C. Reactions were stopped by boiling for 1 min prior to loading onto a 0.6% agarose gel containing 0.5  $\mu\text{g}$  of ethidium bromide per ml.

For Western blot analysis, 25  $\mu\text{l}$  of crude extract was used. Proteins were separated by SDS-PAGE (40), electroeluted onto nitrocellulose membranes (65), and analyzed either by enzymatic chemiluminescence as instructed by the

TABLE 1. Purification scheme for endo VIII<sup>a</sup>

Step	Protein (mg)	Activity (U <sup>b</sup> )	Sp act (U/mg)	Yield (%)	Purification (fold)
Crude extract	4,629	735	0.16	100	1
S-Sepharose	269	871	3.23	119	20.2
Mono S	14	704	51	95.8	317
Phenyl-Superose	1.22	109	90	14.9	562
Hydroxylapatite	0.72	159	221	21.6	1,379
DNA-cellulose <sup>c</sup>	0.024	112	4,710	15.2	29,438

<sup>a</sup> For this purification, 400 g of cell paste of BW434 (*nth xth* mutant) was used.

<sup>b</sup> One unit was defined as the amount of enzyme needed to nick 1 pmol of Tg-pSK DNA at 37°C in 1 min.

<sup>c</sup> Calf thymus DNA was depurinated by incubation in 10 mM sodium citrate (pH 4) at 80°C for 1 h. The AP sites were reduced with 100 mM  $\text{NaBH}_4$  (pH 13). After neutralization, the redAP-DNA was coupled to cellulose.

manufacturer (Amersham Life Science) or by immunostaining with 4-chloro-1-naphthol horseradish peroxidase reagent. In both cases, the secondary antibody was horseradish peroxidase.

**X-ray sensitivity.** Overnight cultures (in stationary phase) of strains SW2-8 ( $\Delta nei::Cm$ ) and SW2-38 (*nth::Kan*  $\Delta nei::Cm$ ) as well as the parental strain for the mutants, BW35 (KL16), and the congenic *nth* mutant strain, BW402 (*nth::Kan*), were diluted 100-fold and grown to mid-log phase (OD at 600 nm of ~0.5). After being washed in PBS three times, cells were incubated in the same volume of PBS (pH 7.4) for 20 min at 37°C with aeration to complete any rounds of DNA replication. An aliquot of a 2-ml cell suspension in PBS was transferred into a 35-by-10-mm culture dish (Corning) and irradiated in air with 45, 90, 135, 180, 225, or 270 Gy (produced by a Philips XRG3000 X-ray generator at 50 kV peak and 2 mA for 0.5, 1, 1.5, 2, 2.5, and 3 min, respectively) with stirring. The irradiated and unirradiated cultures were diluted and plated onto LB-agar plates. The colonies were counted after incubation at 37°C for 10 to 12 h, and the surviving fraction was determined. Survival curves were obtained by plotting the log of the surviving fraction against irradiation time and represent the average of at least three experiments; standard deviations were calculated. SR2263 and its X-ray-sensitive mutant derivative, SR2617 [*nfo-1::Kan*  $\Delta(xth-pnc4)$ ], were used as controls.

**Frequency of spontaneous mutation to Rif<sup>r</sup>.** Overnight cultures of the SW2-8 ( $\Delta nei::Cm$ ) and SW2-38 (*nth::Kan*  $\Delta nei::Cm$ ) strains, as well as BW35 (KL16) and BW402 (*nth::Kan*), were diluted 50- to 100-fold in fresh LB medium and grown at 37°C to mid-log phase (OD at 600 nm of ~0.5). The cultures in mid-log phase were plated onto LB-agar plates with and without 100  $\mu\text{g}$  of rifampin per ml and incubated at 37°C for 15 h. The rifampin-resistant (Rif<sup>r</sup>) mutant fraction for each strain was determined.

**Nucleotide sequence accession number.** The DNA sequence of the *nei* gene has been deposited in the GenBank nucleotide sequence database with accession no. U38616.

## RESULTS

**Purification of endo VIII to apparent homogeneity.** Since *E. coli* Fpg stably binds to an oligodeoxynucleotide containing a redAP site (13), and since endo VIII, like Fpg, is a DNA *N*-glycosylase with an associated AP lyase activity that does not cleave DNA containing redAP sites (43), we examined endo VIII binding to redAP-DNA by using an affinity column made of redAP-DNA-cellulose. Endo VIII eluted at 0.9 M NaCl (data not shown), indicating that endo VIII bound tightly to the affinity column. Table 1 shows a representative purification scheme for endo VIII. Significantly, S-Sepharose and redAP-DNA-cellulose affinity chromatography resulted in about a ~20-fold purification with yields of about 100 and 70%, respectively. The final yield was 15%, and purification was ~30,000-fold. The enzyme activity in the peak fractions appeared to be higher than that of the loading sample after the S-Sepharose and hydroxylapatite columns. The underestimation of the loading activity possibly resulted from reversible inactivation of endo VIII due to inhibitory proteins or to the high concentration or the type of salt in the crude extracts as well as in the phenyl-Superose eluate. Silver staining of the electrophoresed peaks after each chromatographic step (Fig. 1) showed that an apparently homogeneous preparation of

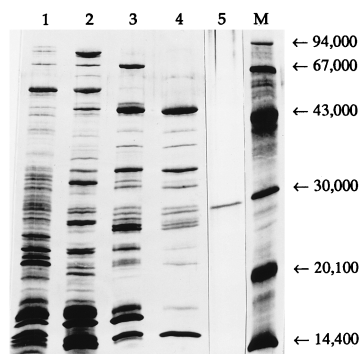


FIG. 1. Silver-stained SDS-polyacrylamide gel of the peak protein fractions after each chromatographic step of endo VIII purification. Twenty microliters of each pooled fraction of peak endo VIII activity was loaded onto an SDS-12% polyacrylamide gel and electrophoresed as described in Materials and Methods. Lane 1, after S-Sepharose; lane 2, after Mono S FPLC; lane 3, after phenyl-Superose FPLC; lane 4, after hydroxylapatite FPLC; lane 5, after redAP-DNA-cellulose affinity chromatography; lane M, protein molecular mass markers (positions indicated in daltons).

endo VIII was obtained and that endo VIII indeed had a molecular mass of 29 to 30 kDa as estimated previously (43). The apparent homogeneity of endo VIII was later supported by the data from the amino acid sequence analysis of the endo VIII preparation.

**Cloning and the sequencing of the *nei* gene.** Since apparently homogeneous endo VIII protein had been obtained, the cloning of the *nei* gene coding for endo VIII was accomplished by reverse genetics. Amino acid sequence analysis of the purified protein gave two internal and the N-terminal amino acid sequences, [D]LNAAQLDALAHALLEIP(R) (int1), GQVDE NKHHGALFR (int2), and PEGPEIRRAADNLEAAIKGK PLT?V(S)FAFPQ(LKP)Y, respectively. To obtain a probe for screening the *E. coli* Kohara library, we carried out two PCRs using degenerate primers whose sequences were determined from the amino acid sequences. A unique DNA product was obtained from each PCR (data not shown). Using a <sup>32</sup>P-labeled ~650-bp fragment as a probe, in situ hybridization showed that two adjacent clones of the Kohara library, 1H5 and 1H6 (representing H5 and H6 in plate 1 of the library stock provided), hybridized with the probe for *nei* (data not shown). After replicating 1H5 and 1H6 phage clones and then dot blotting, with the same probe, 10 plaques of phage DNA from each clone, all 10 plaques from 1H5 were positive whereas only 6 of 10 plaques from 1H6 were positive (data not shown), indicating that clone 1H5 contained the *nei* gene and the 1H6 clone was contaminated by 1H5. Given the fact that the position of each clone from the Kohara library was known, *nei* (1H<sub>5</sub>) was located at approximately 16.7 min on the physical map of the *E. coli* genome. Since this region of the genome has recently been completely sequenced (8, 46a, 48), we can precisely locate *nei* to 16 min.

For sequencing, a smaller subclone of the 40-kb Kohara library clone, pSKNEI, was constructed, and the insert (the fragment from *Pst*I restriction digestion) was shown to contain the entire gene by Southern blot analysis (data not shown). From this subclone of *nei*, a nucleotide sequence of 900 bp which included *nei* was obtained through overlapping sequencing of both strands of DNA by using the Taq DyeDeoxy Terminator sequencing method (Applied Biosystems). As shown in Fig. 2, in this region of 900 bp, there was only one reading frame that resulted in a peptide longer than 50 amino acid residues. It started at position 78 and ended at position 869.

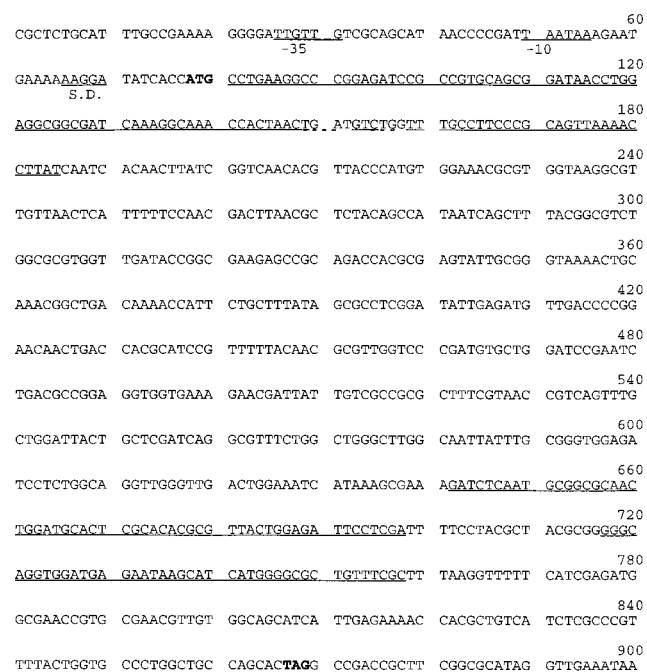


FIG. 2. Nucleotide sequence of DNA containing *nei*. The nucleotide sequence contains 840 bp with the coding region of *nei* from position 78 to 869 (plus strand shown). The start and stop codons are shown in boldface. The translated amino acid sequences encoded by the three underlined nucleotide sequences are identical to those obtained by the amino acid analysis of endo VIII except for the two amino acids encoded by the codons underlined by dashes.

The predicted sequence of 263 amino acids from the coding region included all three sequences obtained from the amino acid sequence analysis of the purified endo VIII protein (underlined in Fig. 2). The total number (263 versus 264) and the composition of amino acid residues of the predicted protein were very close to those obtained from amino acid composition analysis of the purified endo VIII protein. In addition, in the 5' upstream region of *nei*, there were two sequences, positions 26 to 31 and 50 to 55 corresponding to *E. coli* promoter consensus sequences, -35 and -10 hexamers, respectively (51), and a Shine-Dalgarno-like sequence at positions 66 to 70 (61). Taken together, the data show that the nucleotide sequence of 792 bp from position 178 to 969 is the *nei* structural gene coding for endo VIII and that this sequence is identical to the open reading frame in this region submitted by Mori (46a).

**Features of the *nei*-encoded protein, endo VIII.** The calculated molecular mass of the *nei*-encoded protein is 29.7 kDa without counting the first methionine (see below), which is consistent with that of endo VIII estimated from denaturing gel electrophoresis (SDS-PAGE) (Fig. 1). The calculated pI of the protein is 7.87. The predicted amino acid sequence of endo VIII exhibited strong homology to five bacterial Fpg proteins from *E. coli*, *Lactococcus lactis*, *Streptococcus mutans*, *Bacillus firmis*, and *Haemophilus influenzae* (Fig. 3). Specifically, 23 to 27% identity and 35 to 50% similarity of the 76-amino-acid N-terminal sequence and 26 to 28% identity and 43 to 46% similarity of the 113-amino-acid C-terminal sequence were seen between endo VIII and the five Fpg proteins. Like Fpg, endo VIII also contained a single zinc finger motif (Cys-X<sub>2</sub>-Cys-X<sub>16</sub>-Cys-X<sub>2</sub>-Cys) at its C terminus, which is the putative DNA binding domain for Fpg (47, 63). The structure (chelate of zinc) and function (binding to DNA) of the zinc finger motif of endo VIII have yet to be determined.

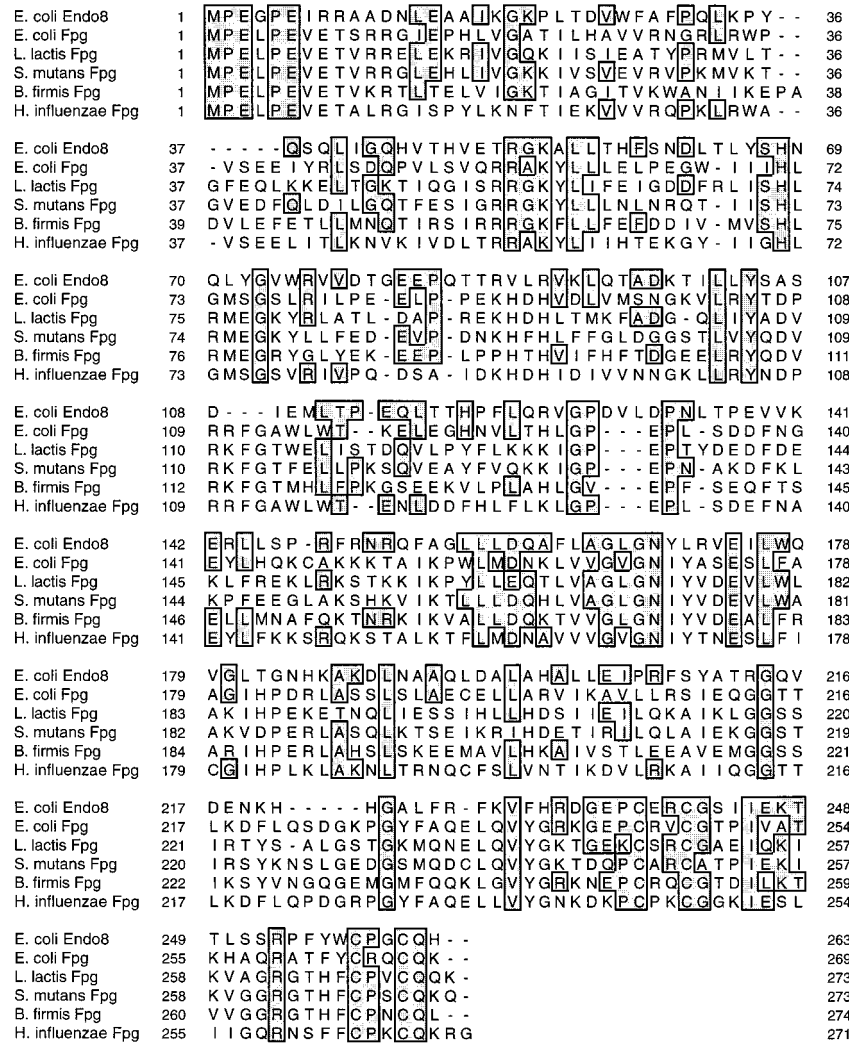


FIG. 3. Alignment of the amino acid sequences of *E. coli* endo VIII and five bacterial Fapy-DNA glycosylases. The amino acid residues of the Fpg proteins that are identical to those of endo VIII are boxed and shaded.

**Overexpression of the *nei* gene.** DNA of the *nei* gene was cloned into an overexpression vector, and a fusion protein with an intein and a chitin binding domain was overproduced. Figure 4 shows a silver-stained SDS gel (lanes 1 to 3) and a nitrocellulose filter immunochemically stained with Fab Nei-31 (lanes 4 to 6) on which were crude cell extracts containing the overexpression vector with the intein-endo VIII fusion protein (lanes 2 and 5) and overexpressed endo VIII protein eluted from a chitin column (lanes 3 and 6). In lanes 2 and 5, Fab Nei-31 staining of the fusion protein (~70 kDa) as well as apparent protein breakdown products can be seen, while a single band was observed (lanes 3 and 6) when purified endo VIII was cleaved from the chitin column. It should be noted that only one band was observed in crude *E. coli* extracts (see Fig. 6B), indicating that Fab Nei-31 does not cross-react with other *E. coli* proteins. Significant amounts of endo VIII activity (data not shown) were found in extracts from cells containing the plasmid construct. These data further confirm that the cloned *nei* gene is the structural gene coding for endo VIII.

**Construction and identification of *nei* single and *nei* *nth* double null mutants.** In vitro studies showed that endo VIII shared substrate specificity with endo III (33, 43), suggesting

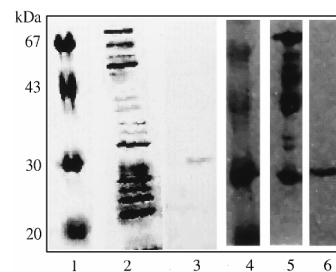


FIG. 4. Overproduction and single-step purification of endo VIII. Endo VIII was overproduced as a fusion protein with an intein and a chitin binding domain, using the IMPACT I system. Lanes 1 to 3 are silver-stained SDS samples of prestained marker proteins, crude cell extract containing the fusion protein construct, and protein eluted from the chitin column, respectively. Lanes 4 to 6 are the same samples in lanes 1 to 3 but electroeluted onto nitrocellulose paper and immunochemically stained with Fab Nei-31 by using 4-chloro-1-naphthol horseradish peroxidase reagent as described in Materials and Methods. For Fig. 4 to 6, the gels were digitized with a Kodak Digital Science DC40 camera. The images were saved as TIFF files, using the software provided with the camera. Images were then cropped and adjusted for brightness and contrast by using Adobe Photoshop and exported into Micromedia Freehand for labeling.

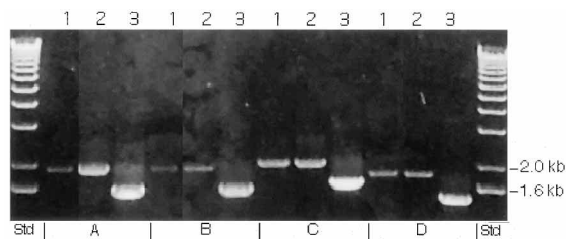


FIG. 5. Genomic PCR analysis of  $\Delta nei::Cm$  mutants. DNA from the *nei* locus of XLI1-Blue (wild type) (lane 3), SW2-8 ( $\Delta nei::Cm$ ) (lane 2), and SW2-38 (*nth::Kan*  $\Delta nei::Cm$ ) (lane 1) was amplified by PCR and electrophoresed on agarose gels as described in Materials and Methods. The primer designation numbers signify the 5' starting nucleotide and relate to the ATG start site (zero) of wild-type *nei*. A, -267 (5') and 1300 (3'); B, -382 (5') and 1204 (3'); C, -382 (5') and 1300 (3'); D, -267 (5') and 1204 (3').

that both enzymes contribute to the repair of oxidative pyrimidine damage in the cell. To determine the biological role of endo VIII and its relationship to endo III, mutants lacking endo VIII activity and lacking both endo III and endo VIII were constructed. By deleting an internal region of ~50% (397 bp) from the middle of the *nei* structural gene (792 bp), which includes the codon for the first cysteine of the zinc finger motif, and replacing it with a chloramphenicol resistance gene (*Cm*) and its promoter (775 bp), a construct of deleted *nei*,  $\Delta nei::Cm$  (which included the flanking ~260 bp upstream and ~400 bp downstream of *nei*), was made and presumed to be incapable of encoding a functional endo VIII. A chromosomal mutant was made by linear transformation, and various mutant constructs were made by P1 transduction. The PCR products obtained from wild-type, SW2-8 (*nei*), and SW2-38 (*nth nei*) genomic DNA confirmed the gene replacement. As can be seen in Fig. 5, for the two sets of primers used, the PCR products from the *nei* and *nei nth* mutants, which contained the *Cm* insert in the *nei* gene, were larger than those obtained from the wild type and in agreement with their expected sizes. These data also show that the mutants did not contain intact copies of the genes which might result from tandem duplications in the recipient.

Analysis of crude extracts for endo VIII and endo III activity on Tg-containing DNA (Fig. 6A) showed that both wild-type (BW35 and KL16) (lane 2) and *nei* mutant (SW2-8) (lane 4) cells contain significant amounts of activity. Wild-type cells contain both endo VIII and endo III activities, while *nei* mutants contain endo III only. Little activity was observed in *nth* mutant (BW402) cells (lane 3). This result is in agreement with previous observations which showed that endo VIII accounted for only about 10% of the cellular Tg glycosylase activity (43). Very little nicking of Tg-containing DNA was observed in the *nth nei* double mutant (lane 5). This low level of nicking might be accounted for by the presence of AP sites in the DNA or by the activity of Fpg, which has also been shown to recognize Tg in vitro (27). Some AP sites were present in the substrate used here and were detected by cleavage with *E. coli* endo IV (lane 6).

Western blot analysis, using a monoclonal Fab antibody against endo VIII, showed that neither the putative *nei* mutant (SW2-8) nor the *nei nth* double mutant (SW2-38) produced immunoreactive endo VIII protein (Fig. 6B). Crude extracts were blotted with monoclonal Fab Nei-31, and endo VIII bands were detected by enzymatic chemiluminescence. Lanes 1 and 2 show that Fab Nei-31 specifically reveals a band of approximately 30,000 Da in wild-type (BW35 and KL16) and *nth* mutant (BW402) extracts, and this band is absent from

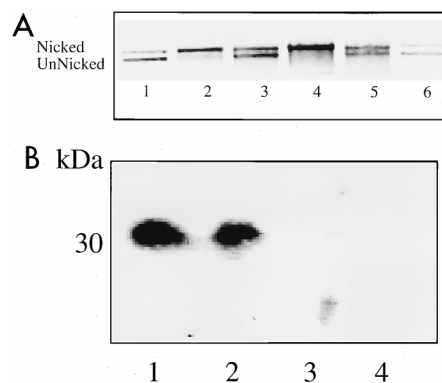


FIG. 6. Detection of endo VIII in crude *E. coli* extracts. (A) Enzymatic activity as measured by nicking of closed circular DNA containing Tg. Lane 1, no enzyme; lane 2, 5  $\mu$ l of extract from BW35 (KL16 [wild type]) cells; lane 3, 5  $\mu$ l of extract from BW402 (*nth*); lane 4, 5  $\mu$ l of extract from SW-8 (*nei*); lane 5, 5  $\mu$ l of extract from SW2-38 (*nth nei*); lane 6, saturating amount of purified endo IV. Reactions were performed as described in Materials and Methods. (B) Detection by immunochemical enzymatic chemiluminescence: Lanes 1 and 2 were loaded with 25  $\mu$ l of crude extracts isolated from BW35 (KL16 [wild type]) and BW402 (*nth*) cells, respectively. Lanes 3 and 4 are samples isolated from SW-8 (*nei*) and SW2-38 (*nth nei*) cells, respectively. The protein bands were detected by immunoreactivity with Fab Nei-31 as described in Materials and Methods.

similarly treated extracts isolated from *nei* (SW2-8) and *nth nei* (SW2-38) strains (lanes 3 and 4). No protein cross-reactive with Fab-Nei 31 was observed in *nei* or *nth nei* double-mutant cells.

**X-ray sensitivity of *nei* single and *nei nth* double null mutants.** Cells lacking endo III are not hypersensitive to DNA-damaging agents that produce lethal substrates recognized by the enzyme. One explanation for this observation is that cellular endo VIII provides a backup activity for endo III. To examine this possibility, we measured the X-ray sensitivity of isogenic wild-type and *nth nei*, and *nth nei* mutant strains. In Fig. 7A, it can be seen that the *nth nei* double mutant was 1.4-fold more sensitive than the wild type, as determined by comparing the mean lethal doses. As previously observed (18), *nth* single mutants lacking endo III were not X-ray sensitive (data not shown). Interestingly, single *nei* mutants consistently showed a slightly increased X-ray sensitivity compared to their isogenic wild-type strain (Fig. 7B). The sensitivity of the *nth nei* double mutant, however, was not nearly as great as the sensitivity of an *xth nfo* mutant which is totally devoid of base excision repair. In agreement with published results of Cunningham et al. (19), the *xth nfo* mutant was about fourfold more sensitive than its isogenic wild-type strain (data not shown). *nth nei* double mutants were also about three- to fourfold more sensitive to hydrogen peroxide than wild-type cells but less sensitive than *xth* mutants (data not shown).

These data support the idea that there are additional lethal lesions produced by X rays or hydrogen peroxide and processed by base excision repair, such as sites of base loss, strand breaks, and lethal purine lesions, that are not recognized by either endo III or endo VIII. Alternatively, in the absence of exo III and endo IV, endo III and endo VIII, due to their lyase activity, may produce lethal breaks from X-ray- or hydrogen peroxide-induced damage.

**Mutator phenotypes of *nei*, *nth*, and *nei nth* strains.** Many of the free radical-modified DNA bases are not blocks to DNA polymerases but miscode and are thus potentially mutagenic. It has previously been established, using several reversion assays, that cells lacking endo III exhibit a mild mutator effect (18, 69), and so it was of interest to examine the role, if any, that endo

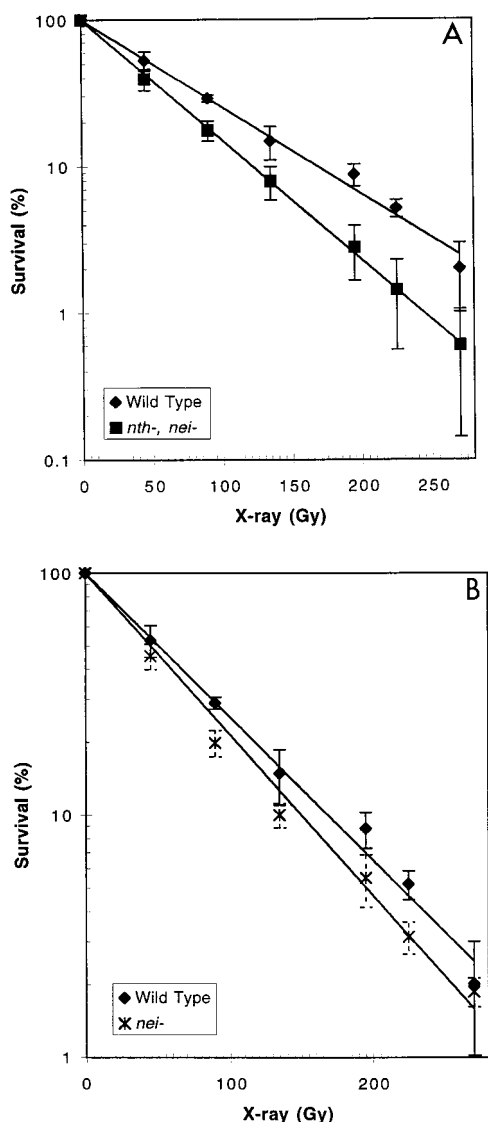


FIG. 7. Sensitivity of *E. coli* mutants to killing by X rays. Cultures of exponentially growing cells of BW35 (KL16 [wild type]) and SW2-38 (*nth nei*) (A) and of BW35 (KL16 [wild type]) and SW2-8 (*nei*) (B) were irradiated in air with 45, 90, 135, 180, 225, or 270 Gy. The log survival of each strain was plotted against the irradiation dose.

VIII might play in this phenotype. Using a rifampin forward mutation assay, the ratio of revertants to total cells plated was used to calculate the mutation frequency. Table 2 corroborates existing data showing that *nth* mutants exhibit a mild mutator

TABLE 2. Spontaneous mutation frequencies in *nth*, *nei*, and *nth nei* mutant strains

Genotype	Spontaneous mutation frequency to Rif <sup>r</sup> (10 <sup>-8</sup> )		
	Expt 1	Expt 2	Expt 3
Wild type	2.2	1.6	2.1
<i>nth</i>	6.0	8.9	4.2
<i>nei</i>	0.8	0.5	2.0
<i>nth nei</i>	33	70	28

phenotype (threefold). *nei* single mutants showed the same mutation frequency as the wild type; however, the *nth nei* double mutant exhibited a strong mutator phenotype, about 20-fold. Thus, in the absence of endo III and endo VIII, a significant number of premutagenic pyrimidine lesions are unrepaired.

## DISCUSSION

In an effort to clone the gene coding for endo VIII and to further characterize the enzyme biochemically and structurally, endo VIII was purified to apparent physical homogeneity. The development of the redAP-DNA-cellulose affinity column was a critical step in the successful purification of endo VIII because cells contain a very small amount of the protein (~95% endo III, compared with 5% endo VIII [43]).

The *nei* structural gene with its flanking regions was cloned and sequenced. The overall predicted protein from the *nei* gene showed significant homology with five bacterial Fpg proteins, especially in N-terminal and C-terminal regions. The proteins also have similar sizes (263 and 269 amino acids), similar molecular masses (29.7 and 30.2 kDa), similar hydrophobicities (-0.32 and -0.26 hydropathy), and similar hydrophobicity patterns (data not shown), especially for the carboxy half (~160 amino acids) of the proteins. (The second number in each comparison is for *E. coli* Fpg.) Endo VIII contains a zinc finger motif (Cys-X<sub>2</sub>-Cys-X<sub>16</sub>-Cys-X<sub>2</sub>-Cys) at the C terminus, as does Fpg. *E. coli* Fpg proteins with mutations in the zinc finger consensus sequences do not bind Zn<sup>2+</sup> as strongly; in addition, these mutant proteins fail to specifically bind to DNA containing a redAP site and exhibit a significantly lower efficiency for all catalytic activities of Fpg (47, 64). Thus, the zinc finger is essential for binding of Fpg to DNA and, probably as a consequence, is essential for Fpg functions. The zinc finger motif of endo VIII is likely to play a similar role. In addition to sequence homology, endo VIII also shares several catalytic activities with Fpg, including cleavage of the DNA backbone by  $\beta,\delta$ -elimination as well as a 5' deoxyribose phosphatase (33). Thus, *nei* and *fpg*, the genes coding for endo VIII and Fpg, may have originated from a common ancestral gene.

In the upstream region of the *nei* structural gene, there are several -35- and -10-like hexamers. Considering the interregion spacing as well as homology to each consensus sequence, TTGTTG and TAATAA, it is possible that positions 26 to 31 and 50 to 55 (Fig. 2) serve as the -35 and -10 hexamers of the *nei* promoter, but this puts the start site of the message fairly close to the putative Shine-Dalgarno sequence. The distance between these two hexamers is 18 nucleotides, which matches the consensus distance (17  $\pm$  1 nucleotides) (26). The sequence of AAGGA (positions 166 to 170 in Fig. 2) overlaps with five of nine nucleotides of the Shine-Dalgarno consensus sequence (UAAGGAGGU) (25), and the spacing from it to the start codon (AUG) is seven nucleotides, which is equal to the average and optimal spacing between this particular partial Shine-Dalgarno sequence and the start codon of an open reading frame for prokaryotes (15). Therefore, AAGGA (positions 66 to 70) is proposed as a potential ribosome binding site for translation of endo VIII.

The *nei* gene also contains an unusually large number of rare codons, 26%, compared to 25 *E. coli* nonregulatory genes which utilize 12% rare codons (35). This use of rare codons by *nei* is high but comparable to that for the products of *fpg*, *uvrC*, *ada*, *alkA*, and *tagA* (25, 21, 26, 34, and 21%, respectively). In contrast, only 12% rare codons are used for endo III. The implication for a significantly increased use of rare codons for some proteins has been controversial (35, 60), but it has been

suggested that the use of rare codons may result in lower translation because of smaller cellular amounts of the corresponding tRNA. Although this idea is consistent with the difference between the amounts of endo III and endo VIII in the cell, it should be pointed out that when *nei* was expressed behind a strong promoter, significant overexpression of the protein resulted (Fig. 4).

The *nei* gene turned out to be upstream of *abrB*, the gene coding for a putative regulator of a DNA repair gene, *aidB*, but with an opposite orientation. In fact, the two genes overlap by 3 bp on their 3' ends (65a). The significance of this, if any, is not clear. The *nei* gene is downstream from two open reading frames, also in the opposite orientation. Thus, it is unlikely that *nei* is part of an operon.

Deletion-substitution mutants of *nei* were constructed by linear transformation, and deletion of *nei* was confirmed by genomic PCR, activity analysis in extracts, and Western blot analysis (Fig. 5 and 6). Hypersensitivity to ionizing radiation (Fig. 7A) and hydrogen peroxide was exhibited by *nth nei* double mutants; thus, these agents produce some lethal pyrimidine lesions subject to repair by endo III and endo VIII. Tg and urea are produced by both ionizing radiation and hydrogen peroxide and are lethal in duplex DNA (41), and so they are likely to be two of the potentially lethal lesions repaired by endo III and endo VIII. Other likely candidates are cytosine- and thymine-derived hydantoin, methyltartronyl urea, and other contraction, fragmentation, and ring-open products. As previously shown (18), wild-type X-ray sensitivity was exhibited by *nth* mutants although a very slight X-ray sensitivity was observed with *nei* mutants (Fig. 7B). Although these data should not be overinterpreted, it might be that endo VIII can repair some potential lesions not repaired by endo III; alternatively, endo III may produce some lethal lesions from otherwise nonlethal lesions.

Double mutants defective in both endo III and endo VIII showed a significant defect in the repair of spontaneous pre-mutagenic lesions (~20-fold) (Table 2) comparable to the mutator effect observed in *fpg* mutants (~15-fold) (12, 44). It is of interest that this mutator phenotype would not have been detected in a screen for spontaneous mutators because both *nei* and *nth* would have had to have been defective. *nei* mutants showed no mutator defect, while *nth* mutants showed a much smaller mutator effect than the double mutant. These data suggest that endo VIII cannot repair all of the pre-mutagenic lesions handled by endo III or that in the absence of endo III, endo VIII fixes otherwise repairable pre-mutagenic lesions. The fact that *nth* and *nei* are not completely epistatic to one another suggests that they are compartmentalized and/or act in independent pathways; however, in certain situations, such as repair of X-ray- or hydrogen peroxide-induced lethal lesions, endo III and endo VIII can substitute for one another, while in others, such as repair of potentially mutagenic lesions, they do not. In the latter case, enzyme abundance may play a role. It is clear from studies with uniquely introduced lesions that there are more oxidized pyrimidine base lesions that are pre-mutagenic than are potentially lethal. Thus, it might be that there is not enough endo VIII present in the cell to compensate in the repair of pre-mutagenic pyrimidine lesions in the absence of otherwise abundant endo III.

There is ample precedence in *E. coli* for both compartmentalization of repair enzyme function as well as effects of enzyme abundance on observed phenotypes. For example, endo IV and exo III are the major *E. coli* AP endonuclease activities. Both enzymes share catalytic activities, and in some cases they can completely substitute for one another while in other cases they cannot (19). Exo III constitutes about 90% of the cellular

hydrolytic AP activity, while endo IV is responsible for only 10% of this activity. Similarly to endo III and endo VIII, exo III and endo IV substitute for one another in the repair of X-ray-induced lethal lesions since single mutants exhibit wild-type sensitivity whereas double mutants are X-ray sensitive (19). However, endo IV, the low-abundance activity, cannot substitute in the repair of hydrogen peroxide-induced lethal lesions since single *xth* mutants lacking exo III are extremely sensitive to hydrogen peroxide (19). Lack of abundance of endo IV has also been attributed to its inability to overcome the lethal phenotype of *dut xth* mutants (70). Endo IV also appears to act in a separate repair pathway since it is induced by redox cyclers (14) and appears to be better able to repair bleomycin-induced lesions than exo III (*nfo* mutants are sensitive to bleomycin, while *xth* mutants are not) (19).

Good candidates for pre-mutagenic lesions recognized by endo III and endo VIII are oxidized cytosines such as 5-OHC, 5-OHU, and Ug. Ug and 5-OHC are the most abundant oxidized pyrimidine lesions (67). 5-OHC mispairs with A in vitro (53) and has been shown to be mutagenic (23). Ug and 5-OHU pair with A (52, 53) and since they are derived from C are expected to be potent pre-mutagenic lesions. All three of these oxidized cytosines would lead to C→T transitions. The thymine ring saturation products all pair correctly with A (22, 29); Tg is at best a poor pre-mutagenic lesion (6). Since the spontaneous pre-mutagenic lesions repaired by endo III and endo VIII are presumed to be the result of oxidative metabolism, the cytosine lesions are the most likely candidates for the spontaneous mutator effect observed in cells lacking these activities. This prediction is currently under investigation.

#### ACKNOWLEDGMENTS

The first four authors contributed equally to this work.

We are grateful to Michael Volkert for communicating his results with *abrB* prior to publication and for helpful discussion and to Yoke Wah Kow and Min Yao for their help and advice. Bernard Weiss generously supplied BW35 (KL16), the parent host for constructing the *nei* mutants. His advice on this strain enabled us to obtain reproducible X-ray survival data. We are also grateful to Yuji Kohara for supplying his *E. coli* ordered library.

This work was supported by National Institutes of Health grant R37 CA33657 awarded by the National Cancer Institute.

#### REFERENCES

1. Achey, P. M., and C. F. Wright. 1983. Inducible repair of thymine ring saturation damage in  $\phi$ X174 DNA. *Radiat. Res.* **93**:609-612.
2. Alberts, B., and G. Herrick. 1971. DNA-cellulose chromatography. *Methods Enzymol.* **21**:198-217.
3. Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* **215**:403-410.
4. Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.). 1994. *Current protocols in molecular biology*. John Wiley & Sons, New York, N.Y.
5. Barbas, C. F., III, and R. A. Lerner. 1991. Combinatorial immunoglobulin libraries on the surface of phage (Phabs): rapid selection of antigen-specific Fabs. *Methods Enzymol.* **2**:119-124.
6. Basu, A. K., E. L. Loechler, S. A. Leadon, and J. M. Essigmann. 1989. Genetic effects of thymine glycols: site-specific mutagenesis and molecular modeling studies. *Proc. Natl. Acad. Sci. USA* **86**:7677-7681.
7. Bespalov, I. A., A. A. Purmal, M. P. Glackin, S. S. Wallace, and R. J. Melamed. 1996. Recombinant phabs reactive with 7,8-dihydro-8-oxoguanine, a major oxidative DNA lesion. *Biochemistry* **35**:2067-2078.
8. Blattner, F. R., G. Plunkett III, G. F. Mayhew, N. T. Perna, and F. D. Glasner. Submitted for publication.
9. Boiteux, S., and O. Huisman. 1989. Isolation of a formamidopyrimidine-DNA glycosylase (*fpg*) mutant of *Escherichia coli* K12. *Mol. Gen. Genet.* **215**:300-305.
10. Boiteux, S., T. R. O'Connor, and J. Laval. 1987. Formamidopyrimidine-DNA glycosylase of *Escherichia coli*: cloning and sequencing of the *fpg* structural gene and overproduction of the protein. *EMBO J.* **6**:3177-3183.
11. Breen, A. P., and J. A. Murphy. 1995. Reactions of oxyl radicals with DNA. *Free Radical Biol. Med.* **18**:1033-1077.



12. Cabrera, M., Y. Nghiem, and J. H. Miller. 1988. *mutM*, A second mutator locus in *Escherichia coli* that generates G · C→T · A transversions. *J. Bacteriol.* **170**:5405–5407.
13. Castaing, B., S. Boiteux, and C. Zelwer. 1992. DNA containing a chemically reduced apurinic site is a high affinity ligand for the *E. coli* formamidopyrimidine-DNA glycosylase. *Nucleic Acids Res.* **20**:389–394.
14. Chan, E., and B. Weiss. 1987. Endonuclease IV of *Escherichia coli* is induced by paraquat. *Proc. Natl. Acad. Sci. USA* **84**:3189–3193.
15. Chen, H., M. Bjerkness, R. Kumar, and E. Jay. 1994. Determination of the optimal aligned spacing between the Shine-Dalgarno sequence and the translation initiation codon of *Escherichia coli* mRNAs. *Nucleic Acids Res.* **22**:4953–4957.
16. Chetsanga, C. J., and T. Lindahl. 1979. Release of 7-methylguanine residues whose imidazole rings have been opened from damaged DNA by a DNA glycosylase from *Escherichia coli*. *Nucleic Acids Res.* **6**:3673–3684.
17. Clark, J. M., and G. P. Beardsley. 1986. Thymine glycol lesions terminate chain elongation by DNA polymerase I *in vitro*. *Nucleic Acids Res.* **14**:737–749.
18. Cunningham, R. P., and B. Weiss. 1985. Endonuclease III (*nth*) mutants of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **82**:474–478.
19. Cunningham, R. P., S. M. Saporito, S. G. Spitzer, and B. Weiss. 1986. Endonuclease IV (*nfo*) mutant of *Escherichia coli*. *J. Bacteriol.* **168**:1120–1127.
20. Demple, B., and L. Harrison. 1994. Repair of oxidative damage to DNA: enzymology and biology. *Annu. Rev. Biochem.* **63**:915–948.
21. Demple, B., and S. Linn. 1980. DNA N-glycosylase and UV repair. *Nature (London)* **287**:203–208.
22. Evans, J., M. Maccabee, Z. Hatahet, J. Courcelle, R. Bockrath, H. Ide, and S. Wallace. 1993. Thymine ring saturation and fragmentation products: lesion bypass, misinsertion and implications for mutagenesis. *Mutat. Res.* **299**:147–156.
23. Feig, D. I., L. C. Sowers, and L. A. Loeb. 1994. Reverse chemical mutagenesis: identification of the mutagenic lesions resulting from reactive oxygen species-mediated damage to DNA. *Proc. Natl. Acad. Sci. USA* **91**:6609–6613.
24. Frenkel, K., A. Cummings, J. Solomon, J. Cadet, J. J. Steinberg, and G. W. Teebor. 1985. Quantitative determination of the 5-(hydroxymethyl)uracil moiety in the DNA of gamma-irradiated cells. *Biochemistry* **24**:4527–4533.
25. Gold, L. 1988. Posttranscriptional regulatory mechanisms in *Escherichia coli*. *Annu. Rev. Biochem.* **57**:199–233.
26. Harley, C. B., and R. P. Reynolds. 1987. Analysis of *E. coli* promoter sequences. *Nucleic Acids Res.* **15**:2343–2361.
27. Hatahet, Z., Y. W. Kow, A. A. Purmal, R. P. Cunningham, and S. S. Wallace. 1994. New substrates for old enzymes. 5-Hydroxy-2'-deoxycytidine and 5-hydroxy-2'-deoxyuridine are substrates for *Escherichia coli* endonuclease III and formamidopyrimidine DNA N-glycosylase, while 5-hydroxy-2'-deoxyuridine is a substrate for uracil DNA N-glycosylase. *J. Biol. Chem.* **269**:18814–18820.
28. Hayes, R. C., and J. E. LeClerc. 1986. Sequence dependence for bypass of thymine glycols in DNA by DNA polymerase I. *Nucleic Acids Res.* **14**:1045–1061.
29. Hayes, R. C., L. A. Petruccio, H. M. Huang, S. S. Wallace, and J. E. LeClerc. 1988. Oxidative damage in DNA. Lack of mutagenicity by thymine glycol lesions. *J. Mol. Biol.* **201**:239–246.
30. Huse, W. D., L. Sastry, S. A. Iverson, A. S. Kang, M. Alting-Mees, D. R. Burton, S. J. Benkovic, and R. A. Lerner. 1989. Generation of a large combinatorial library of the immunoglobulin repertoire in phage lambda. *Science* **246**:1275–1281.
31. Ide, H., Y. W. Kow, and S. S. Wallace. 1985. Thymine glycols and uracil residues in M13 DNA constitute replicative blocks *in vitro*. *Nucleic Acids Res.* **13**:8035–8052.
32. Ide, H., L. A. Petruccio, Z. Hatahet, and S. S. Wallace. 1991. Processing of DNA base damage by DNA polymerases. Dihydrothymine and  $\beta$ -ureidoisobutyric acid as models for instructive and noninstructive lesions. *J. Biol. Chem.* **266**:1469–1477.
33. Jiang, D., Z. Hatahet, R. J. Melamed, and S. S. Wallace. Unpublished data.
34. Kohara, Y., K. Akiyama, and K. Isono. 1987. The physical map of the whole *E. coli* chromosome: application of a new strategy for rapid analysis and sorting of a large genomic library. *Cell* **50**:495–508.
35. Konigsberg, W., and G. N. Godson. 1983. Evidence for use of rare codons in the *dnaG* gene and other regulatory genes of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **80**:687–691.
36. Kow, Y. W., and S. S. Wallace. 1987. Mechanism of action of *Escherichia coli* endonuclease III. *Biochemistry* **26**:8200–8206.
37. Kow, Y. W., G. Faundez, R. J. Melamed, and S. S. Wallace. 1991. Processing of model single-strand breaks in  $\phi$ x-174 RF transfecting DNA by *Escherichia coli*. *Radiat. Res.* **126**:357–366.
38. Kowalski, D. 1979. A procedure for the quantitation of relaxed closed circular DNA in the presence of superhelical DNA: an improved fluorometric assay for nicking-closing enzyme. *Anal. Biochem.* **93**:346–354.
39. Kuwabara, M. 1991. Chemical processes induced by OH attack on nucleic acids. *Radiat. Phys. Chem.* **37**:691–704.
40. Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)* **227**:680.
41. Laspias, M. F., and S. S. Wallace. 1988. Excision repair of thymine glycols, uracil residues, and apurinic sites in *Escherichia coli*. *J. Bacteriol.* **170**:3359–3366.
42. Maccabee, M., J. S. Evans, M. P. Glackin, Z. Hatahet, and S. S. Wallace. 1994. Pyrimidine ring fragmentation products. Effects of lesion structure and sequence context on mutagenesis. *J. Mol. Biol.* **236**:514–530.
43. Melamed, R. J., Z. Hatahet, Y. W. Kow, H. Ide, and S. S. Wallace. 1994. Isolation and characterization of endonuclease VIII from *Escherichia coli*. *Biochemistry* **33**:1255–1264.
44. Michaels, M. L., C. Cruz, A. P. Grollman, and J. H. Miller. 1992. Evidence that MutY and MutM combine to prevent mutations by an oxidatively damaged form of guanine in DNA. *Proc. Natl. Acad. Sci. USA* **89**:7022–7025.
45. Michaels, M. L., L. Pham, C. Cruz, and J. H. Miller. 1991. MutM, a protein that prevents G · C→T · A transversions, is formamidopyrimidine-DNA glycosylase. *Nucleic Acids Res.* **19**:3629–3632.
46. Miller, J. H. 1974. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- 46a. Mori, H. Accession no. D90710.
47. O'Connor, T. R., R. J. Graves, G. de Murcia, B. Castaing, and J. Laval. 1993. Fpg protein of *Escherichia coli* is a zinc finger protein whose cysteine residues have a structural and/or functional role. *J. Biol. Chem.* **268**:9063–9070.
48. Oshima, T., H. Aiba, T. Baba, K. Fujita, K. Hayashi, A. Honjo, K. Ikemoto, T. Inada, T. Itoh, M. Kajihara, K. Kanai, K. Kashimoto, S. Kimura, M. Kitagawa, K. Makino, S. Masuda, T. Miki, K. Mizobuchi, H. Mori, K. Motomura, Y. Nakamura, H. Nashimoto, Y. Nishio, N. Saito, G. Sampei, Y. Seki, H. Tagami, K. Takemoto, C. Wada, Y. Yamamoto, M. Yano, and T. Horiuchi. 1996. A 718-kb DNA sequence of the *Escherichia coli* K-12 genome corresponding to the 12.7–28.0 min region on the linkage map. *DNA Res.* **3**:137–155.
49. Parnley, S. F., and G. P. Smith. 1988. Antibody-selectable filamentous fd phage vectors: affinity purification of target genes. *Gene* **73**:305–318.
50. Poehling, H.-M., and V. Neuhoff. 1981. Visualization of proteins with a silver "stain": a critical analysis. *Electrophoresis* **2**:141–147.
51. Pribnow, S. D. 1975. Nucleotide sequence of an RNA polymerase binding site at an early T7 promoter. *Proc. Natl. Acad. Sci. USA* **72**:784–788.
52. Purmal, A. A., J. Bond, Y. W. Kow, and S. S. Wallace. Klenow fragment (exonuclease free) incorporates dAMP opposite 2'-deoxyuridine glycol and bypasses it more effectively than 2'-deoxythymidine glycol. Submitted for publication.
53. Purmal, A. A., Y. W. Kow, and S. S. Wallace. 1994. Major oxidative products of cytosine, 5-hydroxycytosine and 5-hydroxyuracil, exhibit sequence context-dependent mispairing *in vitro*. *Nucleic Acids Res.* **22**:72–78.
54. Radman, M. 1976. An endonuclease from *Escherichia coli* that introduces single polynucleotide chain scissions in ultraviolet-irradiated DNA. *J. Biol. Chem.* **251**:1438–1445.
55. Riek, R. P., M. D. Handschumacher, S. S. Sung, M. Tan, M. J. Glynias, M. D. Schluchter, J. Novotny, and R. M. Graham. 1995. Evolutionary conservation of both the hydrophilic and hydrophobic nature of transmembrane residues. *J. Theor. Biol.* **172**:245–258.
56. Rouet, P., and J. M. Essigmann. 1985. Possible role for thymine glycol in the selective inhibition of DNA synthesis on oxidized DNA templates. *Cancer Res.* **45**:6113–6118.
57. Russell, C. B., D. S. Thaler, and F. W. Dahlquist. 1989. Chromosomal transformation of *Escherichia coli* *recD* strains with linearized plasmids. *J. Bacteriol.* **171**:2609–2613.
58. Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
59. Sastry, L., M. Alting-Mees, W. D. Huse, J. M. Short, J. A. Sorge, B. N. Hay, K. D. Janda, S. J. Benkovic, and R. A. Lerner. 1989. Cloning of the immunological repertoire in *Escherichia coli* for generation of a heavy chain variable region-specific cDNA library. *Proc. Natl. Acad. Sci. USA* **86**:5728–5732.
60. Sharp, P. M., and W. Li. 1986. Codon usage in regulatory genes in *Escherichia coli* does not reflect selection for 'rare' codons. *Nucleic Acids Res.* **14**:7737–7749.
61. Shine, J., and L. Dalgarno. 1975. Determinant of cistron specificity in bacterial ribosomes. *Nature (London)* **254**:34–38.
62. Strniste, G., and S. S. Wallace. 1975. Endonucleolytic incision of x-irradiated deoxyribonucleic acid by extracts of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **72**:1997–2001.
63. Tchou, J., H. Kasai, S. Shibutani, M. H. Chung, J. Laval, A. P. Grollman, and S. Nishimura. 1991. 8-Oxoguanine (8-hydroxyguanine) DNA glycosylase and its substrate specificity. *Proc. Natl. Acad. Sci. USA* **88**:4690–4694.
64. Tchou, J., M. L. Michaels, J. H. Miller, and A. P. Grollman. 1993. Function of the zinc finger in *Escherichia coli* Fpg protein. *J. Biol. Chem.* **268**:26738–26744.
65. Towbin, H., T. Staehelin, and J. Gordon. 1979. Electrophoretic transfer of

- proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl. Acad. Sci. USA* **76**:4350–4354.
- 65a. **Volkert, M.** Personal communication.
66. **von Sonntag, C.** 1987. The chemical basis of radiation biology. Taylor and Francis, London, Great Britain.
67. **Wagner, J. R., C. C. Hu, and B. N. Ames.** 1992. Endogenous oxidative damage of deoxycytidine in DNA. *Proc. Natl. Acad. Sci. USA* **89**:3380–3384.
68. **Wallace, S. S.** 1997. Oxidative damage to DNA and its repair, p. 49–90. *In* J. G. Scandalios (ed.), *Oxidative stress and the molecular biology of antioxidant defenses*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
69. **Weiss, B., R. P. Cunningham, E. Chan, and I. R. Tsaneva.** 1988. AP endonucleases of *Escherichia coli*, p. 133–142. *In* E. C. Friedberg and P. Hanawalt (ed.), *Mechanisms of consequences of DNA damage processing*. Alan R. Liss, Inc., New York, N.Y.
70. **Weiss, B., S. G. Rogers, and A. F. Taylor.** 1978. The endonuclease activity of exonuclease III and the repair of uracil-containing DNA in *Escherichia coli*, p. 191–192. *In* P. C. Hanawalt, E. C. Friedberg, and C. F. Fox (ed.), *DNA repair mechanisms*. Academic Press, Inc., New York, N.Y.
71. **Yacoub, A., L. Augeri, M. R. Kelley, P. W. Doetsch, and W. A. Deutsch.** 1996. A *Drosophila* ribosomal protein contains 8-oxoguanine and abasic site DNA repair activities. *EMBO J.* **15**:2306–2312.
72. **Zhou, M., and Z. Hatahet.** 1995. An improved ligase-free method for directional subcloning of PCR amplified DNA. *Nucleic Acids Res.* **23**:1089–1090.