Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 5197–5202, May 1996
Genetics

Cloning and expression in *Escherichia coli* of the *OGG1* gene of *Saccharomyces cerevisiae*, which codes for a DNA glycosylase that excises 7,8-dihydro-8-oxoguanine and 2,6-diamino-4-hydroxy-5-*N*-methylformamidopyrimidine

PATRICIA AUFFRET VAN DER KEMP†, DOMINIQUE THOMAS‡, RÉGINE BARBEY‡, RÉGINA DE OLIVEIRA†, AND SERGE BOITEUX†§

†Unité de Recherche Associée 147 Centre National de la Recherche Scientifique, Groupe "Réparation des lésions radio et chimioinduites," PR2-Institut Gustave Roussy, 94800-Villejuif, France; and ‡Centre de Génétique Moléculaire, Centre National de la Recherche Scientifique, 91000-Gif sur Yvette, France

Communicated by Maurice S. Fox, Massachusetts Institute of Technology, Cambridge, MA, December 26, 1995 (received for review October 30, 1995)

ABSTRACT **A spontaneous mutator strain of *Escherichia coli* (*fpg mutY*) was used to clone the *OGG1* gene of *Saccharomyces cerevisiae*, which encodes a DNA glycosylase activity that excises 7,8-dihydro-8-oxoguanine (8-OxoG). *E. coli* (*fpg mutY*) was transformed by a yeast DNA library, and clones that showed a reduced spontaneous mutagenesis were selected. The antimutator activity was associated with pYSB10, an 11-kbp recombinant plasmid. Cell-free extracts of *E. coli* (*fpg mutY*) harboring pYSB10 possess an enzymatic activity that cleaves a 34-mer oligonucleotide containing a single 8-oxoG opposite a cytosine (8-OxoG/C). The yeast DNA fragment of 1.7 kbp that suppresses spontaneous mutagenesis and overproduces the 8-OxoG/C cleavage activity was sequenced and mapped to chromosome XIII. DNA sequencing identified an open reading frame, designated *OGG1*, which encodes a protein of 376 amino acids with a molecular mass of 43 kDa. The *OGG1* gene was inserted in plasmid pUC19, yielding pYSB110. *E. coli* (*fpg*) harboring pYSB110 was used to purify the Ogg1 protein of *S. cerevisiae* to apparent homogeneity. The Ogg1 protein possesses a DNA glycosylase activity that releases 8-OxoG and 2,6-diamino-4-hydroxy-5-*N*-methylformamidopyrimidine. The Ogg1 protein preferentially incises DNA that contains 8-OxoG opposite cytosine (8-OxoG/C) or thymine (8-OxoG/T). In contrast, Ogg1 protein does not incise the duplex where an adenine is placed opposite 8-OxoG (8-OxoG/A). The mechanism of strand cleavage by Ogg1 protein is probably due to the excision of 8-OxoG followed by a β-elimination at the resulting apurinic/apyrimidinic site.**

Reactive oxygen species formed in cells as by-products of aerobic metabolism or during oxidative stress have been suggested to play a role in biological processes such as mutagenesis, carcinogenesis, and aging (1, 2). Hydroxyl radical ($OH^{\cdot}$) and singlet oxygen ($^1O_2$) are highly reactive and produce complex patterns of DNA modifications (3, 4). Several lines of evidence suggest that an oxidatively damaged form of guanine, 7,8-dihydro-8-oxoguanine (8-OxoG), is highly mutagenic *in vitro* and *in vivo* (5, 6). *Escherichia coli* possesses two DNA glycosylase activities that prevent spontaneous mutagenesis by 8-OxoG: the Fpg protein, which excises 8-OxoG residues in DNA (4, 7–9), and the MutY protein, which excises the adenine residues incorporated opposite 8-OxoG (7, 10, 11). The Fpg protein also releases imidazole ring-opened purines such as 2,6-diamino-4-hydroxy-5-*N*-methylformamidopyrimidine (Fapy) (9). Inactivation of both the *fpg* (*mutM*) and the *mutY* (*micA*) genes of *E. coli* results in a strong GC → TA spontaneous mutator phenotype (7, 12, 13). These results in *E. coli* strongly suggest that endogeneous oxidative stress is a major cause of DNA damage and that excision of 8-OxoG in DNA contributes to the maintenance of genetic stability. Enzyme activities that repair oxidized bases in DNA have been partially purified from the yeast *Saccharomyces cerevisiae* (14, 15), but genes coding for these DNA repair activities have not yet been identified.

In this work, we report the cloning and sequencing of the *OGG1* gene from *S. cerevisiae*. Expression of *OGG1* partially suppresses spontaneous mutagenesis of the *E. coli* (*fpg mutY*) mutant. The *OGG1* gene codes for a 43-kDa protein, which was purified to apparent homogeneity and shown to release both 8-OxoG and Fapy residues from DNA. These results suggest that *OGG1* is a functional eukaryotic homologue of the bacterial *fpg* gene.

MATERIALS AND METHODS

Bacterial Strains, Enzymes, and Yeast Genomic DNA Library. *E. coli* K12 strains used are CC104 [*ara*, Δ(*gpt-lac*)$_5$, *rpsL*/F′ (*lacI*$_{378}$, *lacZ*$_{461}$, *proA*$^+$*B*$^+$)] (16), BH1190 [same as CC104 but *fpg*::Kan^R (*fpg*), *micA* (*mutY*)::Kan^R, *recA*::Cm^R; Kan^R = kanamycin resistance and Cm^R = chloramphenicol resistance] (refs. 17 and 18 and this study), JM105 [*supE, endA, sbcB15, hsdR4, rpsL, thi*, Δ(*lac-proAB*)/F′ (*traD36, proA*$^+$*B*$^+$, *lacI*^{q}, *lacZ* ΔM15)], and BH410 (same as JM105 but *fpg*) (19). The *S. cerevisiae* library was prepared by inserting the product of a partial *Sau*3A digest of chromosomal DNA from the wild-type strain FL100 in the *Bam*HI site of pFL44L (20, 21). Homogeneous Fpg protein of *E. coli* was purified from an overproducing strain as described (22).

Selection for Antimutators. Mutator strain BH1190 (*fpg mutY*) was transformed by 200 ng of the yeast genomic DNA library using electroporation. The transformation mixture was plated onto M9 minimal medium containing glucose (2%), phenyl β-D-galactoside (500 μg/ml), 5-bromo-4-chloro-3-indolyl β-D-galactoside (40 μg/ml), and ampicillin (100 μg/ml) at a density of 200–500 colonies per plate. After 4–6 days at 37°C, colonies with three or less blue papillae were selected and individually grown overnight at 37°C in 2 ml of Luria–Bertani (LB) broth medium containing ampicillin (100 μg/ml). The resulting cultures were analyzed for the generation of

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

Abbreviations: 8-OxoG, 7,8-dihydro-8-oxoguanine; Fapy, 2,6-diamino-4-hydroxy-5-*N*-methylformamidopyrimidine; Rif^R, rifampicin resistant; AP, apurinic/apyrimidinic; MB-light DNA, DNA exposed to methylene blue plus visible light.

§To whom reprint requests should be sent at the present address: Département de Radiobiologie et Radiopathologie, Commissariat à l'Energie Atomique, BP6 Bat.05, Fontenay aux Roses, France.

¶The sequence reported in this paper has been deposited in the GenBank data base (accession no. U44855).

rifampicin-resistant (Rif^R) mutants after plating on LB broth plates containing ampicillin (100 μg/ml) and rifampicin (100 μg/ml).

DNA Manipulations. The sequence of the *OGG1* gene was determined after subcloning of restriction fragments into M13mp18 and M13mp19 vectors. Both strands were sequenced using single-stranded templates by the dideoxy chain-termination method with T7 DNA polymerase (Sequenase). Restriction enzymes, DNA ligase, and DNA polymerases were from commercial sources and used as recommended by manufacturers. Homology searches were performed using the BLAST network service at the National Center for Biotechnology.

Assay for the Repair of 8-OxoG. The 34-mer oligonucleotides containing a single 8-OxoG or a G residue at position 16 were synthesized as described (23). The sequences used in this study are the following:

Oligo1:

5′-GGCTTCATCGTTATT(8-OxoG)ATGACCTGGTGGATACCG-5′*

Oligo2:

5′-GGCTTCATCGTTATT(G)ATGACCTGGTGGATACCG-5′*

Complementary sequences with C, T, G, or A opposite 8-OxoG were also synthesized. The nucleotide at the 3′ end of these 34-mer oligonucleotides was inverted, yielding a 5′-$(N)_{33}$-3′-P-3′-N-5′* sequence (24). Both ends of these 34-mers were ^{32}P-labeled by T4 polynucleotide kinase and [γ-^{32}P]ATP (3000 Ci/mmol; 1 Ci = 37 GBq Amersham) (24, 25). The ^{32}P-labeled 8-OxoG-containing strand was annealed with a complementary sequence at a 1:1.5 molar ratio. The assay mixture (25 μl final volume) contained 25 mM Tris·HCl (pH 7.6), 50 mM KCl, 5 mM Na_2EDTA, 50 fmol of ^{32}P-labeled duplex, and 1–5 μl of cell-free extract (26) or column fraction. The products of the reaction were analyzed by 20% PAGE containing 7 M urea (25).

Assay for the 8-OxoG DNA Glycosylase Activity. To generate 8-OxoG residues, calf thymus DNA was exposed to methylene blue plus visible light (MB-light DNA) as described (4). The reaction mixture (50 μl final volume) contained 25 mM Tris·HCl (pH 7.6), 2 mM Na_2EDTA, 50 mM KCl, 5 μg of MB-light DNA, and 1–10 μl of column fraction. The assay mixture was incubated at 37°C for 30 min. Aliquots of the enzyme reaction were analyzed by HPLC with electrochemical detection by using a C_{18} μBondapak column (Waters). The column was coupled to a Waters model 460 electrochemical detector set at a potential of +700 mV for the detection of 8-OxoG. The column was eluted with 50 mM citrate buffer (pH 4.5) containing 5% methanol at a flow rate of 1 ml/min. Authentic 8-OxoG (Chemical Dynamics) and guanine marker molecules were eluted at 8.9 min and 7.5 min, respectively.

Assay for the Fapy DNA Glycosylase Activity. The activity was measured using [^3H]Fapy-poly(dG-dC) as substrate as described (22).

Purification of the Ogg1 Protein. LB broth medium (10 liters) containing ampicillin (50 μg/ml) was inoculated with 50 ml of an overnight culture of *E. coli* BH410 (*fpg*) hosting the pYSB110 plasmid. The cells were grown at 37°C until OD_{600} = 1.0 and supplemented with 0.5 mM isopropyl β-D-thiogalactopyranoside and incubated for 4 hr at 37°C. The cells (21 g) were lysed in buffer A (25 mM Tris·HCl, pH 7.6/2 mM Na_2EDTA/5% glycerol/1 mM phenylmethylsulfonyl fluoride) containing 300 mM NaCl and lysozyme (Sigma) at 1 mg/ml. The lysate was centrifuged at 30,000 rpm (130,000 *g*) for 45 min at 2°C. The supernatant was the crude extract fraction. The purification procedure of the Ogg1 protein includes a QMA anion-exchange column (Waters-ACELL) equilibrated with buffer A containing 300 mM NaCl to separate nucleic acids from proteins. The Ogg1 protein was further purified by Phospho-Ultrogel cation exchange (IBF-LKB), AcA54 gel filtration (IBF–LKB), phenyl-Sepharose CL-4B (Pharmacia), DNA cellulose (Sigma), and FPLC cationic Mono S HR 5/5 (Pharmacia) chromatography (22).

Table 1. Spontaneous mutagenesis of the *E. coli* (*fpg mutY*) mutant expressing the *OGG1* gene of the yeast *S. cerevisiae* from plasmid pYSB10

Bacterial strain	Plasmid	$Rif^R/10^9$	$Lac^+/10^9$
CC104 (wild type)	pFL10	10	2
BH1190 (*fpg mutY*)	pFL10	10,400	11,934
BH1190 (*fpg mutY*)	pYSB10	880	985

E. coli strain CC104 possesses an altered *lacZ* gene, which can only revert to wild type via a GC → TA transversion event (16). Plasmid pFL10 is a random plasmid from the genomic yeast library in pFL44L. Ten independent overnight cultures of *E. coli* CC104 (wild type) or BH1190 (*fpg mutY*) transformed with either pFL10 or pYSB10 were analyzed for Rif^R and Lac^+ mutation events. Aliquots were plated on LB agar containing ampicillin (100 μg/ml) and rifampicin (100 μg/ml) or on minimal glucose or minimal lactose plates containing ampicillin (100 μg/ml).

RESULTS

Isolation of a Yeast Gene That Suppresses the Spontaneous Mutagenesis of the *E. coli* BH1190 (*fpg mutY*) Mutant. The aim

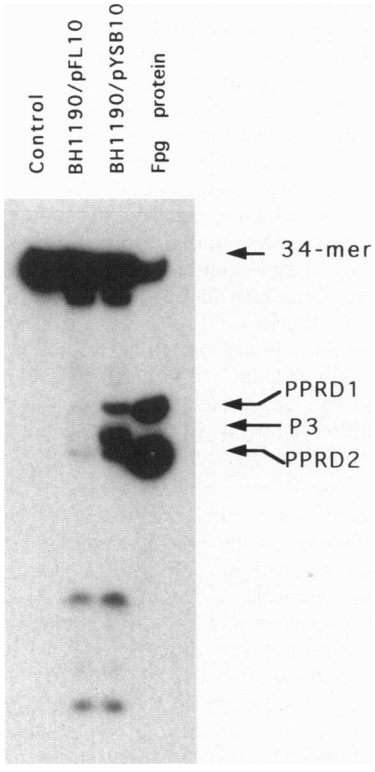

FIG. 1. Cleavage of the 8-OxoG/C duplex by crude extracts of *E. coli* BH1190 (*fpg mutY*) hosting pFL10 or pYSB10 plasmid. The 8-OxoG/C duplex 50 fmol was incubated either with 2 ng of purified Fpg protein or with BH1190 cell-free extracts (30 μg of protein). The two extremities of the 8-OxoG-containing strand are ^{32}P-labeled and consequently two labeled oligonucleotides are generated after cleavage. The products of the reaction were analyzed by denaturing 20% PAGE containing 7 M urea. PPRD1 is the 3′ oligonucleotide (5′-P-(N_{17})-3′-P-3′-N-[^{32}P]-5′) and PPRD2 is the 5′ oligonucleotide (5′-[^{32}P]-(N_{15})-3′-P) generated by chemical cleavage of the 34-mer oligonucleotide at 8-OxoG by 1 M piperidine at 90°C for 30 min or by the *E. coli* Fpg protein. The P3 product is probably PPRD2 with a deoxyribose residue attached at its 3′ end.

of this work was to isolate yeast genes that encode enzymes implicated in the repair of 8-OxoG in DNA. The cloning strategy relied on the ability of yeast DNA fragments cloned in a plasmid vector to suppress spontaneous GC → TA mutagenesis of the *E. coli* strain BH1190 (*fpg mutY*) (Table 1). The spontaneous mutator character of BH1190 (*fpg mutY*) is visualized by the formation of numerous blue papillae growing on bacterial colonies plated on indicator plates containing a combination of β-galactosides (27). A yeast genomic library cloned into pFL44L vector was introduced into BH1190 (*fpg mutY*), and 30,000 transformants were analyzed. Bacterial clones that exhibited a reduced number of blue papillae were selected and individually analyzed for spontaneous formation of Rif^R and Lac^+ mutants in liquid culture. The two clones that were finally selected both contained the same 11-kbp recombinant plasmid, pYSB10. Table 1 shows that pYSB10 partially suppresses the spontaneous mutagenesis of BH1190 (*fpg mutY*), since the formation of either Rif^R or Lac^+ revertants was reduced ≈10-fold compared to the same strain hosting a control plasmid.

The Cloned Yeast Gene Encodes an Enzyme That Cleaves DNA Containing 8-OxoG. The suppression of the spontaneous mutagenesis could be either due to the expression of a gene product that inactivates endogeneous reactive oxygen species or due to the expression of a DNA repair enzyme. Thus, we have measured the capacity of crude extracts of BH1190 (*fpg mutY*) hosting the pYSB10 plasmid to cleave a 34-mer double-stranded DNA fragment that contains a single 8-OxoG placed opposite a cytosine (8-OxoG/C). Fig. 1 shows that crude extracts of BH1190 hosting the control plasmid pFL10 do not incise the 8-OxoG/C duplex. In contrast, crude extracts of BH1190 hosting pYSB10 possess an activity that nicks the 8-OxoG/C duplex, yielding three major bands on the autoradiograph (Fig. 1). Two of the three products comigrate with the products generated by the Fpg protein of *E. coli* that cleaves DNA at 8-OxoG (Fig. 1). These results strongly suggest that the yeast DNA insert cloned in pYSB10 contains a gene that codes for a DNA repair protein that incises DNA at 8-OxoG residues. We named this yeast gene *OGG1* for 8-*O*xo-*G*uanine DNA *G*lycosylase.

The *OGG1* Gene Codes for a 43-kDa Protein and Maps to Chromosome XIII. A restriction map of the 5-kbp yeast DNA insert cloned in pYSB10 was used to make a series of subclones to localize the *OGG1* gene (Fig. 2). The subclones were tested for their ability to produce the 8-OxoG/C nicking activity and to suppress spontaneous mutagenesis of the BH1190 (*fpg mutY*) mutant. Plasmid pYSB110 was the shortest subclone retaining both biological and biochemical activity (Fig. 2). Deletion of the 100-bp *Xho* I–*Xho* I fragment, yielding plasmid pYSB120, abolished both the nicking and the complementation activities (Fig. 2). These data localized the *OGG1* gene within a 1.7-kbp *Eco*RI–*Nsp* I fragment. When this 1.7-kbp DNA fragment was inverted with respect to the *lacZ* promoter, the enzyme activity was greatly reduced, which suggests that the transcription of the *OGG1* gene is driven from the *lacZ* promoter in pYSB10 and pYSB110 (Fig. 2).

The sequence of the 1.7-kbp *Eco*RI–*Nsp* I yeast DNA fragment that contained the *OGG1* gene was determined on both strands (Fig. 3). This DNA sequence contains an open reading frame with the potential to encode a protein of 376 amino acid residues with a calculated molecular mass of 42,781 Da and an isoelectric point of 8.83. Comparison of the DNA sequence of the *OGG1* region with sequences in the GenBank data base revealed that *OGG1* is adjacent to the *PIF1* gene (28). While this work was in progress, the systematic yeast genome sequencing project submitted to the European Molecular Biology Laboratory data base a fragment of chromosome XIII that contains the *OGG1* gene (accession number Z46729). Fig. 2 shows that the *OGG1* and *PIF1* genes are divergently transcribed; their respective translational start

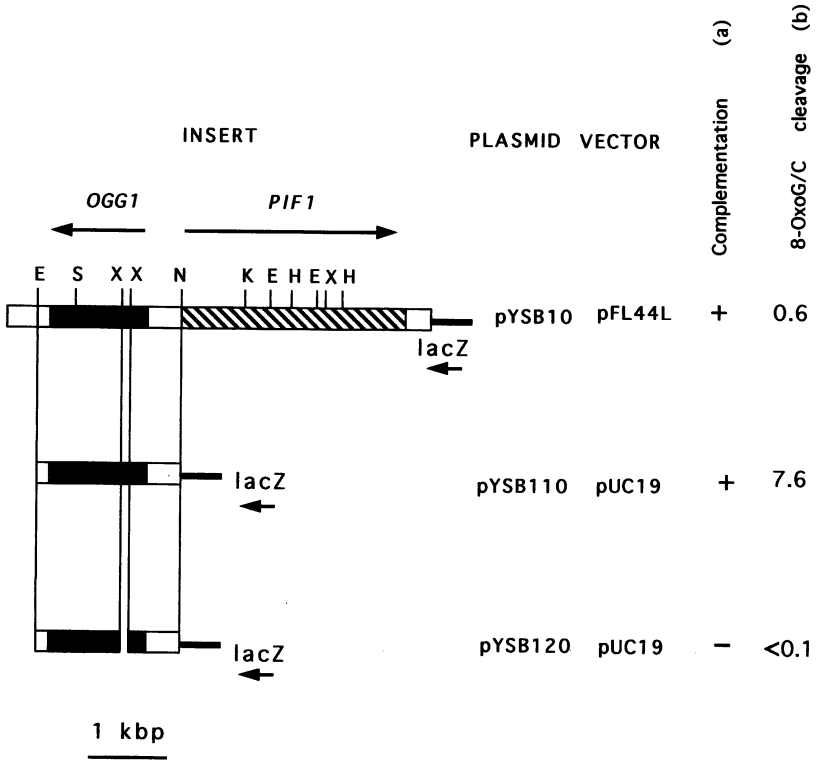

FIG. 2. Restriction map and subcloning of the yeast DNA fragment that contains the *OGG1* gene. Experiment a: complementation of the spontaneous mutagenesis of the BH1190 *E. coli* (*fpg mutY*) mutant. Experiment b: overproduction of an enzyme activity that cleaves the 8-OxoG/C duplex in BH1190 cell-free extracts. Specific activities are reported (units/mg). One unit cleaves 1 pmol of 8-OxoG/C in 15 min at 37°C. The black bar shows the *OGG1* gene, and the hatched bar shows the *PIF1* gene. Restriction sites: E, *Eco*RI; *S*, *Sac*I; *X*, *Xho* I; *N*, *Nsp* I; *K*, *Kpn* I; *H*, *Hind*III.

```
-459                        TCATAGAAAACGATAGTTTGGCGTGCGATACGTTTTTGAGTAAAAGAAAACTGTTGAAA -401
-400 CTACAGATAAATGGCCTTCTTGGTATAATATGATTCAATGTTGATCTTATCCACTTTGGCATGTAAAATTGATTGATACA -321
-320 AGTAAGCTAATCGCTCAATGGATAATATTACAAATAAAAATATCGAGGCAATACCAGTATTTCAATGTCTGGCCAAATAA -241
-240 ATAGTTTAGTACAGCAATTTGATTAACGTTAAAGTTACTCTCCTCTTTGAATAATGTCGTGGTTAAAAAGTAAAGTATTT -161
-160 GTTTATACATTAAGTGAAGAAAATTAAGACGCGCCTTTCTTAATGTAACGCCAAAAATTTTTAAGAGATGAGCTGAACAG -81
 -80 CATCACAAAACTACTATTTCCAGCGGAAGAAGGCATTTGAAGCGTCCTGATTCATAATTGCGATTTTATTTATCAACCAG -1

   1 ATG TCT TAT AAA TTC GGC AAA CTT GCC ATT AAT AAA AGT GAG CTA TGT CTA GCA AAT GTG  60
   1 M   S   Y   K   F   G   K   L   A   I   N   K   S   E   L   C   L   A   N   V   20

  61 CTA CAG GCT GGC CAA TCC TTT CGA TGG ATT TGG GAT GAA AAG TTA AAC CAG TAC AGC ACT 120
  21 L   Q   A   G   Q   S   F   R   W   I   W   D   E   K   L   N   Q   Y   S   T   40

 121 ACA ATG AAG ATA GGC CAG CAA GAA AAG TAT TCA GTA GTG ATT TTG AGG CAA GAT GAA GAG 180
  41 T   M   K   I   G   Q   Q   E   K   Y   S   V   V   I   L   R   Q   D   E   E   60

 181 AAT GAG ATC CTG GAA TTT GTT GCT GTC GGC GAC TGT GGC AAT CAG GAC GCC TTA AAA ACT 240
  61 N   E   I   L   E   F   V   A   V   G   D   C   G   N   Q   D   A   L   K   T   80

 241 CAT CTA ATG AAA TAT TTT AGA TTG GAC GTG TCA CTG AAG CAC TTA TTC GAT AAT GTT TGG 300
  81 H   L   M   K   Y   F   R   L   D   V   S   L   K   H   L   F   D   N   V   W  100

 301 ATT CCA AGC GAT AAA GCA TTT GCG AAA CTT TCT CCA CAA GGC ATC CGT ATC TTG GCC CAG 360
 101 I   P   S   D   K   A   F   A   K   L   S   P   Q   G   I   R   I   L   A   Q  120

 361 GAA CCA TGG GAG ACA TTG ATT TCC TTT ATT TGC TCG AGT AAT AAT AAC ATT TCA AGA ATT 420
 121 E   P   W   E   T   L   I   S   F   I   C   S   S   N   N   N   I   S   R   I  140

 421 ACG AGA ATG TGC AAT AGC CTT TGC TCT AAT TTC GGG AAC TTA ATC ACA ACA ATA GAT GGG 480
 141 T   R   M   C   N   S   L   C   S   N   F   G   N   L   I   T   T   I   D   G  160

 481 GTC GCC TAC CAT TCG TTT CCT ACA AGT GAG GAG CTA ACT TCT CGA GCC ACT GAA GCC AAG 540
 161 V   A   Y   H   S   F   P   T   S   E   E   L   T   S   R   A   T   E   A   K  180

 541 TTG CGT GAG CTG GGC TTT GGA TAT AGA GCT AAG TAT ATT ATC GAA ACC GCG AGA AAA CTG 600
 181 L   R   E   L   G   F   G   Y   R   A   K   Y   I   I   E   T   A   R   K   L  200

 601 GTG AAT GAT AAA GCT GAA GCT AAT ATT ACT TCT GAC ACC ACA TAC CTT CAA AGT ATC TGT 660
 201 V   N   D   K   A   E   A   N   I   T   S   D   T   T   Y   L   Q   S   I   C  220

 661 AAA GAT GCT CAA TAT GAA GAC GTT AGA GAG CAC CTC ATG TCC TAC AAT GGT GTA GGC CCC 720
 221 K   D   A   Q   Y   E   D   V   R   E   H   L   M   S   Y   N   G   V   G   P  240

 721 AAA GTT GCT GAT TGC GTT TGT CTA ATG GGA CTT CAC ATG GAT GGC ATC GTC CCC GTA GAT 780
 241 K   V   A   D   C   V   C   L   M   G   L   H   M   D   G   I   V   P   V   D  260

 781 GTC CAT GTT AGT AGA ATT GCG AAG AGG GAC TAC CAA ATA TCT GCG AAC AAG AAC CAT CTC 840
 261 V   H   V   S   R   I   A   K   R   D   Y   Q   I   S   A   N   K   N   H   L  280

 841 AAA GAA TTG AGG ACA AAA TAC AAC GCC TTA CCT ATC TCA AGA AAA AAA ATC AAC CTA GAG 900
 281 K   E   L   R   T   K   Y   N   A   L   P   I   S   R   K   K   I   N   L   E  300

 901 CTC GAC CAT ATT AGA TTA ATG CTT TTT AAA AAA TGG GGA TCA TAC GCT GGT TGG GCC CAA 960
 301 L   D   H   I   R   L   M   L   F   K   K   W   G   S   Y   A   G   W   A   Q  320

 961 GGT GTC CTA TTC TCT AAA GAA ATC GGA GGG ACT AGT GGT AGC ACT ACA ACT GGT ACT ATA 1020
 321 G   V   L   F   S   K   E   I   G   G   T   S   G   S   T   T   T   G   T   I  340

1021 AAG AAG AGA AAA TGG GAT ATG ATA AAG GAA ACG GAA GCA ATT GTT ACG AAA CAA ATG AAA 1080
 341 K   K   R   K   W   D   M   I   K   E   T   E   A   I   V   T   K   Q   M   K  360

1081 CTG AAA GTT GAA TTG TCT GAT CTT CAC ATC AAA GAA GCA AAA ATA GAT TAG CTTAAAAAGTCA 1143
 361 L   K   V   E   L   S   D   L   H   I   K   E   A   K   I   D   *              377

1144 TAGTAAATACCACGATAAAAGCACGCGACCGAAAAGGCGATACATATTCTTTAAAGGAGCTTGAAAGGTAATGATTGAAA 1223
1224 TATTAGAGATATTTACAGTATAAAACACAAATATTTAGAACGATAAATAATACAGTAAAAAACAATGCAATACCTAAATC 1303
1304 TATCTACAAAAAACGGTCACCAAGGCGTTGGAAGATAGAAATCTATCAAATACTATTTCTTCTATGTAATAAAAACTCAG 1383
1384 GAACTTGAGAATTCAAAGTGGTTTCTTGTGAACCAGGTATATATGGCATCTTGTCGTCAATAAGCCCTTGCAAAAAG 1463
1464 ATACGACCATAGTCTACACCCACATGATAAATTTCTTCGAACTTGCTAAAGTCTAATGTGGCATATTCTTCAATAGGTGG 1543
1544 TCTCACATAAACGACACCTGGTGTATTTTTCGCTTTCTCTAAAGCATTTACTGATGCGACGTAACCCAATCTAACTTGAA 1623
1624 TCTCTGCCATATTGGGTATGTTTGGATGAGAAGAAAATGGATTCCATCTATTGAAAATGATCCCCGGG              1691
```

FIG. 3. Nucleotide sequence of the *OGG1* gene of *S. cerevisiae*. The nucleotide sequence between *Nsp* I (−347) and *Eco*RI (+1392) restriction sites was determined on both strands. The bipartite nuclear localization sequence is underlined.

codons are separated by 339 bp. Southern blotting analysis confirms that *OGG1* and *PIF1* genes are adjacent (data not shown). Analysis of the predicted amino acid sequence of the Ogg1 protein showed no obvious structural motifs and no significant homology to protein sequences present in the National Biomedical Research Foundation/Protein Identification Resource data base. The Ogg1 protein, however, shows a bipartite nuclear localization sequence at residues 341–362 [KKRK-X_{12}-KQMKL] (Fig. 3). The carboxyl-terminal part of the protein is rich in basic residues as demonstrated by a calculated isoelectric point of 10 for the last 125 amino acids of the Ogg1 protein.

Purification and Substrate Specificity of the Ogg1 Protein. The Ogg1 protein was purified by taking advantage of the overproduction in BH410 (*fpg*) cells hosting the pYSB110 plasmid (Fig. 2). The cleavage of the 8-OxoG/C duplex was used as an activity assay during the Ogg1 purification procedure. Strain BH410 possesses a disrupted chromosomal *fpg* gene, and crude extracts of this mutant have no detectable 8-OxoG/C cleavage activity using our assay conditions. The Ogg1 protein was purified to apparent homogeneity as indicated by a single band on SDS/PAGE with an apparent molecular mass of 43 kDa, which is consistent with the 42.7 kDa predicted from the nucleotide sequence of the *OGG1* gene (data not shown). To elucidate the mechanism of 8-OxoG repair, calf thymus DNA containing 8-OxoG was incubated with purified Ogg1 protein, and the products of the reaction were analyzed using HPLC with electrochemical detection (8).

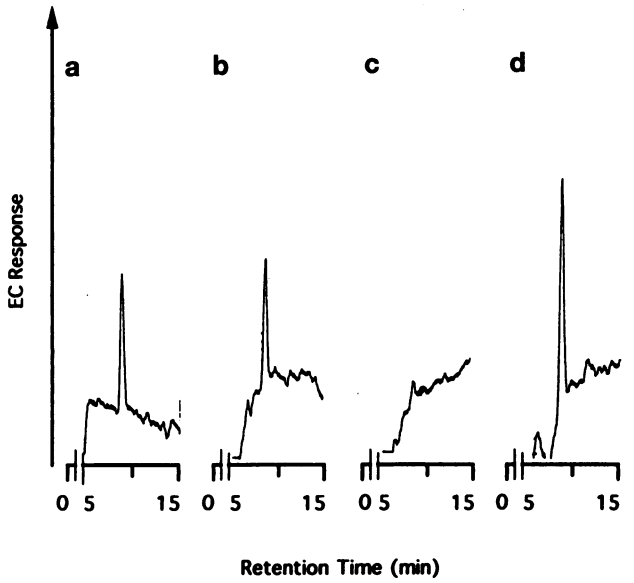

FIG. 4. HPLC with electrochemical detection analysis of products released by the Ogg1 protein from MB-light-treated DNA. Five micrograms of MB-light DNA was incubated with 20 ng of Ogg1 protein for 30 min at 37°C, and aliquots of the reaction products were analyzed. Trace a, standard 8-OxoG (1 pmol); trace b, aliquot of the reaction mixture containing the Ogg1 protein and MB-light DNA; trace c, aliquot of the reaction mixture containing heat-inactivated Ogg1 protein and MB-light DNA; trace d, coinjection of samples used in traces a and b. EC, electrochemical.

Fig. 4 shows that the Ogg1 protein releases a product that coelutes with the 8-OxoG marker molecule. This product was not observed when 8-OxoG-containing DNA was incubated in the presence of heat-inactivated Ogg1 protein (Fig. 4). These results show that the Ogg1 protein acts as a DNA glycosylase that releases 8-OxoG as a free base. The ability of the Ogg1 protein to excise Fapy residues was also analyzed. The results show that Ogg1 protein releases free Fapy residues from [^{3}H]Fapy-poly(dG-dC) substrate. A comparison of the DNA glycosylase activities of the Fpg protein of *E. coli* and of the Ogg1 protein reported in Table 2 shows that Ogg1 protein repairs 8-OxoG 12-fold faster than Fapy.

The capacity of the Ogg1 protein to repair various mismatches was analyzed using 34-mer DNA duplexes that contain an 8-OxoG or a G residue placed opposite each of the four DNA bases. Fig. 5 shows that 8-OxoG/C and 8-OxoG/T duplexes are readily incised by the Ogg1 protein. In contrast, the 8-OxoG/G duplex is slowly incised and the 8-OxoG/A duplex is not incised at a detectable rate (Fig. 5). Control experiments show that the Ogg1 protein does not incise duplexes containing G/A, G/T, or G/G mismatches (Fig. 5). Comparison of the oligonucleotide products generated by the Ogg1 protein of *S. cerevisiae* and the Fpg protein of *E. coli* shows that the two proteins generate the same 3′-oligonucleotide product (PPRD1) but different 5′-

Table 2. Comparison of the 8-OxoG and Fapy DNA glycosylase activities of the Fpg protein of *E. coli* and of the Ogg1 protein of *S. cerevisiae*

Protein	8-OxoG DNA glycosylase, kilounits/mg	Fapy DNA glycosylase, kilounits/mg	Ratio*
Fpg	30.2	29.3	1.03
Ogg1	14.6	1.2	12.20

8-OxoG and Fapy DNA glycosylase activities were measured using MB-light-treated calf thymus DNA and [^{3}H]Fapy-poly(dG-dC) as substrate, respectively. One unit releases 1 pmol of 8-OxoG or Fapy in 5 min at 37°C.

*8-OxoG DNA glycosylase activity/Fapy DNA glycosylase activity.

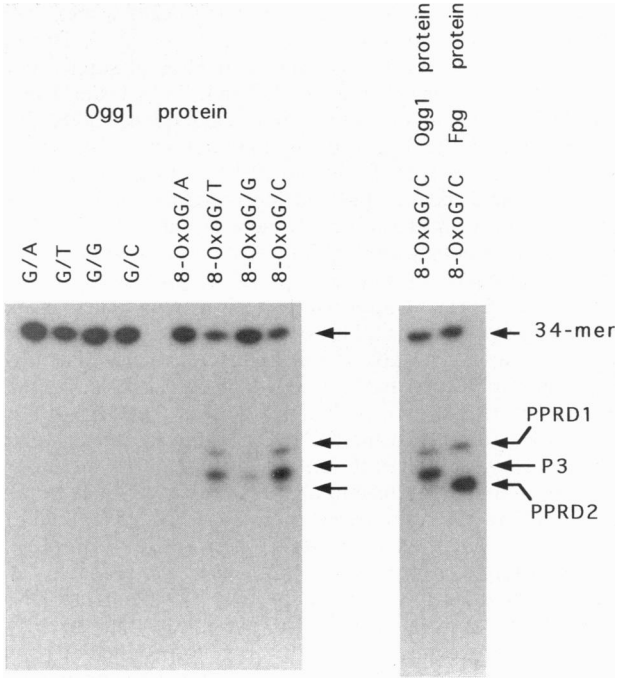

FIG. 5. Cleavage of DNA duplexes containing a single 8-OxoG mismatched with one of the four DNA bases by the Ogg1 protein. The 34-mer oligonucleotide sequences used are reported in *Materials and Methods*. The 8-OxoG or the G_{16}-containing strand was ^{32}P-labeled and annealed with one of the four possible complementary sequences. The ^{32}P-labeled DNA duplexes were incubated with 5 ng of the Ogg1 protein of *S. cerevisiae* or 1 ng of Fpg protein of *E. coli*. (*Left*) Cleavage of the various duplexes by the Ogg1 protein of *S. cerevisiae*. (*Right*) Comparison of the products generated by the Ogg1 protein of *S. cerevisiae* and the Fpg protein of *E. coli*. The analysis of the products was as described in the legend of Fig. 1.

oligonucleotide products (P3 or PPRD2) (Fig. 5). Postincubation of enzyme assays in the presence of an excess of endonuclease III, which nicks DNA at apurinic/apyrimidinic (AP) sites, does not enhance strand cleavage (data not shown). These results strongly suggest that the cleavage of the DNA duplex results from the removal of 8-OxoG and the incision at the AP site; both reactions are catalyzed by the Ogg1 protein. The cleavage pattern at the AP site is consistent with an enzyme cleaving DNA at the 3′ side of an AP site through a β-elimination reaction (29).

DISCUSSION

In the yeast *S. cerevisiae*, several genes involved in base excision repair such as *UNG1, MAG*, and *APN1* have recently been cloned (30–33). Here we report the cloning of the *OGG1* gene of *S. cerevisiae*, which encodes a DNA glycosylase activity that excises 8-OxoG and Fapy residues from DNA. The *OGG1* gene is localized on chromosome XIII immediately upstream and divergently transcribed from the *PIF1* gene, which encodes a protein endowed with a DNA helicase activity (28).

Our work shows that the *OGG1* gene of *S. cerevisiae* codes for a protein of 376 amino acids, the Ogg1 protein, which has been purified to apparent homogeneity. The Ogg1 protein repairs 8-OxoG when paired with pyrimidines. In contrast, the 8-OxoG/G duplex is slowly repaired, and the 8-OxoG/A duplex is not repaired at a detectable rate. The Ogg1 protein is endowed with an activity that nicks DNA at AP sites, probably via a β-elimination mechanism. Taken together, these results suggest that *OGG1* is an eukaryotic functional analogue of the *fpg* gene of *E. coli* (12, 26). However, the enzymatic properties of the two proteins display some differences. Whereas Fpg carries out a β-elimination followed by a δ-elimination, leaving a mono-

nucleoside gap, the Ogg1 reaction seems to yield a nick at the apurinic site as the final product, P3 (Figs. 1 and 5). The Fpg protein of *E. coli* releases 8-OxoG and Fapy at similar rates, whereas the Ogg1 protein releases 8-OxoG 12-fold faster than the Fapy lesion. Comparison of the amino acid sequence of the Ogg1 protein and of the Fpg protein does not reveal an obvious conserved region (12). Both, the PELPEVE sequence at the amino-terminal ends and the zinc finger-structure at the carboxy-terminal ends of bacterial Fpg proteins are absent in the Ogg1 protein, whereas they are highly conserved in bacteria (12).

We previously purified a Fapy DNA glycosylase activity from *S. cerevisiae* with approximately the same molecular mass as the Ogg1 protein (15). However, the substrate specificity of the yeast Fapy DNA glycosylase is different from that of the Ogg1 protein. The yeast Fapy DNA glycosylase cleaves most efficiently the 8-OxoG/G duplex, whereas the Ogg1 protein cleaves most efficiently the 8-OxoG/C duplex. The yeast Fapy DNA glycosylase acts preferentially at Fapy residues, and the Ogg1 protein acts preferentially at 8-OxoG residues. Finally, the yeast Fapy DNA glycosylase activity is strongly inhibited at 100 mM KCl, and the Ogg1 protein is activated at 100 mM KCl. These differences suggest that *S. cerevisiae* possesses two enzyme activities to release Fapy and 8-OxoG from DNA. However, we cannot exclude the possibility that these two activities are two forms of the same gene product. A recent report has suggested that human cells possess two enzymes that repair 8-OxoG (34). The definitive answer to this question awaits the analysis of enzyme activities that repair 8-OxoG and Fapy in Ogg1 mutants of *S. cerevisiae*.

Oxidative stress has been associated with major degenerative syndromes in man including cancer and aging (1, 2). However demonstration of the deleterious action of endogeneous reactive oxygen species in eukaryotes remains unclear. This is mostly due to the lack of mutants that are specifically defective in the repair of oxidized bases in DNA. Therefore, the cloning of the *OGG1* gene of *S. cerevisiae* is an important step toward understanding the biological impact of oxidative stress in eukaryotes.

The authors thank Dr. F. Lacroute for the kind gift of the yeast *S. cerevisiae* genomic library and Dr. P. Nehls and J. P. Radicella for oligonucleotides and bacterial strains. We thank Dr. T. R. O'Connor and J. Laval for discussions and critical reading of the manuscript. The authors gratefully acknowledge support from the Centre National de la Recherche Scientifique, the Association pour la Recherche sur le Cancer, and the European Community (Grant EV5V-CT92-0223). R.d.O. was supported by a fellowship from the Ligue Nationale contre le Cancer.

1. Breimer, L. H. (1990) *Mol. Carcinogenesis* **3,** 188–197.
2. Ames, B. N., Gold, L. S. & Willet, W. C. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 5255–5258.
3. Dizdaroglu, M. (1991) *Free Radical Biol. Med.* **10,** 225–242.
4. Boiteux, S., Gajewski, E., Laval, J. & Dizdaroglu, M. (1992) *Biochemistry* **31,** 106–110.
5. Michaels, L. M. & Miller, J. H. (1992) *J. Bacteriol.* **174,** 6321–6325.
6. Grollman, A. P. & Moriya, M. (1993) *Trends Genet.* **9,** 246–249.
7. Michaels, L. M., Cruz, C., Grollman, A. P. & Miller, J. H. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 7022–7025.
8. Tchou, J., Kasai, H., Shibutani, S., Chung, M. H., Laval, J., Grollman, A. P. & Nishimura, S. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 4690–4696.
9. Boiteux, S. (1993) *Photochem. Photobiol. B* **19,** 87–96.
10. Au, K. G., Clarck, S., Miller, J. H. & Modrich, P. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 8777–8781.
11. Michaels, L. M., Tchou, J., Grollman, A. P. & Miller, J. H. (1992) *Biochemistry* **31,** 10964–10968.
12. Duwat, P., De Oliveira, R., Ehrlich, S. D. & Boiteux, S. (1995) *Microbiology* **141,** 411–417.
13. Tajiri, T., Maki, H. & Sekiguchi, M. (1995) *Mutat. Res.* **336,** 257–267.
14. Gosset, L., Lee, K., Cunningham, R. P., Doetsch, P. W. (1988) *Biochemistry* **27,** 2629–2634.
15. De Oliveira, R., Auffret van der Kemp, P., Thomas, D., Geiger, A., Nehls, P. & Boiteux, S. (1994) *Nucleic Acids Res.* **22,** 3760–3764.
16. Cupples, C. G. & Miller, J. H. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 5345–5349.
17. Boiteux, S. & Huisman, O. (1989) *Mol. Gen. Genet.* **215,** 300–305.
18. Radicella, J. P., Clark, E. A. & Fox, M. S. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 9674–9678.
19. Tudek, B., Laval, J. & Boiteux, S. (1993) *Mol. Gen. Genet.* **236,** 433–439.
20. Bonneaud, N., Ozier-Kalegoropoulos, O., Li, G., Labouesse, M., Minvielle-Sabasta, L. & Lacroute, F. (1991) *Yeast* **7,** 609–615.
21. Stettler, S., Chiannilkulchai, N., Hermann-Le Denmat, S., Lalo, D., Lacroute, F., Santenac, A. & Thuriaux, P. (1993) *Mol. Gen. Genet.* **239,** 169–176.
22. Boiteux, S., O'Connor, T. R., Lederer, F., Gouyette, A. & Laval, J. (1990) *J. Biol. Chem.* **265,** 3916–3922.
23. Geiger, A., Seliger, H. & Nehls, P. (1993) *Nucleosides Nucleotides* **12,** 463–477.
24. Ramalho-Ordigao, J. F., Rosch, H., Selter, H., Frohlich, A., Lorenz, A., Montenach, M. & Seliger, H. (1992) *Antisense Res. Dev.* **2,** 129–146.
25. Castaing, B., Geiger, A., Nehls, P., Seliger, H., Laval, J., Zelwer, C. & Boiteux, S. (1993) *Nucleic Acids Res.* **21,** 2899–2905.
26. Boiteux, S., O'Connor, T. R. & Laval, J. (1987) *EMBO J.* **6,** 3177–3183.
27. Nghiem, Y., Cabrera, M., Cupples, C. G. & Miller, J. H. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 2709–2713.
28. Foury, F. & Lahaye, A. (1987) *EMBO J.* **6,** 1441–1449.
29. Doestsch, P. W. & Cunningham, R. P. (1990) *Mutat. Res.* **236,** 173–201.
30. Percival, K. J., Klein, M. B. & Burgers, P. M. J. (1989) *J. Biol. Chem.* **264,** 2593–2598.
31. Chen, J., Derfler, B., Maskati, A. & Samson, L. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 7961–7965.
32. Berdal, K. G., Bjoras, M., Bjelland, S. & Seeberg, E. (1990) *EMBO J.* **9,** 4563–4568.
33. Popoff, S. C., Spira, A. I., Johnson, A. W. & Demple, B. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 4193–4197.
34. Bessho, T., Tano, K., Kasai, H., Ohtsuka, E. & Nishimura, S. (1993) *J. Biol. Chem.* **268,** 19416–19421.