Molecular cloning and functional expression of a human cDNA encoding the antimutator enzyme 8-hydroxyguanine-DNA glycosylase

TERESA ROLDÁN-ARJONA*, YING-FEI WEI†§, KENNETH C. CARTER§, ARNE KLUNGLAND*, CATHERINE ANSELMINO*, RUI-PING WANG§, MEENA AUGUSTUS§, AND TOMAS LINDAHL*¶

ABSTRACT The major mutagenic base lesion in DNA caused by exposure to reactive oxygen species is 8-hydroxyguanine (8-oxo-7,8-dihydroguanine). In bacteria and Saccharomyces cerevisiae, this damaged base is excised by a DNA glycosylase with an associated lyase activity for chain cleavage. We have cloned, sequenced, and expressed a human cDNA with partial sequence homology to the relevant yeast gene. The encoded 47-kDa human enzyme releases free 8-hydroxyguanine from oxidized DNA and introduces a chain break in a double-stranded oligonucleotide specifically at an 8-hydroxyguanine residue pair bound with cytosine. Expression of the human protein in a DNA repair-deficient E. coli mutM mutY strain partly suppresses its spontaneous mutator phenotype. The gene encoding the human enzyme maps to chromosome 3p25. These results show that human cells have an enzyme that can initiate base excision repair at mutagenic DNA lesions caused by active oxygen.

DNA bases are susceptible to damage in vivo by lipid peroxidation and other endogenous cellular processes that generate reactive oxygen species, as well as by low-level ionizing radiation and oxidative effects of near-ultraviolet (320–380 nm) light. At present there is no direct method to measure the in vivo rate of generation of such endogenous oxidative DNA damage, so its relative importance remains uncertain; recent estimates for mammalian cells have varied from a few hundred to 4 × 10^3 oxidatively altered DNA base residues being introduced per day in each cell (1, 2). Major expected variations between different cell types and subcellular compartments further complicate the issue. The existence of specific DNA repair enzymes that remove oxidized bases from DNA strongly indicates that this form of endogenous damage is physiologically relevant. DNA pyrimidine residues can be oxidized to noncoding ring-saturated and ring-fragmented derivatives (e.g., thymine glycol, cytosine glycol, and N-substituted uracil). A DNA glycosylase with an associated apurinic/apyrimidinic (AP) lyase activity can initiate DNA repair of such lesions, and an Escherichia coli enzyme of this type, endonuclease III, has been extensively characterized (3). A homology search using the cDNA of an E. coli endonuclease III, and the cDNA encoding the enzyme has been cloned and functionally expressed (4). In purines, the imidazole ring is more susceptible to oxidative damage than the pyrimidine ring, so DNA purine residues are converted to the ring-saturated and ring-fragmented derivatives, 8-hydroxyguanine and formamidopyrimidines. A distinct DNA glycosylase is present to remove such damage, and the E. coli enzyme excises both types of lesions at similar rates (5). This formamidopyrimidine–DNA glycosylase (Fpg, MutM) was initially discovered by its action on the cytotoxic ring-fragmented purines (6, 7), but the most important substrate for the enzyme appears to be the mutagenic base 8-hydroxyguanine, which usually occurs in its keto form, 8-oxo-7,8-dihydroguanine (8-oxoG) (8).

In spite of the continuous cellular repair of base derivatives generated by endogenous DNA decay, minute amounts of unrepaired DNA damage (up to 1,000 residues per genome), susceptible to treatment with E. coli Fpg or endonuclease III in vitro, have been reported to be present in the DNA of mammalian cells (9). The 8-oxoG residues generate transversion mutations, because they can base pair with adenine residues in an anti-syn arrangement, so active repair serves to reduce spontaneous mutation frequencies (10, 11). A Saccharomyces cerevisiae functional homolog of the E. coli Fpg gene, OGG1 (8-oxoguanine DNA glycosylase), has been cloned and characterized (12, 13). In the present work, we describe a human homolog of the microbial genes which encodes an enzyme that removes 8-oxoG residues from DNA. It should now be possible to construct or identify mammalian cells deficient in this form of DNA repair, which will allow improved and quantitative determinations of the rate of endogenous DNA base damage by active oxygen in vivo.

MATERIALS AND METHODS

Bacterial Strains and Reagents. E. coli strains used for complementation assays were AB1157 and its derivative, MK611 (mutY mutM::Kat) from M. Sekiguchi (Kyushu University, Japan). E. coli Fpg protein was prepared from the overproducer strain JM105/pFPG230 (5). Plasmid pSE420 was from Invitrogen; the expression strain BL21(DE3), plasmid pET28c(+), and T7-Tag Antibody HRP conjugate were from Novagen. 8-Oxo-2′-deoxyguanosine-CE phosphoramidite was purchased from Glen Research (Sterling, VA).

Sequence Homology Search and Screening for Full-Length cDNA. A homology search using the S. cerevisiae OGG1 amino acid sequence against a database containing ~500,000 human expressed sequence tags (ESTs; Genomic Research and Human Genome Sciences, Rockville, MD) (14, 15) was performed. Candidate ESTs were identified, and approximately 10^6 recombinants of a human cDNA library were screened.

Abbreviations: AP, apurinic/apyrimidinic; EST, human expressed sequence tag; IPTG, isopropyl β-D-thiogalactopyranoside; LB medium, Luria–Bertani medium; 8-oxoG, 8-oxo-7,8-dihydroguanine. Data deposition: The sequence reported in this paper has been deposited in the GenBank database (accession no. Y13277).

†Present address: Departamento de Genética, Facultad de Ciencias, Universidad de Córdoba, San Alberto Magnó s/n, 14071-Córdoba, Spain.

‡T.R.-A. and Y.-F.W. contributed equally to this work.

To whom reprint requests should be addressed. e-mail: lindahl@icrf.icnet.uk.
with a radiolabeled partial cDNA probe by standard procedures (16). Apparently full-length hOGG1 cDNAs were isolated and subcloned as a SalI–NotI restriction fragment into the multiple cloning site of pSPORT (Life Technologies, Gaithersburg, MD) with the 5' end proximal to the T7 promoter. The insert size of positive clones was verified by PCR with gene-specific and vector-specific primers. DNA sequencing was carried out on both strands by using an automated ABI 373 DNA Analysis System (Applied Biosystems). The hOGG1 sequence has been deposited in the GenBank database.

**Protein Overexpression and Purification.** The 1.3-kb open reading frame coding for the hOGG1 protein was amplified by PCR using Pfu DNA polymerase (Stratagene) and the oligonucleotides 5'-TATGGATCCGATGCCTCGCCTCCGCGCCGTTC-3' and 5'-GAGTCTGTCGACGAGACCACTCAACAAGGAGACT-3' to engineer BamHI and XhoI restriction sites at the 5' and 3' ends of the hOGG1 cDNA, respectively. The amplified fragment was inserted into the pET-28c (+) vector by digestion with BamHI and XhoI restriction enzymes and ligation, resulting in the addition of a 6-histidine tag at the amino-terminal end of the hOGG1 protein. The insert was sequenced on both strands, and products of two independent PCRs were subcloned and sequenced.

The plasmid carrying hOGG1 cDNA was used to transform E. coli BL21(DE3) (17) and transformant colonies inoculated into 3 liters of Luria–Bertani (LB) medium containing kanamycin (30 μg/ml). The culture was incubated with shaking, first at 37°C until A600 reached 0.250 and then at 15°C. After 24 h of incubation at 15°C, the cells were collected by centrifugation at 4,000 g for 30 min at 4°C, and resuspended in 20 ml of buffer (HSB: 20 mM Tris-HCl, pH 8.0/0.1 M KCl, 0.5 mM EDTA, 0.2 mg/ml BSA, and 75 mM of 32P-labeled 8-oxoG-containing duplex was incubated at 37°C for 1 h with purified hOgg1 (1 μg). Reactions were stopped by adding an equal volume of 95% formamide/dyes. The products were separated by gel electrophoresis and visualized by autoradiography or analyzed on a PhosphorImager (Molecular Dynamics). Electrochemical detection of free 8-oxoG in dried ethanol supernatants of 100-μl reaction mixture (containing phosphate buffer, pH 7, instead of Hepes, and no DTT) was performed with the Beckman HPLC Gold System using an octadeclisilane ABZ plus column (Supelco) and an EG & G (Salem, MA) electrochemical detector model 400.

**Assay for Mutation Suppression in E. coli.** The 1.3-kb cDNA encoding hOGG1 was subcloned as a SalI–NotI restriction fragment into the multiple cloning site of the pSE420 vector to place it under the control of an isopropyl β-D-thiogalactopyranoside (IPTG)-inducible promoter. E. coli strains AB1157 (wild type) or MK611 (mutM mutY) were transformed with either the vector or the plasmid containing hOGG1 cDNA. A single transformant was inoculated into 10 ml of LB medium containing carbenicillin (50 μg/ml) and incubated at 37°C for 7 h. The culture was diluted in fresh medium and 150–250 cells were inoculated in 5 ml of LB medium containing carbenicillin in the presence or absence of 1 mM IPTG. Five independent cultures were grown for 24 h at 15°C with shaking and then analyzed for generation of rifampicin-resistant (RifR) mutants after plating on LB agar containing carbenicillin (50 μg/ml) and rifampicin (100 μg/ml).

**Chromosomal Mapping.** The hOGG1 cDNA was labeled with digoxigenin-dUTP by nick-translation (Boehringer Mannheim), and fluorescence in situ hybridization was performed according to Johnson et al. (18). Individual chromosomes were counterstained with DAPI (4',6-diamidino-2-phenylindole), and color digital images, containing both DAPI and gene signals, were recorded using a triple-band pass filter set (Chroma Technology, Brattleboro, VT) in combination with a charged coupled-device camera (Photometrics, Tucson, AZ) and variable excitation wavelength filters (19). Images were analyzed by using the ISITE software package (Inovision, Durham, NC).

**RESULTS**

**Human cDNA Encoding 8-oxoG-DNA Glycosylase.** A database search of amino acid sequences conserved between several bacterial formamidopyrimidine-DNA glycosylases did not reveal a similar human sequence. However, when the amino acid sequence of S. cerevisiae OGG1 was used, an EST derived from human frontal cortex was identified that showed partial homology with a region of the yeast OGG1. Other positive ESTs were identified from human keratinocytes, liver, prostate, fetal heart, and the Burkitt lymphoma Raji cell line. Several human cDNA libraries were screened by hybridization with the first EST clone to obtain a full-length human OGG1 (hOGG1) cDNA. Two identical, apparently full-length cDNA sequences were isolated from libraries of a teratocarcinoma cell line and frontal cortex (mainly dendritic cells). The nucleotide sequence of the hOGG1 cDNA revealed a 1,275-bp open reading frame encoding a predicted 47.2-kDa protein of 424 amino acids. The amino acid sequence of hOGG1 showed 33% identity and 54% similarity to S. cerevisiae OGG1, and several highly conserved regions were present (Fig. 1). A notable feature of the hOGG1 sequence is the presence of a putative DNA-binding helix–hairpin–helix motif, previously identified in several different DNA glycosylases (3, 13, 21). Northern blot analysis of several tissues gave very weak signals, and direct enzyme assays of cell extracts also indicated that hOGG1 is expressed at a low level in many cell types (data not shown).

**Overexpression and Purification of hOGG1.** A polyhistidine (His6) affinity tag was attached to the amino terminus of hOGG1 by subcloning into the pET-28c vector. The protein was expressed at a high level following IPTG induction of E. coli BL21(DE3) transformed with the plasmid that carries the hOGG1 cDNA (Fig. 2, lanes 1–4). When the induction was performed at 37°C, 30°C, or 20°C, the protein was insoluble and found in inclusion bodies. However, growing the cells at
15°C for 24 h in the absence of IPTG resulted in low-level expression, with part of the protein remaining soluble. The hexahistidine-tagged hOGG1 protein was purified by Ni²⁺ affinity chromatography. Elution with 500 mM imidazole revealed a major protein band by SDS-PAGE migrating with an apparent molecular mass of approximately 50 kDa (Fig. 2, lanes 5–10).

**Substrate Specificity of the hOGG1 Protein.** *E. coli* Fpg protein and *S. cerevisiae* OGG1 act as DNA glycosylases/AP lyases at 8-oxoG residues base-paired with C, but only very poorly when base-paired with A (10, 12, 13). An 8-oxoG–A base pair is instead a substrate for the distinct MutY DNA glycosylase, which excises the adenine residue (10). The hOGG1 protein was found to cleave an 8-oxoG-containing oligonucleotide with a complementary strand that had a C opposite the oxidized residue, although activity was low (Fig. 3). The observed activity was shown to be directly proportional to enzyme concentration over a 10-fold range (data not shown). When the complementary strand instead had an A opposite the 8-oxoG residue, no enzyme activity was detected (Fig. 3).

In parallel experiments with *E. coli* Fpg protein, there was a 100-fold difference in the rate of enzyme attack on these two substrates (data not shown); a similar reduced activity on the substrate with an 8-oxoG–A pair by hOGG1 would not have been detected here. When hOGG1 activity on the substrate with an 8-oxoG–C base pair was assessed by HPLC with electrochemical detection of 8-oxoG released in free form, the enzyme was 5–10 times more active, indicating incomplete AP lyase activity.

**Fig. 1.** Sequence homology between human 8-oxoG-DNA glycosylase and *S. cerevisiae* OGG1. Amino acid sequences were aligned by the PILEUP and BESTFIT programs in the Genetics Computer Group package (Version 7; ref. 20). Amino acids identical in both sequences are boxed. The position of the possible helix–hairpin–helix motif is indicated. The Lys and Asp residues that have been associated with dual DNA glycosylase/AP lyase activity (13) are shown with an arrow and a circle, respectively.

**Fig. 2.** Purification of recombinant hOGG1 protein. The hOGG1 was overexpressed in *E. coli*, and proteins were visualized on a 10% SDS/polyacrylamide gel by Coomassie blue staining. Lanes 1–4 show overexpression of the protein at 37°C by IPTG induction and contain whole-cell lysate (5 μl) from bacteria harboring the vector before (lane 1) and after induction (lane 2), or the plasmid carrying hOGG1 cDNA before (lane 3) and after induction (lane 4). Lanes 5–10 show peak fractions (10 μl) eluted with 0.5 M imidazole from a Ni²⁺-nitrilotriacetic acid column loaded with a crude extract from cells grown at 15°C (see Materials and Methods). The few protein bands migrating below hOGG1 represent degradation or premature termination products of the overexpressed protein, as determined by immunoblotting with the T7-Tag Antibody (data not shown). Positions of protein size markers in kDa (Bio-Rad) are indicated on the left.

**Fig. 3.** Comparison of hOGG1 activity on double-stranded oligonucleotides containing an 8-oxoG–C or an 8-oxoG–A mismatch. Oligonucleotides were ³²P-labeled at the 5′ end of the 8oxoG-containing strand and were incubated at 37°C for 1 h with purified hOGG1 as described in Materials and Methods. Reaction products were analyzed by autoradiography after electrophoretic separation in a denaturing 20% polyacrylamide gel.
lyase activity to generate chain breaks at abasic sites (data not shown). 8-oxo-G was also released in free form by hOGG1 from plasmid DNA treated with methylene blue and light to introduce 8-oxoG residues. The enzyme was at least 10-fold less active on a DNA substrate with formamidopyrimidine instead of 8-oxoG, whereas *E. coli* Fpg protein was similarly active with both substrates (data not shown). The *S. cerevisiae* OGG1 was observed previously to act 12 times slower at formamidopyrimidines than at 8-oxoG residues (12). Data on the *S. cerevisiae* homolog of endonuclease III indicate that this enzyme, rather than OGG1, may be responsible for excision of oxidatively fragmented purine residues from yeast DNA (22). Our hOGG1 preparation was not detectably active on a double-stranded polynucleotide containing pyrimidines fragmented by KMnO₄ treatment (4), which is the preferred substrate for the human endonuclease III (data not shown). The hOGG1 activity differed markedly from that of *E. coli* Fpg with regard to salt dependence; when KCl was excluded from the standard reaction mixture with an 8-oxoG-containing substrate, hOGG1 was slightly (1.3-fold) stimulated, whereas *E. coli* Fpg activity was reduced 3-fold.

**Suppression of the Mutator Phenotype of *E. coli* mutM mutY by hOGG1.** An *E. coli* mutM(fpg) mutY double mutant showed a high level of spontaneous transversion mutations (10). The hOGG1 cDNA was expressed in *E. coli* MK611 (mutM mutY) and tested for its ability to suppress the mutator phenotype. The hOGG1 cDNA had been subcloned under control of a tightly repressed promoter that was inducible by IPTG. MK611 cells transformed with hOGG1 showed marked but incomplete suppression of the mutator phenotype (Fig. 4). This decrease was not observed in the absence of IPTG, so the antimutator effect was correlated with expression of the human protein. Similar partial complementation of *E. coli* mutM mutY with *S. cerevisiae* OGG1 has been observed previously (12).

**Chromosomal Localization of hOGG1.** To determine the chromosomal location of the hOGG1 gene, single-gene fluorescence in situ hybridization to human chromosome metaphase spreads was performed (23). Approximately seven spreads were analyzed by eye, most of which had a doublet signal characteristic of genuine hybridization on at least one chromosome 3. Doublet signals were not consistently detected on any other chromosome. Detailed analysis of 12 individual chromosomes, using fluorescence banding combined with high-resolution image analysis, indicated that the hOGG1 gene is positioned within band 3p25. The majority of signals seen were positioned within the distal portion of this band, indicating that the gene may fall within subbands 3p25.2–25.3 (Fig. 5). The DNA repair gene *XPC* (xeroderma pigmentosum, complementation group C) also is located in this band (24) but is distinct from hOGG1.

**DISCUSSION**

DNA damage caused by exposure to reactive oxygen species such as hydroxyl radical (OH·), singlet oxygen (¹O₂), superoxide radical (O₂⁻), and hydrogen peroxide (H₂O₂) is believed to be one of the most important reasons for endogenous DNA decay. The widespread occurrence of relevant repair enzymes (i.e., DNA glycosylases specific for oxidized bases) reinforces this view. Here we reported on the cloning of a human cDNA...
encoding an enzyme that acts specifically at mutagenic 8-oxoG residues in DNA. This DNA glycosylase shares some important features with the well characterized E. coli Fpg protein but is more closely related to the recently described S. cerevisiae OGG1 protein (12, 13), so it is called hOGG1. A human cDNA encoding hOGG1, also located to chromosome 3p25, has been independently isolated by J. P. Radicella and S. Boiteux (personal communication). We have shown that the specificity of the hOGG1 enzyme is similar to that of the bacterial and yeast enzymes, cleaving a double-stranded oligonucleotide at an 8-oxoG residue base paired with a C residue, but not detectably at an 8-oxoG base paired with an A (10, 12, 13). By analogy with the known reaction mechanisms of several DNA glycosylases (3–8, 13), the action of hOGG1 apparently involves hydrolytic excision of the free oxidized base, followed by inefficient DNA chain cleavage, presumably by a β-elimination reaction. Moreover, expression of the protein in E. coli partially suppresses the spontaneous mutator phenotype of a mutM(fpg) mutY mutant, which is deficient in repair of endogenously produced 8-oxoG.

Ames et al. (25, 26) have proposed that oxidative DNA damage may be a significant causative factor for cancer and aging in mammals. To evaluate the possible genotoxic effects of endogenously produced active oxygen, it would be desirable to generate knockout mice deficient in DNA repair enzymes that counteract oxidative damage. Cloning of mammalian cDNAs encoding such enzymes is a necessary first step in this direction, and the antimutator factor characterized here should be of major relevance in this regard. Previous attempts to generate knockout mice that are deficient in the later steps of base excision repair have resulted in embryonic lethal phenotypes (reviewed in ref. 27), but the properties of microbial mutants deficient in individual DNA glycosylases suggest that loss of such a function may not be a lethal event in mammals.

The hOGG1 gene maps to the short arm of chromosome 3, within the distal portion of band p25 (Fig. 5). Chromosomal abnormalities in this region, including allele loss and deletions, have been reported for several carcinomas, including lung cancer (28). A search for the possible absence of a functional hOGG1 gene in tumors and tumor cell lines with 3p25 alterations should now be undertaken. Such cells might be expected to exhibit an increased spontaneous mutation frequency and genomic instability, which could contribute to malignancy.

We thank Iain Goldsmith and the oligonucleotide synthesis laboratory at the Imperial Cancer Research Fund for the synthesis of oligonucleotide substrates, and Graham Daly for purified E. coli Fpg protein. The nucleotide sequence of hOGG1 was determined at the sequencing core facility at Human Genome Sciences Inc. The human tissues used for cDNA library construction were provided by the Cooperative Human Tissue Network, funded by the National Cancer Institute. T.R.A. and A.K. were supported by European Community training fellowships under the Human Capital and Mobility Programme, and C.A. had a long-term European Molecular Biology Organization fellowship.