Isolation and characterization of the human uracil DNA glycosylase gene

(DNA repair/cell proliferation/gene regulation/genomic organization)

THOMAS M. VOLLBERG*, KATHERINE M. SIEGLER, BARBARA L. COOL†, AND MICHAEL A. SIROVER‡

Fels Institute for Cancer Research and Molecular Biology and the Department of Pharmacology, Temple University School of Medicine, Philadelphia, PA 19140

Communicated by Sidney Weinhouse, August 17, 1989 (received for review March 31, 1989)

ABSTRACT  A series of anti-human placental uracil DNA glycosylase monoclonal antibodies was used to screen a human placental cDNA library in phage Agt11. Twenty-seven immunopositive plaques were detected and purified. One clone containing a 1.2-kilobase (kb) human cDNA insert was chosen for further study by insertion into pUC8. The resultant recombinant plasmid selected by hybridization a human placental mRNA that encoded a 37-kDa polypeptide. This protein was immunoprecipitated specifically by an anti-human placental uracil DNA glycosylase monoclonal antibody. RNA blot-hybridization (Northern) analysis using placental poly(A)+ RNA or total RNA from four different fibroblast cell strains revealed a single 1.6-kb transcript. Genomic blots using DNA from each cell strain digested with either EcoRI or Pst I revealed a complex pattern of cDNA-hybridizing restriction fragments. The genomic analysis for each enzyme was highly similar in all four human cell strains. In contrast, a single band was observed when genomic analysis was performed with the identical DNA digests with an actin gene probe. During cell proliferation there was an increase in the level of glycosylase mRNA that paralleled the increase in uracil DNA glycosylase enzyme activity. The isolation of the human placental uracil DNA glycosylase gene permits an examination of the structure, organization, and expression of a human DNA repair gene.

Human cells contain two major multienzyme DNA excision repair pathways to remove critical lesions from DNA. The nucleotide excision pathway excises bulky DNA adducts (1–3), while the base excision repair pathway removes alkylated bases and spontaneous DNA damage (1–3). As an initial step in base excision repair, the uracil DNA glycosylase functions to remove uracil residues from DNA by cleaving the base-sugar glycosyl bond. Uracil can be formed in DNA by the mutagenic deamination of cytidine (4, 5) or by the utilization of 5′ dUTP during DNA synthesis (6, 7). Recent evidence showed that DNA repair genes are regulated actively in human cells. This intrinsic gene control includes, at a minimum, the temporal expression of DNA repair genes during the defined pattern of gene expression observed during cell proliferation (8–12) and the selective repair of specific lesions within defined genes (13–16).

To gain insight into the basic molecular mechanisms through which human cells control DNA repair enzymes and pathways, we used recombinant DNA technology to isolate the human uracil DNA glycosylase gene. A series of monoclonal antibodies was prepared against the partially purified human placental uracil DNA glycosylase (17). Subsequent analysis showed that each antibody recognized determinants on the homogeneous placental enzyme (18). One antibody was then used to screen a human placental cDNA library in phage Agt11. We report now the isolation and characterization of the human uracil DNA glycosylase gene. Twenty-seven immunopositive Agt11 colonies were isolated and plaque-purified. One clone that contained a 1.2-kilobase (kb) human cDNA insert was chosen for further analysis. The 1.2-kb human cDNA was isolated and inserted into pUC8. This plasmid (pChug-20.1) hybrid-selected a 1.6-kb mRNA that coded for the 37-kDa uracil DNA glycosylase protein. This mRNA was actively transcribed during cell proliferation. Southern blot analysis demonstrated a complex chromosomal organization of the glycosylase gene. The isolation and characterization of this DNA repair gene provides a mechanism to analyze the inherent organization, structure, and regulation of the human genes that encode DNA repair proteins.

MATERIALS AND METHODS

Identification of the Uracil DNA Glycosylase Gene. Hybrid selection was performed as described by Ploegh et al. (19). DNA was denatured by boiling, dot-blotted to nitrocellulose filters, and fixed by baking at 80° C under vacuum. Prehybridization was performed for 2 hr at 42° C in a buffer that contained 50% denionized formamide, 10 mM Pipes (pH 6.4), 0.2% sodium dodecyl sulfate (SDS), 0.4 M NaCl, 5 μg of yeast tRNA per ml, and 10 μg of poly(dA) per ml. The prehybridization buffer was removed by aspiration. Hybridization was performed with 1 μg of human placental poly(A)+ DNA per ml dissolved in the buffer described above minus the yeast tRNA. The filters were incubated for 4 hr at 42° C and washed 10 times with 10 mM Tris-HCl, pH 7.6/0.15 M NaCl/1 mM Na3EDTA. The filters were then washed twice in that buffer containing 0.5% SDS. The poly(A)+ RNA bound to the filter was removed by boiling and snap-freezing. The RNA was extracted with phenol/chloroform/isoamyl alcohol, 25:24:1 (vol/vol). In vitro translation with rabbit reticulocyte lysates, immunoprecipitation with anti-human placental uracil DNA glycosylase monoclonal antibody 40.10.09, SDS gel electrophoresis, and autoradiography of the immunoprecipitated in vitro translated products were successively performed as described (20).

Transcriptional Expression of the Human Uracil DNA Glycosylase Gene. Total RNA (15 μg) was separated on a 1.5% agarose gel containing 6.6% formaldehyde in an electrophoresis buffer consisting of 20 mM Mops (3-(N-morpholino)propanesulfonic acid), 5 mM sodium acetate, and 1 mM Na3EDTA. The gel was electrophoresed at a constant voltage of 30 V for 18 hr at room temperature. The RNA was transferred to a nylon sheet by capillary action and then

*Present address: Laboratory of Pulmonary Pathobiology, National Institute of Environmental Health Sciences, P.O. Box 12233, Research Triangle Park, NC 27709.
†Present address: Abbott Laboratories, Department 90D, Building L-3, North Chicago, IL 60064.
‡To whom reprint requests should be addressed.

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.
cross-linked to the membrane by a 5-min exposure of short-wave (254 nm) UV transillumination (21). Prehybridization was performed for 2 hr at 55°C in a solution that contained 50% deionized formamide, 0.5% Denhardt’s solution (0.1% Ficoll/0.1% polyvinylpyrrolidone/0.1% bovine serum albumin in water), 5 x SSPE (0.9 M NaCl, 0.05 M sodium phosphate (pH 7.7), and 5 mM Na2EDTA), 0.1% SDS, 0.2 mg of denatured salmon sperm DNA per ml, and 10% dextran sulfate. Heat-denatured pChug-20.1 (5 x 10^6 cpm per lane) was added, and the membrane was hybridized for 12-24 hr. The hybridized membrane was washed under moderate stringency (two 20-min washes in 1 x SSC/1% SDS at room temperature, followed by two 30-min washes in 0.1 x SSC/1% SDS at 60°C). The dried membrane was autoradiographed with Kodak XAR film at 70°C for 1 week with Dupont Cronex Lightning Plus screens.

**Genomic Analysis of the Human Uracil DNA Glycosylase Gene.** Genomic DNA (20 μg) was isolated as described (22) and digested with 5 units of either Pst I or EcoRI per μg overnight at 37°C and then applied to a 1% agarose gel. Electrophoresis was carried out at a constant voltage of 45 V overnight in TBE buffer (89 mM Tris borate/89 mM boric acid/2 mM EDTA). Ethidium bromide was added to the buffer to a final concentration of 0.5 μg/ml. The gel was prepared for transfer as described by Maniatis et al. (23). After transfer to a 0.45-μm nylon membrane (Micron Separations, Westboro, MA) by capillary action with 20 x SSC for 24 hr at room temperature, DNA fragments were cross-linked by a 5-min exposure to a short-wave (254 nm) UV transiluminator. Prehybridization was performed for 2 hr at 65°C in 6 x SSC/5 x Denhardt’s solution/0.5% SDS/0.25 mg of heat-denatured salmon sperm DNA per ml/10% dextran sulfate. Hybridization was performed by the addition of 5 x 10^6 cpm of heat-denatured 32P-labeled pChug-20.1 per lane. The recombinant plasmid was 32P-labeled in a nick-translation reaction (Amersham). Typical probes had a specific activity of 4 x 10^8 cpm/μg of DNA. Hybridization was allowed to proceed for 18-24 hr at 65°C. The nylon membrane was washed under high stringency (twice for 15 min with 2 x SSC at room temperature, followed by two 1-hr washes with 0.2% SSC/1% SDS at 70°C). The air-dried membrane was autoradiographed with Kodak XAR film at -70°C for 1 week with Dupont Cronex Lightning Plus screens. Southern blot analysis using the cardiac actin gene probe (pHRL3-IVS4; American Type Culture Collection) was performed as described (24).

**RESULTS**

**Isolation and Identification of a Human Uracil DNA Glycosylase cDNA Clone.** To verify the antigenic recognition of the uracil DNA glycosylase in unpurified cell preparations, immunoblot analysis of crude extracts was performed with human placental uracil DNA glycosylase monoclonal antibody 40.10.09. This antibody detected only a 37-kDa protein in a purified human placental preparation or in crude cellular extracts from a normal human fibroblast cell strain (Fig. 1). This result was in accord with previous immunoblot analyses with this monoclonal antibody (18, 27-29). The 40.10.09 glycosylase antibody was then used to immunoscreen an oligo(dT)-primed human placental cDNA library in phase Agt11. In the initial analysis of 1 x 10^6 plaques, 40 recombinant plaques were identified as producing an immunoreactive fusion protein. Further immunological screening through five successive plaque isolations resulted in the purification of 27 Agt11 recombinant clones. EcoRI restriction analysis of DNA purified from each clone demonstrated the presence of human cDNA inserts that ranged in size from 0.8 to 1.2 kb. One clone, Agt11 20.1, contained a cDNA insert of 1.2 kb. This insert cross-hybridized to a number of smaller human cDNA inserts in the immunopositive colonies. Accordingly, this insert was cloned into EcoRI-digested, phosphatase-treated pUC8 plasmid and was designated as pChug-20.1.

To verify the identity of the cDNA in the pChug-20.1 clone, hybrid selection was performed with the pChug-20.1 plasmid and with the pUC8 plasmid itself as a negative control. Human placental poly(A)+ RNA was selected by hybridization to the respective plasmid bound to nitrocellulose and was eluted by boiling and snap-freezing. The recovered RNA was precipitated, resolubilized, and translated in vitro. The newly synthesized protein was analyzed by SDS gel electrophoresis (Fig. 2). Hybridization of human placental poly(A)+ RNA with the pChug-20.1 plasmid selected for a human mRNA that encoded a 37-kDa protein (Fig. 2, lane A). Although other radiolabeled proteins were detected, they were also observed when hybrid selection was performed with the pUC8 plasmid by itself as the selecting DNA (Fig. 2, lane B). In addition, the identical pattern was observed in a mock-hybrid selection in which no poly(A)+ RNA was added when the pChug-20.1 plasmid was bound to the nitrocellulose filter (Fig. 2, lane C). Thus, the additional 35S-labeled products represent endogenous products in the lysate mixture. Further, the 37-kDa protein synthesized by the mRNA isolated by hybridization with the pChug-20.1 plasmid was selectively immunoprecipitable by the 40.10.09 anti-human placental uracil DNA glycosylase monoclonal antibody (results not shown). As the pChug-20.1 plasmid selectively hybridized the mRNA encoding the immunoreactive glycosylase protein, these findings showed that the 1.2-kb human cDNA insert in the pChug 20.1 plasmid contained the human uracil DNA glycosylase gene.

**Genomic Organization of the Human Uracil DNA Glycosylase Gene.** To examine the expression of the human uracil DNA gene, RNA blot-hybridization (Northern) analysis was performed with human placental poly(A)+ RNA and total RNA prepared from four separate human fibroblast cell strains collected at confluence (20). Nick-translated pChug-20.1 plasmid was used as the probe. The radiolabeled pChug-
results suggest that the 1.2-kb cDNA insert in the pChug-20.1 plasmid contains at least a significant portion of the human uracil DNA glycosylase gene.

The organization of the human uracil DNA glycosylase gene within the human genome was then examined by Southern blot analysis. Genomic DNA was isolated from each of the four human fibroblast cell strains and digested with either EcoRI or Pst I, and the restriction fragments were separated by agarose gel electrophoresis (Fig. 4). Surprisingly, probing of the Southern blot with the nick-translated pChug-20.1 DNA revealed a complex pattern of DNA hybridization in each digest. However, the hybridizing-restricting fragments produced by each endonuclease were identical for the four human cell strains. Approximately 7 EcoRI and 10 Pst I fragments were detectable in the digest from each cell strain. These data suggest that, at this level of detection, the genomic organization of the human glycosylase gene was identical in the four cell strains.

Four separate controls were performed to demonstrate the validity of the Southern blot-hybridization analysis. First, digestion of each genomic DNA sample was continued for an additional 8 hr with additional restriction enzyme that was added after the initial overnight digestion. An identical hybridization pattern was observed. Second, the Southern blot of the restricted DNA was analyzed with a nick-translated actin probe. Previous results showed that the actin gene was present as a single-copy gene (30). In each of the Pst I- or EcoRI-digested DNA samples, only one hybridizing band was observed. This result indicated the complete digestion of the actin genomic sequences by both of the restriction enzymes. Third, the 1.2-kb human cDNA insert was isolated and digested with Sal I. A 0.5-kb fragment was recovered and used in the Southern blot analysis. No change in the hybridization pattern was observed. Fourth, alteration of the stringency of the Southern blot did not greatly affect the genomic analysis. In particular, although approximately five minor bands were removed, the homology of these repeated se-

Fig. 2. Hybrid selection of uracil DNA glycosylase mRNA. Human placental poly(A)⁺ RNA was prepared as described (20). Hybrid selection was performed as described. Lanes: A, hybrid selection with pChug-20.1; B, hybrid selection with pUC8; C, hybrid selection with pChug-20.1 (no RNA added).

20.1 plasmid hybridized to a single RNA band in each of the five samples (Fig. 3). The size of the RNA transcript was between 1.55 and 1.60 kb, in accord with a previous finding that the human placental uracil DNA glycosylase was synthesized in vitro from a 16S poly(A)⁺ RNA (20). As identical RNA transcripts were observed with either poly(A)⁺ RNA or total RNA, the transcribed gene did not appear to require extensive processing from a precursor RNA. Further, these

Fig. 3. Transcriptional expression of the human uracil DNA glycosylase gene. Human cells were obtained from either the American Type Culture Collection or the Human Genetic Mutant Cell Repository (Camden, NJ). Normal human fibroblasts (CRL-1222; WI-38) or Bloom syndrome fibroblasts (GM-1492; GM-2548) were cultured as described (8-11). Total RNA was isolated by cesium chloride gradient centrifugation after cell lysis by guanidinium isothiocyanate (20). Northern blot analysis was performed as described. Lanes: 1, human placental poly(A)⁺ RNA (5 μg); 2-5, total RNA (15 μg) isolated from CRL-1222, WI-38, GM-1492, and GM-2548 cell strains, respectively.

Fig. 4. Genomic analysis of the human uracil DNA glycosylase gene. Genomic DNA (20 μg) was digested with 5 units of either Pst I or EcoRI per μg overnight at 37°C. Digestion with Pst I was performed in 10 mM Tris-HCl, pH 7.5/100 mM NaCl/10 mM MgCl₂/100 μg of bovine serum albumin per ml. Digestion with EcoRI was carried out in 100 mM Tris-HCl, pH 7.6/50 mM NaCl/10 mM MgCl₂. Lanes: 1-4, EcoRI-digested DNA; 5-8, Pst I-digested DNA; 1 and 5, CRL-1222 DNA; 2 and 6, WI-38 DNA; 3 and 7, GM-1492 DNA; 4 and 8, GM-2548 DNA.
quences with the glycosylase gene probe was quite extensive, since the hybridizing bands were observed even after washing under high-stringency conditions (0.2 x SSC/0.1% SDS at 70°C for 1 hr). These four separate control experiments demonstrated: (i) the complex genomic organization of the glycosylase gene was not an artifact based on incomplete digestion of the genomic DNA; and (ii) the sequences of the human uracil DNA glycosylase gene represented within the pChug-20.1 probe were repeated within the human genome.

**Regulation of the Human Uracil DNA Glycosylase Gene.** To initially consider the regulation of the human uracil DNA glycosylase gene, the transcriptional expression of the gene was examined during cell proliferation. Confluent normal human cells were replated at a lower density to initiate cell growth. The rate of cell proliferation was then examined at three intervals within the growth curve (Fig. 5A). The maximal rates of cell proliferation and DNA synthesis were observed in the 0- to 48-hr interval. Glycosylase regulation was examined by quantitating the increase in enzyme activity and by examining the level of glycosylase mRNA transcription by dot-blot analysis (Fig. 5B).

Glycosylase activity increased 3.0-fold, with maximal levels observed at 48 hr. Enzyme activity then decreased and approached basal levels at 96 hr. This is in accordance with previous studies that showed that cell growth ceased at later intervals (8-12). Similarly, with the nick-translated pChug-20.1 probe there was a 4.5-fold increase in the level of the glycosylase mRNA that paralleled the increase in glycosylase enzyme activity and the induction of DNA synthesis.

To determine whether a new mRNA species was transcribed during cell growth, Northern blot analysis was performed with total RNA collected at the three intervals examined in the growth curve. No detectable radioactivity could be observed in the Northern blot (Fig. 6, lane 1).

**DISCUSSION**

Normal human cells actively regulate DNA repair enzymes and pathways during the defined pattern of gene expression observed during cell proliferation (8-12). In particular, this laboratory has demonstrated a specific temporal relationship between the induction of DNA repair and DNA replication. This conclusion is in accord with other studies that showed that the proliferative-dependent regulation of DNA repair is a general phenomenon common to mammalian cells (12). In this report, we provide evidence to begin to examine this regulation at the molecular level. Using the pChug-20.1 recombinant DNA plasmid in Northern blot analysis, we documented that the uracil DNA glycosylase gene is actively transcribed in normal human cells during cell growth. This result shows that the proliferation-dependent induction of this DNA repair enzyme activity can be directly related to molecular events that control the level of expression of the glycosylase mRNA.

The data presented in this report document that the pChug-20.1 insert contains a significant portion of the human uracil DNA glycosylase gene. This conclusion was based on the hybrid-selection of the glycosylase mRNA and the selective immunoprecipitation of the immunoreactive 37-kDa glyco-
sylase polypeptide. The isolation of the nearly complete gene is in accord with extensive investigations showing that this 37-kDa protein is the only nonmitochondrial uracil DNA glycosylase detected in human cells. Our background studies...
Biochemistry: Vollberg et al.

included: identification of the 37-kDa human placental uracil DNA glycosylase purified to molecular size homogeneity, as defined by SDS/PAGE and Sephadex G-100 gel filtration (18); further verifications of this homogeneity by reactions with glycosylase antibodies 37.04.12, 40.10.09, and 42.08.07, as monitored by ELISA and by enzyme inhibition analysis (18, 32); immunoprecipitation of that protein from [35S]-methionine-labeled human cells (20); and, finally, the immuno-
precipitation of the protein synthesized in vitro by a 16S human placental poly(A)" mRNA (20). On the other hand, another laboratory has proposed that four smaller species of the human placental uracil DNA glycosylase may exist (33). Different N-terminal sequences were postulated for two of those putative species. It is unclear whether those sequences were derived from degradative products of the 37-kDa glycosylase, the mitochondrial enzyme, or an unrelated protein in the glycosylase preparation.

Genomic analysis using the pChug-20.1 plasmid showed multiple hybridizing bands. As the 1.2-kb insert does not contain EcoRI or Pst I sites, several possibilities may initially explain this complex genomic organization. (i) Human cells may contain a uracil DNA glycosylase multigene family in human cells. Basal levels of uracil DNA glycosylase activity were observed consistently in noncyling human cells. One can postulate the existence of two different uracil DNA glycosylase genes. The first might be constitutively transcribed to provide basal levels of glycosylase activity, while the second might be transcribed during cell growth and encode a cell cycle-induced glycosylase. (ii) Instead of a human uracil DNA glycosylase gene family, the multiple hybridizing genomic DNA fragments observed in the Southern blot analysis may represent the presence of a uracil DNA glycosylase pseudogene. In addition, the genomic data may indicate the presence of a repeated intron or introns that contain EcoRI or Pst I restriction sites within a single-copy uracil DNA glycosylase gene. In this regard, the aggregate size of the hybridizing fragments in the Southern blot could compose a single gene > 50 kb. (iii) The uracil DNA glyco-
sylase may be part of a human base-excision repair multigene family. Recent evidence shows that highly purified mammalian DNA glycosylases and apurinic/apyrimidinic acid (AP) endonucleases have molecular sizes of approximately 32-37 kDa (18, 34-38). Thus, the genes that encode these proteins may share common domains that have a high degree of sequence homology. As each enzyme would be of similar molecular weight, it is possible that each gene transcript may also be of similar size. This would be in accord with the detection of a single class of RNA transcripts by the uracil DNA glycosylase clone. However, the Northern blot analysis of the four different human fibroblast total RNA samples does not preclude the presence of multiple RNA transcripts. Present evidence does not permit elimination of any of these three possibilities. However, the isolation and identification of the human uracil DNA glycosylase gene provides the basis for further examination of the organization, structure, and expression of the human DNA repair gene system.

We thank Dr. Sidney Weinhouse and Dr. Sam Sorof for generous counsel. This research project was funded by a grant to M.A.S. from the National Institutes of Health (CA-29414), a grant to the Fels Institute for Cancer Research and Molecular Biology from the National Institutes of Health (CA-12227), and support by the American Cancer Society (SIG-6). K.M.S. is a postdoctoral trainee of the National Institutes of Health (5-T32-CA00214).