A human nuclear uracil DNA glycosylase is the 37-kDa subunit of glyceraldehyde-3-phosphate dehydrogenase

(DNA repair/multifunctional protein/Bloom syndrome)

KATHERINE MEYER-SIEGLER, DAVID J. MAURO, GITA SEAL, JAMES WURZER, JON K. DERIEL, 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, July 3, 1991

ABSTRACT We have isolated and characterized a plasmid (pChug 20.1) that contains the cDNA of a nuclear uracil DNA glycosylase (UDG) gene isolated from normal human placenta. This cDNA directed the synthesis of a fusion protein (M, 66,000) that exhibited UDG activity. The enzymatic activity was specific for a uracil-containing polynucleotide substrate and was inhibited by a glycosylase antibody or a β-galactosidase antibody. Sequence analysis demonstrated an open reading frame that encoded a protein of 335 amino acids of calculated Mr, 36,050 and pl 8.7, corresponding to the Mr, 37,000 and pl 8.1 of purified human placental UDG. No homology was seen between this cDNA and the UDG of herpes simplex virus, Escherichia coli, and yeast; nor was there homology with the putative human mitochondrial UDG cDNA or with a second human nuclear UDG cDNA. Surprisingly, a search of the GenBank data base revealed that the cDNA of UDG was completely homologous with the 37-kDa subunit of human glyceraldehyde-3-phosphate dehydrogenase. Human erythrocyte glyceraldehyde-3-phosphate dehydrogenase was obtained commercially in its tetrameric form. A 37-kDa subunit was isolated from it and shown to possess UDG activity equivalent to that seen for the purified human placental UDG. The multiple functions of this 37-kDa protein as here and previously reported indicate that it possesses a series of activities, depending on its oligomeric state. Accordingly, mutation(s) in the gene of this multifunctional protein may conceivably result in the diverse cellular phenotypes of Bloom syndrome.

Human cells contain two major DNA excision-repair pathways to remove DNA lesions (1). Bulky DNA adducts are eliminated by the nucleotide-excision pathway, whereas most alkylated bases and alterations due to spontaneous damage are removed by the base-excision pathway. DNA glycosylases remove modified bases in the latter pathway by cleaving the base–sugar bond. Uracil present in DNA as a result of utilization of dUTP during DNA synthesis (2) or by deamination of existing cytosine residues (3, 4) is removed by the uracil DNA glycosylase (UDG).

In an examination of the molecular mechanisms involved in expression of human nuclear DNA-repair genes, we isolated a normal human placental cDNA that hybrid-selected the mRNA encoding the nuclear UDG (5). Northern (RNA) blot analysis revealed the presence of a 1.6-kilobase (kb) RNA transcript. In this study we report that the nucleotide sequence of this human glycosylase cDNA* and the deduced amino acid sequence of the glycosylase are identical to those reported for the 37-kDa subunit of the glycolytic enzyme glyceraldehyde-3-phosphate dehydrogenase (G3PD). This finding reveals an unusual and different biochemical function of the monomeric form of G3PD in normal human cells as that of a UDG that functions in the base-excision repair of DNA. A variety of other activities have also been attributed to different oligomeric forms of the 37-kDa subunit of G3PD. Thus, the possibility is considered that mutation(s) in the gene encoding this multifunctional protein may be responsible for various pleiotropic phenotypes associated with Bloom syndrome.

MATERIALS AND METHODS

Purification and Analysis of E. coli Recombinant Proteins. Recombinant plasmid pChug 20.1 was prepared and introduced into Escherichia coli DH1 as described (5). Cells transformed with pUC8 were used as a negative control. Overnight transformed E. coli cultures were grown at 37°C in LB medium supplemented with ampicillin (50 µg/ml), diluted 1:10 in 1 liter of fresh LB medium, and incubated until the absorbance at 600 nm reached 0.6. Cells were pelleted by centrifugation at 3000 × g, resuspended in buffer I [20 mM Tris·HCl, pH 8.0/1 mM dithiothreitol/10% (vol/vol) glycerol], and sonicated three times at 15 s. Debris was removed by centrifugation at 3500 × g. Total protein was precipitated from the supernatant fluid at 75% saturated ammonium sulfate, collected at 5000 × g, resuspended in 1 ml of buffer I, and dialyzed overnight at 0–4°C against buffer I, which contained 100 mM NaCl.

Gel Electrophoresis and Electroelution. Human erythrocyte G3PD (Sigma) and human placental UDG, fraction III (6), were purified by 10% SDS/PAGE (7). A protein in the unstained gel in the region of 37 kDa was excised and eluted by using an LKB model 2014 Extraphor electrophoretic concentrator (Pharmacia LKB) in elution buffer (50 mM Tris, pH 8.9/50 mM glycine/0.1% SDS). Protein was recovered in a high-salt buffer (elution buffer containing 3 M NaCl) by applying 100 V for ~45 min. Each protein was dialyzed overnight against 50 mM Tris, pH 8.0, to remove the NaCl and stored at ~20°C. Each 37-kDa protein was then used in experiments examining UDG activity of G3PD.

Determination of Glycosylase Activity. Two polynucleotide substrates, poly(dA·dT) and poly(dA·dT), were prepared in a DNA polymerase reaction with E. coli DNA polymerase I and [3H]dUTP or [3H]dTMP, as described (8). The release of radioactivity with either polynucleotide as a substrate was examined in the following reaction mixture (total volume 100 µl): 50 mM Tris·HCl, pH 8.0/10 mM K2·EDTA/5 mM dithiothreitol/20 µg of bovine serum albumin/1 µg of either polynucleotide substrate. Limiting amounts of

Abbreviations: UDG, uracil DNA glycosylase; G3PD, glyceraldehyde-3-phosphate dehydrogenase; mAb, monoclonal antibody.

*The sequence reported in this paper has been deposited in the GenBank data base (accession no. X53778).

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.
enzyme were used. Analysis of the fusion protein that eluted in the void volume of the Sephadex G-100 column was done by using 0.1–1 μg of protein. Enzyme activity of the electroeluted human proteins was measured with 0.2–6.0 μg of protein. After incubation at 37°C for 60 min, the reaction was terminated by the sequential addition of 300 μl of EtOH (−20°C), 2 M NaCl (60 μl), and heat-denatured calf thymus DNA (1 mg/ml, 100 μl) at 4°C. After a minimum of 60 min at −20°C, the suspension was centrifuged at 2300 × g for 10 min. An aliquot (200 μl) of the ethanol-soluble supernatant was removed, and its radioactivity was determined.

RESULTS

Identification of a UDG Fusion Protein. To determine whether the pChug 20.1 plasmid encoded an additional glycosylase activity, soluble proteins from E. coli DH1 transformed with either pUC8 or pChug 20.1 were resolved on Sephadex G-100. The proteins from pUC8-transformed cells provided a single peak of UDG activity (Fig. 1A), corresponding to an approximate molecular mass of 30 kDa. This activity was comparable to the protein of M, 25,664 encoded by the E. coli UNG gene (9). However, when the proteins from pChug 20.1-transformed E. coli were chromatographed under identical conditions two peaks of glycosylase activity were detected (Fig. 1B). In addition to the UDG activity at 30 kDa, a second peak of activity was also detected in the void volume. As total 30-kDa glycosylase was undiminished in the pChug 20.1-transformed cells, this additional glycosylase activity at the void volume could not result from formation of a large-molecular mass complex that contained the bacterial glycosylase.

The properties of this additional glycosylase activity found in pChug 20.1-transformed cells were determined. (i) This enzyme displayed a defined substrate specificity (Fig. 2). It released radioactivity from a poly(dA–[^3H]dT) substrate in a manner proportional to the protein concentration in the glycosylase assay. In contrast, no radioactivity was released from a poly(dA[^3H]dU) substrate. (ii) Antibody inhibition studies were done to identify whether this additional glycosylase activity was due to a fusion protein (Fig. 3). Preincubation of the additional activity with conformation-specific glycosylase, monoclonal antibody (mAb) 40.10.09, reduced enzyme catalysis by 20% as compared with that seen in control nonreactive antibody 1.05 (10). Preincubation of the enzyme with a β-galactosidase antibody reduced UDG activity by 70%. Further, as detected by immunoblot analysis, a protein with a relative molecular weight of 66,000 was detected when the Sephadex G-100-purified fractions exhibiting UDG activity were immunoprecipitated with either mAb 40.10.09 or β-galactosidase antibody. Purification of this protein through DEAE-, phospho- and DNA cellulose chromatography showed identical substrate specificity and immunological characteristics to that seen for the activity initially isolated by Sephadex G-100 column chromatography. These results directly demonstrate that the pChug 20.1 plasmid produced a bacterial fusion protein of M, 66,000, recognized by both glycosylase mAb 40.10.09 and the β-galactosidase mAb, and which contained a human UDG activity encoded by the 1.3-kb cDNA.

Analysis of the Human Nuclear UDG cDNA. The restriction map of the 1.3-kb human placental cDNA was then examined. The cDNA had unique restriction sites for both HindIII and Xba I and did not contain any Pst I or EcoRI sites. The restriction map of the cDNA is in accord with our Southern blot analysis of human genomic DNA digested with either EcoRI or Pst I (5). In contrast, a recently described putative mitochondrial human UDG cDNA and a second human nuclear UDG cDNA did not contain Xba I sites and did contain multiple Pst I sites (11, 12). The complete nucleotide sequence of the cDNA was then determined (Fig. 4). The cDNA was 1289 nucleotides in length, which corresponded closely to the size previously calculated by gel electrophoresis of the EcoRI fragment from pChug 20.1. This cDNA contained an open reading frame starting with an initiation codon at position 77 that encoded a protein of 335 amino acids. This ATG codon is within a region (positions 54–60) matching the consensus sequence of a utilized translation initiation site (14). The TAA stop codon is located at position 1062. There is a single polyadenylation signal (AATAAA)
at position 1242 followed by a poly(A)⁺ region 20 nucleotides downstream. The presence of the translational start and stop sites, the polyadenylation signal, and the 25-bp poly(A)⁺ tail confirms isolation of the entire coding sequence.

The calculated molecular mass of the glycosylase at 36,050 Da corresponded to 37 kDa for the purified human placental enzyme isolated by SDS/PAGE and electroelution. The calculated pI of 8.5 agrees with the pI of the purified human placental UDG (G.S. & M.A.S., unpublished work). Comparison of this sequence with all other known UDG sequences demonstrated no significant homology (9, 11, 12, 15, 16). However, a search of the GenBank data base revealed complete (>99%) homology with nucleic acid and amino acid sequences for the 37-kDa subunit of the glycosylate enzyme G3PD (17).

**UDG Activity of the 37-kDa Subunit of Human G3PD.** Human erythrocyte G3PD (Sigma) was obtained in its tetrameric form and resolved by Sephadex G-100 chromatography. As defined by absorbance at 280 nm, a single protein peak was observed eluting in the void volume. No absorbance was detected in any other column fractions. Further, no glycosylase activity was seen in any portion of the column, including the void volume and those fractions corresponding to a protein with a molecular mass equal to 37 kDa. These results demonstrate that the commercially available G3PD did not contain a contaminant UDG.

The 37-kDa monomer of G3PD was isolated by SDS/PAGE. The protein was then reblotted onto nitrocellulose paper, and the UDG-specific mAb 40.10.09 was used to detect immunoreactive proteins. The mAb 40.10.09 detected a single immunoreactive 37-kDa band (Fig. 5, lane 1), corresponding to the monomeric 37-kDa subunit of G3PD, which was the only protein detected by Coomassie blue staining (Fig. 5, lane 2). Accordingly, this marker glycosylase mAb recognizes an antigenic determinant on the G3PD molecule. These results also demonstrate further the purity of the commercially available G3PD.

To determine whether the monomeric form of G3PD possessed UDG activity, the 37-kDa protein was electrophoresed from the SDS/polyacrylamide gel and tested in the standard UDG assay. Noteworthy, this 37-kDa protein released radioactivity from the poly(dA•dT)HdU) substrate proportional to increased protein concentration (Fig. 6). The activity of

**Fig. 3.** Antibody inhibition of UDG activity isolated from pChug 20.1-transformed cells. Void volume fractions containing UDG activity were collected and concentrated by ammonium sulfate precipitation. After dialysis, the protein was concentrated against buffer I (20 mM Tris-HCl, pH 8.0/1 mM dithiorthreitol/10% (vol/vol) glycerol) using a Centricon 10 (Amicon). Glycerol-gradient separations were performed as described (10). Protein (76 μg) was preincubated with equal amounts of either negative control antibody 1.05, uracil glycosylase mAb 40.10.09, or β-galactosidase antibody, as described (10). Antibody–protein complexes were sedimented at 40,000 rpm for 16 hr through a 5-ml 10–35% linear glycerol gradient in a SW50 rotor for 18 hr at 4°C (10), and glycosylase activity was assayed (8). Incubation with control antibody 1.05 (○); UDG mAb 40.10.09 (●), or β-galactosidase antibody (□).

**Fig. 4.** Sequence analyses of human UDG cDNA. The nucleotide sequence of the glycosylase cDNA was determined using three independent strategies: (i) EcoRI–HindIII fragments of the cDNA insert were subcloned into M13mp18 and M13mp19. Single-strand M13 fragments were generated as described (13); (ii) specific synthetic oligonucleotide primers were used to direct sequencing of the positive and negative strands of CsCl gradient-purified pChug 20.1; (iii) Erase-a-Base (Promega) was used to generate pChug 20.1 subclones with progressive deletions. Data sequence were analyzed, and the amino acid sequence was predicted by using MacVector (IBI). Kosak consensus sequence, ATG codon, single polyadenylation signal, and poly(A) region are underlined.
This enzyme was specific for uracil-containing DNA, as no release of radioactivity was seen by using the polynucleotide substrate (Fig. 6). The specific activity of the electrophoresed 37-kDa subunit of G3PD was virtually identical to that seen for the 37-kDa human placental UDG purified first by conventional column chromatography (fraction III, ref. 6), then isolated by SDS/PAGE, and recovered from the gel by electrophoresis.

To identify this additional enzyme activity of the 37-kDa monomer, the radioactivity recovered in the ethanol-soluble supernatant from poly(dA[3H]dU) substrate was analyzed by Sephadex G-10 column chromatography (Fig. 7). Radioactivity released by this protein cochromatographed solely with an authentic uracil marker. These results show that the 37-kDa monomer functions as a UDG. These findings reveal another surprising function of the monomer of G3PD—namely, removal of uracil from DNA.

**DISCUSSION**

These results show that a normal human UDG cDNA encodes the 37-kDa subunit of G3PD. This unexpected conclusion rests on three findings: (i) complete identity between the sequence of the 1.3-kb human nuclear UDG cDNA and the reported sequence for G3PD (17); (ii) isolation of a catalytically active fusion protein that released radioactivity solely from a labeled uracil-containing polynucleotide substrate; and (iii) exhibition of UDG activity by the 37-kDa monomer of human G3PD. Accordingly, these findings indicate the existence of a genetic linkage between a nuclear UDG and G3PD.

G3PD is an abundant cellular protein that is also found in the nucleus. Previous studies have attributed a variety of biological and biochemical activities to the 37-kDa subunit of G3PD that may exist in an equilibrium mixture of tetramers, dimers, and monomers (21). As the tetramer, the G3PD monomer has been indicated to function in glycolysis to convert glyceraldehyde-3-phosphate to 1,3-diphosphoglycerate (22, 23). Other activities claimed include the following: (i) regulation of microtubule bundling/unbundling in vitro (24); (ii) interaction of this protein with the inner red cell membrane and the outer membrane of most other hemopoietic cell types (25, 26); (iii) catalytic formation of the triad junction from transverse tubules and terminal cisternae (27); (iv) protein kinase activity phosphorylating transverse-tubule proteins (28); and (v) a single-stranded DNA-binding activity that appears to regulate transcription similarly to that seen for nonhistone proteins (29, 30).

The degree of oligomerization of the G3PD monomer is influenced by local concentrations of ATP, NAD⁺, and protein (21, 30, 31). G3PD interacts with several different proteins at highly acidic protein domains (27). These affinities may allow it to regulate a number of diverse cellular functions. NAD⁺ strongly inhibits binding of the 37-kDa monomer of G3PD to DNA, suggesting that binding of the protein to DNA occurs at the NAD⁺-binding sites (32). It may be noteworthy that another glycolytic enzyme, M chain lactate dehydrogenase (also called lactate dehydrogenase 5),
was reported to be identical to a helix-destabilizing protein from rat liver (33). NAD$^+$-binding sites are highly conserved in a number of different glycolytic dehydrogenases (34). Further studies seem in order to determine whether certain glycolytic enzymes may have multiple biochemical functions.

Bloom syndrome is an autosomal recessive human genetic disorder characterized, in part, by a high incidence of neoplasia and by multiple cellular alterations (for review, see ref. 35). These include several separate DNA metabolic defects that appear independent of each other (19, 36–43). Although complementation analysis has suggested the presence of a single-genetic-locus for Bloom syndrome (20), attributing these diverse cellular abnormalities to only one gene has been difficult. In view of this study and previous reports describing multiple activities of the 37-kDa protein, mutation(s) in this single gene could have multiple consequences, including possibly those phenotypes characteristic of Bloom syndrome.

The generous counsel of Drs. S. Weinhouse, S. Sorof, L. A. Loeb, B. Singer, and T. M. Vollberg is deeply appreciated. This study was funded by a grant to M.A.S. from the National Institutes of Health (CA-29414), a grant to the Fels Research Institute for Cancer Research and Molecular Biology from the National Institutes of Health (CA-12227) and from the American Cancer Society (SIG-6). K.M.-S. was a postdoctoral trainee of the National Institutes of Health (T-32-CA-09214). D.J.M. was the recipient of a Daniel Swern Memorial Fellowship.