Ugene, a Newly Identified Protein That Is Commonly Overexpressed in Cancer and Binds Uracil DNA Glycosylase

Chunguang Guo,1,4 Xiaodong Zhang,3,4 Stephen P. Fink,2,4 Petra Platzer,2,4 Keith Wilson,7 James K.V. Willson,7 Zhenghe Wang,2,4 and Sanford D. Markowitz2,1,4,5,6

Departments of Molecular Biology and Microbiology,7 Medicine, and 7Genetics; Case Comprehensive Cancer Center, Case Western Reserve University; Case Medical Center; Howard Hughes Medical Institute, Cleveland, Ohio; PDL BioPHARMA, Redwood City, California; and Harold C. Simmons Comprehensive Cancer Center, University of Texas Southwestern Medical Center, Dallas, Texas

Abstract

Expression microarrays identified a novel transcript, designated as Ugene, whose expression is absent in normal colon and colon adenomas, but that is commonly induced in malignant colon cancers. These findings were validated by real-time PCR and Northern blot analysis in an independent panel of colon cancer cases. In addition, Ugene expression was found to be elevated in many other common cancer types, including breast, lung, uterus, and ovary. Immunofluorescence of V5-tagged Ugene revealed it to have a nuclear localization. In a pull-down assay, uracil DNA glycosylase 2 (UNG2), an important enzyme in the base excision repair (BER) pathway, was identified as a partner protein that binds to Ugene. Coimmunoprecipitation and Western blot analysis confirmed the binding between the endogenous Ugene and UNG2 proteins. Using deletion constructs, we find that Ugene binds to the first 25 amino acids of the UNG2 NH2 terminus. We suggest that Ugene induction in cancer may contribute to the cancer phenotype by interacting with the BER pathway.

Introduction

Cancers of the colon and rectum (CRC) are the second leading cause of cancer deaths among adult Americans. Colon cancer develops as a result of the progressive accumulation of genetic and epigenetic alterations in key oncogenes and tumor suppressor genes that lead to the transformation of normal colonic epithelium to adenocarcinoma (1). In addition to primary genetic alterations, the cancer phenotype is also importantly modulated by downstream alterations in the levels of expression of different effector genes (2–4). Expression microarrays enable comprehensive profiling of the cancer transcriptome and identification of cancer-specific changes in gene expression (5, 6).

In this study, using a global gene expression profiling array, we identified a previously uncharacterized gene, called Ugene, which is overexpressed in malignant colon cancers. We further show that Ugene is frequently elevated in most malignant tumor types. In addition, we provide experimental evidence showing that Ugene protein is localized within the nucleus and forms a complex with uracil DNA glycosylase 2 (UNG2), a base excision repair (BER) enzyme.

Materials and Methods

Cell lines and tissues. VACO cell lines were established and maintained as previously described (7). DLD1 and SW480 cell lines were obtained from the American Type Culture Collection. Normal colons, primary colon cancers, and liver metastasis tissues were obtained from the archives of University Hospitals of Cleveland under an institutional review board–approved protocol. Total RNA and genomic DNA were prepared as described (8).

DNA expression microarray analysis. As described previously (9), we designed custom expression monitoring microarrays using Affymetrix GeneChip technology (10). Preparation of samples, hybridization to GeneChip expression microarrays, and data analysis were all performed as described previously (9).

Rapid amplification of cDNA ends PCR. 5′ and 3′ Rapid amplification of cDNA ends (RACE)–ready cDNAs were generated from 2 μg of total RNA (V241 cell line) using the 5′/3′ RACE kit (Roche). The gene-specific primers used for 5′ RACE were as follows: SP1, 5′-CGGGGACCTAGAGC-CTTTTCT-3′; SP2, 5′-GAGGCAAGTGAGTTGAAAG-3′; and SP3, 5′-ATCCCTTCCCAGACATTAAG-3′. The gene-specific primer for 3′ RACE was 5′-ACC-TCTACCTCTCCGAG-3′. Full-length Ugene was PCR amplified from RACE-ready cDNA using the forward 5′-CGGACTGACCTTCAAAGCGAC-3′ and reverse 5′-CTCTGATTCAAACTCTGTC-3′ primers.

Northern blot analysis. Northern blot analysis was performed as previously described (11) using the entire Ugene coding region as the probe.

Southern blot analysis. Total genomic DNA from cell lines and normal tissues were digested with PstI, separated by electrophoresis on a 0.8% agarose gel, and transferred onto a Zeta-Probe blotting membrane. 32P-labeled DNA probes were prepared by random primer extension of a fragment containing the Ugene coding sequence. Equal loading of DNA was confirmed by rehybridizing blots with a probe designated to the TGFβRII gene, which is relatively copy number invariant in CRC.

Human cancer dot blots. Radioactively labeled cDNA probes were synthesized from human Ugene or ubiquitin control cDNA using random primer labeling followed by probe purification on CHROMA SPIN+ STE-100 columns (BD Biosciences). Hybridization of the cancer profiling array with human Ugene probes and washings of the array were done according to the manufacturer’s recommendations (BD Biosciences). The hybridized cancer profiling arrays were then exposed to the phosphorimaging screens and scanned with a Storm 840 PhosphorImager. We then stripped this same membrane and hybridized it with human ubiquitin cDNA probe to show equal sample loading.

Ugene real-time PCR. Primers and a fluorogenic hybridization probe were designed using the Primer3 software (12). Ugene was amplified using 400 nmol/L of forward primer 5′-CTGGTCCTTTCTGTCGGCAAC-3′ and reverse primer 5′-TAGGACGTATACCTGCTGGAG-3′ and detected using fluorogenic hybridization probe 5′-6FAM/ATAAATGCTGTTGTTTGAAC-3′ and reverse primer 5′-TCTCTCCTTCCAATCTTCACTC-3′.

Note: Supplementary data for this article are available at Cancer Research Online (http://cancerres.aacrjournals.org/).

Z. Wang and S.D. Markowitz contributed equally to this work.

Requests for reprints: Sanford D. Markowitz, Howard Hughes Medical Institute Laboratory, Wolstein Research Building, 3rd Floor, Mailstop 7285, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, OH 44106-7285. Phone: 216-368-1976; Fax: 216-368-8928; E-mail: sxm10@cwru.edu.

©2008 American Association for Cancer Research. doi:10.1158/0008-5472.CAN-08-1259

Cancer Res 2008;68(15):6118–26

www.aacrjournals.org

Research Article


6118

Downloaded from cancerres.aacrjournals.org on July 13, 2017. © 2008 American Association for Cancer Research.
followed by 50 cycles of 95°C for 15 s, 57°C for 30 s, and 72°C for 30 s. The level of Ugene expression was determined as the ratio of Ugene: B2M = 2\(^{\Delta\Delta CT}\) (Ugene).

**Construction of expression/deletion vectors.** The coding sequence of Ugene (Ugene-p/Ugene-q, XM_001133365) and UNG2 (NM_080911) was PCR amplified and cloned into the eukaryotic expression vector pcDNA3.1/V5/His-TOPO (Invitrogen) to generate COOH-terminal V5-tagged Ugene/ UNG2 expression vectors. The primer sequences for constructing the vectors are as follows: for Ugene, forward 5'-ACCTCATCCTTCTCGAGC-3' and reverse 5'-TCATACACTCCTGCTGAG-3'; for UNG2, forward 5'-ATGGCGTGCTCTGCTGGTG-3' and reverse 5'-CAGCTCTGTCCAGT-3'. FLAG-tagged constructs were similarly made by adding the complimentary FLAG tag sequence with a stop codon (5'-TTACTTGT- CATTCTGTCCCTTTGTC-3') at the 5' end of the reverse primer. UNG2 deletion constructs were generated by blunt ligation of the PCR products, amplified using V5-tagged UNG2 expression vector as a template, with the

---

**Figure 1.** Ugene mRNA expression in normal and cancer samples. A, expression of U4638 on GeneChip microarrays. Shown for comparison are analyses of RNA samples from normal colon epithelium; colon adenomas; primary colon cancers of stages II, III, and IV; colon cancer hepatic metastases; and colon cancer cell lines. Horizontal bars, median expression values within each group. Average intensity of 25 units corresponds to null expression. B, real-time PCR measurement of Ugene transcript expression in 11 normal colon epithelial samples versus 13 colon cancer cell lines. Ugene values are normalized against expression of the housekeeping gene, B2M. Black bars, mean value for each group. C, the ratio of Ugene expression in colon cancer versus matched normal colon mucosa, as measured by real-time PCR in 20 patients. Ugene values are again normalized against B2M. Values represent averages of six replicates. D, cDNA dot blot analysis of Ugene expression level in matched tumor and normal tissue blots from the various organs. N, normal; T, tumor. Numbers below each blot represent the number of tumors with >2-fold increase in Ugene expression relative to their matched normal [Ugene (†)] versus the number of total samples analyzed (total #). Each L-shaped box of three dots represents normal tissue (left), primary tumor (top right), and metastases (bottom right) from the same patient. Equal sample loadings over all lanes were confirmed by rehybridizing the blot to a ubiquitin probe (data not shown).
primer sets as listed in Supplementary Table SI. The decoy fusion protein 1-25-UNG2-green fluorescence protein (GFP) expression vector was constructed by ligating the following three fragments: (a) GFP DNA [PCR amplified from pEGFP-N1 template (Clontech) using forward 5'-TTGAAATT-CTCATGGTAGAAGGGGAGGAGG-3' and reverse 5'-TTCTCAGCCGCTCCTG-3'] and reverse 5'-TTCTCAGCCGCTCCTG-3' and reverse 5'-TTGAAATT-CTCATGGTAGAAGGGGAGGAGG-3'; and (b) nucleotides corresponding to 1 to 25 amino acids of UNG2 (amplified from pEGFP-N1 template (Clontech) using forward 5'-TTGAAATT-CTCATGGTAGAAGGGGAGGAGG-3'). The Ugene nucleotide sequence is provided in the upper reading frame and numbered in roman type. The deduced amino acid sequence is provided underneath the nucleotide sequence, numbered in italic. The in-frame stop codon (TGA) of the start codon is indicated in boldface. Underlined letters represent codons that differ in Ugene-q.

**Figure 2.** Structure of the Ugene locus. A, numbered black boxes denote the four Ugene exons, with locations of initiator ATG and termination TAA designated. Arrowheads, sites that differentiate Ugene-p and Ugene-q. B, nucleotide and deduced amino acid (aa) sequence of complete Ugene (Ugene-p) coding region. The Ugene nucleotide sequence is provided in the upper reading frame and numbered in roman type. The deduced amino acid sequence is provided underneath the nucleotide sequence, numbered in italic. The in-frame stop codon (TGA) of the start codon is indicated in boldface. Underlined letters represent codons that differ in Ugene-q.

Endogenous Ugene epitope tagging. 3' FLAG tagging of the endogenous Ugene-p gene was performed by somatic cell knock-in vectors, as described by Zhang and colleagues (14). The primers were as follows: left arm 5'-GGCTCGAGGAGCATGAGAT-3', 5'-CCGAAATTCTGAGTGCATCCTTCTGAGAT-3'; right arm 5'-GGACTAGTATGCTTCTGAGAT-3'. Cells expressing only UNG2 were generated by UNG1-specific knockout of one allele combined with the UNG locus exon 2 using left targeting arms amplified with the following primers: left arm, 5'-GGCTCGAGGAGCATGAGAT-3', 5'-CCGAAATTCTGAGTGCATCCTTCTGAGAT-3'; right arm 5'-GGACTAGTATGCTTCTGAGAT-3'. Cells expressing only UNG2 were generated by UNG1-specific knockout of one allele combined with the UNG locus exon 2 using left targeting arms amplified with the following primers: left arm 5'-GGCTCGAGGAGCATGAGAT-3', 5'-CCGAAATTCTGAGTGCATCCTTCTGAGAT-3'; right arm 5'-GGACTAGTATGCTTCTGAGAT-3'.

**Somatic cell knockout.** Somatic cell knockout was performed as described (13). Knockout of both UNG1 and UNG2 transcription units was accomplished by disrupting the UNG locus exon 2 using left and right targeting arms amplified with the following primers: left arm, 5'-GGCTCGAGGAGCATGAGAT-3', 5'-CCGAAATTCTGAGTGCATCCTTCTGAGAT-3'; right arm 5'-GGACTAGTATGCTTCTGAGAT-3'. Immunofluorescence. SW-480 cells were seeded at 1.0 × 10^4/100-mm dish and transplanted the next day with 2 μg of V5-tagged Ugene expression vector using 12 μL of Fugene 6 (Roche Applied Sciences) as per the manufacturer's protocols. Immunofluorescence was performed 48 h after transfection using V5 antibody (Invitrogen) at 1:200, followed by Alexa Fluor 488 goat anti-mouse IgG antibody (Invitrogen) at 1:400.

Immunoprecipitation and Western blot analysis. HEK 293T cells were seeded at 4.0 × 10^4/75 flask and transfected the next day with a total of 6 μg plasmids using 36 μL of Lipofectamine 2000 (Invitrogen). Cell lysates were prepared 48 h after transfection using the lysis buffer [50 mmol/L Tris-HCl (pH 7.5), 1 mM EDTA (pH 8.0), 150 mmol/L NaCl, 1% Triton X-100] supplemented with the protease inhibitor mixture (Roche Applied Sciences). FLAG immunoprecipitation was performed with anti-FLAG M2 affinity gel as described in the manufacturer's protocol. After elution by either FLAG peptide (Sigma) or 3′× FLAG peptide (Sigma), eluates were used for biochemical activity assay or for Western blot. Western blots were prepared using anti-FLAG M2 (Sigma, 1:1,000, anti-V5 (1:1,000), anti-UNG (1:500, Abcam), anti-UNG (1:5,000, Invitrogen), or anti-UNG (1:100,000, Invitrogen) antibodies, followed by horseradish peroxidase–conjugated donkey anti-mouse secondary antibody (1:1,500; Jackson ImmunoResearch Laboratories) and visualized by using an Enhanced Chemiluminescence Plus detection kit (Amersham Biosciences).
Small interfering RNA–mediated Ugene silencing. The Ugene-specific and control small interfering RNAs (siRNA) were synthesized by Dharmacon. For siRNA transfection experiments, DLD1 cells were seeded on 100-mm culture dishes and transfected with 30 µL of 20 µmol/L siRNA stock using 30 µL LipofectAMINE-2000 (Invitrogen). Cell lysates were collected 48 h after transfection and knockdown was validated by Western blot. The sequences of Ugene siRNA-1017 were as follows: sense 5′-GGAGAUGGCUAUUUCACCAUU-3′, antisense 5′-UGGUGAAAGUCAUU CUUCCUU-3′.

In vitro UNG biochemical activity assay. The uracil-containing oligonucleotide (5′-CTTCGCTTGGACGCTGTTGG-3′; RoD Systems) was annealed to an equimolar amount of its complementary strand (5′-CCACAGGCTCAGGGGAA-3′ or 5′-CCACAGGCTGCCAGGGCAGG-3′) for U-A and U-G pairs, respectively, mixed and heated to 95°C in annealing buffer [20 mmol/L Tris-HCl (pH 8.0), 1 mmol/L EDTA, 1 mmol/L DTT, and 0.1 mg/mL bovine serum albumin], and allowed to slowly cool to room temperature. The DNA was then end labeled with γ32P-dATP by T4 polynucleotide kinase. In vitro UNG biochemical activity assay was performed as per manufacturer's instruction (RoD Systems). In brief, after exposure to UNG2, which deglycosylates uracil, the deglycosylated oligonucleotide was split in half by incubation in alkaline buffer (300 mmol/L NaOH, 97% formamide) at 100°C for 10 min.

Results

Identification of a novel gene, Ugene, overexpressed in most cancer types. To identify genes potentially involved in colon tumorigenesis, we used GeneChip gene expression microarrays to compare patterns of gene expression in RNA samples extracted from colon cancers versus normal colon epithelium. One transcript, an expressed sequence tag (EST) corresponding to Genbank accession number NT_086586 and that mapped to within 20 kb of U46258. Connection to this EST, we identified by a BLAST search all ESTs in the dbEST database that exist, likely all located on chromosome 1. Because Ugene-p is conserved across all mammalian species, we focused on Ugene-p in this study.

As shown in Supplementary Fig. S1, Northern blot analysis corroborated that Ugene transcripts are expressed by malignant but not normal colon tissues, detecting a single 2.4-kb Ugene mRNA with moderate to strong intensity in 10 of 11 colon cancer cell lines, but in none of four normal colon epithelial tissue samples. To provide a more quantitative measurement of Ugene induction, we extended this analysis by using real-time PCR. Real-time PCR showed only barely detectable Ugene expression in 11 of 11 normal colon epithelial samples (mean value 2.6, range 1.1–4.7), whereas colon cancer cell lines showed an average of 56-fold increased level of expression (mean value 147, range 3.3–503), with 11 of 13 colon cancer cell lines showing a >15-fold increase in expression (Fig. 1B). To determine expression of Ugene-p versus Ugene-q–encoded transcripts, we sequenced individual Ugene cDNA clones from two colon cancer cell lines, SW480 and VACO241. We found that in both cell lines, Ugene-p represents 60% of Ugene expression, whereas Ugene-q represents 40% (data not shown).

Overexpression of Ugene in malignant colon cancer was also confirmed by real-time PCR analysis of Ugene mRNA in primary colon cancers versus matched normal colon mucosa from the same individuals. A median increase of 6.8-fold in Ugene expression was observed in cancers versus matched colon normals, with >2-fold increase exhibited by 18 of the 20 tumors (Fig. 1C). These 20 colon cancers examined by real-time PCR constituted a “validation set” of samples completely independent of those that had been previously characterized on the GeneChip expression microarrays.

Southern blotting of 12 colon cancer cell lines did not show any increase in Ugene gene copy number as an explanation for Ugene overexpression (data not shown).

To examine whether Ugene might be overexpressed in other cancer types, we probed a cDNA Cancer Profiling Array (BD Biosciences), comparing Ugene expression level in matched tumor and normal tissue from a variety of organs. As expected, a high proportion of colon cancer samples were observed to have elevated expression of Ugene (Fig. 1D). Using densitometry, the intensity of the radioactive probe signal from each cDNA sample was quantitated. Twenty-two of 34 colon cancer cases (65%) showed >2-fold increased expression of Ugene. Furthermore, we also found Ugene expression elevated in multiple other common cancer types,
including breast (56% of cases), lung (52% of cases), stomach (64% of cases), uterus (67% of cases), and ovary (79% of cases).

**Ugene encodes a nuclear protein.** To investigate the subcellular localization of Ugene-encoded protein, a construct expressing V5 epitope–tagged Ugene-p protein was transfected into SW480 cells. Figure 3 shows the immunofluorescent staining for the V5 tag (green) in Ugene-transfected cells. Results show that tagged Ugene protein accumulates in nuclei, which were defined by 4',6-diamidino-2-phenylindole (DAPI) staining (red). As Ugene is a small protein (16.9 kDa) and lacks a nuclear localization signal, this accumulation suggested that Ugene might be held in the nucleus by interacting with other nuclear proteins.

**Ugene binds to UNG2, a base excision DNA repair protein.** To look for a potential Ugene partner, we performed a pull-down assay using FLAG-tagged Ugene-p protein overexpressed in SW480 cells. Using mass spectrometry, a protein band that precipitated in the Ugene-p pull down was identified as UNG2 (data not shown), an enzyme that is involved in BER by catalysis of uracil excision from DNA and is constitutively located in the nucleus (15, 16).

To confirm the interaction of Ugene-p and UNG2, we first cotransfected tagged Ugene-p and UNG2 constructs, followed by immunoprecipitation of either protein, and then performed Western blot analysis of the immunoprecipitates to detect the presence of potential partners. We found Ugene-p and UNG2 coimmunoprecipitated together in assays in which either of the proteins was first pulled down (Fig. 4A).

To prove that this binding of Ugene-p and UNG2 was not an artificial result due to protein overexpression, we tagged the endogenous Ugene-p locus in DLD1 cells by somatic cell knock-in of a 3× FLAG epitope at exon 4, corresponding to the COOH terminus of the protein (Fig. 4B; ref. 14). We then coimmunoprecipitated Ugene-p with antibodies against the FLAG epitope. Western blot analysis confirmed that endogenous UNG2 coimmunoprecipitated with the tagged endogenous Ugene-p protein (Fig. 4C). PCR analysis showed no change in total Ugene transcript expression in cells bearing the 3× FLAG knock-in epitope and also showed the knock-in Ugene allele to be expressed at essentially the same level as the nontargeted alleles (Supplementary Fig. S4).

**Ugene-p binds to the NH2 terminus of UNG2.** To determine the UNG2 motif responsible for binding to Ugene-p, we made a series of constructs expressing V5 epitope–tagged nested UNG2 deletions (Fig. 5A). After cotransfecting each of these V5 epitope–tagged UNG2 deletion constructs with FLAG epitope–tagged Ugene-p, we immunoprecipitated Ugene-p and performed Western blot analysis to test for coimmunoprecipitation of each of the UNG2 deletion constructs (Fig. 5B). One UNG2 deletion that lacked only sequences between codons 3 and 33 showed complete loss of the capacity to bind to Ugene-p. This result suggested that Ugene-p binds to the NH2 terminus of UNG2.

To further test if the UNG2 NH2 terminus is able to bind to Ugene-p, we artificially synthesized peptides encoding the NH2-terminal amino acids 1 to 25 of UNG2 (1-25-UNG2). We then tested if this peptide could competitively block the binding of endogenous UNG2 to FLAG epitope–tagged endogenous Ugene-p. Figure 5C shows that adding the 1-25-UNG2 peptide into cell lysates blocked UNG2 binding to Ugene-p in a dose-dependent fashion. 1-25-UNG2 (5 μmol/L) competing peptide could compete out almost all Ugene-p binding to UNG2.
To test if the NH2-terminal 1 to 25 amino acids of UNG2 is sufficient for binding to Ugene-p, we expressed a fusion protein with 1-25-UNG2 fused to GFP (1-25-UNG2-GFP) under the regulatory control of doxycycline. We performed this in cells already containing the 3× FLAG epitope–tagged endogenous Ugene-p. Serial immunoprecipitation and Western blot analysis confirmed that Ugene-p bound to the (1-25-UNG2-GFP) protein (Fig. 5D). Indeed, induction of the 1-25-UNG2-GFP decoy protein could completely outcompete and block coimmunoprecipitation of endogenous UNG2 with endogenous Ugene-p (Fig. 5D). Therefore, the NH2-terminal 1 to 25 amino acids of UNG2 are sufficient in vivo for the interaction with Ugene-p.

Interestingly, despite having only two amino acid differences, Ugene-q was found not to interact with UNG2 and did not coimmunoprecipitate with it (data not shown). Introducing a single Ugene-q–specific codon change, changing tryptophan-125 to arginine was also sufficient to abolish Ugene-p binding to UNG2 (data not shown). Therefore, Trp125 of Ugene-p is required for binding of Ugene-p to UNG2.

Ugene binding does not directly alter UNG2 enzymatic activity or localization. To examine potential functional effects of Ugene-p binding to UNG2, we performed a coimmunoprecipitation to collect UNG2 bound to Ugene-p (pulled down by antibodies against the FLAG epitope). A biochemical assay showed that UNG2 bound to Ugene-p was an active enzyme, as indicated by initiating a cascade causing cleavage of a uracil-containing oligonucleotide from the parental 21-nucleotide (nt) size down to 10 nt (Fig. 6A, lane 2). To ensure the activity in the Ugene-p (FLAG) immunoprecipitates derived from captured UNG2, we repeated the assay in DLD1 cells rendered UNG null by somatic cell knockout (as described in Materials and Methods; Supplementary Fig. S3). No activity was detected in Ugene-p immunoprecipitates from UNG null cells. Thus, we conclude that the biochemical activity detected in Ugene-p precipitates from parental DLD1 cells derives from active UNG2 bound to Ugene-p.

To examine whether binding to Ugene-p can alter UNG2 subcellular localization, we expressed V5 epitope–tagged UNG2 (UNG2-V5) in the cells conditionally expressing the 1-25-UNG2-GFP fusion protein under doxycycline regulation. Immunofluorescence against the V5 epitope showed that UNG2 was localized in the nucleus irrespective of expression of the 1-25-UNG2-GFP decoy protein (Supplementary Fig. S2). Therefore, expressing a competitor for Ugene-p binding did not alter UNG2 nuclear localization.

The UNG locus encodes both a nuclear protein UNG2 and a mitochondrial isoform UNG1 that both share the same catalytic domain but are of different sizes (17). In repeated assays, only a UNG2-sized protein was ever detected in Ugene-p immunoprecipitates (data not shown).

To further assay the effects of Ugene-p expression on UNG2 activity, we generated cells null for UNG1. This was done by...
selective knockout of the UNG1-specific exon 1 from the UNG locus. In these cells expressing UNG2 only, we again introduced the 1-25-UNG2-GFP decoy protein under doxycycline regulation. These cells were used to determine UNG2 enzymatic activity under two experimental conditions (Fig. 6C). First, we compared UNG2 activity in cell lysates without (dox−) and with (dox+) induced expression of the 1-25-UNG2-GFP decoy protein (top). As shown in Fig. 5D, the highly expressed decoy protein totally abolished the interaction of Ugene-p and UNG2, but did not alter UNG2 biochemical activity in the lysates, as shown in Fig. 6C (top). Specifically, an equal signal intensity of the 10-nt cleavage product of the uracil-containing oligonucleotides was seen in both dox− and dox+ conditions. Second, we compared UNG2 activity in lysates prepared from cells without and with suppression of Ugene expression by siRNA. As shown in Fig. 6B, Ugene-specific siRNA (siRNA1017) could efficiently suppress Ugene expression by >90% at 48 h after transfection. However, Ugene knockdown did not change the enzymatic activity of UNG2 as shown in Fig. 6C (bottom). These findings were equally true whether UNG2 activity was analyzed with a 21-bp oligonucleotide containing a U-A or a U-G mispair at position 10, which, respectively, modeled uracil misincorporation into DNA and uracil arising from spontaneous deamination of cytosine. These results suggest that under the experimental condition used, changing Ugene-p expression did not alter UNG2 biochemical activity.

Discussion

We report here the identification of a novel gene, Ugene-p, the expression of which is broadly induced in many human cancer

---

**Figure 5.** Mapping of the UNG2 domain for binding to Ugene-p. A, schematic diagram of constructs expressing nested UNG2 deletions. B, HEK 293T cells were transfected with FLAG-tagged Ugene-p and a cDNA encoding either V5-tagged wild-type UNG2 or the indicated UNG2 deletion mutants. At 48 h after transfection, immunoprecipitates prepared by immunoprecipitation with FLAG antibodies were analyzed by Western blot (IB) for UNG2 (V5) and Ugene-p (FLAG). Box, deletion of UNG2 amino acid 3-33 abolished the binding to Ugene-p. Bottom, expression of V5-tagged wild-type UNG2 and UNG2 deletion proteins in whole-cell lysates. C, cell lysates of DLD1 expressing endogenous 3-33 FLAG-tagged Ugene-p were mixed with artificially synthesized competing peptides (1-25-UNG2, amounts as indicated), then immunoprecipitated with FLAG antibodies. Immunoprecipitates were analyzed by Western blot (IB) for the presence of Ugene-p (FLAG) and for coimmunoprecipitation of UNG2 (V5). D, DLD1 cells expressing endogenous 3-33 FLAG-tagged Ugene-p were transfected with pcDNA6/TR and pcDNA4-1-25-UNG2-GFP and selected by blasticidin (10 μg/mL) and zeocin (200 μg/mL) to derive clones conditionally expressing 1-25-UNG2-GFP fusion protein under doxycycline (dox) regulation. These clones are designated as DLD/Ugene-p-3-33-FLAG/1-25-UNG2-GFP. The interaction of Ugene-p with either UNG2 or the 1-25-UNG2-GFP decoy protein was assayed by immunoprecipitation for Ugene-p with the FLAG antibody, followed by Western blot detection of Ugene-p, GFP, and UNG2, in cells without (dox−) and with (dox+) induced expression of the 1-25-UNG2-GFP decoy protein.
Figure 6. Assays of UNG enzymatic activity. A, wild-type or UNG null DLD1 cells were transfected with a plasmid expressing FLAG-tagged Ugene-p or the corresponding empty vector. Cell lysates were then immunoprecipitated with FLAG antibodies. Immunoprecipitates were subjected to a UNG biochemical activity assay, as indicated by the presence of a 10-nt product (arrow) generated by cleavage of a 21-nt input double-stranded DNA. Input oligonucleotides contain a single U-A base pair. Bottom, Western blot of UNG2 in Ugene-p immunoprecipitates. B, DLD1 cells expressing endogenous 3′ -FLAG-tagged Ugene-p were transfected with Ugene-specific siRNA (siRNA-1017) or control siRNA (siGLO). Ugene-p expression in cell lysates was analyzed by Western blot against the FLAG epitope. C, using an AAV-mediated somatic knockout technique, cells expressing UNG2 only were constructed in the DLD1/Ugene-p-3′-FLAG/1-25-UNG2-GFP background. The top two panels show the UNG2 activity in the cell lysates (amount as indicated) from cells without (-dox) and with (+dox) induced expression of the 1-25-UNG2-GFP decoy protein. The bottom two panels compare UNG2 activity in lysates (amount as indicated) from cells without (siGLO) and with (siRNA-1017) the suppression of Ugene expression by siRNA. In the left two panels, the input oligonucleotides for the assay contain a single U-A base pair; for the right two panels, the input oligonucleotides contain a single U-G base pair. UNG2 activity is indicated by autoradiography of the 10-nt oligonucleotides.

Types. Moreover, we show the interaction of Ugene-p with UNG2, a BER enzyme. The frequent overexpression of Ugene-p in cancer suggests that this gene may participate in the cancer phenotype. A direct assay to test Ugene-p for oncogenic activity, however, revealed no transforming or focus-forming activity when tested in epithelial (HMEC) or fibroblast (WI-38, IMR-90, REF-52, NIH 3T3) cells (data not shown).

The interaction of Ugene-p with UNG2 in particular, is highly intriguing, as multiple DNA repair pathways are now recognized as targets for alteration in cancers, including inactivation of genes in the mismatch repair pathway in colon cancers (1) and inactivation of the BRCA1/2 proteins in breast cancers (18). Despite this intriguing association, we have not yet been able to show a direct regulation of UNG2 repair activity by Ugene-p in vitro. It is likely, however, that the in vivo activity of UNG2 is more complicated than we have been able to model in in vitro assays, as UNG2 in vivo activity involves recognition of misincorporated uracil at the replication fork and involves recognition of uracils that are spontaneously generated through cytosine deamination in native chromatin, in addition to involving interactions with other members of the BER complex. The fact that immunoprecipitation of overexpressed FLAG-tagged Ugene-p pulled down only a subpopulation of total UNG2 protein (data not shown) suggests that Ugene-p could promote a specialized function of UNG2. Of note, the NH2 terminus of UNG2, to which Ugene-p binds, has been shown to also bind to the PPM1D phosphatase that dephosphorylates Thr-6, effecting a protein modification that is suggested to play an important role in the regulation of UNG2 activity under some circumstances (19). Further analysis of the effect of Ugene-p on UNG2 in these native contexts will be undertaken in future studies.

Genome comparisons show that Ugene-p arose as a feature of mammalian cells, in which it is highly conserved, suggesting an important role for the protein in higher organisms. It is, however, unclear from the current genome assemblies whether there are two copies of the Ugene-p on chromosome 1 or whether the chromosome 1 assembly remains in need of revision. In contrast, Ugene-q, which is unable to bind UNG2, is specific to humans and absent in other mammals. It is tempting to speculate that Ugene-q may act as a competitor of Ugene-p interactions with other proteins, but testing of this model awaits additional clarification of the functional activities of Ugene-p. Additional future study will also be needed to clarify the role of a second protein, NOSIP, that in our initial pull-down experiment was also identified as binding to Ugene-p, but that on further evaluation proved not to coprecipitate with UNG2 (data not shown).

In summary, we report Ugene-p as a novel gene commonly overexpressed in human cancers and participating in a nuclear complex with the BER gene, UNG2.

Disclosure of Potential Conflicts of Interest

K. Wilson: ownership interest, PDL BioPharma. The other authors disclosed no potential conflicts of interest.

Acknowledgments

Received 4/3/2008; revised 5/27/2008; accepted 6/1/2008.

Grant support: Howard Hughes Medical Institute (S.D. Markowitz) and USPHS grant CA116867 and CA120237 (S.D. Markowitz).

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

We thank James D. Lutterbaugh for the excellent technical support and Kishore Guda for helpful discussions.
References
Ugene, a Newly Identified Protein That Is Commonly Overexpressed in Cancer and Binds Uracil DNA Glycosylase

Chunguang Guo, Xiaodong Zhang, Stephen P. Fink, et al.


Updated version
Access the most recent version of this article at:
http://cancerres.aacrjournals.org/content/68/15/6118

Supplementary Material
Access the most recent supplemental material at:
http://cancerres.aacrjournals.org/content/suppl/2008/07/31/68.15.6118.DC1

Cited articles
This article cites 19 articles, 11 of which you can access for free at:
http://cancerres.aacrjournals.org/content/68/15/6118.full#ref-list-1

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.