Nucleic Acids Research Advance Access published October 1, 2015

doi: 10.1093/nar/gkv977

Detection of uracil within DNA using a sensitive labeling method for *in vitro* and cellular applications

Gergely Róna[1,2,*], Ildikó Scheer[1,2], Kinga Nagy[1,2], Hajnalka L. Pálinkás[1,3], Gergely Tihanyi[1,2], Máté Borsos[1], Angéla Békési[1] and Beáta G. Vértessy[1,2,*]

[1]Institute of Enzymology, RCNS, Hungarian Academy of Sciences, Magyar Tudósok Str. 2, H-1117 Budapest, Hungary, [2]Department of Applied Biotechnology and Food Sciences, Budapest University of Technology and Economics, Szt Gellért Square 4, H-1111 Budapest, Hungary and [3]Doctoral School of Multidisciplinary Medical Science, University of Szeged, H-6720 Szeged, Hungary

Received May 17, 2015; Revised September 10, 2015; Accepted September 15, 2015

ABSTRACT

The role of uracil in genomic DNA has been recently re-evaluated. It is now widely accepted to be a physiologically important DNA element in diverse systems from specific phages to antibody maturation and *Drosophila* development. Further relevant investigations would largely benefit from a novel reliable and fast method to gain quantitative and qualitative information on uracil levels in DNA both *in vitro* and *in situ*, especially since current techniques does not allow *in situ* cellular detection. Here, starting from a catalytically inactive uracil-DNA glycosylase protein, we have designed several uracil sensor fusion proteins. The designed constructs can be applied as molecular recognition tools that can be detected with conventional antibodies in dot-blot applications and may also serve as *in situ* uracil-DNA sensors in cellular techniques. Our method is verified on numerous prokaryotic and eukaryotic cellular systems. The method is easy to use and can be applied in a high-throughput manner. It does not require expensive equipment or complex know-how, facilitating its easy implementation in any basic molecular biology laboratory. Elevated genomic uracil levels from cells of diverse genetic backgrounds and/or treated with different drugs can be demonstrated also *in situ*, within the cell.

INTRODUCTION

Uracil, one of the nucleic acid bases present in RNA, is usually considered to be a mistake when appearing in DNA (1). Two independent pathways may lead to the presence of uracil in DNA. On the one hand, hydrolytic deamination of cytosine within the DNA, a rather frequent event, results in numerous uracil moieties, which are mutagenic since these replacements, if left unrepaired, will lead to an exchange of a G:C base pair to a A:U (A:T) base pair (2). On the other hand, most polymerases cannot distinguish between deoxyuridine and deoxythymidine and will readily incorporate either of these two building blocks depending on the ratio of cellular dUTP and dTTP pools (3). The nucleotide pool is usually sanitized by dUTPases (in most cases, these enzymes are encoded by the *dut* gene), an enzyme family conserved from bacteria to human, to avoid such thymine-replacing uracil incorporation events (4–7). Genomic uracil is specifically recognized by representatives of the uracil-DNA glycosylase superfamily (UDG), cleaving the N-glycosidic bond between the pyrimidine ring and deoxyribose and resulting in apyrimidinic (AP) sites that are further processed by base excision repair (8,9). UNG, one of the four UDGs found in mammalian cells, specifically excise uracil bases from both double-stranded (dsDNA) and single-stranded DNA (ssDNA). *In vitro* the enzyme removes uracil in the order of preference ssU > U:G >> U:A (10–13). UNG activity is somewhat affected by the sequence context, having a slightly different affinity for uracil in A/T rich regions compared to G/C rich environment (12,14–16). To a lesser extent, bases formed from cytosine oxidation are also substrates of UNG (5-hydroxyuracil, isodialuric acid and alloxan; the latter only recognized by the human enzyme) (17,18). With a slower rate, 5-fluorouracil is also processed, however other larger 5-halouracils (like BrdU) are not recognized (19,20). A growing number of results show that UNG is somewhat capable of binding AP sites, but with a lower affinity compared to genomic uracil (14,21–23). No activity has been detected against normal DNA bases or against uracil in RNA (13) since RNA is excluded from the DNA-binding pocket due to unfavorable steric reasons (24,25).

Fine-tuned regulation of nucleotide pools is also of key importance for genomic stability. Inhibitors targeting path-

*To whom correspondence should be addressed. Tel: +36 13 826 707; Fax: +36 14 665 465; Email: vertessy@mail.bme.hu
Correspondence may also be addressed to Gergely Róna. Tel: +36 13 826 762; Fax: +36 14 665 465; Email: rona.gergely@ttk.mta.hu

ways involved in proper dUTP/dTTP pool maintenance, such as the *de novo* thymidylate biosynthesis pathways, induce thymine-less cell death and are a focus of cancer treatment (26). Importantly, genomic uracil may also appear under normal physiological conditions. In the most extreme cases of specific bacteriophages, such as in *Bacillus subtilis* PBS1 and PBS2 phages, and the *Yersinia enterocolitica* ΦR1–37 phage, the phage DNA contains deoxyuridine but no deoxythymidine (27–30). Uracil in DNA was implicated as a key factor in B lymphocyte function during somatic hypermutation and class-switch recombination (31–33). The surprisingly high uracil content (estimated as >25 000 uracil/million bases) of reverse-transcribed HIV genomic DNA has been suggested to play an important role in the viral life cycle (34). DNA from fruit fly larvae and pupae also contains highly elevated levels of uracil (200–2000 uracil/million bases) (35,36).

As summarized above, several different fields in biology from phage genetics to lentiviral infection mechanisms, from antibody maturation and *Drosophila* development to chemotherapeutic approaches in cancer treatment heavily rely on genomic uracil occurrence. Hence, a reliable, fast, cheap and easy method to gain quantitative and qualitative information on uracil levels in DNA is of high importance for *in vitro* and *in vivo* studies. Currently available genomic uracil quantification methods vary in specificity, sensitivity and price. Even though LC/MS/MS based methods are sensitive, they need laborious, excessive sample preparation that involves nucleotide or uracil hydrolysis of the samples (37–42). Chemical modification of uracil moieties to enhance detection also provides a highly sensitive method but needs several steps in sample preparation (43,44). Real-time polymerase chain reaction (PCR)-based techniques reflect the uracil content only on a limited DNA fragment, and total genomic uracil content is calculated as an extrapolation based on the assumption that uracil residues are evenly distributed throughout the genome (45) — although this may not be always the case.

Most techniques reported to date that aim to quantify genomic uracil levels excise uracil from DNA during the process and do not allow *in situ* cellular detection. Interestingly, for detection of numerous other non-orthodox DNA bases such as 5-methylcytosine (46), 5-hydroxymethylcytosine (47), 5-hydroxymethyluracil (48), thymine dimers (49), 8-oxo-guanine (50) and 8-nitroguanine (51), antibodies have been described. To our knowledge, no such method has yet been reported for the uracil moieties in DNA.

In the present work, we therefore aimed at designing a uracil sensor by applying a catalytically inactive UNG uracil-DNA glycosylase, which is capable of binding to but not excising uracil (22). The uracil-recognizing UNG sensor was designed in such a way that it can be detected either with conventional antibodies in dot-blot applications or also *in situ* using an immunocytochemical approach. Our method is a relative quantification approach that delivers the sensitivity of MS based approaches, reaching a femtomol uracil detection limit. It may also be developed further as a ChiP like approach to gain position- and sequence-specific information on genomic uracil content. Performance of the herein described uracil sensor has been analyzed in dot-blot and immunocytochemical approaches, using prokaryotic CJ236 *Escherichia coli* [*dut−, ung−*] or BL21(DE3) *ung-151 E. coli* [*ung-*] and also eukaryotic (*Drosophila* and human) cell lines with altered base excision repair background (36,52). We have also analyzed cells treated with several different chemotherapeutic drugs, known to interfere with thymidylate biosynthesis and leading to increased uracil content in DNA. Our results are in good agreement with current reports from the literature verifying that our method is sensitive, cost-effective and adequate.

MATERIALS AND METHODS

Plasmid constructs and cloning

Human uracil-DNA glycosylase 2 (UNG2) cDNA was a generous gift of Professor Salvatore Caradonna (Department of Molecular Biology, University of Medicine and Dentistry of New Jersey) and was cloned into the XhoI/KpnI sites of the pDsRed-Monomer-N1 vector (Clontech, Mountain View, CA, USA), as described previously (53). DsRed-fused UNG2 was further PCR amplified and cloned into NdeI/XhoI sites of the vector pET-20b (Novagen, Merck Millipore, Billerica, MA, USA) with primers UNG_F and UNG_R (Supplementary Table S1). Point mutations (D154N and H277N) were created by the QuickChange mutagenesis method (Stratagene, Santa Clara, CA, USA) (with primers D154N_F, D154N_R and H277N_F, H277N_R respectively). UNG2 (D154N and H277N) lacking the first 84 amino acid from its N-terminus (ΔUNG) was PCR amplified (with primers 1×FLAG_F, 3×FLAG_F and 1×/3×FLAG_R) and was cloned into the NdeI/XhoI sites of the vector pET-15b (Novagen) FLAG-tagged (1× or 3×) yielding the constructs 1×FLAG-ΔUNG and 3×FLAG-ΔUNG respectively. Constructs ΔUNG-DsRed and FLAG-ΔUNG-DsRed were PCR amplified (with primers 1×FLAG_Ds_F, Ds_F and Ds_R) and cloned into the NdeI/XhoI sites of the vector pET-20b. The vector expressing the human codon optimalized UGI along with GFP (pLGC-hUgi) (54) was a kind gift of Michael D. Wyatt (South Carolina College of Pharmacy, University of South Carolina). Primers used in this study were synthesized by Eurofins MWG GmbH (Ebersberg, Germany) and are summarized in Supplementary Table S1. All constructs were verified by sequencing at Eurofins MWG GmbH.

Cell culture and transfection

The MLH1-deficient (a mismatch repair deficient) human colorectal adenocarcinoma cell line, HCT116, was purchased from the European Collection of Cell Cultures (ECACC, Salisbury, UK). Mouse embryonic fibroblast (MEF) cells lacking functional UNG (55) were a generous gift from Dr Hilde Nilsen (University of Oslo). HCT116 cells were cultured in McCoy's 5A medium (Gibco, Life Technologies, Carlsbad, CA, USA) while MEF cells in DMEM/F12 HAM (Sigma); supplemented with 50 μg/ml Penicillin-Streptomycin (Gibco) and 10% FBS (Gibco) in a humidified 37°C incubator with 5% CO_2 atmosphere. Schneider S2 cells (derived from *Drosophila melanogaster*)

were purchased from Gibco and were cultured in Schneider Insect Medium (Sigma) supplemented with 10% FBS (Gibco) and 50 μg/ml Penicillin-Streptomycin (Gibco) and kept in a 26°C incubator.

HCT116 cells were transfected with FuGENE® HD (Promega, Madison, WI, USA) in T25 tissue culture flasks according to the manufacturer's recommendation. For immunocytochemistry MEF cells were transfected in a 6-well plate with 4 μg of normal pEGFP-N1 (purified from XL1-Blue [*dut+, ung+*] *E. coli* cells) or uracil-rich pEGFP-N1 vector (purified from CJ236 [*dut−, ung−*] *E. coli* cells) and 12 μl FuGENE® HD transfection reagent according to the recommendation of the manufacturer. After 16 h cells were thoroughly trypsinized and washed with phosphate buffered saline (PBS) extensively (to wash away extracellular plasmid aggregates attached to the cell surface) and finally splitted onto 24-well plates containing poly-L-lysine coated cover glasses.

Treatment of cells, DNA isolation and purification

Plasmid DNA. pEGFP-N1 plasmid (Clontech) was transformed into XL1-Blue [*dut+, ung+*] (Stratagene) or CJ236 [*dut−, ung−*] (NEB, Ipswich, MA, USA) *E. coli* strains. Cell cultures were grown for 16 h in Luria broth (LB) media supplemented with kanamycin at 37°C, and the plasmids were purified using PureYield™ Plasmid Midiprep Kit (Promega) according to the instructions of the manufacturer.

Genomic DNA. XL1-Blue, BL21(DE3) *ung-151* (56) and CJ236 *E. coli* strains were propagated in LB media at 37°C and were harvested at log-phase ($A_{600nm} = 0.5$). BL21(DE3) *ung-151* cells were also grown either in the presence of 30.7 μM 5-fluoro-2′-deoxyuridine (5FdUR) or 200 μM 2′-deoxyuridine (dUR) or in the presence of both drugs. Genomic DNA was purified with MasterPure™ DNA Purification Kit (Epicentre, Madison, WI, USA), followed by an additional purification with the Genomic DNA Clean & Concentrator Kit (ZYMO Research, Irvine, CA, USA) using the recommendations of the manufacturer. *Drosophila* Schneider S2 cells were grown either in the absence or presence of 100 μM 5FdUR, 500 μM dUR or 10 μM methotrexate (MTX), 100 nM raltitrexed (RTX), 500 μM dUR for 48 h. Genomic DNA of S2 cells was purified as above. Forty hours before treatment, HCT116 cells were transfected either with the UGI-GFP expressing vector or with an empty vector expressing GFP alone (pEGFP-N1). Both transfected and non-transfected cells were grown for an additional 48 hours in the presence or absence of 20 μM 5FdUR before collecting them for genomic DNA purification described as above.

Recombinant protein production

All UNG constructs were expressed in the *E. coli* BL21(DE3) *ung-151* strain and purified using Ni-NTA affinity resin (Qiagen, Hilden, Germany). Transformed cells growing in LB medium were induced at $A_{600nm} = 0.6$ with 0.6 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG) for 24 h at 18°C. Cells were harvested and lyzed in lysis buffer (50 mM TRIS·HCl, pH = 8.0, 300 mM NaCl, 0.5 mM ethylenediaminetetraacetic acid (EDTA), 0.1% Triton X-100, 10 mM β-mercaptoethanol, 1 mM phenylmethylsulfonyl fluoride, 5 mM benzamidine, 1×cOmplete ULTRA™ EDTA free protease inhibitor cocktail tablet (Roche), 0.1 mg/ml lysozyme, 0.1 mg/ml DNase (Sigma, St. Louis, MO, USA) and 0.01 mg/ml RNAse A (Invitrogen, Carlsbad, CA, USA)) assisted with sonication. Cell debris was pelleted by centrifugation at 20 000 × *g* for 30 min. Supernatant was applied onto a Ni-NTA column and washed with a set of washing buffers: low salt buffer (50 mM HEPES, pH = 7.5, 30 mM KCl, 5 mM β-mercaptoethanol), high salt buffer (50 mM HEPES, pH = 7.5, 300 mM KCl, 5 mM β-mercaptoethanol) and very high salt buffer (50 mM HEPES, pH = 7.5, 500 mM NaCl, 40 mM Imidazole, 5 mM β-mercaptoethanol). UNG constructs were finally eluted with elution buffer (50 mM HEPES, pH = 7.5, 30 mM KCl, 300 mM imidazole, 5 mM β-mercaptoethanol) and dialyzed against the following buffer: 30 mM Tris·HCl, pH = 7.4, 140 mM NaCl, 0.01% Tween-20, 1 mM EDTA, 15 mM β-mercaptoethanol.

Assay for testing UNG activity

A total of 150 ng of SmaI (NEB) linearized pEGFP-N1 or uracil-rich pEGFP-N1 vector was incubated with 0.02 μg of each UNG constructs (hUNG2-DsRed WT, hUNG2-DsRed, 1×-Flag-ΔUNG, 3×-Flag-ΔUNG, Flag-ΔUNG-DsRed, ΔUNG-DsRed) for 1 h at 37°C in a final volume of 11 μl, in Endo IV buffer (50 mM Tris-acetate (pH 7.5), 50 mM KCl, 1 mM EDTA, 0.05% Triton X-100; Fermentas, Waltham, MA, USA). One unit of Endonuclease IV (Fermentas) was added to the reaction mixture and was further incubated for 1 h at 37°C. Endonuclease IV (Endo IV) is apurinic/AP endonuclease that will hydrolyse AP sites in DNA. AP sites are cleaved at the phosphodiester bond that is 5′ to the lesion leaving a hydroxyl group at the 3′ terminus and a deoxyribose 5′-phosphate at the 5′ terminus. UNG and Endo IV treatment leads to nicks in the phosphodiester backbone of the DNA, resulting in extensive fragmentation of uracil-rich DNA. Reaction was stopped by adding 4 μl of inactivation mixture containing 2.5% sodium dodecyl sulfate, 2.5 mg/ml Proteinase K (Sigma) and 1.5× concentrated DNA Loading Dye (Fermentas). Standard agarose gel electrophoresis was performed in a 1% gel.

Electrophoretic mobility shift assay (EMSA)

A total of 100 ng of SmaI (NEB) linearized pEGFP-N1 or uracil-rich pEGFP-N1 vector was incubated with a series of two-fold dilution of the different UNG constructs (hUNG2-DsRed WT, hUNG2-DsRed, 1×-Flag-ΔUNG, 3×-Flag-ΔUNG, Flag-ΔUNG-DsRed, ΔUNG-DsRed) starting with 1 μg of protein, for 5 min at room temperature in UNG buffer (30 mM TRIS·HCl; 140 mM NaCl; 0,01% Tween-20; 1 mM EDTA; 15 mM β-mercaptoethanol; pH = 7.4). Standard agarose gel electrophoresis was performed in a 0.75% gel.

Dot-blot based assay for quantification of DNA

Genomic DNA isolated from CJ236 *E. coli* strain [*dut−, ung−*] (in log phase) served as a uracil standard. A total of 5 ng of this genomic DNA was diluted into 2 μg of ultrapure salmon sperm DNA as an inert carrier (Invitrogen), which was kept constant during the two-third dilution series of this standard. The two-third serial dilutions for XL1-Blue, BL21(DE3) and BL21(DE3) *ung-151 E. coli* samples started with 1 μg of DNA mixed into 1 μg of carrier salmon sperm DNA. In case of samples derived from S2 and HCT116 cells 0.6 μg of their genomic DNA was diluted into 1.4 μg of carrier salmon sperm DNA. In every case the 2 μg total DNA content was kept constant with salmon sperm DNA. DNA samples were spotted onto a prewetted positively charged nylon membrane (Amersham Hybond-Ny+; GE Healthcare, Little Chalfont, UK) using a vacuum-driven microfiltration apparatus (Bio-Dot, Bio-Rad, Hercules, CA, USA). After 10 min of air-drying, immobilization of DNA was performed with 2 h of incubation at 80°C. Membrane was blocked by a 15 min incubation in blocking buffer: ETBS-T (25 mM Tris·HCl, pH = 7.4, 2.7 mM KCl, 137 mM NaCl, 1 mM EDTA, 0.05% Tween-20) containing 100 μg/ml salmon sperm DNA, 5% non-fat milk powder and 10 mM β-mercaptoethanol. Membrane was incubated with the 3×-Flag-ΔUNG construct (18.1 μg/ml) in blocking buffer overnight at 4°C. After several washing steps with ETBS-T, anti-Flag M2 antibody (Sigma) was added for 1 h at room temperature (1:2000 dilution in ETBS-T with 5% non-fat milk powder). After washing the membranes, horseradish peroxidase coupled secondary antibody was applied (Sigma). Immunoreactive bands were visualized by enhanced chemiluminescence reagent (GE Healthcare, Buckinghamshire, UK) and 16 bit images were captured by a BioRad ChemiDoc™ MP Imaging system. Densitometry was done using ImageJ 1.48p software (National Institutes of Health, Bethesda, MD, USA). Normalized intensity values were calculated by adjusting the raw intensity values for the background originating from the carrier salmon sperm DNA.

The average molar mass of a nucleotide in *E. coli* ($M_{E.coli}$) or *D. melanogaster* ($M_{D.melanogaster}$) genomic DNA was calculated with the following form:

$$M\left[\frac{\mathrm{g}}{\mathrm{mol}}\right] = \frac{GC\%}{2} \times (M_{dGMP} + M_{dCMP} - 2M_{H_2O}) + \frac{1-GC\%}{2} \times (M_{dAMP} + M_{dTMP} - 2M_{H_2O})$$

where M_{dGMP} = 347.2 g/mol, M_{dCMP} = 307.2 g/mol, M_{dAMP} = 331.2 g/mol, M_{dTMP} = 322.2 g/mol, M_{H2O} = 18.0 g/mol, are the respective molecular weights of the given compounds; dGMP-dCMP percentage (GC%) of the *E. coli* genome is 50.7% and is 42.1% for *D. melanogaster* (based on NCBI genome database, average of reference genomes). Calculated values were 308.95 g/mol and 308.90 g/mol for *E. coli* ($M_{E.coli}$) and *D. melanogaster* ($M_{D.melanogaster}$), respectively. The very slight difference between $M_{E.coli}$ and $M_{D.melanogaster}$ was neglected during further analysis.

In each dot of the standard samples, the mass of non-carrier uracil-containing DNA was known and termed as $m_{\mathrm{dot,\ stand}}$. The number of DNA nucleotides present in each dot of the standard samples were calculated using the following form: $n_{E.coli} = \frac{m_{\mathrm{dot,\ stand}}}{M_{E.coli}}$. The number of deoxyuridine nucleotides were calculated using the previously determined uracil content of genomic DNA isolated from log phase culture of CJ236 [*dut−, ung−*] *E. coli*, i.e. 6580±174 deoxyuridine/million nucleotide (45). The following equation gives the amount of uracil in each dot of the standard sample ($n_{U,\ \mathrm{standard}}$):

$$n_{U,\mathrm{standard}} = 6580/1000000 \times n_{E.coli}$$

Calibration curve from the dilution of the standard was visualized the following way for quantification: the amount of uracil in each dot of the standard samples ($n_{U,\ \mathrm{standard}}$) were plotted against the corresponding normalized intensity values ($I_{\mathrm{norm,\ standard}}$). Values were fitted with a polynomial with the least order that provided a fit with $R^2 \geq 0.99$. The $\frac{\text{number of uracil}}{\text{million bases}}$ in the 'unknown' genomic DNAs were determined by interpolating their normalized intensities ($I_{\mathrm{norm,\ unknown}}$) in the calibration plot based on the amount of DNA applied ($m_{\mathrm{dot,\ unknown}}$).

Statistics

Statistical analysis was carried out by InStat 3.05 software (GraphPad Software, San Diego CA, USA) using the non-parametric Kruskal–Wallis test or one-way ANOVA test with Student-Newman-Keuls multiple comparison post-hoc test when samples passed equal variance (Bartlett's test) and normal distribution tests (Kolmogorov–Smirnov test). Differences were considered statistically significant at $P < 0.05$.

Western blot

Cells were collected, washed with PBS, and resuspended in lysis buffer (50 mM TRIS·HCl pH = 7.4; 140 mM NaCl; 0,4% NP-40; 2 mM dithiothreitol (DTT); 1 mM EDTA, 1 mM phenylmethylsulfonyl fluoride; 5 mM benzamidin, 1×cOmplete ULTRA™ EDTA free protease inhibitor cocktail tablet (Roche)). Cell lysis was assisted with sonication. Insoluble fraction was removed by centrifugation (20 000 × *g* × 15 min at 4°C). Protein concentration was measured with BioRad Protein Assay to ensure equivalent total protein load per lane. Proteins were resolved under denaturing and reducing conditions on a 12% polyacrylamide gel and transferred to PDVF membrane (Immobilon-P, Merck Millipore, Billerica, MA, USA). Membranes were blocked with 5% non-fat dried milk and were developed against GFP (1:2000, Molecular probes, Life Technologies, Carlsbad, CA, USA) and actin (1:500, Sigma) for loading control. After applying horseradish peroxidase coupled secondary antibodies (Amersham Pharmacia Biotech), immunoreactive bands were visualized by enhanced chemiluminescence reagent (GE Healthcare) and images were captured by a BioRad ChemiDoc™ MP Imaging system.

Staining uracil residues in *E. coli* genomic DNA

Immunofluorescence staining was done based on the work of (57) with modifications. Briefly, 500 μl of XL1-Blue [*dut+, ung+*] and CJ236 [*dut−, ung−*] *E. coli* cells were collected in log-phase (A_{600nm} = 0.5), by centrifuging them at

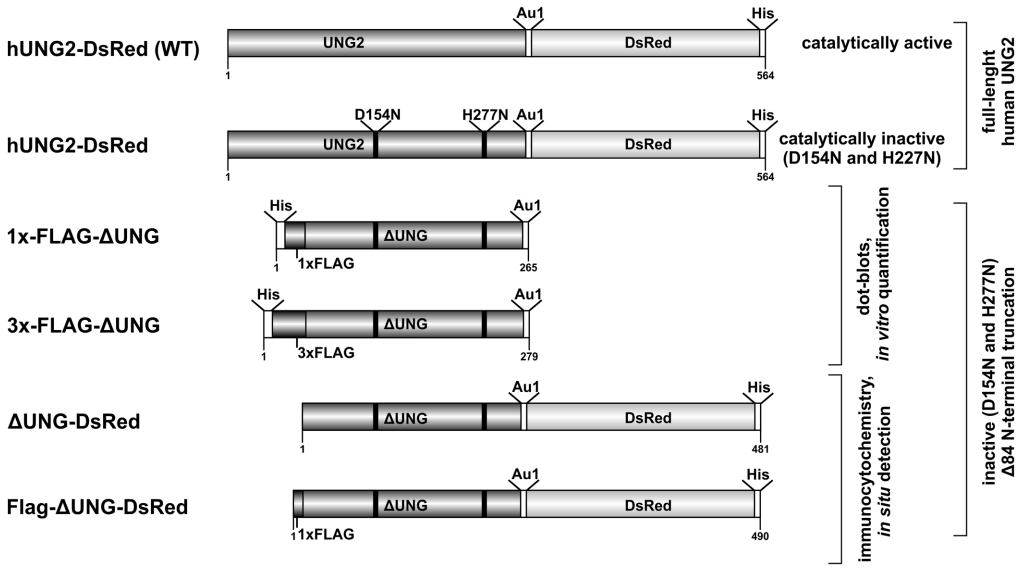

Figure 1. Schematics of the used constructs for uracil detection. In our constructs, human UNG2 was used as the uracil sensor core domain. During *in vitro* quantification and for *in situ* detection, a double mutant UNG2 was created (D154N and H277N, mutated sites indicated with black lines within the schematics of the protein domains). This mutant is catalytically inactive but is still capable of binding uracil moieties in DNA (cf text for more details). The N-terminal 84 residues, responsible for the binding to RPA and PCNA, were also removed (ΔUNG) to prevent non-specific binding. The ΔUNG uracil recognizing core was fused to epitope tags (1×/3×FLAG, Au1) for immunodetection, DsRed-monomer for direct fluorescent detection and to His-tag for affinity purification.

7000 × *g* for 5 min and were washed with PBS. Cells were fixed with Carnoy's fixative (ethanol:acetic acid:chloroform = 6:3:1) for 20 min at 4°C. Rehydration was performed as following: washing with 1:1 ethanol:PBS, 3:7 ethanol:PBS and finally incubating in PBS containing 0.05% Triton X-100 (PBST) for 5 min. Cells were washed once with GTE buffer (50 mM glucose, 20 mM Tris, pH = 7.5 and 10 mM EDTA), and finally resuspended in GTE buffer containing 10 mg/ml lysozyme (Sigma) for 5 min. The suspension was applied onto poly-L-lysine coated cover glasses for an additional 5 minutes, excess fluid was drained and cells were left to air-dry. Cells were washed with PBST for 10 min and were blocked in blocking buffer (5% BSA, in PBST) for 15 min. Uracil residues were visualized by applying 4.64 μg/ml of the Flag-ΔUNG-DsRed construct in blocking buffer, overnight at 4°C. After several washing steps with PBST, anti-FLAG M2 antibody (Sigma) was applied (1:2000 dilution) for 1 h in a blocking buffer. FLAG epitope was visualized by applying Alexa 488 conjugated secondary antibody (1:1000, Molecular Probes). Cells were counterstained with 1 μg/ml DAPI (4′,6-diamidino-2-phenylindole, Sigma) and embedded in FluorSave™ Reagent (Calbiochem, Merck Millipore, Billerica, MA, USA). Images were acquired with a Zeiss LSCM 710 microscope using a 63× NA = 1.4 Plan Apo objective.

Staining uracil residues of extrachromosomal plasmids in MEF cells

Cells were washed with prewarmed (37°C) phosphate buffered saline (PBS, pH = 7.4) and were fixed with ice cold Carnoy's fixative (ethanol:acetic acid:chloroform = 6:3:1) for 20 minutes at 4°C. Rehydration was performed as following: washing step with 1:1 ethanol:PBS, 3:7 ethanol:PBS, finally with PBS for 5 min. Epitope unmasking was done by applying 1 N HCl, 0.5% Triton X-100 for 15 min, after which 0.1 M $Na_2B_4O_7$ (pH = 8.5) was used for neutralization for 5 min followed by PBS washing. HCl was used to denature DNA, allowing our UNG construct to have better access to genomic uracil. This method is also routinely applied when using antibodies against BrdU, doing cell proliferation assays (see (58)). Blocking was done at room temperature for 1 h in blocking buffer: 200 μg/ml salmon sperm DNA, 5% fetal goat serum (FGS), 3% fetal bovine serum albumin (BSA) and 0.05% Triton X-100 in PBS. Uracil residues were visualized by applying 4.64 μg/ml of the Flag-ΔUNG-DsRed construct in blocking buffer, overnight at 4°C. After several washing steps with blocking buffer, anti-FLAG M2 antibody (Sigma) was applied (1:2000 dilution) for 1 h in a blocking buffer not containing salmon sperm DNA. FLAG epitope was visualized by applying Alexa 488 conjugated secondary antibody (1:1000, Molecular Probes). Cells were counterstained with 1 μg/ml DAPI (Sigma) and embedded in FluorSave™ Reagent (Calbiochem). Images were acquired with a Zeiss LSCM 710 microscope using a 63× NA = 1.4 Plan Apo objective.

RESULTS AND DISCUSSION

Construction and analysis of catalytically inactive uracil-sensor proteins

Wild-type human UNG2 possesses a highly selective substrate binding site for uracil and is specific for excising uracil from DNA, with a negligible activity toward the natural DNA bases cytosine or thymine (59). The catalytically inactive double mutant (D145N, H268N) human UNG2 preserves this highly specific and strong binding interac-

tion with uracil-containing DNA similar to the wild-type enzyme. (22,60). Similar K_d values were observed for the wild-type and the double mutant enzyme for non-cleavable substrate analogs (22). We therefore aimed at employing this double mutant protein as a specific uracil sensor that strongly binds to the uracil base but does not excise it from DNA. Since our sensor relies on the characteristics of the human UNG2, it is capable of recognizing a few derivatives of uracil, and could also bind abasic sites (APs) but with a lower affinity compared to genomic uracil (see details in the Introduction). This concept, if successful, could be used as a novel labelling method capable of recognizing genomic uracil both *in vitro* using a dot-blot based method and *in situ*, similarly to an immunocytochemical approach.

To obtain an even more specific UNG-based uracil-sensor construct, we eliminated the N-terminal 84 residues from the human UNG enzyme that comprises a recognition site for PCNA and RPA proteins (61–63), resulting in the construct termed as ΔUNG in the present work (cf Figure 1, note that all constructs termed ΔUNG harbor the double mutations D145N, H268N). This truncation was deemed to be highly desirable in order to erase non-specific protein binding while retaining similar specificity and binding characteristics as the full length form while also being more resistant to proteolysis (10,12,64). Moreover, the lack of the N-terminal 84 residues diminishes the need of Mg^{2+} for proper UNG function which is useful to lower any residual nuclease activity during the assays by applying EDTA (10,64). We also equipped this truncated construct with different tags: His-tag for purification and Flag- and Au1-tags for antibody-based detection. Further, in order to achieve a low background through direct detection in immunocytochemistry, we have attached a red fluorescent protein, the monomeric form of DsRed, to the C-terminal end of our construct.

In order to check the required and expected functionality of our constructs, we have carried out enzyme activity assays (as described in 'Materials and Methods' section). Figure 2A confirms that the only construct showing catalytic activity on uracil-rich DNA is the wild type UNG, whereas all of our ΔUNG constructs lack any detectable excising activity. On the other hand, as shown on Figure 2B and Supplementary Figure S1, all of our constructs show preferential binding to uracil-rich as compared to normal DNA.

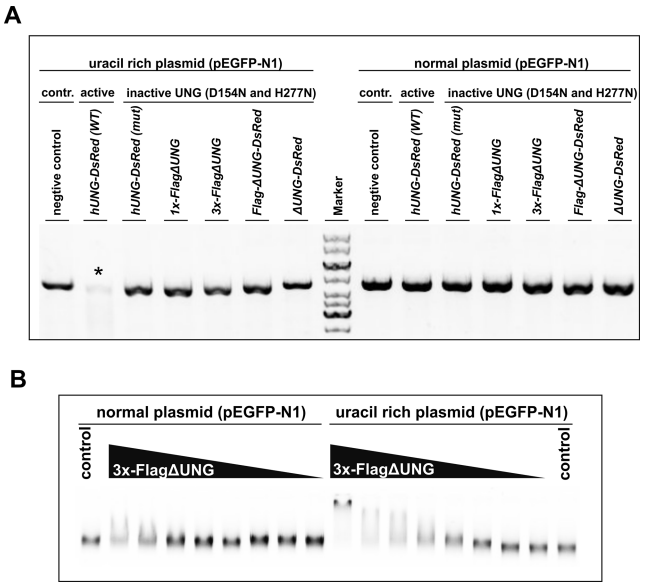

Figure 2. Activity and uracil binding capability of the used constructs. (**A**) Agarose gel electrophoresis based assay was applied to detect UNG activity. Only the hUNG2-DsRed WT construct was active on uracil-rich plasmid (indicated with an asterisk), which did not harbor the two point mutations (D154N and H277N). All other constructs used in the study do not excise uracil from DNA. (**B**) Uracil binding capability of the 3xFLAG-ΔUNG construct was addressed with Electrophoretic Mobility Shift Assay (EMSA). Increasing amount of the construct clearly shifts the position of the linearized vector, which is more prominent in case of uracil-rich template, indicating that the construct is capable of binding genomic uracil moieties. We have experienced similar result with the other tested constructs (Supplementary Figure S1).

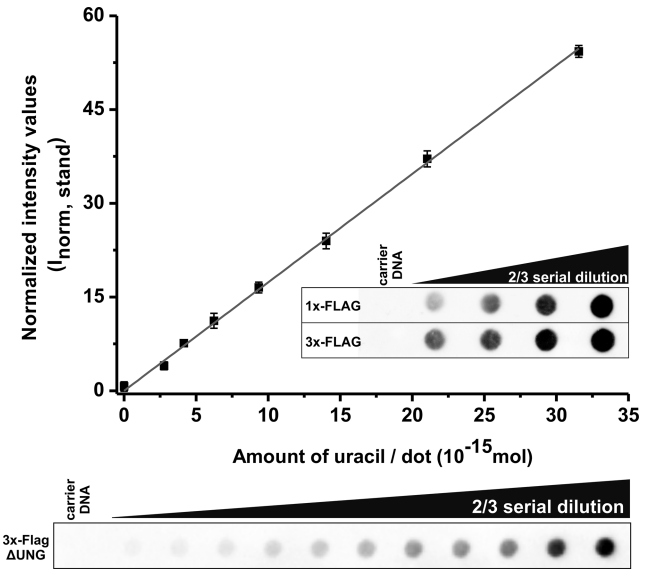

Figure 3. Design of a standard curve for *in vitro* quantification of genomic uracil levels. Genomic DNA isolated from log phase growing CJ236 [*dut−, ung−*] *Escherichia coli* strain was used as a standard with well-defined uracil-content during quantification. Applying a serial dilution of this standard provides a wide and reproducible range for uracil quantification. The normalized calibration curve is from four independent datasets ($n = 4$), where error bars show standard errors of the mean (SEM). The inset shows that the 3×FLAG-ΔUNG construct is slightly more sensitive under similar conditions, compared to the 1×FLAG-ΔUNG construct (as based on a 4 point two-third serial dilution, starting with 100 ng of standard genomic DNA).

Development of a dot-blot based quantitative assay for detection of genomic uracil *in vitro*

In testing our different constructs, we have used either just a single (1×) or a triplicated (3×) Flag-tag form and employed them on our standard, namely the genomic DNA isolated from log phase CJ236 [*dut−, ung−*] *E. coli* cells for which uracil content has been previously quantified (43,45). Figure 3 shows that the triplicated Flag-tag containing ΔUNG sensor is more sensitive in the dot-blot assay. This construct was therefore used in all further experiments. Figure 3 also indicates that a linear response could be reached on a series of dilutions in wide dynamic range. For each experiment, a calibration curve was always recorded

which was highly reproducible thus served as a firm basis for our further studies.

It has been described in the literature that treatments with drugs interfering with thymidylate biosynthesis (e.g. 5-fluorodeoxyuridine (5FdUR) (65)) can lead to considerably elevated uracil levels in genomic DNA (66–68). Accordingly, Figure 4 summarizes the *de novo* thymidylate biosynthesis pathway and indicates the enzymatic steps known to be perturbed by different drugs (frequently used in anti-cancer chemotherapy). We therefore applied such drug treatments on *E. coli* cells. Figure 5 shows that the dot-blot assay adequately reflects the expected increase in uracil content for genomic DNA isolated from the BL21(DE3) *ung-151 E. coli* strain. Although a basal higher uracil content could have been expected, the lack of the major uracil-DNA glycosylase in the *ung-151* cells in the absence of drug treatment did not lead to any significant increase in the genomic uracil content (4.66 ± 2.24 uracil/million bases) compared to XL1-Blue (3.29 ± 1.74 uracil/million bases) (cf Figure 5. This indicates that the difference, if exists, has to be below the detection limit. Because the dUTPase enzyme is still present in the *ung-151* cells, the dNTP pool is not perturbed to an extent where erroneous uracil incorporation through replication is expected to occur. Thus, the sole lack of UNG might not be enough to create detectable increase in genomic uracil.

Also, treating the BL21(DE3) *ung-151* cells with deoxyuridine (dUR) alone did not change the genomic uracil content (4.28 ± 2.05 uracil/million bases) (Figure 5). In contrast, applying 5FdUR on exponentially growing *ung-151* cells, the dot-blot assay clearly showed an ∼10-fold increase in the uracil content (39.89 ± 7.72 uracil/million bases) (Figure 5). The same ∼10-fold increase was observed when the two drugs were applied together (35.61 ± 4.54), arguing that the effect induced by 5FdUR is not enhanced by dUR. The *ung-* background is a major factor allowing the increase of genomic uracil content upon drug treatment, since drug treatment of cells with a [*dut+, ung+*] genetic background (either XL1 Blue or BL21(DE3) cells) did not lead to increased uracil levels (Supplementary Figure S2).

For further experiments, we applied a eukaryotic cell line with a *ung−* background. As described previously, the *ung* gene is interestingly absent from *D. melanogaster* (69), hence we used the *Drosophila* embryo-derived Schneider S2 (termed as S2 in the present study) cell line. We have shown that this cell line can tolerate the increased genomic uracil content well (36), making it an optimal object for our present work.

Figure 6 shows that dual treatment of S2 cells with 5FdUR and dUR (25.54±2.98 uracil/million bases) leads to a significantly elevated genomic uracil content compared to non-treated cells (15.68 ± 3.02 uracil/million bases). Another drug combination of raltitrexed (RTX) and methotrexate (MTX), targeting thymidylate synthase (TYMS) (70) and dihydrofolate reductase (DHFR) (71), respectively (cf Figure 4), also induced a significant, two-fold increase in the genomic uracil content in S2 cells (32.18 ± 3.24 uracil/million bases). High methotrexate tolerance was already suggested for the S2 cell line, with potential perturbation of the relevant pathways (72). Hence, these results

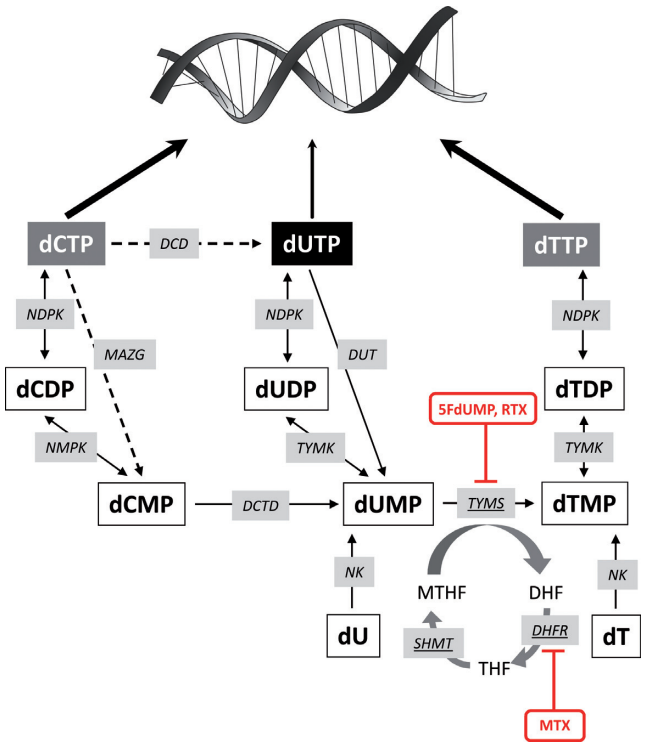

Figure 4. Pathways involved in thymidine synthesis in *Escherichia coli*, *Drosophila melanogaster* and humans. Key steps in dNTP synthesis focusing on the *de novo* thymidylate biosynthesis are shown (directly involved enzymes underlined). Dashed arrow shows pathways only present in *E. coli*. Inhibitors of the pathway are shown in red. 5FdUMP (5-fluorodeoxyuridylate), the metabolite of 5FU and 5FdUR, along with raltitrexed (RTX) inhibits thymidylate synthase (TYMS) while methotrexate (MTX) inhibits dihydrofolate reductase (DHFR). The enzyme responsible for dCTP-dCMP conversion in mammals is DCTPP1 (dCTP pyrophosphatase 1) (78), however, using a BLAST search, no clearcut homologue could be identified in *D. melanogaster*, hence we did not include it in this Figure. In *E. coli*, the nucleoside triphosphate pyrophosphohydrolase MAZG is responsible for this activity (79). Abbreviations are as follows: DCD: dCTP deaminase, DCTD: dCMP deaminase, DUT: dUTPase, DHF: dihydrofolate, DHFR: dihydrofolate reductase, MTHF: 5,10-methylene tetrahydrofolate, NDPK: nucleoside-diphosphate kinase, NK: nucleoside kinase, MAZG: nucleoside triphosphate pyrophosphohydrolase, NMPK: nucleoside monophosphate kinase, SHMT: serine hydroxymethyltransferase, THF: tetrahydrofolate, TYMK: dTMP kinase, TYMS: thymidylate synthase.

are in agreement with previous literature and attest to the applicability of our assay.

In order to test our method on human derived samples we used HCT116, a colon carcinoma cell line. UNG deficiency was achieved by applying a specific UNG inhibitor, UGI (73,74). The expression of human codon optimalized UGI inhibits endogenous UNG activity thus resembling a ung−/− phenotype (68). Figure 7 shows that treating HCT116 cells with 5FdUR after transfection with UGI (GFP is also expressed from the vector for monitoring purposes) significantly elevates genomic uracil levels (347.87 ± 84.62 uracil/million bases) compared to non-treated cells (7.82 ± 2.82 uracil/million bases). On its own neither 5FdUR treatment nor UNG inhibition elevates genomic uracil levels. 10.47 ± 2.82 uracil/million bases was measured for the non-treated and 10.92 ± 1.97 uracil/million

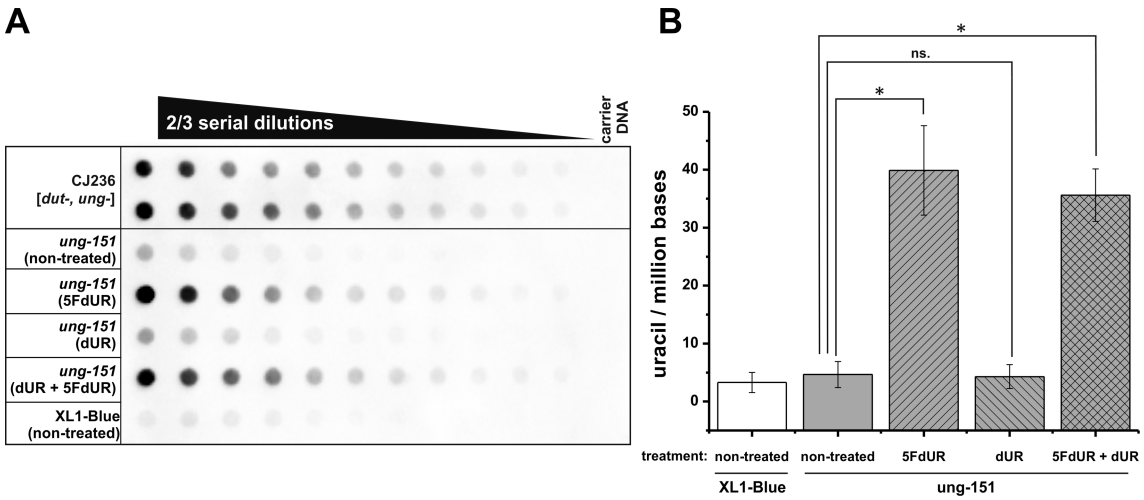

Figure 5. Dot-blot assay for measuring genomic uracil levels of 5FdUR, dUR treated *Escherichia coli* cells. (**A**) CJ236 [*dut*−, *ung*−] *E. coli* genomic DNA was used as standard for the dot-blot assay. Quantity of genomic uracil of different drug-treated (5FdUR and dUR or both) and non-treated *E. coli* BL21(DE3) *ung-151* samples were measured along with XL1-Blue [*dut*+, *ung*+], applied as a negative control. (**B**) Bar graph shows the uracil moieties/million bases of each sample (mean values ± the standard errors of the mean). Significant incrase (*) in uracil-DNA content was only observed using 5FdUR treatment, or the combined 5FdUR and dUR treatment as compared to non-treated cells ($P < 0.05$). Calculations were based on six independent datasets ($n = 6$).

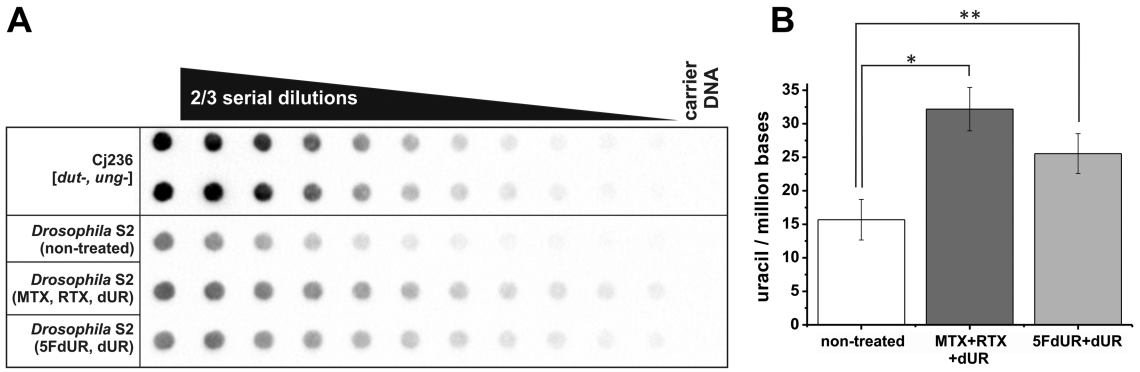

Figure 6. Dot-blot assay for measuring genomic uracil levels of *Drosophila* Schneider S2 cells after treatment with *de novo* thymidylate biosynthesis pathway inhibitors. (**A**) CJ236 [*dut*−, *ung*−] *Escherichia coli* genomic DNA was used as standard for the dot-blot assay. Genomic uracil content of different drug-treated (5FdUR, dUR or MTX, RTX, dUR) and non-treated *Drosophila* S2 cells were measured. (**B**) Bar graph shows the uracil moieties/million bases of each sample (mean values ± the standard errors of the mean). Both types of treatments led to significantly elevated genomic uracil levels as compared to non-treated cells (* = $P < 0.01$, ** = $P < 0.05$). Calculations were based on six independent datasets ($n = 6$).

bases for the 5FdUR treated. Similarly no significant difference was measured in case of cells that were transfected with an empty vector, expressing only GFP as control (4.17 ± 1.19 for the non-treated and 7.39 ± 2.98 uracil/million bases for the 5FdUR treated cells, respectively). These results are in good agreement with previously reported data (cf (68)) and also highlight the importance of simultaneous UNG inhibition along with drug treatments targeting the *de novo* thymidylate biosynthesis pathway for effective cancer therapy.

Application of the catalytically inactive UNG constructs for detection of uracil in DNA *in situ*

Figure 8 presents that the highly uracil-rich character of the CJ236 [*dut*−, *ung*−] *E. coli* cells, containing 6580 uracil/million bases, can be readily visualized via immunocytochemistry using our presently developed uracil sensor construct. Relying on the Flag-tag in our ΔUNG constructs, the uracil-DNA staining is easy to detect and colocalizes, as expected, with the DAPI signal for DNA. It is also important to note that the DsRed-tag may also be used for direct visualization. Specificity of the signal is adequately corroborated by lack of the staining in the XL1-Blue [*dut*+, *ung*+] cells. Application of the UNG inhibitor UGI (obtained from NEB) protein erases the signal, showing again the specific character of our assay (Supplementary Figure S3). The well-described UNG-UGI interaction has been documented to prevent binding of UNG to DNA (74).

We have also attempted to use similar strategy for staining of uracil-DNA in a mammalian cellular background. To this end, we have transfected [*ung*−/−] MEF cells with plasmid DNA produced by CJ236 [*dut*−, *ung*−] *E. coli* cells. Such plasmid DNA contains ∼6580 uracil/million bases (45) and its tolerance in the MEF cells is ensured by lack of

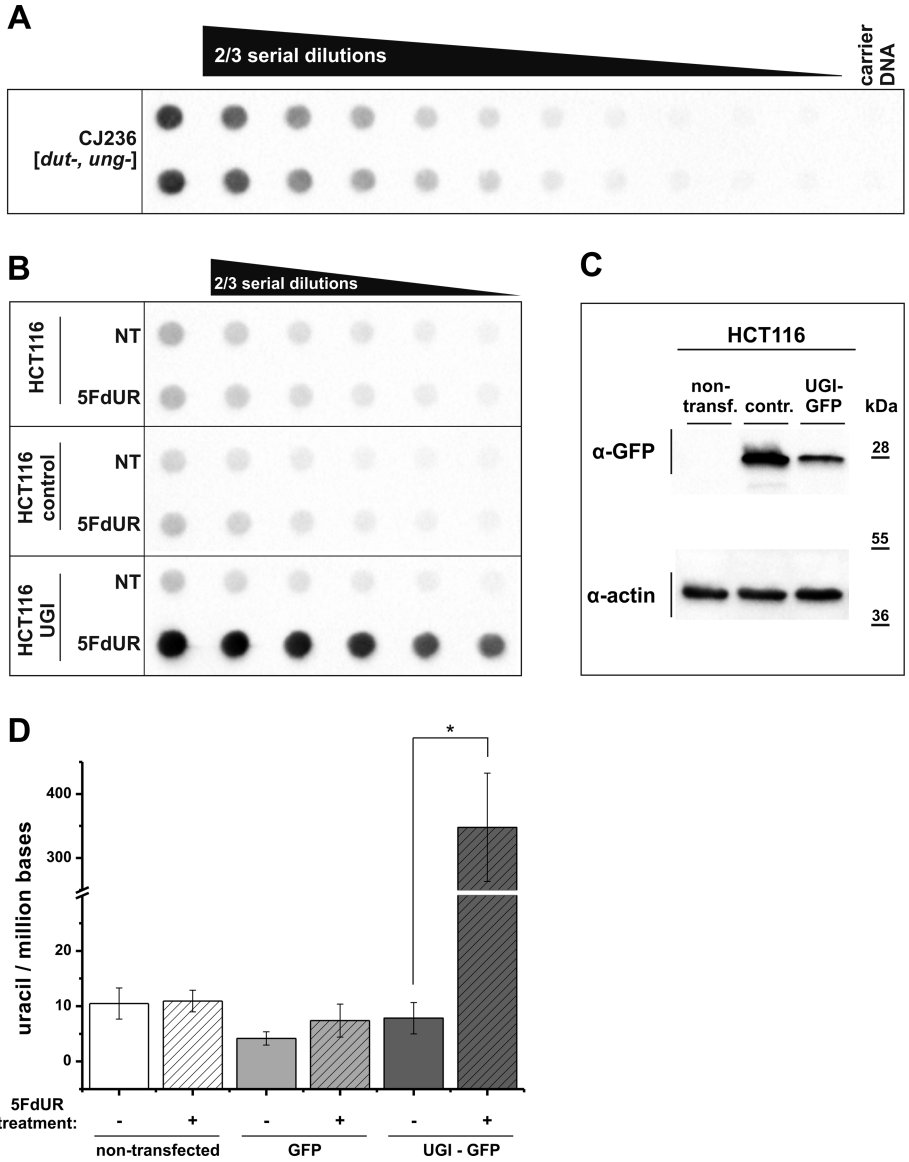

Figure 7. Dot-blot assay for measuring genomic uracil levels of HCT116 cells after treatment with *de novo* thymidylate biosynthesis pathway inhibitors and UNG inhibition. (**A**) CJ236 [*dut−, ung−*] *Escherichia coli* genomic DNA was used as standard for the dot-blot assay. (**B**) Genomic uracil levels of 5FdUR treated and non-treated HCT116 cells were measured in the contex of endogenous UNG inhibition with UGI expression. (**C**) Bar graph shows the uracil moieties/million bases of each sample (mean values ± the standard errors of the mean). 5FdUR treatment led to significantly elevated uracil levels only in cells also expressing UGI (* = $P < 0.05$) when compared to non-treated cells. Calculations were based on four independent datasets ($n = 4$). (**D**) Western blot showing GFP expression of cells transfected by the UGI-GFP vector and the empty vector used as a control (only expressing GFP). The membrane were also developed against actin as a loading control.

the UNG enzyme in this mouse cell line. In this experiment, we make use of a well-known artifactual effect of plasmid transfection, namely that upon using high amount of plasmid DNA during transfection, plasmid aggregates can occur within the cells (75) (Figure 9A). We therefore worked out our transfection experimental conditions such that to allow the potential accumulation of intracellular plasmid aggregates (cf 'Materials and Methods' section). Figure 9B leftmost panels clearly show DAPI staining of these plasmid aggregates indicated by white asterisks. In case of non-transfected cells (negative control), no plasmid aggregates can be observed (Supplementary Figure S4). On Figure 8, upper panels with cells transfected with uracil-rich plasmid show positive reaction with our uracil sensor molecules, either via Flag-tag or via direct DsRed detection. Additionally as a negative control, cells transfected by normal plasmid (produced in XL1-Blue cells) show no staining. Further controls are also shown in Supplementary Figure S4 where the lack of uracil-specific staining can be observed when the UNG-inhibitor UGI was applied.

Although these results clearly show the possibility of using our method for *in situ* microscopic detection, the sensitivity of this staining approach needs to be developed further to allow detection of lower uracil levels. Additional enhancement of the signal might be expected from the application of brighter and more stable fluorophores, such as

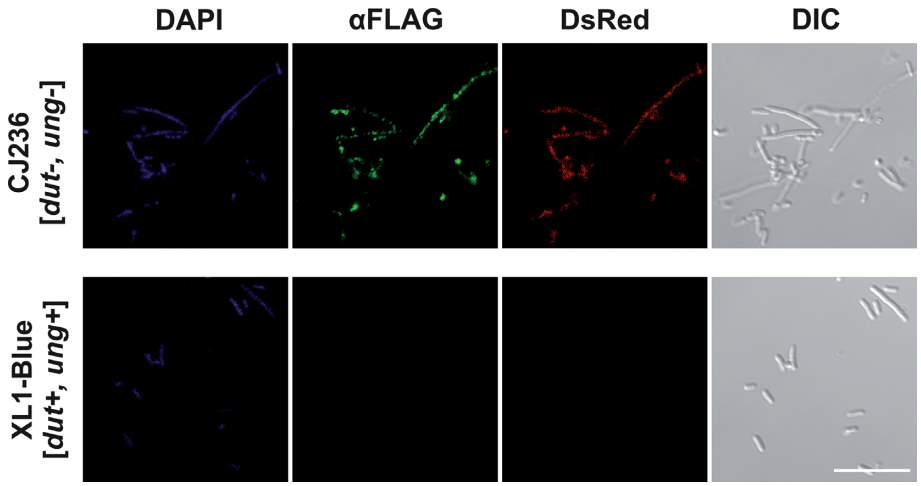

Figure 8. *In situ* genomic uracil detection in *Escherichiacoli* using an immunocytochemistry approach. The genomic uracil content of CJ236 [*dut*−, *ung*−] *E. coli* cells was visualized with the Flag-ΔUNG-DsRed construct. As a negative control, XL1-Blue cells [*dut*+, *ung*+] were also used in the same staining procedure. Only the CJ236 [*dut*−, *ung*−] *E. coli* sample showed staining, either detected directly through the signal of DsRed (red) or through the FLAG epitope tag (green). DAPI was used to counterstain DNA. Scale bar represents 10 μm.

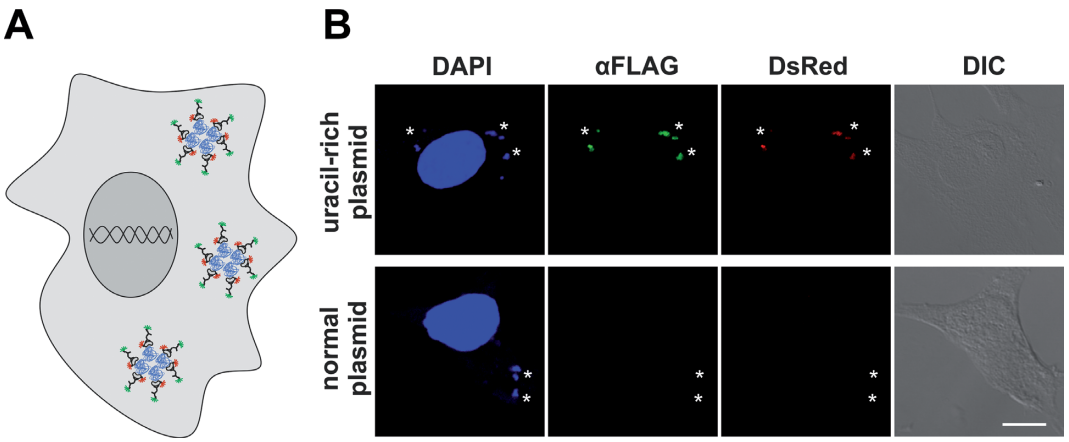

Figure 9. Detection of uracil-rich and normal plasmid DNA aggregates in MEF [ung-/-] cells. (**A**) Schematic image of the cytoplasmic plasmid aggregates visualized by the Flag-ΔUNG-DsRed construct. (**B**) Asterisks (*) show plasmid aggregates. Only cells transfected with uracil-rich plasmids could be visualized both through the DsRed (red) tag and the FLAG epitope tag (green). The DAPI staining is oversaturated to show the faint DAPI positive plasmid aggregates in the cytoplasm. Scale bar represents 10 μm.

nanocrystals or quantum dots (cf (76)). Our present method forms the basis for such further developments.

CONCLUSION

The workflow of the novel method described in this study is shown schematically on Figure 10, both for the *in vitro* dot-blot based detection (Figure 10A) and the *in situ* immunocytochemical approach used for cellular detection (Figure 10B). Uracil moieties are recognized by a catalytically inactive UNG sensor protein and the readout signal is enhanced by using primary and secondary sets of antibodies. The recently described novel mycobacterial uracil-DNA binding protein may also serve as an alternative sensor framework (77). The method is straightforward, easy to use and can be applied in a high-throughput manner to analyze DNA from any organism. It does not require expensive instruments or complex know-how, facilitating its easy implementation in any basic molecular biology laboratory. Elevated genomic deoxyuridine content of cells from diverse genetic background and/or treated with different drugs can be demonstrated *in situ*, within the cell. Direct detection is possible through the DsRed-tagged construct, or antibodies may be used for signal enhancement through the different epitope tags.

Direct comparisons between values for genomic uracil content obtained by different methods are far from straightforward (cf e.g. data for genomic DNA from CJ236 [*dut*−, *ung*−] *E. coli* cells range between 3000–18 000/million bases (38,43,45)). Hence, the truly reliable approach is to compare relative differences induced by different cellular stages, environment and/or drug treatments using the same method. Our present method is optimal for detecting such differences due to its ease of application, robustness and amenity to high-throughput studies. With the dot-blot based assay, comparative data between different organisms and different cellular conditions are obtained fast and in a quantita-

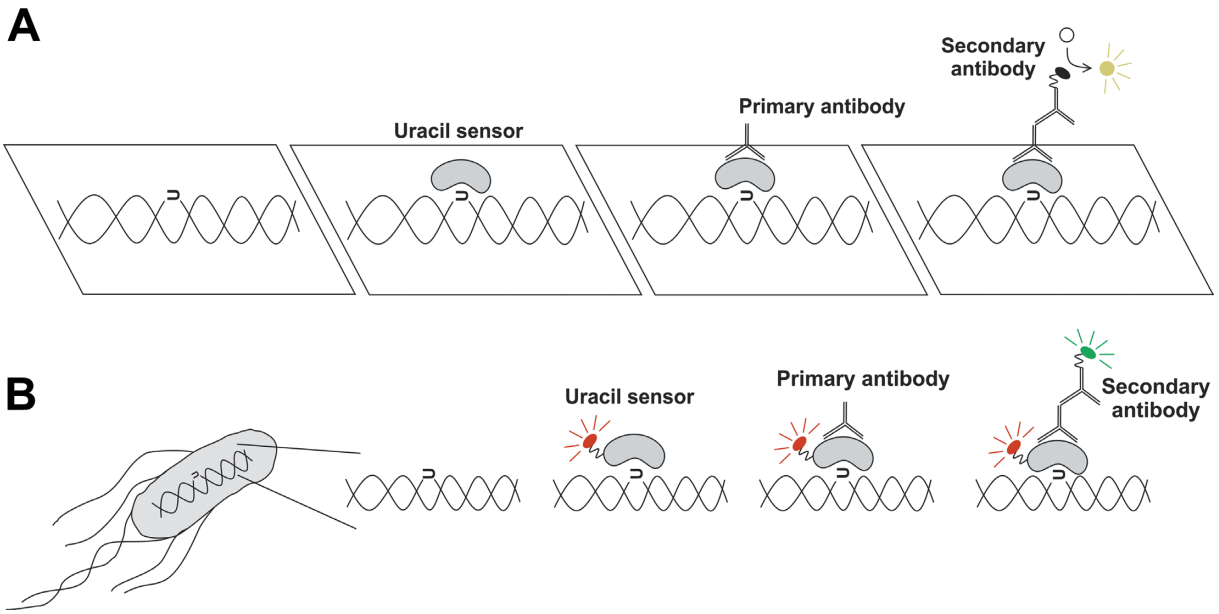

Figure 10. Schematics summarizing the developed *in vitro* quantification and *in situ* detection method. (**A**) Schematic image of the applied *in vitro* quantification through a dot-blot approach. (**B**) Schematic image of the applied immunocytochemical approach for *in situ* uracil detection.

tive manner. Finally, a major further significance and novelty in our presently proposed study is that it provides potential for *in situ* detection of uracil-DNA within cells. To our knowledge at least, such *in situ* detection has not yet been described for uracil in DNA. Considering the numerous physiological events and pathological conditions where uracil levels in DNA are modified, our method presents a biologically relevant tool for assessing the composition of genomic DNA and its putative alterations during different cellular conditions. In this respect, we wish to point out that our results obtained with UGI-expressing HCT116 cells simultaneously treated with 5FdUR indicates that the combined perturbation of base-excision repair and *de novo* thymidylate biosynthesis leads to a synergistic cellular response elevating genomic uracil levels. The significance of our proposed technique is further enhanced by the fact that uracil detection has yet escaped the highly powerful single-molecule real-time sequencing (SMRT) technology, as well.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENT

Human uracil-DNA glycosylase 2 (UNG2) cDNA was a generous gift of Professor Salvatore Caradonna (Department of Molecular Biology, University of Medicine and Dentistry of New Jersey). The vector expressing the human codon optimalized UGI along with GFP (pLGC-hUgi) was a kind gift of Michael D. Wyatt (South Carolina College of Pharmacy, University of South Carolina).

FUNDING

Hungarian Scientific Research Fund OTKA [NK 84008, K109486]; Baross Program of the New Hungary Development Plan [3DSTRUCT, OMFB-00266/2010 REG-KM-09-1-2009-0050]; Hungarian Academy of Sciences (MedinProt program) [TTK IF-28/2012]; ICGEB Research Grant [CRP/HUN14-01]; New Hungary Development Plan KMR_12-1-2012-0140. Funding for open access charge: Hungarian Academy of Sciences; Hungarian Scientific Research Fund OTKA.
Conflict of interest statement. None declared.

REFERENCES

1. Visnes,T., Doseth,B., Pettersen,H.S., Hagen,L., Sousa,M.M., Akbari,M., Otterlei,M., Kavli,B., Slupphaug,G. and Krokan,H.E. (2009) Uracil in DNA and its processing by different DNA glycosylases. *Philos. Trans. R Soc. Lond. B Biol. Sci.*, **364**, 563–568.
2. Krokan,H.E., Drablos,F. and Slupphaug,G. (2002) Uracil in DNA–occurrence, consequences and repair. *Oncogene*, **21**, 8935–8948.
3. Wardle,J., Burgers,P.M., Cann,I.K., Darley,K., Heslop,P., Johansson,E., Lin,L.J., McGlynn,P., Sanvoisin,J., Stith,C.M. *et al.* (2008) Uracil recognition by replicative DNA polymerases is limited to the archaea, not occurring with bacteria and eukarya. *Nucleic Acids Res.*, **36**, 705–711.
4. Vertessy,B.G. and Toth,J. (2009) Keeping uracil out of DNA: physiological role, structure and catalytic mechanism of dUTPases. *Acc. Chem. Res.*, **42**, 97–106.
5. Takacs,E., Grolmusz,V.K. and Vertessy,B.G. (2004) A tradeoff between protein stability and conformational mobility in homotrimeric dUTPases. *FEBS Lett.*, **566**, 48–54.
6. Vertessy,B.G., Zalud,P., Nyman,P.O. and Zeppezauer,M. (1994) Identification of tyrosine as a functional residue in the active site of Escherichia coli dUTPase. *Biochim. Biophys. Acta*, **1205**, 146–150.
7. Nagy,G.N., Leveles,I. and Vertessy,B.G. (2014) Preventive DNA repair by sanitizing the cellular (deoxy)nucleoside triphosphate pool. *FEBS J.*, **281**, 4207–4223.
8. Wallace,S.S. (2014) Base excision repair: a critical player in many games. *DNA Repair (Amst)*, **19**, 14–26.
9. Krokan,H.E. and Bjoras,M. (2013) Base excision repair. *Cold Spring Harb. Perspect. Biol.*, **5**, a012583.
10. Kavli,B., Sundheim,O., Akbari,M., Otterlei,M., Nilsen,H., Skorpen,F., Aas,P.A., Hagen,L., Krokan,H.E. and Slupphaug,G. (2002) hUNG2 is the major repair enzyme for removal of uracil from

U:A matches, U:G mismatches, and U in single-stranded DNA, with hSMUG1 as a broad specificity backup. *J. Biol. Chem.*, **277**, 39926–39936.
11. Panayotou,G., Brown,T., Barlow,T., Pearl,L.H. and Savva,R. (1998) Direct measurement of the substrate preference of uracil-DNA glycosylase. *J. Biol. Chem.*, **273**, 45–50.
12. Slupphaug,G., Eftedal,I., Kavli,B., Bharati,S., Helle,N.M., Haug,T., Levine,D.W. and Krokan,H.E. (1995) Properties of a recombinant human uracil-DNA glycosylase from the UNG gene and evidence that UNG encodes the major uracil-DNA glycosylase. *Biochemistry*, **34**, 128–138.
13. Lindahl,T., Ljungquist,S., Siegert,W., Nyberg,B. and Sperens,B. (1977) DNA N-glycosidases: properties of uracil-DNA glycosidase from Escherichia coli. *J. Biol. Chem.*, **252**, 3286–3294.
14. Bellamy,S.R. and Baldwin,G.S. (2001) A kinetic analysis of substrate recognition by uracil-DNA glycosylase from herpes simplex virus type 1. *Nucleic Acids Res.*, **29**, 3857–3863.
15. Eftedal,I., Guddal,P.H., Slupphaug,G., Volden,G. and Krokan,H.E. (1993) Consensus sequences for good and poor removal of uracil from double stranded DNA by uracil-DNA glycosylase. *Nucleic Acids Res.*, **21**, 2095–2101.
16. Nilsen,H., Yazdankhah,S.P., Eftedal,I. and Krokan,H.E. (1995) Sequence specificity for removal of uracil from U.A pairs and U.G mismatches by uracil-DNA glycosylase from Escherichia coli, and correlation with mutational hotspots. *FEBS Lett.*, **362**, 205–209.
17. Dizdaroglu,M., Karakaya,A., Jaruga,P., Slupphaug,G. and Krokan,H.E. (1996) Novel activities of human uracil DNA N-glycosylase for cytosine-derived products of oxidative DNA damage. *Nucleic Acids Res.*, **24**, 418–422.
18. Hatahet,Z., Kow,Y.W., Purmal,A.A., Cunningham,R.P. and Wallace,S.S. (1994) New substrates for old enzymes. 5-Hydroxy-2′-deoxycytidine and 5-hydroxy-2′-deoxyuridine are substrates for Escherichia coli endonuclease III and formamidopyrimidine DNA N-glycosylase, while 5-hydroxy-2′-deoxyuridine is a substrate for uracil DNA N-glycosylase. *J. Biol. Chem.*, **269**, 18814–18820.
19. Liu,P., Burdzy,A. and Sowers,L.C. (2002) Substrate recognition by a family of uracil-DNA glycosylases: UNG, MUG, and TDG. *Chem. Res. Toxicol.*, **15**, 1001–1009.
20. Warner,H.R. and Rockstroh,P.A. (1980) Incorporation and excision of 5-fluorouracil from deoxyribonucleic acid in Escherichia coli. *J. Bacteriol.*, **141**, 680–686.
21. Bharti,S.K. and Varshney,U. (2010) Analysis of the impact of a uracil DNA glycosylase attenuated in AP-DNA binding in maintenance of the genomic integrity in Escherichia coli. *Nucleic Acids Res.*, **38**, 2291–2301.
22. Krusong,K., Carpenter,E.P., Bellamy,S.R., Savva,R. and Baldwin,G.S. (2006) A comparative study of uracil-DNA glycosylases from human and herpes simplex virus type 1. *J. Biol. Chem.*, **281**, 4983–4992.
23. Pettersen,H.S., Sundheim,O., Gilljam,K.M., Slupphaug,G., Krokan,H.E. and Kavli,B. (2007) Uracil-DNA glycosylases SMUG1 and UNG2 coordinate the initial steps of base excision repair by distinct mechanisms. *Nucleic Acids Res.*, **35**, 3879–3892.
24. Mol,C.D., Arvai,A.S., Slupphaug,G., Kavli,B., Alseth,I., Krokan,H.E. and Tainer,J.A. (1995) Crystal structure and mutational analysis of human uracil-DNA glycosylase: structural basis for specificity and catalysis. *Cell*, **80**, 869–878.
25. Savva,R., McAuley-Hecht,K., Brown,T. and Pearl,L. (1995) The structural basis of specific base-excision repair by uracil-DNA glycosylase. *Nature*, **373**, 487–493.
26. Wilson,P.M., Danenberg,P.V., Johnston,P.G., Lenz,H.J. and Ladner,R.D. (2014) Standing the test of time: targeting thymidylate biosynthesis in cancer therapy. *Nat. Rev. Clin. Oncol.*, **11**, 282–298.
27. Skurnik,M., Hyytiainen,H., Happonen,L.J., Kiljunen,S., Datta,N., Mattinen,L., Williamson,K., Kristo,P., Szeliga,M., Kalin-Manttari,L. *et al.* (2012) Characterization of the genome, proteome and structure of Yersiniophage ΦR1-37. *J Virol.*, **86**, 12625–12642.
28. Kiljunen,S., Hakala,K., Pinta,E., Huttunen,S., Pluta,P., Gador,A., Lonnberg,H. and Skurnik,M. (2005) Yersiniophage phiR1–37 is a tailed bacteriophage having a 270 kb DNA genome with thymidine replaced by deoxyuridine. *Microbiology*, **151**, 4093–4102.
29. Langridge,R. and Marmur,J. (1964) X-ray diffraction study of a DNA which contains uracil. *Science*, **143**, 1450–1451.
30. Lozeron,H.A. and Szybalski,W. (1967) Incorporation of 5-fluorodeoxyuridine into the DNA of Bacillus subtilis phage PBS2 and its radiobiological consequences. *J. Mol. Biol.*, **30**, 277–290.
31. Pettersen,H.S., Galashevskaya,A., Doseth,B., Sousa,M.M., Sarno,A., Visnes,T., Aas,P.A., Liabakk,N.B., Slupphaug,G., Saetrom,P. *et al.* (2015) AID expression in B-cell lymphomas causes accumulation of genomic uracil and a distinct AID mutational signature. *DNA Repair (Amst)*, **25**, 60–71.
32. Maul,R.W. and Gearhart,P.J. (2010) AID and somatic hypermutation. *Adv. Immunol.*, **105**, 159–191.
33. Liu,M. and Schatz,D.G. (2009) Balancing AID and DNA repair during somatic hypermutation. *Trends Immunol.*, **30**, 173–181.
34. Yan,N., O'Day,E., Wheeler,L.A., Engelman,A. and Lieberman,J. (2011) HIV DNA is heavily uracilated, which protects it from autointegration. *Proc. Natl. Acad. Sci. U.S.A.*, **108**, 9244–9249.
35. Horvath,A., Bekesi,A., Muha,V., Erdelyi,M. and Vertessy,B.G. (2013) Expanding the DNA alphabet in the fruit fly: uracil enrichment in genomic DNA. *Fly (Austin)*, **7**, 23–27.
36. Muha,V., Horvath,A., Bekesi,A., Pukancsik,M., Hodoscsek,B., Merenyi,G., Rona,G., Batki,J., Kiss,I., Jankovics,F. *et al.* (2012) Uracil-containing DNA in Drosophila: stability, stage-specific accumulation, and developmental involvement. *PLoS Genet.*, **8**, e1002738.
37. Galashevskaya,A., Sarno,A., Vagbo,C.B., Aas,P.A., Hagen,L., Slupphaug,G. and Krokan,H.E. (2013) A robust, sensitive assay for genomic uracil determination by LC/MS/MS reveals lower levels than previously reported. *DNA Repair (Amst)*, **12**, 699–706.
38. Atamna,H., Cheung,I. and Ames,B.N. (2000) A method for detecting abasic sites in living cells: age-dependent changes in base excision repair. *Proc. Natl. Acad. Sci. U.S.A.*, **97**, 686–691.
39. Blount,B.C. and Ames,B.N. (1994) Analysis of uracil in DNA by gas chromatography-mass spectrometry. *Analyt. Biochem.*, **219**, 195–200.
40. Blount,B.C., Mack,M.M., Wehr,C.M., MacGregor,J.T., Hiatt,R.A., Wang,G., Wickramasinghe,S.N., Everson,R.B. and Ames,B.N. (1997) Folate deficiency causes uracil misincorporation into human DNA and chromosome breakage: implications for cancer and neuronal damage. *Proc. Natl. Acad. Sci. U.S.A.*, **94**, 3290–3295.
41. Chango,A., Abdel Nour,A.M., Niquet,C. and Tessier,F.J. (2009) Simultaneous determination of genomic DNA methylation and uracil misincorporation. *Med. Princ. Pract.*, **18**, 81–84.
42. Mashiyama,S.T., Courtemanche,C., Elson-Schwab,I., Crott,J., Lee,B.L., Ong,C.N., Fenech,M. and Ames,B.N. (2004) Uracil in DNA, determined by an improved assay, is increased when deoxynucleosides are added to folate-deficient cultured human lymphocytes. *Analyt. Biochem.*, **330**, 58–69.
43. Lari,S.U., Chen,C.Y., Vertessy,B.G., Morre,J. and Bennett,S.E. (2006) Quantitative determination of uracil residues in Escherichia coli DNA: Contribution of ung, dug, and dut genes to uracil avoidance. *DNA Repair (Amst)*, **5**, 1407–1420.
44. Shalhout,S., Haddad,D., Sosin,A., Holland,T.C., Al-Katib,A., Martin,A. and Bhagwat,A.S. (2014) Genomic uracil homeostasis during normal B cell maturation and loss of this balance during B cell cancer development. *Mol. Cell. Biol.*, **34**, 4019–4032.
45. Horvath,A. and Vertessy,B.G. (2010) A one-step method for quantitative determination of uracil in DNA by real-time PCR. *Nucleic Acids Res.*, **38**, e196.
46. Wu,D., Chen,L., Sun,Q., Wu,X., Jia,S. and Meng,A. (2014) Uracil-DNA glycosylase is involved in DNA demethylation and required for embryonic development in the zebrafish embryo. *J. Biol. Chem.*, **289**, 15463–15473.
47. Ladopoulos,V., Hofemeister,H., Hoogenkamp,M., Riggs,A.D., Stewart,A.F. and Bonifer,C. (2013) The histone methyltransferase KMT2B is required for RNA polymerase II association and protection from DNA methylation at the MagohB CpG island promoter. *Mol. Cell. Biol.*, **33**, 1383–1393.
48. Cliffe,L.J., Hirsch,G., Wang,J., Ekanayake,D., Bullard,W., Hu,M., Wang,Y. and Sabatini,R. (2012) JBP1 and JBP2 proteins are Fe2+/2-oxoglutarate-dependent dioxygenases regulating hydroxylation of thymidine residues in trypanosome DNA. *J. Biol. Chem.*, **287**, 19886–19895.
49. Moriel-Carretero,M. and Aguilera,A. (2010) A postincision-deficient TFIIH causes replication fork breakage and uncovers alternative

Rad51- or Pol32-mediated restart mechanisms. *Mol. Cell*, **37**, 690–701.
50. Soultanakis,R.P., Melamede,R.J., Bespalov,I.A., Wallace,S.S., Beckman,K.B., Ames,B.N., Taatjes,D.J. and Janssen-Heininger,Y.M. (2000) Fluorescence detection of 8-oxoguanine in nuclear and mitochondrial DNA of cultured cells using a recombinant Fab and confocal scanning laser microscopy. *Free Radic. Biol. Med.*, **28**, 987–998.
51. Terasaki,Y., Akuta,T., Terasaki,M., Sawa,T., Mori,T., Okamoto,T., Ozaki,M., Takeya,M. and Akaike,T. (2006) Guanine nitration in idiopathic pulmonary fibrosis and its implication for carcinogenesis. *Am. J. Respir. Crit. Care Med.*, **174**, 665–673.
52. Bekesi,A., Zagyva,I., Hunyadi-Gulyas,E., Pongracz,V., Kovari,J., Nagy,A.O., Erdei,A., Medzihradszky,K.F. and Vertessy,B.G. (2004) Developmental regulation of dUTPase in Drosophila melanogaster. *J. Biol. Chem.*, **279**, 22362–22370.
53. Rona,G., Borsos,M., Ellis,J.J., Mehdi,A.M., Christie,M., Kornyei,Z., Neubrandt,M., Toth,J., Bozoky,Z., Buday,L. *et al.* (2014) Dynamics of re-constitution of the human nuclear proteome after cell division is regulated by NLS-adjacent phosphorylation. *Cell cycle*, **13**, 3551–3564.
54. Kaiser,S.M. and Emerman,M. (2006) Uracil DNA glycosylase is dispensable for human immunodeficiency virus type 1 replication and does not contribute to the antiviral effects of the cytidine deaminase Apobec3G. *J. Virol.*, **80**, 875–882.
55. Andersen,S., Ericsson,M., Dai,H.Y., Pena-Diaz,J., Slupphaug,G., Nilsen,H., Aarset,H. and Krokan,H.E. (2005) Monoclonal B-cell hyperplasia and leukocyte imbalance precede development of B-cell malignancies in uracil-DNA glycosylase deficient mice. *DNA Repair (Amst)*, **4**, 1432–1441.
56. Bennett,S.E., Chen,C.Y. and Mosbaugh,D.W. (2004) Escherichia coli nucleoside diphosphate kinase does not act as a uracil-processing DNA repair nuclease. *Proc. Natl. Acad. Sci. U.S.A.*, **101**, 6391–6396.
57. Sun,Q., Yu,X.C. and Margolin,W. (1998) Assembly of the FtsZ ring at the central division site in the absence of the chromosome. *Mol. Microbiol.*, **29**, 491–503.
58. Dolbeare,F. (1995) Bromodeoxyuridine: a diagnostic tool in biology and medicine, Part I: Historical perspectives, histochemical methods and cell kinetics. *Histochem. J.*, **27**, 339–369.
59. Pearl,L.H. (2000) Structure and function in the uracil-DNA glycosylase superfamily. *Mutat. Res.*, **460**, 165–181.
60. Drohat,A.C., Jagadeesh,J., Ferguson,E. and Stivers,J.T. (1999) Role of electrophilic and general base catalysis in the mechanism of Escherichia coli uracil DNA glycosylase. *Biochemistry*, **38**, 11866–11875.
61. Otterlei,M., Warbrick,E., Nagelhus,T.A., Haug,T., Slupphaug,G., Akbari,M., Aas,P.A., Steinsbekk,K., Bakke,O. and Krokan,H.E. (1999) Post-replicative base excision repair in replication foci. *EMBO J.*, **18**, 3834–3844.
62. Nagelhus,T.A., Haug,T., Singh,K.K., Keshav,K.F., Skorpen,F., Otterlei,M., Bharati,S., Lindmo,T., Benichou,S., Benarous,R. *et al.* (1997) A sequence in the N-terminal region of human uracil-DNA glycosylase with homology to XPA interacts with the C-terminal part of the 34-kDa subunit of replication protein A. *J. Biol. Chem.*, **272**, 6561–6566.
63. Torseth,K., Doseth,B., Hagen,L., Olaisen,C., Liabakk,N.B., Graesmann,H., Durandy,A., Otterlei,M., Krokan,H.E., Kavli,B. *et al.* (2012) The UNG2 Arg88Cys variant abrogates RPA-mediated recruitment of UNG2 to single-stranded DNA. *DNA Repair (Amst)*, **11**, 559–569.
64. Scaramozzino,N., Sanz,G., Crance,J.M., Saparbaev,M., Drillien,R., Laval,J., Kavli,B. and Garin,D. (2003) Characterisation of the substrate specificity of homogeneous vaccinia virus uracil-DNA glycosylase. *Nucleic Acids Res.*, **31**, 4950–4957.
65. Santi,D.V., McHenry,C.S., Raines,R.T. and Ivanetich,K.M. (1987) Kinetics and thermodynamics of the interaction of 5-fluoro-2′-deoxyuridylate with thymidylate synthase. *Biochemistry*, **26**, 8606–8613.
66. Wyatt,M.D. and Wilson,D.M. 3rd. (2009) Participation of DNA repair in the response to 5-fluorouracil. *Cell. Mol. Life Sci.*, **66**, 788–799.
67. Andersen,S., Heine,T., Sneve,R., Konig,I., Krokan,H.E., Epe,B. and Nilsen,H. (2005) Incorporation of dUMP into DNA is a major source of spontaneous DNA damage, while excision of uracil is not required for cytotoxicity of fluoropyrimidines in mouse embryonic fibroblasts. *Carcinogenesis*, **26**, 547–555.
68. Luo,Y., Walla,M. and Wyatt,M.D. (2008) Uracil incorporation into genomic DNA does not predict toxicity caused by chemotherapeutic inhibition of thymidylate synthase. *DNA Repair (Amst)*, **7**, 162–169.
69. dos Santos,G., Schroeder,A.J., Goodman,J.L., Strelets,V.B., Crosby,M.A., Thurmond,J., Emmert,D.B., Gelbart,W.M. and FlyBase Consortium. (2015) FlyBase: introduction of the Drosophila melanogaster Release 6 reference genome assembly and large-scale migration of genome annotations. *Nucleic Acids Res.*, **43**, D690–D697.
70. Blackledge,G. (1998) New developments in cancer treatment with the novel thymidylate synthase inhibitor raltitrexed ('Tomudex'). *Br. J. Cancer*, **77**(Suppl. 2), 29–37.
71. Rajagopalan,P.T., Zhang,Z., McCourt,L., Dwyer,M., Benkovic,S.J. and Hammes,G.G. (2002) Interaction of dihydrofolate reductase with methotrexate: ensemble and single-molecule kinetics. *Proc. Natl. Acad. Sci. U.S.A.*, **99**, 13481–13486.
72. Rancourt,S.L. and Walker,V.K. (1990) Kinetic characterization of dihydrofolate reductase from Drosophila melanogaster. *Biochem. Cell Biol.*, **68**, 1075–1082.
73. Bennett,S.E., Schimerlik,M.I. and Mosbaugh,D.W. (1993) Kinetics of the uracil-DNA glycosylase/inhibitor protein association. Ung interaction with Ugi, nucleic acids, and uracil compounds. *J. Biol. Chem.*, **268**, 26879–26885.
74. Lundquist,A.J., Beger,R.D., Bennett,S.E., Bolton,P.H. and Mosbaugh,D.W. (1997) Site-directed mutagenesis and characterization of uracil-DNA glycosylase inhibitor protein. Role of specific carboxylic amino acids in complex formation with Escherichia coli uracil-DNA glycosylase. *J. Biol. Chem.*, **272**, 21408–21419.
75. Zabner,J., Fasbender,A.J., Moninger,T., Poellinger,K.A. and Welsh,M.J. (1995) Cellular and molecular barriers to gene transfer by a cationic lipid. *J. Biol. Chem.*, **270**, 18997–19007.
76. Beke,D., Szekrényes,Z., Pálfi,D., Róna,G., Balogh,I., Maák,P.A., Katona,G., Czigány,Z., Kamarás,K., Rózsa,B. *et al.* (2013) Silicon carbide quantum dots for bioimaging. *J. Mater. Res.*, **28**, 205–209.
77. Sang,P.B., Srinath,T., Patil,A.G., Woo,E.J. and Varshney,U. (2015) A unique uracil-DNA binding protein of the uracil DNA glycosylase superfamily. *Nucleic Acids Res.*, doi:10.1093/nar/gkv854.
78. Requena,C.E., Perez-Moreno,G., Ruiz-Perez,L.M., Vidal,A.E. and Gonzalez-Pacanowska,D. (2014) The NTP pyrophosphatase DCTPP1 contributes to the homoeostasis and cleansing of the dNTP pool in human cells. *Biochem. J.*, **459**, 171–180.
79. Zhang,J. and Inouye,M. (2002) MazG, a nucleoside triphosphate pyrophosphohydrolase, interacts with Era, an essential GTPase in Escherichia coli. *J. Bacteriol.*, **184**, 5323–5329.