DNA Repair 25 (2015) 60–71

Contents lists available at ScienceDirect

DNA Repair

journal homepage: www.elsevier.com/locate/dnarepair

AID expression in B-cell lymphomas causes accumulation of genomic uracil and a distinct AID mutational signature

Henrik Sahlin Pettersen[a,c,d], Anastasia Galashevskaya[a], Berit Doseth[a], Mirta M.L. Sousa[a,e], Antonio Sarno[a], Torkild Visnes[a,1], Per Arne Aas[a], Nina-Beate Liabakk[a], Geir Slupphaug[a,e], Pål Sætrom[a,b], Bodil Kavli[a,*], Hans E. Krokan[a,*]

[a] Department of Cancer Research and Molecular Medicine, Norwegian University of Science and Technology, NO-7491 Trondheim, Norway
[b] Department of Computer and Information Science, Norwegian University of Science and Technology, NO-7491 Trondheim, Norway
[c] Liaison Committee between the Central Norway Regional Health Authority (RHA) and the Norwegian University of Science and Technology (NTNU), Trondheim, Norway
[d] St. Olav's Hospital, Trondheim University Hospital, NO-7006 Trondheim, Norway
[e] The Proteomics and Metabolomics Core Facility (PROMEC) at NTNU, NO-7491 Trondheim, Norway

ARTICLE INFO

Article history:
Received 10 September 2014
Received in revised form 6 November 2014
Accepted 17 November 2014
Available online 24 November 2014

Keywords:
Activation-induced cytidine deaminase (AID)
Genomic uracil
B-cell lymphoma
Base excision repair
Kataegis
Mutational signature

ABSTRACT

The most common mutations in cancer are C to T transitions, but their origin has remained elusive. Recently, mutational signatures of APOBEC-family cytosine deaminases were identified in many common cancers, suggesting off-target deamination of cytosine to uracil as a common mutagenic mechanism. Here we present evidence from mass spectrometric quantitation of deoxyuridine in DNA that shows significantly higher genomic uracil content in B-cell lymphoma cell lines compared to non-lymphoma cancer cell lines and normal circulating lymphocytes. The genomic uracil levels were highly correlated with AID mRNA and protein expression, but not with expression of other APOBECs. Accordingly, AID knockdown significantly reduced genomic uracil content. B-cells stimulated to express endogenous AID and undergo class switch recombination displayed a several-fold increase in total genomic uracil, indicating that B cells may undergo widespread cytosine deamination after stimulation. In line with this, we found that clustered mutations (*kataegis*) in lymphoma and chronic lymphocytic leukemia predominantly carry AID-hotspot mutational signatures. Moreover, we observed an inverse correlation of genomic uracil with uracil excision activity and expression of the uracil-DNA glycosylases UNG and SMUG1. In conclusion, AID-induced mutagenic U:G mismatches in DNA may be a fundamental and common cause of mutations in B-cell malignancies.

1. Introduction

The only sources of uracil in DNA were previously thought to be misincorporation of dUMP during DNA replication and spontaneous deamination of DNA cytosine. The discovery of activation-induced cytidine deaminase (AID, also called AICDA) and several other APOBEC-family enzymes as probable DNA-cytosine deaminases introduced a third possible source (reviewed in [1]).

AID was first identified following induction of class switch recombination (CSR) in the CH12 mouse B-cell lymphoma cell line and initially thought to be an RNA-editing enzyme [2]. However, evidence that AID was a DNA mutator in *Escherichia coli* [3] and its functional interaction with uracil-DNA glycosylase UNG in adaptive immunity [4–6], indicated that AID is a DNA-cytosine deaminase. Later several of the other known APOBEC-family enzymes were also found to be DNA-cytosine deaminases *in vitro* [7,8]. DNA cytosine deamination by APOBEC-family enzymes is a natural event in both the adaptive and innate immune systems, through targeted deamination of immunoglobulin (Ig) genes by AID and deamination of viral DNA by APOBEC enzymes, respectively [7]. Despite their important physiological functions, these host defense mechanisms entail a high risk of potentially carcinogenic off-target genomic mutagenesis. Recent high-throughput sequencing of large numbers of human cancer genomes showed that mutations at cytosine residues, particularly C to T transitions, are the most prevalent

* Corresponding authors. Tel.: +47 72 57 30 74/+47 72 573221; fax: +47 72 57 64 00.
E-mail addresses: bodil.kavli@ntnu.no (B. Kavli), hans.krokan@ntnu.no (H.E. Krokan).
[1] Present address: Science for Life Laboratory, Division of Translational Medicine and Chemical Biology, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, S-17121 Stockholm, Sweden.

http://dx.doi.org/10.1016/j.dnarep.2014.11.006

mutations in human cancer, highlighting enzymatic deamination of cytosine to uracil as a potential source of mutagenesis [9–11]. However, the actual uracil level in normal and various cancer genomes has remained elusive.

Here, a sensitive LC/MS/MS-based method for quantification of genomic 2′-deoxyuridine (dUrd) was applied to demonstrate that B-cell lymphoma cell lines contain several-fold increased levels of genomic uracil compared to normal human lymphocytes and non-lymphoma cell lines. Genomic uracil content correlated with AID protein expression but not with other APOBEC enzymes. In accordance with AID-generated uracil, we found that regions of clustered mutations (*kataegis*) in lymphoma and chronic lymphocytic leukemia (CLL) have a distinct AID-hotspot mutational signature. Importantly, we also show that uracil excision capacity and expression of the uracil-DNA glycosylases UNG and SMUG1 correlated negatively with genomic uracil levels and to some extent diminished the effect of AID. This study provides direct mechanistic evidence for genomic uracil accumulation due to enzymatic DNA cytosine deamination in human cancers.

2. Materials and methods

2.1. Primary cells, cell lines, cultivation, and reagents

Human cell lines HeLaS3 (ATCC CCL-2.2™), HEK293T (ATCC CRL-11268™), and U2OS (ATCC HTB-96™) were from ATCC. L428 (DSMZ ACC 197), DU145 (DSMZ ACC 261), KARPAS422 (DSMZ ACC 32), T24 (DSMZ ACC 376), DOHH2 (DSMZ ACC 47), SUDHL4 (DSMZ ACC 4956), JJN3 (DSMZ ACC 541), SUDHL5 (DSMZ ACC 571), SUDHL6 (DSMZ ACC 572 6), RAMOS (DSMZ ACC 603), RL (DSMZ ACC 613), DAUDI (DSMZ ACC 78 5), A431 (DSMZ ACC 91) were from DSMZ. OCILY3 was a gift from Dr. L.M. Staudt, Metabolism Branch, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA. Peripheral blood mononuclear cells (PMBCs) were purified from buffy coats from three healthy blood donors using the Lymphoprep™ (Progen) kit according to the manufacturer's protocol. Human B-lymphocytes were purified from buffy coats from three healthy blood donors using a negative selection kit from StemCell Technologies according to the manufacturer's protocol. HeLaS3, HEK293T, T24, A431, DU145, and U2OS cells were cultured in DMEM (4500 mg/l glucose) with 10% FCS, 0.03% L-glutamine, 0.1 mg/ml gentamicin and 2.3 μg/ml fungizone at 37 °C and 5% CO_2. DAUDI, DOHH-2, KARPAS, RAMOS, SU-DHL-4, SU-DHL-6, OCILY-3, L-428, RL, SU-DHL-5, and JJN3 cells were cultured in RPMI-1640 with 4500 mg/l glucose, 0.03% L-glutamine, Pen-Strep (1× final), 0.1 mg/ml gentamicin, and 2.3 μg/ml fungizone, and 20% heat inactivated (56 °C, 20 min) FCS at 37 °C and 5% CO_2. For quantitative rtPCR and uracil measurements cells were harvested at densities between 750 000 and 2 million cells/ml.

Cell doubling times for suspension cells were measured using a *Countess*® cell counter (Invitrogen) by two parallel daily measurements for three to five day periods from cell densities of 50 000–200 000 cells/ml to one to three million cells/ml. For adherent cells, doubling time was measured in 96 well plates (3–6 parallel wells; starting density 50 000 cells/ml) for a three day period by daily fluorescent measurement of resazurin (Sigma) metabolism according to the manufacturer's protocol. Doubling times were calculated by exponential regression.

SUDHL5 AID knockdown and control cells were made using Open Biosystem TransLenti Viral Packaging Mix, pTRIPZ AICDA shRNA (RHS4741-EG57379; vectors V2THS_58282, 58283, and 58319) or pTRIPZ non-silencing control vector according to the manufacturer's protocol. Briefly, lentiviruses were produced in HEK293T cells, and then supernatant from three consecutive days 48 h after HEK293T transfection were used to infect SUDHL5 cells. Infected SUDHL5 cells were amplified for another 48 h and then selected with 2 μg/ml puromycin for 30 days. Expression was induced with 1 μg/ml doxycycline.

CH12F3 AID-EYFP and EYFP stable transfectants, confocal microscopy, and stimulation experiments were described previously [12]. CH12F3 cells (2×10^6 cells/ml) were cultured in RPMI medium, with 10% heat-inactivated fetal calf serum, 0.03% L-glutamine, 50 μM β-mercaptoethanol, 1 mM Na-pyruvate, 0.1 mg/ml penicillin/streptomycin, 2.3 μg/ml fungizone, and 1.0 mg/ml G418. CH12F3 cells were stimulated to undergo class switch recombination by adding 10 ng/ml mouse recombinant IL-4 (Peprotech), 2 μg/ml anti-mouse CD40 monoclonal antibody (BD Biosciences) and 1 ng/ml human TGF-β1 (Peprotech) and harvested 48 h post stimulation for DNA and protein isolation. Western analysis of AID protein expression was performed using mouse anti-AID monoclonal antibody no. 39-2500, clone ZA001, 500 μg/ml (Invitrogen). Nuclear extracts from synchronized HeLa cells were prepared essentially as described [13,14].

2.2. RNA isolation and quantitative real-time PCR (qRT-PCR)

Total RNA for mRNA analysis was prepared using the mirVana miRNA isolation kit (Ambion) according to the manufacturer's instructions. RNA concentration and quality was measured on a NanoDrop ND-1000 UV–vis spectrophotometer. Total RNA (770 ng) was reverse transcribed for gene expression analysis using TaqMan reverse transcription reagents (Applied Biosystems). The following TaqMan gene expression assays (Applied Biosystems) were used: AID (Hs00757808_m1), UNG (Hs00422172_m1), SMUG1 (Hs04274951_m1), TDG (Hs00702322_s1), MBD4 (Hs00187498_m1), APOBEC1 (Hs00242340_m1), APOBEC2 (Hs00199012_m1), APOBEC3A (Hs00377444_m1), APOBEC3B (Hs00358981_m1), APOBEC3C (Hs00819353_m1), APOBEC3D (Hs00537163_m1), APOBEC3G (Hs00222415_m1), APOBEC3F (Hs01665324_m1), APOBEC3H (Hs00419665_m1), APOBEC4 (Hs00378929_m1), and GAPDH (Hs99999905_m1). Quantitative PCR was carried out on a Chromo4 (BioRad) real-time PCR detection system. Relative expression of mRNA was calculated by the ΔCt method using GAPDH as endogenous control. Regression analyses were done using GraphPad Prism where data were fitted by linear regression (log/linear(X) *vs.* log/linear(Y)) as indicated.

2.3. Quantification of uracil in DNA by LC/MS/MS

Genomic uracil was quantified as previously described [15]. Briefly, DNA was isolated by phenol:chloroform:isoamyl extraction, treated with alkaline phosphatase to remove free deoxyribonucleosides, and then enzymatically hydrolyzed to deoxyribonucleosides. Deoxyuridine (dU) was then separated from deoxycytidine (dC) by HPLC fractionation using a reverse-phase column with embedded weak acidic ion-pairing groups (2.1 mm × 150 mm, 5 μm, Primesep 200, SIELC technologies), using a water/acetonitrile gradient containing 0.1% formic acid. The dU fraction was finally analyzed by ESI-LC/MS/MS using a reverse phase column (2.1 mm × 150 mm, 3.5 μm, Zorbax SB-C18, Agilent Technologies), using a water/methanol gradient containing 0.1% formic acid on an API5000 triple quadrupole mass spectrometer (Applied Biosystems) in positive ionization mode. A small fraction of the hydrolyzed deoxyribonucleosides were quantified by LC/MS/MS in parallel and used to determine the amount of dU per 10^6 deoxyribonucleosides.