DNA REPAIR 6 (2007) 505–516

available at www.sciencedirect.com

ScienceDirect

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

Uracil in DNA—General mutagen, but normal intermediate in acquired immunity

Bodil Kavli, Marit Otterlei, Geir Slupphaug, Hans E. Krokan*

Department of Cancer Research and Molecular Medicine, Faculty of Medicine, Norwegian University of Science and Technology, N-7006 Trondheim, Norway

ARTICLE INFO

Article history:
Published on line 20 November 2006

Keywords:
Uracil in DNA
Uracil-DNA glycosylase
Somatic hypermutation
Class switch recombination
B cell lymphoma

ABSTRACT

Deamination of cytosine in DNA results in mutagenic U:G mispairs, whereas incorporation of dUMP leads to U:A pairs that may be genotoxic directly or indirectly. In both cases, uracil is mainly removed by a uracil-DNA glycosylase (UDG) that initiates the base excision repair pathway. The major UDGs are mitochondrial UNG1 and nuclear UNG2 encoded by the *UNG*-gene, and nuclear SMUG1. TDG and MBD4 remove uracil from special sequence contexts, but their roles remain poorly understood. UNG2 is cell cycle regulated and has a major role in post-replicative removal of incorporated uracils. UNG2 and SMUG1 are both important for prevention of mutations caused by cytosine deamination, and their functions are non-redundant. In addition, SMUG1 has a major role in removal of hydroxymethyl uracil from oxidized thymines. Furthermore, UNG-proteins and SMUG1 may have important functions in removal of oxidized cytosines, e.g. isodialuric acid, alloxan and 5-hydroxyuracil after exposure to ionizing radiation. UNG2 is also essential in the acquired immune response, including somatic hypermutation (SHM) required for antibody affinity maturation and class switch recombination (CSR) mediating new effector functions, e.g. from IgM to IgG. Upon antigen exposure B-lymphocytes express activation induced cytosine deaminase that generates U:G mispairs at the Ig locus. These result in GC to AT transition mutations upon DNA replication and apparently other mutations as well. Some of these may result from the generation of abasic sites and translesion bypass synthesis across such sites. SMUG1 can not complement UNG2 deficiency, probably because it works very inefficiently on single-stranded DNA and is down-regulated in B cells. In humans, UNG-deficiency results in the hyper IgM syndrome characterized by recurrent infections, lymphoid hyperplasia, extremely low IgG, IgA and IgE and elevated IgM. *Ung*$^{-/-}$ mice have a similar phenotype, but in addition display dysregulated cytokine production and develop B cell lymphomas late in life.

1. Introduction

The RNA base uracil may be present in DNA in small amounts as a result of cytosine deamination or misincorporation of dUMP during DNA replication. Such lesions are normally faithfully repaired by base excision repair (BER) initiated by a uracil-DNA glycosylase [1]. Contrary to expectations, mice deficient in the major uracil-DNA glycosylase encoded by the *Ung*-gene are viable, fertile and develop normally [2]. This is probably due to the existence of at least three alternative uracil-DNA glycosylases named SMUG1 (single strand selective monofunctional uracil-DNA glycosylase), TDG (thymine/uracil mismatch DNA glycosylase) and MBD4 (methyl binding domain 4 protein) [1]. These apparently have specialized functions but

* *Corresponding author*. Tel.: +47 72 573221; fax: +47 72 576400.
E-mail address: bodil.kavli@ntnu.no (B. Kavli).
1568-7864/$ – see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.dnarep.2006.10.014

Table 1 – Deaminations of cytosine in DNA per 24 h in human genome[a]

Reference	Events per C in ssDNA	Events per C in dsDNA	Human genome		
			All ssDNA	All dsDNA	0.1% ssDNA
[14]	1.7×10^{-5}	6.9×10^{-8}	4.4×10^{4}	180	224
[15]	0.9×10^{-5}	6.0×10^{-8}	2.3×10^{4}	156	179
[16]	Not done	2.75×10^{-8}	–	72	–

[a] Assuming human genome with 20% cytosine (2.6×10^{9} cytosines per diploid genome). Genetic data [15,16] assume no uracil repair in strain.

may also serve as backups for each other. Surprisingly, uracil in DNA and uracil-DNA glycosylase encoded by the *UNG* gene also have functions in the acquired immune system [3,4], as well as in the innate immune system in defense against retroviral infections [5]. Thus, there is important crosstalk between the ancient DNA repair mechanisms and the innate immune system, as well as the much younger acquired immune system that came with vertebrates. For the acquired immune system mismatch repair proteins and proteins required for double strand break repair are also involved. This demonstrates that the interactions between these defense systems are in fact quite extensive [6]. While Ung-deficient mice have no overt phenotype macroscopically, they develop B-cell lymphomas late in life [7,8]. In this review we will present the current knowledge about the origins of uracil in DNA and discuss functional characteristics of the quantitatively dominating glycosylases removing uracil and some closely related oxidative lesions. Finally, we will discuss the role of uracil and uracil-DNA glycosylase UNG2 in the Ig diversification process.

2. Uracil in DNA

Uracil in DNA is in several ways a rather special case among DNA lesions. First, it is a normal constituent in RNA but not in cellular DNA. Second, template uracil in U:G mispairs is a non-blocking, but 100% mutagenic lesion that results in mutation in one of the two daughter cells if it is replicated across. Incorporation of dUMP from the normal nucleotide dUTP results in U:A pairs that themselves are not mutagenic. However, chromosomal abasic sites resulting from uracil-removal are mutagenic and cytotoxic in eukaryotic cells [9]. Third, recent research has revealed that U:G mispairs are key intermediates in the modulation of Ig genes during somatic hypermutation (SHM) and class switch recombination (CSR) [3,4]. These processes are responsible for affinity maturation of antibodies and for generating a repertoire of antibodies with different effector functions, respectively [10]. Thus, cytosine in DNA is target for general spontaneous chemical deamination, and gene-specific enzymatic deamination as part of a regulated response. Finally, enzymatic deamination of cytosine to uracil in retroviral first strand cDNA may be an important defense mechanism against retroviral infection [11]. Whether targeted deaminations of cytosine in DNA may take place in other physiological processes is not known.

2.1. *Spontaneous deamination of cytosine to uracil*

Cytosine deamination has been estimated to take place at a rate of 60–500 events per human genome per day [1,12]. The various estimates of the deamination rate from earlier studies are presented in Table 1 (recalculated to deaminations per 24 h in the human genome). The uncertainty related to the rate is due to the 140–300-fold higher rate of deamination in single-stranded compared with double-stranded DNA, and the lack of quantitative knowledge about the proportion of single-stranded DNA resulting from replication, transcription and possibly other processes. These data therefore must be considered somewhat soft, but as outlined below chemical measurements and genetic assays are in reasonable agreement and indicate a deamination rate most likely in the range of 70–200 events/day in the human genome. In early studies, the rate of chemical deamination of cytosine at 95 °C was measured, extrapolated to 37 °C and found to be $\sim 1.7 \times 10^{-5}$ per cytosine per day in single-stranded DNA and $\sim 6.9 \times 10^{-8}$ per day in double-stranded DNA [13,14]. Using a mutant reversion assay in *ung*$^{-}$ *E. coli* cells the cytosine deamination rate constants at 37 °C for single- and double-stranded DNA at pH 7.4 were found to be 0.9×10^{-5} and about 6×10^{-8} per day, respectively [15]. A tRNA suppressor mutation assay gave similar, or somewhat lower estimates of cytosine deamination, 2.75×10^{-8} in dsDNA [16]. Thus, chemical and biological data are in fairly good agreement with only some 1.2–2.5-fold variation between chemical and biological measurements. The human genome contains approximately 20% cytosine and thus some 2.6×10^{9} cytosines per diploid human genome that contain 12.8×10^{9} bases. In summary, this amounts to about 70–180 deaminations per day in a double-stranded human genome, and some 224 in total if 0.1% of the genome is single-stranded at any time, probably a high estimate. The fraction of single-stranded DNA mainly represents DNA at the lagging strand of replication forks, the non-transcribed strand during transcription and to a smaller extent repair patches. Based on the estimates of the number of replicons (20,000 or less) and transcription bubbles (20,000 or less) active at any time in a cell, as well as a size of single-stranded regions at replication forks of 110–320 nt [17,18], the fraction of single-stranded DNA is most likely below 0.1%. Single-stranded DNA in the genome therefore does not contribute much to cytosine deamination, unless the genome contains single-stranded regions resulting from processes not considered by us. Our best estimate would therefore be that ~70–200 deaminations occur per cell per day.

2.2. *Deamination of cytosine to uracil by gamma radiation*

Gamma irradiation to DNA results in a large number of base lesions largely caused by hydroxyl radicals (•OH) from radiolysis of water [19]. Recently, generation of uracil by gamma