ACKNOWLEDGEMENTS

ABSTRACT

Uracil, a nitrogen base usually found in RNA, may be found in small quantities in genomic DNA. Genomic uracil may arise as a result of deamination of cytosine or misincorporation of deoxyuridine monophosphate (dUMP) instead of deoxythymidine monophosphate (dTMP) by replicative polymerases. Uracil is normally detected as a lesion in DNA and repaired by base excision repair. Failure to do so may lead to genetic mutations. Conversely, the presence of uracil in the genome can be beneficial: it is an important step in the adaptive immunity processes of class switch recombination and somatic hypermutation. In this sense, a lack of genomic uracil results in immunodeficiency. Its importance in these pathways merits genomic uracil as an interesting research topic, and several attempts have been made to quantitate uracil levels in DNA; however, there are discrepancies in the reported basal genomic uracil levels as well as some inherent problems in the methods employed to determine them.

In this thesis, an assay is introduced to measure genomic uracil levels. The assay involves pretreatment with alkaline phosphatase to minimize intracellular dUMP contamination, followed by DNA hydrolysis to deoxynucleotides by treatment with DNase I, snake venom diesterase, and alkaline phosphatase. Deoxyuridine is then separated from dC in a precursory HPLC step prior to the final separation and quantitation by a second HPLC coupled to an electrospray ionization-based triple quadrupole mass spectrometer in multiple reaction monitoring mode. The limits of detection and quantitation for this assay are comparable to the most sensitive assay in the literature: 2 and 5 fmol, respectively. Furthermore, deoxyuridine detection is linear across three orders of magnitude. The drawbacks found in similar methods, namely the use of uracil N-glycosylase to excise uracil from DNA or a 95 °C DNA denaturation step, are circumvented. Thus it is concluded that the assay presented is the most accurate absolute quantitation of genomic uracil.

TABLE OF CONTENTS

ABBREVIATIONS

AP	– apurinic/apyrimidinic site
A	– adenine
AID	– activation induced deaminase
BER	– base excision repair
bp	– base pairs
C	– cytosine
CV	– coefficient of variance
CSR	– class switch recombination
dA	– deoxyadenosine
dAMP	– deoxyadenosine monophosphate
dC	– deoxycytidine
dCMP	– deoxycytidine monophosphate
dG	– deoxyguanosine
dGMP	– deoxyguanosine monophosphate
dN	– deoxynucleoside
dNMP	– deoxynucleotide monophosphate
ds	– double-stranded
dT	– deoxythymidine
dTMP	– thymidine monophosphate
dU	– deoxyuridine
dUMP	– deoxyuridine monophosphate
ε	– absorption coeffient
G	– guanine
GC/MS	– gas chromatography coupled to mass spectrometry
h	– hour(s)
HIGM	– hyper-immunoglobulin M (syndrome)
HPLC	– high-performance liquid chromatography
Ig	– immunoglobulin
KO	– knock out
LC/MS(/MS)	– liquid chromatography coupled to (tandem) mass spectrometry
LIG	– ligase
m	– minute(s)

MRM	– multiple reaction monitoring
MS/MS	– tandem mass spectrometry
MW	– molecular weight
Neo	– neomycin resistance gene/polyadenylation signal
NFκB	– nuclear factor kappaB
NP1	– nuclease P1
RT	– room temperature
POL	– polymerase
s	– second(s)
SAM	– S-adenosylmethionine
SHM	– somatic hypermutation
SMUG	– single-strand selective monofunctional uracil DNA glycosylase
ss	– single-stranded
SVPD	– snake venom phosphodiesterase
T	– thymine
TS	– thymidylate synthase
U	– uracil
UNG	– uracil DNA glycosylase
UV	– ultraviolet
WT	– wild type

1. INTRODUCTION

The canonical components of DNA are adenine (A), thymine (T), cytosine (C), and guanine (G). In double-stranded DNA, nucleotides form hydrogen bonds, forming Watson-Crick base pairs A:T and C:G. Uracil (U), normally found in RNA in place of T, may also be present in DNA. U is structurally similar to T, which has a methyl group at the 5-carbon, and usually forms base pairs with A, but can also base pair with G (figure 1-1). Similarly, U differs from C only in having a proton at the 3-nitrogen and a keto group at the 4-carbon in place of a amino group, which can be readily deaminated [1,2]. Were U the normal base in DNA instead of T, the deamination of C would be highly mutagenic because there would be no indication of whether the U:G mismatch came from C:G or U:A. The incorporation of T as the normal base allows U in correctly-replicated DNA to be identified as a deaminated C and be repaired. Despite this, small quantities of U can be detected in DNA (reviewed in [3-5]). Here the different origins, consequences, and methods employed to measure genomic U will be discussed.
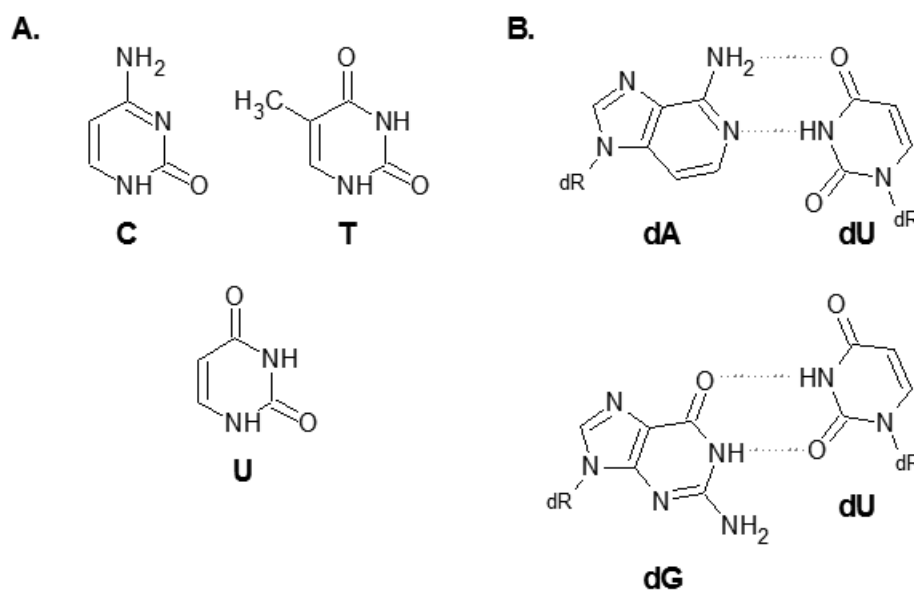
Figure 1-1: Structure and base pairing of U. A. Chemical structures of C, T, and U. B. dU can form a base pair with dG as well as dA.

1.1 Origins of genomic uracil

1.1.1 Chemical cytosine deamination

The two known main sources of genomic uracil are cytosine deamination and dUMP incorporation [3]. The former can arise both chemically and enzymatically. Although less vulnerable than RNA, DNA is subject to spontaneous decomposition [6]. At neutral pH, cytosine can be spontaneously converted to uracil by direct hydrolytic deamination via alkali-

catalyzed hydrolysis or by attack by water on the protonated base (figure 1-2) [7]. Both pathways are feasible at physiological conditions, although the rate of deamination depends on several factors such as pH and temperature. In addition, whether the DNA exists in double-stranded (ds) or single-stranded (ss) conformation has major impact upon the rate of deamination [2]. Due to biological pH and temperature regulation, whether DNA is single- or double-stranded is arguably the most crucial factor. Indeed, *in vitro* studies have demonstrated that the half-life of a single cytosine moiety is about 200 years for single-stranded (ss) and 30,000 years for double-stranded (ds) DNA [6]. In agreement with this, the rate of cytosine deamination in the yeast *Saccharomyces cervisiae* has been observed to be 40-fold higher than in *Escherichia coli*, which has been explained by eukaryotic transcription being slower than prokaryotic, thus keeping DNA in transient ss form for a longer period [8].
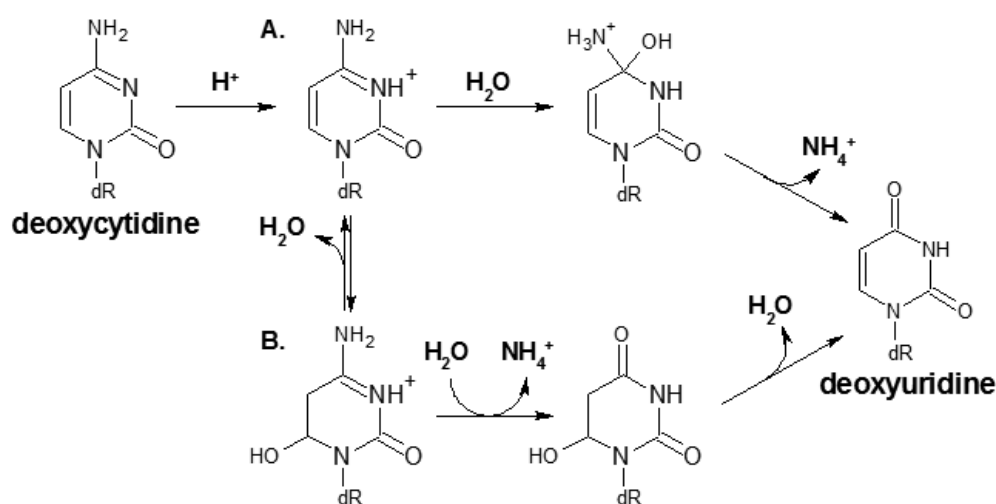
Figure 1-2: Pathways to spontaneous chemical deamination of dC. Modified from Seeberg and Kleppe [9]. Deamination can occur by (A.) direct hydrolytic deamination via alkali-catalyzed hydrolysis or (B.) water attack on the protonated base.

Exposure to several chemicals also leads to cytosine deamination. Sodium bisulfite is found in many beverages and has been found to lead to cytosine deamination, though it is thought unlikely to be a major contributer to genomic uracil content because of the low concentrations at which it is usually consumed [10,11]. Nitrous anhydride (N_2O_3) also leads to cytosine deamination and has been observed to be formed *in vivo* upon oxidation of nitric oxide (NO•), a product of enzymatic arginine oxidation [12]. A study in which human lymphoblastoid cells were treated with sub-cytotoxic levels of NO• and O_2 demonstrated up to a 2-fold increase in genomic uracil after treatment [13]. Given the already low levels of uracil in DNA, this was a modest increase; however, selective deamination of e.g. ssDNA could lead to the targeting of specific genes for deamination [10].

Irradiation by ultraviolet (UV) light can lead to the formation of genomic uracil. Cyclopyrimidine dimers are major UV photoproducts and their repair rate can be quite variable, depending on nucleotide position [10,14]. Deamination of these dimers has been shown in lambda phages to be 7-8 orders of magnitude faster than in regular cytosine [15]. Photoproducts are not equivalent to normal deaminated cytosines because they remain crosslinked after deamination, but they nevertheless present a possible source of genomic uracil that may yield similar consequences.

1.1.2 Enzymatic cytosine deamination

Methyltransferases normally methylate DNA at CpG islands during epigenetic DNA regulation [16], but can also lead to introduction of genomic uracil. Cytosine methylation is traditionally thought of as mutagenic because of its propensity to undergo deamination from 5-methylcytosine to thymine [17-19]. These enzymes work by transferring a methyl group to cytosine from S-adenosylmethionine (SAM), but it has been demonstrated that they can also catalyze cytosine deamination when SAM concentration is low or if methyltransferase carries a mutation that impairs functional SAM-binding [19-21]. As such, dysregulation of these enzymes can contribute to genomic uracil.

The discovery of activation induced deaminase (AID) in 1999 introduced yet another mechanism by which genomic cytosine may be deaminated [22]. AID is involved in the adaptive immunity mechanisms of class switch recombination (CSR) and somatic hypermutation (SHM, discussed below and reviewed in [20,23]). The chemical mechanism employed by AID is not yet clear, but high sequence homology with a bacterial cytidine deaminase has led to a proposed mechanism in which position 4 (adjacent the amino group) of the pyrimidine ring is nucleophilically attacked by AID [20]. AID is only expressed in activated B-cells as well as in some cancer cells and actively deaminates cytosines in the immunoglobulin (Ig) loci. Sequence analysis has shown that deamination and subsequent mutation occurs primarily in 2-7 kbp-long "hot spots" within the Ig loci, a necessary step in adaptive immunity. Nevertheless, recent data demonstrates that AID also deaminates cytosines in up to 25% of the expressed genes in mouse germinal center B-cells, albeit to a much lower extent [3,23,24]. AID therefore presents a possibly substantial source of genomic uracil in the entire genome.

1.1.3 dUMP misincorporation

Most known replicative polymerases incorporate deoxyuridine monophosphate (dUMP) and deoxythymidine monophosphate (dTMP) with similar efficiency and lack the ability to differentiate between the two [25]. It has therefore been suggested that low dUMP

incorporation by polymerases depends on a high deoxythymidine triphosphate (dTTP) to deoxyuridine triphosphate (dUTP) concentration ratio. One study estimates physiological dUTP and dTTP concentrations to be 0.2 and 37 ± 30 µM, respectively, making the intracellular ratio considerably lower than 1 % in normal cells [26]. Intracellular dUTP levels are regulated by the enzyme dUTPase, which hydrolyzes dUTP [27]. This enzyme is especially important in regulation of the dUTP/dTTP ratio because the product of this reaction, dUMP, is also a precursor for dTTP synthesis by thymidylate synthase (TS) [3,28]. Thus it regulates the dUTP/dTTP ratio by both decreasing dUTP production and increasing dTTP production. Although the estimated dUTP concentration is ~1 % that of dTTP, it was estimated that about one dUMP is incorporated per 10^4 dTMPs per genome per day, possibly due to differential shuttling of the nucleotides to different parts of the nucleus [29]. Nevertheless, this constitutes a constant and significant source of genomic uracil.

1.2 Consequences of genomic uracil

1.2.1 Excision/Repair

Genomic uracil is repaired by the base excision repair (BER) pathway. BER may processed via two pathways, depending in part on cell cycle stage, initiating enzyme, and nature of DNA damage [30,31]: short patch BER exclusively replaces the damaged nucleotide and long patch BER replaces 2-8 nucleotides (figure 1-3). Both pathways are initiated by a damage specific glycosylase such as uracil N-glycosylase (UNG), thymine-DNA glycosylase (TDG), or single-strand selective monofunctional uracil DNA glycosylase (SMUG). The glycosylase recognizes and excises U, resulting in a free base and an apyrimidinic/apurinic (AP) site. The AP site is subsequently processed by AP endonuclease 1, which cleaves the 5′-phosphodiester bond. In short patch BER, DNA polymerase β (POLβ) removes the 5′ AP deoxyribophosphate and replaces the excised nucleotide, and DNA ligase III (LIG3) ligates the strands. In long patch BER, POLβ may insert the first nucleotide, but the others are inserted by POLδ/ε. The resulting "flap" is removed by flap-endonuclease 1 and the DNA strand ligated by LIG1 [4,5].
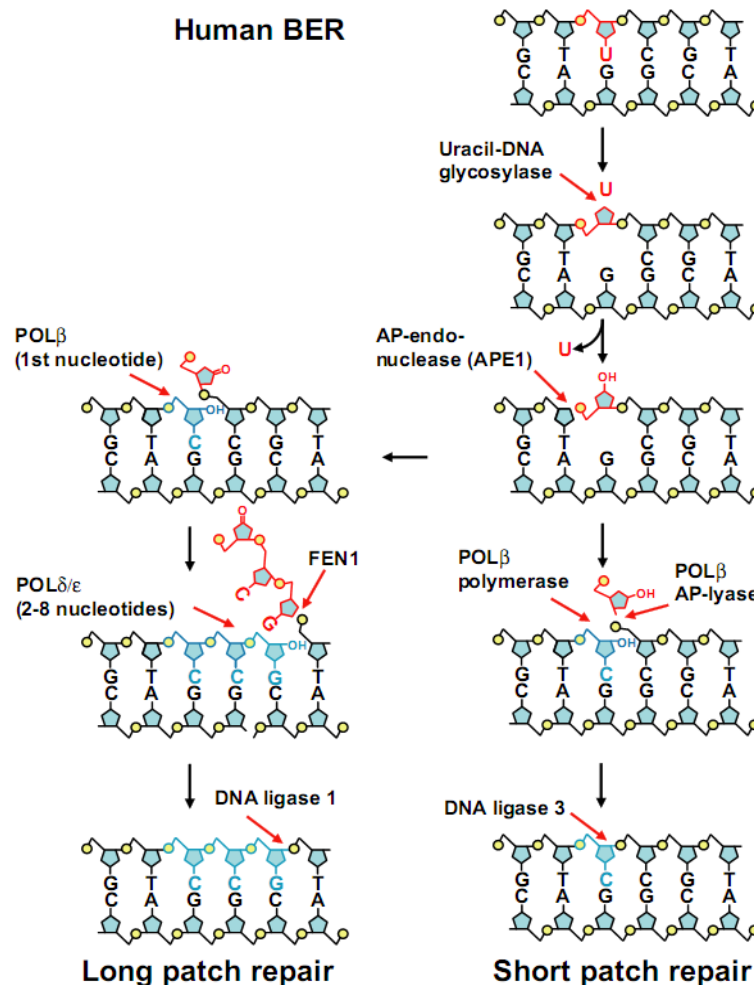
Figure 1-3: Visual summary of short-and long-patch BER pathways. UNG recognizes and excises U and AP endonuclease 1 cleaves the DNA at the resulting AP site. In short patch BER, POLβ replaces the excised nucleotide and LIG3 ligates the strand. In long patch BER, POLδ/ε insert the rest of the nucleotides, the flap is removed by flap-endonuclease 1, and the DNA is ligated by LIG1. Obtained from Sousa *et al.* [10].

Whether uracil is removed from DNA depends on several factors. The maximum UNG protein level was observed during S phase in a human cervical cancer cell line [32]. UNG is considered the major glycosylase involved in BER and has been reported as responsible for initiation of all of U:A match repair and 90% of U:G mismatch repair [33]. In accordance with this, it has been reported that uracil is less efficiently excised in non-proliferating mouse embryonic fibroblasts as compared to proliferating cells [34]. Furthermore, the rate of dUMP misincorporation by polymerases involved in BER depends on cellular dUTP/dTTP levels, so dysregulation of TS or dUTPase could lead to reincorporation of dUMP. Finally, UV-induced cyclopyrimidine dimers are more readily deaminated than regular cytosine and inaccessible to UNG [10]. Most organisms rely on light-activated photolyase activity to repair photodimers, but placental mammals lack these enzymes and rely on the less efficient

nucleotide excision repair pathway [35]. Thus U in photodimers may be repaired with less fidelity, although the repair mechanism of U photodimers has not been well characterized.

1.2.2 DNA damage

Misincorporation of dUMP is presumably the most dominant source of genomic uracil and results in a U:A match, which is not considered mutagenic [34]. Nevertheless, this has not been fully explored and there exists the possibility that U:A matches can contribute to mutagenicity or cytotoxicity when improperly repaired. On the other hand, deamination of genomic cytosine converts a C:G pair into a U:G mismatch, resulting in a C:G and a mutagenic U:A pair after replication [10]. This probably represents the major cause of uracil-derived mutagenicity, although the sum of single insertions and deletions were as common in mice [36]. UNG repair of U:G mismatches has been observed as faster than repair of U:A matches, a characteristic that may help counteract U:G mismatch mutagenicity [33]. Furthermore, AP sites in DNA can be highly cytotoxic because they irreversibly trap topoisomerase I in complex with DNA [1]. This rationalizes why AP sites have been observed to remain bound to UNG in solution [37]. Nevertheless, if the cell's capacity to process AP sites is lower than their generation by glycosylases, cytotoxicity may ensue. Although genomic uracil by both deamination and misincorporation should be repaired, the repair fidelity of BER is variable. A study by Akbari *et al.* reported that non-proliferating cell extracts from a keratinocyte cell line produced two- to three-fold higher levels of synthesis errors as a result of BER compared to proliferating cell extracts [31]. It was found that this may be attributed to the prominence of POLβ in non-proliferating cells. POLβ lacks 3′-5′ endonuclease proofreading ability and exhibits an error rate of roughly one per 4,000 nucleotides inserted [38]. Besides adding to the evidence that repair of genomic uracil is more mutagenic in non-proliferating cells than proliferating cells, this observation presents a possible U:A match mutagenicity through improper repair.

1.2.3 Role in adaptive immunity

The deamination of C to generate U by AID plays a crucial role in adaptive immunity. B-cells produce antibodies against a large number of epitopes despite containing a limited number of Ig genes. This is accomplished by recombination of the variable, diverse, and joining gene segments of Ig-genes (commonly referred to as V(D)J recombination) as well as SHM of variable regions of Ig-logi (IgV) [39]. In the current model for SHM, AID deaminates a C residue, creating a U:G mismatch and leading to several possible consequences [40]. First, DNA can be replicated, interpreting U as a T and resulting in a C-to-T transition mutation. Secondly, the mismatch can be repaired by error-prone short-patch

BER, leading to a spectrum of mutations. Finally, the U lesion can be repaired by long-patch BER or the mismatch repair pathway, in which the U-bearing strand is excised and error-prone polymerases recruited to fill the gap.

Class switch recombination is the pathway by which antibodies are modified to yield isotypes with differing functions [39,41,42]. The deamination of C to U in the genome is also an important first step in this process. Here AID introduces U in both strands in AID "hot spots" at switch (S) regions of the Ig-loci. The majority of the U residues are excised by UNG, resulting in AP-sites, which lead to DNA strand breaks. The mismatch repair complex of proteins have a probable role in converting single-strand breaks to double-strand breaks [43,44]. The breaks in two separate S-regions are linked by non-homologous end joining. Both SHM and CSR are briefly summarized in figure 1-4.
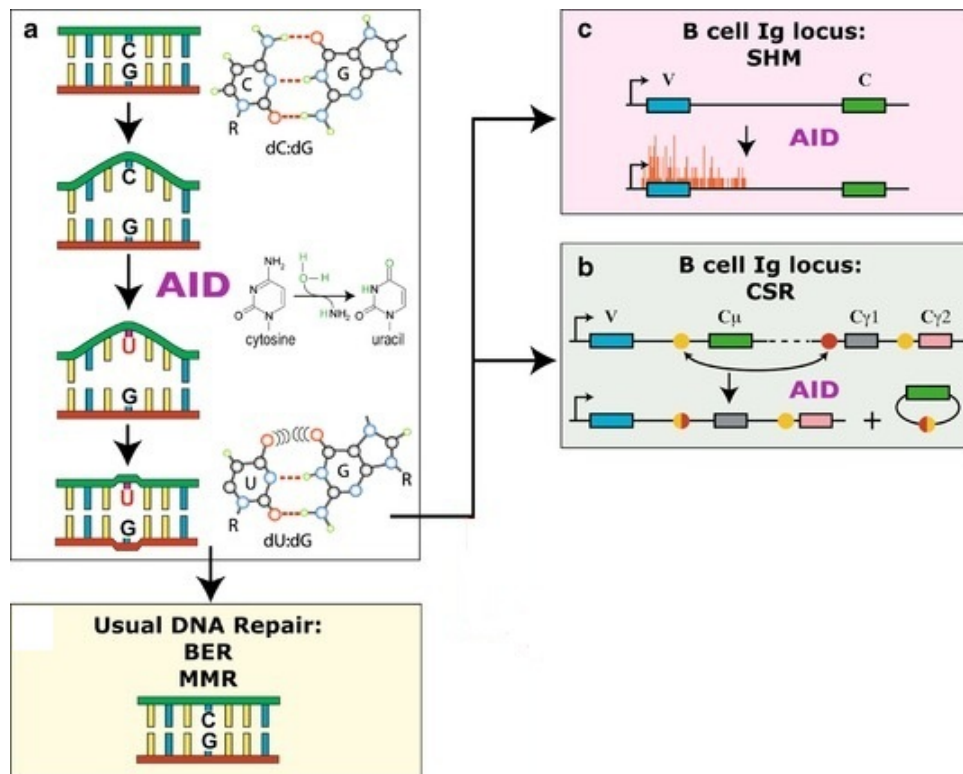
Figure 1-4: The consequences of AID-mediated dC deamination. A. AID deaminates C to U, resulting in a U:G mismatch, which are normally repaired by BER. In the Ig-loci of activated B-cells: B. Repair of U lesion by UNG leaves AP-site, leading to a strand break and CSR. C. Low-fidelity repair of U lesion leads to SH. Figure obtained from Petersen-Mahrt *et al.* [42].

1.2.4 Pathology

Genomic U can be both beneficial and dangerous. As mentioned above, antibody gene diversification relies of the presence of U in Ig-loci. Hyper-immunoglobulin M (HIGM) syndrome is a condition in which defective CSR leads to a deficiency of IgG, IgA, and IgE

with preserved or elevated IgM levels [44,45]. The resulting clinical outcome is immunodeficiency. One recognized cause for HIGM syndrome is AID deficiency, most often in the form of mutation to the AID gene [46]. The complete or impaired functionality as well as improper translocation of the AID enzyme result in a shortage or complete lack of C deamination to U in the Ig-loci [44]. Both CSR and SHM are therefore impaired. Deficiency of UNG has been described as another cause for HIGM, though in a smaller number of patients [47]. The clinical outcome of UNG deficiency is similar to AID deficiency, although SHM is not completely impaired, but rather skewed towards an increase in G:C$\rightarrow$ A:T transitions [48]. In this case, UNG fails to excise AID-generated U, so AP-dependent strand breaks in AID hotspots are less frequent. HIGM syndrome therefore demonstrates the crucial role of genomic U as an intermediate in adaptive immunity.

The consequence of UNG deficiency in HIGM syndrome stems from a halt in a DNA-rearranging process (CSR), but it can also result in impaired BER and possibly mutagenesis outside of the Ig-loci leading to carcinogenesis. Although little evidence exists linking UNG deficiency to human cancers, it was observed that UNG knock-out mice developed lymphoid hyperplasia at three months and exhibited a 22-fold increased risk of B-cell lymphoma development late in life [4,49]. Similarly, C deamination to U by AID is a central part of adaptive immunity and aberrant AID activity may lead to both DNA mutagenesis and tumorigenesis [50]. As previously mentioned, AID has been shown to target genes outside of the Ig-loci, albeit to a lower extent. Dysregulated AID has been reported as a source of mutation in AID hot spot sequences in several proto-oncogenes and tumor suppressor genes [24,51]. Moreover, transgenic mice that constitutively and ubiquitously express AID developed lymphomas and tumors in the lung, liver, and stomach [3,50,52-54].

Chronic inflammation e.g. as a result of constitutive expression of nuclear factor kappaB (NFκB, a transcription factor involved in immunity and inflammation) has been recognized as a risk factor for a variety of cancers [55-57]. This is relevant for the following reasons: first, NO•, which can be oxidized to yield the aforementioned cytosine-deaminating N_2O_3, is formed in substantial quantities in macrophages, cell lines, and tissues during inflammation [10]. The over-production of NO• and N_2O_3 during inflammation could increase C deamination to U and thus lead to mutagenesis. Secondly, it has been suggested that *Helicobacter pylori* infection directly activates proinflammatory cytokines via the NFκB pathway, possibly triggering AID expression [50,54]. Aberrant AID expression could then lead directly to mutation of the *p53* tumor suppressor gene and thus eventually oncogenesis.

1.3 Analysis of genomic uracil

Its role in so many aspects of biology make genomic U an important research topic, but its low abundance relative to A, T, C, and G makes it difficult to directly quantitate. There have been several attempts to measure genomic dU by both direct and indirect means [13,34,58-62].

1.3.1 Indirect measurement of genomic uracil

An indirect assay was described to quantitate dUMP incorporation to DNA in mouse embryonic fibroblasts [34]. Here genomic DNA was pretreated with UNG to excise uracil and subsequently treated with AP-endonuclease to cleave at abasic sites. Alkaline elution was then employed to monitor DNA fragment size. Alkaline elution works by pumping cell lysate through a filter under alkaline conditions. The rate at which the DNA elutes is size-dependent [63]. Thus the fragments eluted from an UNG treated sample can be compared to those from an untreated sample to ascertain genomic uracil content. This indirect measurement suffers from the drawback that DNA may be sheared by other forces besides endonucleolytic cleavage at AP-sites [1]. The high pH used in the assay denatures DNA and may lead to or facilitate strand breakage. Moreover, Kohn *et al.* observed that disturbing the elution apparatus introduced shearing, and even routine laboratory practices like vortexing have been reported to shear genomic DNA [63,64].

1.3.2 Absolute genomic uracil quantitation by UNG excision

There are two main strategies for direct quantitation of total genomic uracil: enzymatic excision of uracil from DNA by UNG and hydrolysis of DNA to nucleosides by nuclease, phosphodiesterase, and phosphatase treatment. Both approaches may employ mass spectrometry to detect U or dU, though they differ in their initial chromatographic separation steps. The strategies are summarized in figure 1-5.
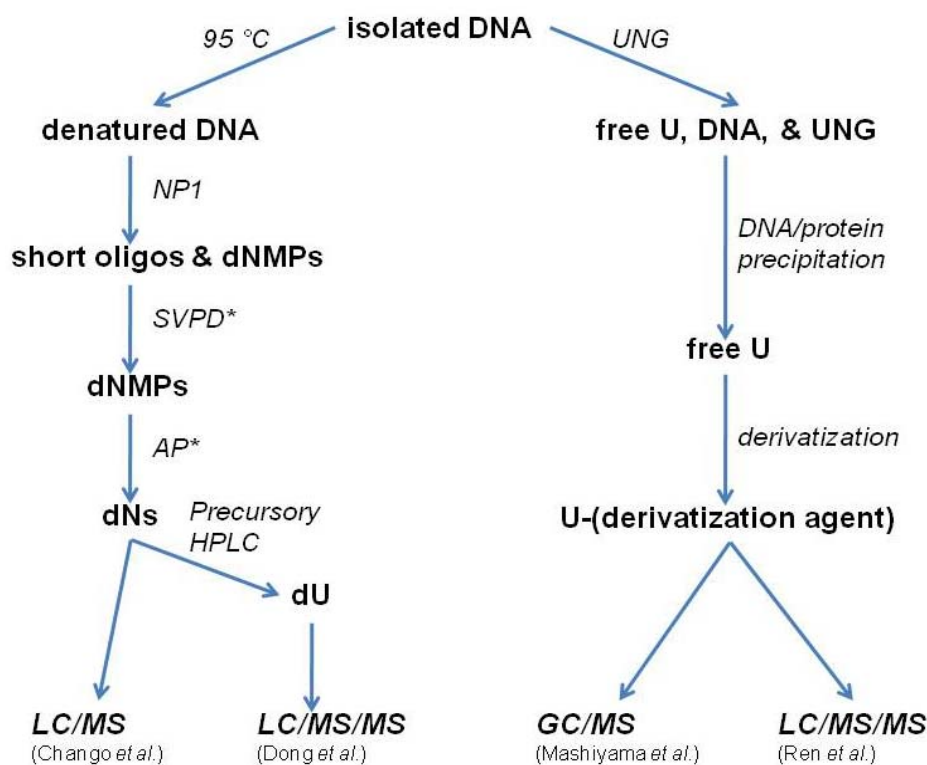
Figure 1-5: Brief visual overview of genomic uracil quantitation methods. Chango et al. and Dong et al. treated heat-denatured DNA with nucleases and phosphatase to yield dU, which was separated from dC by an initial HPLC step (Dong et al.) and quantitated by LC/MS or LC/MS/MS, respectively [13,58,60]. Mashiyama et al. and Ren et al. treated DNA with UNG to excise U, derivatize it, and quantitated it with CG/MS or LC/MS/MS, respectively [59,61]. *SVPD and AP were added in a single step, but are shown as two for clarity.

The assays involving UNG treatment of DNA are followed by derivatization of the released free uracil [59,61]. Whereas Mashiyama et al. employed gas chromatography coupled to mass spectrometry (GC/MS) for sample preparation and uracil detection, Ren et al. employed high performance liquid chromatography (HPLC) coupled to tandem mass spectrometry (LC/MS/MS). Tandem mass spectrometry in multiple reaction monitoring (MRM) mode works by selecting precursor ions by mass in an initial quadrupole, inducing fragmentation of these ions by collision with an inert gas in the second quadrupole, and filtering one specific fragment ion in the third quadrupole prior to detection by an electron multiplier [65,66]. A given compound's fragment ions depend on its structure and bond strengths, so fragment ions are different from molecule to molecule. Thus the detection of a fragment ion specific to a compound adds detection specificity (both approaches are summarized in figure 1-6).
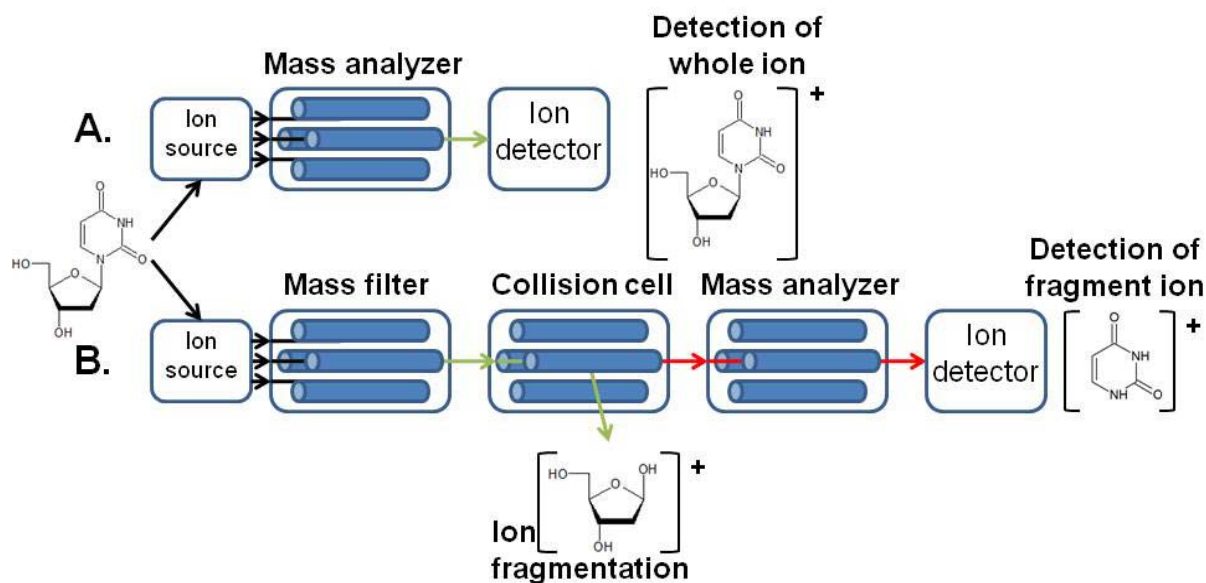
Figure 1-6: Visual comparison of single vs. tandem mass spectrometry. In both methods after chromatographic separation of sample, the analyte is ionized in the spectrometer's ion source. In a single mass spectrometer (A.), all ions isobaric with the analyte are analyzed. In tandem mass spectrometry (B.), ions isobaric with the analyte pass to a second quadropole, where they are fragmented by collision with an inert gas. A third quadropole allows ions isobaric with the expected product ion to pass to the ion detector.

The use of UNG to excise uracil for absolute quantitation of total DNA-uracil has several disadvantages. Experimental design plays a large part in the accuracy and intra-experimental variability of UNG-based assays. For example, magnesium (II) concentration has been observed to increase UNG activity *in vitro* [3,67,68]. Buffer composition altered by sample impurities could also change the enzyme efficacy and lead to inaccuracy. Furthermore, intrinsic enzyme preferences may lead to biased results: it has been reported that UNG exhibits preference for ssDNA over dsDNA and U:G pairs over U:A pairs, as well as sequence preference [4,67,68]. Consequently, UNG may introduce bias by not excising all substrate in the samples [3]. To control for this, Ren *et al.* used a 198-bp long DNA fragment that had been PCR amplified using dUTP in place of dTTP; however, the uracil content in the oligo is radically higher (about one fourth of all bases) than that of genomic DNA. The kinetics of UNG recruitment to substrate sites has not been compared in these contexts, so it may be questionable to assume the UNG can detect uracil in genomic DNA with the same efficiency as in a short DNA fragment. Mashiyama *et al.* validate their assay using a "linearity test," consisting of measuring uracil content in samples with different concentrations of the same genomic DNA. It is noteworthy that this method also employs restriction digestion of DNA prior to UNG treatment, which fragments the DNA and possibly

increases enzyme substrate detection by minimizing supercoiling and lowering the DNA melting temperature, thereby increasing access to otherwise obscured uracil. Nevertheless, the same fundamental drawback applies in that UNG may detect substrate differently in DNA from different sources (lymphocytes were used for DNA extraction) or with different uracil content levels.

Derivatization is used by both Ren *et al.* and Mashiyama *et al.* and adds additional complexity, analysis time, and (most notably) more sources of error for the assays. Derivatization is a method whereby the analyte is covalently bound to a chemical agent that increases the ionization efficiency and chromatographic separation of the product compound while introducing a labile bond for tandem mass spectrometry [65,69]. Generally, derivatization must be well controlled in so that no impurities are introduced and the assay remains quantitative. That is, the intra-experiment variability of percentage of sample derivatized must be constant and the percentage of derivatized sample must be high [69]. This was a problem for Ren *et al.* in that they encountered difficulty derivatizing uracil from DNA samples and therefore employed different derivatization reaction conditions for the standard curve and analytes. Although the problem was rectified by slight modification of the reaction, the differential treatment of analytes and standards negates true absolute quantitation because the samples cannot be considered equally derivatized. Furthermore, the extent or uniformity of derivitization between samples remains untested by both groups, leaving a possible error in quantitation unchecked.

1.3.3 Genomic deoxyuridine quantitation by enzymatic DNA hydrolysis

Instead of using UNG to excise uracil from DNA, Chango *et al.* and Dong *et al.* hydrolyzed DNA to deoxynucleotides and measured deoxyuridine (dU) [13,58,60]. This strategy is based on a method by Crain for DNA and RNA adduct quantitation by mass spectrometry [70]. The first step is treatment with nuclease P1 (NP1) from *Penicillium citrinum*, which cleaves both ssDNA and RNA at every position to yield 5′-dNMPs and 5′-phosphooligo deoxynucleotides [71,72]. Importantly, this enzyme does not recognize dsDNA, necessitating an additional heating step at 95 °C for 5 min to denature DNA before hydrolysis. The DNA is then further hydrolyzed by the addition of snake venom phosphodiesterase (SVPD), which actively hydrolyzes the 5′-oligonucleotide products produced by NP1 [73]. The combination of these two enzymes ensures a high degree of hydrolysis to 5′-dNMPs. These are dephosphorylated by the addition of alkaline phosphatase, an enzyme that catalyzes the removal of 5′ phosphate groups from nucleotides [74]. The products of this triple-enzyme reaction, deoxynucleosides (dN), are then separated by reverse-phase chromatography and analyzed by mass

18

spectrometry. Measuring dU is advantageous to measuring U using tandem mass spectrometry because dU (as is the case with all nucleosides) contains an N-glycosylic bond. The N-glycosylic bond breaks upon fragmentation of the molecule, resulting in U as a product ion. The product ions are smaller when U undergoes fragmentation, which results in a higher background because more unspecific contaminants are usually more abundant with decreasing molecular weight [75].

The isotope of deoxycytidine containing one carbon-13 (^{13}C-dC) is isobaric with and therefore mass spectrometrically indistinguishable from monoisotopic dU. Given that the natural abundance of carbon-13 is 1.109%, every carbon in dC has about 1% likelihood to contain an extra neutron [76,77]. Chromatographic separation of dC and dU may be insufficient for mass spectrometric analysis because the relative abundance of ^{13}C-dC is so much greater than dU that the dU signal is obscured due to "tailing" of ^{13}C-dC. To compensate for this, Dong *et al.* introduced a precursory HPLC step in which dU was separated from dC and concentrated prior to LC/MS/MS analysis [58].

The heating step necessary for denaturation of DNA prior to NP1 treatment can lead to dC deamination [6,7] and thus falsely high dU signals. In accordance with this, it was observed that heating DNA at 95 °C led to a 3-fold increase in uracil signal after 5 min and up to 40-fold increase after 20 min [59]. Heat-induced *in situ* deamination represents a major complication with this method [60].

The differing methods have likely contributed significantly to the discrepancy in genomic uracil content reported in the literature (table 1-1) [61]. The reported values range from 0.20 to 129 pg U per µg DNA, which corresponds to 0.55 to 355 U molecules per million bases. This enormous discrepancy may be due to the various cell types from which DNA was isolated as well as study design [3]. Nevertheless, the technical shortcomings detailed above need to be addressed before genomic uracil measurements are deemed reliable.

Table 1-1: Genomic uracil contents reported in the literature. Modified from Mashiyama et al. [61].

Reference	pg U/μg DNA		U/10^6 N		Cell type
	Min	Max	Min	Max	
Chango et al. [60]	3.19*		8.78		human liver carcinoma cells
Mashiyama et al. [61]	0.10	0.145	0.27	0.40	human lymphocytes
Dong and Dedon [13]	4.73	14.55	13.00	40.00	human lymphoblasts
Mashiyama et al. [62]	0.33	6.35	0.91	17.46	human lymphocytes
Choi et al. [78]	3.89	5.09	10.70	14.00	rat colon cells
Ren et al. [59]	0.20	1.94	0.55	5.34	human lymphocytes
Crott et al. [79]	0.76	4.84	2.09	13.31	human lymphocytes
Koury et al. [80]	0.46	1.77	1.27	4.87	mouse erythroblasts
Blount et al. [81]	1.66	129.02	4.57	354.81	human peripheral leukocytes
Ramsahoye et al. [82]	41.41	42.93	113.88	118.06	human bone marrow cells
Blount and Ames [83]	0.70	3.00	1.93	8.25	rat hepatocytes

*Basal amount measured in control cells. Originally calculated in pg dU.

1.4 Thesis aims

The aim of this thesis was to establish a method to achieve reliable quantitation of genomic uracil. To this end, an LC/MS/MS-based absolute total deoxyuridine quantitation assay was developed and validated. The method employs nuclease/phosphodiesterase/phosphatase treatment while minimizing in vitro genomic C deamination as well as dUMP contamination. The shortcomings of the assays described above are addressed, and the steps in assay development as well as the controls employed will be discussed in detail. Notably, pretreatment with UNG to excise genomic U revealed an inability of the enzyme to locate and excise all uracil in a genomic context. Furthermore, it was observed that heating DNA at the temperature used in some contemporary assays introduces genomic U linearly with time, producing more than three-fold more dU signal after only five minutes of heating.

2. MATERIALS AND METHODS

2.1 Transgenic mouse genotyping

2.1.1 Transgenic mouse genotyping theoretical background

Transgenic C57/B6 mice were generated by insertion of a 1.1-kbp-long neomycin resistance gene/polyadenylation signal (*Neo*) cassette in exon 4 of the *Ung* gene [84]. To confirm whether a mouse was knock-out (KO), wild-type (WT), or heterozygous for this gene, PCR was employed using primers located in the third and fourth exons of the UNG gene. The PCR products were 1.1 kbp larger in KO mice than in WT because of the *Neo* cassette insert.

2.1.1 Method for genotyping transgenic mice

Prior to PCR, DNA was isolated from mouse ear or tail clips using Qiagen Blood and Tissue kit (Qiagen, Germantown, MD, USA) according to the manufacturer's instructions. DNA concentration and purity was determined by absorption spectrometry using a NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). To a final volume of 20 µl, 2 µl of 2 to 50 ng/µl DNA were added to a reaction mixture containing 1 U PlatinumTaq polymerase (Invitrogen, Eugene, OR, USA), 1x PlatinumTaq buffer, 1.5 mM MgCl$_2$, 0.2 mM dNTP mix (Finnzymes, Espoo, Finnland), and 0.2 µM sense (5′-GGCCACCCTGACAAATCCCC) and antisense (5′-CACGGACCTAATCAAGCTCACG) primers (Sigma-Aldrich, St. Louis, MO, USA). The thermal cycler program used was as follows: an initial denaturation step at 98 °C for 30 s, 30 cycles of 98 °C for 10 s, 65 °C for 30 s, and 72 °C for 2 min, and a final elongation step at 72 °C for 5 min. PCR products were analyzed by agarose gel electrophoresis using a 2 % (w/v) agarose gel in TAE buffer (40 mM Tris-HCl, pH 7.5, 20 mM acetate, 1 mM EDTA). Prior being loaded on the gel, 5 µl PCR products was mixed with 1 µl 6x loading dye (0.25 % bromophenol blue and 30 % glycerol). The gel was electrophoresed in TAE buffer at 100 V for 25 min, stained in 0.5 µg/ml ethidium bromide for 5 min, destained in water for 30 s, and analyzed on a Gel Logic 200 imaging system using Kodak Molecular Imaging Software (Kodak, Rochester, NY, USA).

2.2 UNG activity assay

2.2.1 UNG activity measurement theoretical background

UNG activity was measured as previously described [85]. The protocol, which could also be employed to measure UDG activity in cell lysate, involved incubating DNA containing tritiated dUMP ([^{3}H]-dUMP) with UNG, stopping the reaction, and precipitating the DNA and protein from the reaction. The supernatant contains [^{3}H]-U, which can be measured by scintillation counting due to the radioactive decay of tritium. Scintillation works by

dissolving samples in a cocktail that contains solvents and scintillants. In the case of measuring UNG activity, the radioactive decay from tritium constantly emits beta particles, the energy of which is transferred to the scintillant by the solvent in the scintillation cocktail. Once excited, the scintillant emits light that can be measured with photometry [86].

2.2.2 Method for UNG activity measurement

In a total assay volume of 20 µl, 5 µl of purified UNG was incubated in (final) 1.8 µM [³H]-dUMP-containing calf thymus DNA ([³H]-dUMP-DNA, specific activity 0.5 mCi/µmol) in 20 mM Tris-HCl (pH 7.5), 10 mM NaCl, 1 mM EDTA, 1 mM DTT, and 0.5 mg/ml bovine serum albumin. Reactions were incubated for 10 min at 30 °C and stopped on ice. DNA was precipitated by the addition of 50 µl (1 mg/ml) salmon sperm DNA carrier (in 10 mM Tris-HCl, ph 7.5, and 1 mM EDTA) and 500 µl ice-cold 5 % trichloroacetic acid, incubated on ice for 10 min, and centrifuged at 16,100 x *g* for 10 min at 4 °C. Finally, 500 µl of the supernatant was added to 5 ml Ready Protein+ liquid scintillation cocktail (Beckman Coulter). Acid-soluble [³H]-uracil was then quantitated by scintillation counting using a Tri-Carb 2900TR liquid scintillation analyzer using Quantasmart v.2.02 software (Perkin Elmer, Waltham, MA, USA).

2.3 DNA isolation and purification

2.3.1 DNA isolation theoretical background

DNA used for genomic deoxyuridine quantitation was isolated by phenol-chloroform isoamyl extraction [87] followed by isopropanol precipitation. This method works by addition of solution containing water-saturated phenol, chloroform, and isoamyl alcohol (P:C:I) to the cell lysate [88,89]. DNA is a polar molecule and therefore more soluble in water than it is in phenol. Chloroform is an antifreeze and is added to increases phenol phase density. Isoamyl alcohol is added to the chloroform to reduce foaming. Upon mixing and centrifugation of P:C:I and cell lysate, the proteins in the water phase are denatured to expose their hydrophobic residues and partition in the organic phase along with the non-polar lipids, while the nucleic acids stay in the aqueous phase. The phenol is removed by precipitating the DNA with chloform:isoamyl and then isopropanol. Co-purification of RNA with DNA is minimized by the addition of RNase A prior to extraction.

2.3.2 Cellular dUMP removal theoretical background

Cellular dUMPs could potentially be co-isolated with DNA. These dUMPs would be dephosphorylated during DNA hydrolysis and result in a falsely high dU measurement. Purified DNA was therefore treated with alkaline phosphatase to dephosphorylate dUMP and repurified with spin-columns prior to DNA hydrolysis. Spin-columns contain a solid phase

containing silica resin. DNA is mixed with a chaotropic buffer that denatures DNA by disrupting the shell of hydration around it and provides optimal solid phase binding condition. The buffer must contain positively-charged ions that bind to the negatively-charged phosphodiester DNA backbone and bind to negatively-charged oxygen on the silica resin, forming a salt bridge between DNA and silica. Other particles not bound to the column are eluted. The column is washed with ethanol-containing buffers to remove the chaotropic salts from the column and eluted with water. Pure water is added to the column to rehydrate DNA, disrupting the salt bridge that bound it to silica, and eluting it from the column [90,91].

2.3.3 Method for DNA isolation and purification

Isolated cells or tissue were suspended in 400 µl extraction buffer (10 mM Tris-HCl, pH 8.0; 10 mM NaCl; 0.5 % SDS; 25 mM DTT, 0.25µg/µl proteinase K, Worthington Biochemical, Lakewood, NJ, USA; 0.1 µg/µl RNase A from bovine pancreas, Sigma Aldrich) per 30 mg of tissue or 5×10^7 cells, with a minimum working volume of 400 µl. Tissue was homogenized using a Dounce homogenizer. Tissue and isolated cells were incubated at 60 °C for 1 h with 250 rpm shaking. The homogenate was centrifuged at 9,100 x *g* for 10 min at room temperature (RT) to remove cell debris and the supernatant transferred to a new tube.

One volume of phenol-chloroform isoamyl mix (25:24:1) was added to the homogenate supernatant, mixed by vortexing, and centrifuged at 9,100 x *g* for 10 min at RT. The top aqueous phase was transferred to a new tube and the bottom organic phase was discarded. One volume of chloroform isoamyl alcohol mix (24:1) was added to the aqueous phase, mixed by vortexing, and centrifuged at 9,100 x *g* for 10 min at RT. The top aqueous phase was transferred to a new tube, the organic bottom phase was discarded, and this step was repeated once.

The DNA was precipitated by addition of 0.3 volumes of 10 M ammonium acetate (pH 7.9) and 1 volume 100% isopropanol to the aqueous phase, which was mixed by vortexing and centrifuged at 16,100 x *g* for 20 min at 4 °C. The supernatant was gently aspirated without disturbing the DNA pellet, 1 ml of 70 % ethanol was added, and the mixture was centrifuged at 16,100 x *g* for 15 min at 4 °C. The supernatant was gently aspirated and the pellet was allowed to air dry in an open tube at RT for 10 min. Finally, the pellet was dissolved in water and the DNA purity analyzed by spectrophotometry using a NanoDrop ND-1000 spectrophotometer (Thermo Fisher).

Purified DNA was buffered with 100 mM ammonium bicarbonate and 1 mM $MgCl_2$. Next, 0.2 U/µl alkaline phosphatase from *Escherichia coli* (Sigma-Aldrich) was added to the solution, mixed by vortexing, and incubated at 37 °C for 2 h. DNA was then further purified

with spin-columns using a Qiagen Blood and Tissue kit (Qiagen) according to the manufacturer's instructions using 100 µl water twice for the final elution step. The eluted DNA was further purified and concentrated by isopropanol precipitation as described above.

2.4 UNG excision

UNG was used to excise U from DNA. A version of UNG lacking the N-terminal regulatory domain (UNG-Δ84) was used. This version of the enzyme is more easily overexpressed and purified and has been shown to be more catalytically active than full-length UNG [68]. Prior to DNA hydrolysis, 5 µg of DNA was added to a solution containing 1 µg recombinant UNG-Δ84 (produced in-house) in 20 mM Tris-HCl (pH 7.5), 60 mM NaCl, 1 mM EDTA, and 1 mM DTT in a total volume of 100 µl. The solution was incubated at 37 °C for 1 h and the reaction stopped by placing the tube on ice. Following U excision by UNG, the DNA was precipitated using isopropanol as described above.

2.5 DNA hydrolysis and sample preparation

2.5.1 DNA hydrolysis theoretical background

DNA samples were enzymatically hydrolyzed to nucleosides using a protocol modified from Crain [70] and the resulting dU separated from dC by reverse phase HPLC. The hydrolysis works by treating DNA with DNase I (cleaving DNA to short oligomers and single dNMPs), phosphodiesterase (breaking the oligomers to single dNMPs), and alkaline phosphatase (dephosphorylating dNMPs to dNs). Notably, DNA denaturation by heating to 95 °C is not required. This method is summarized in figure 2-1.
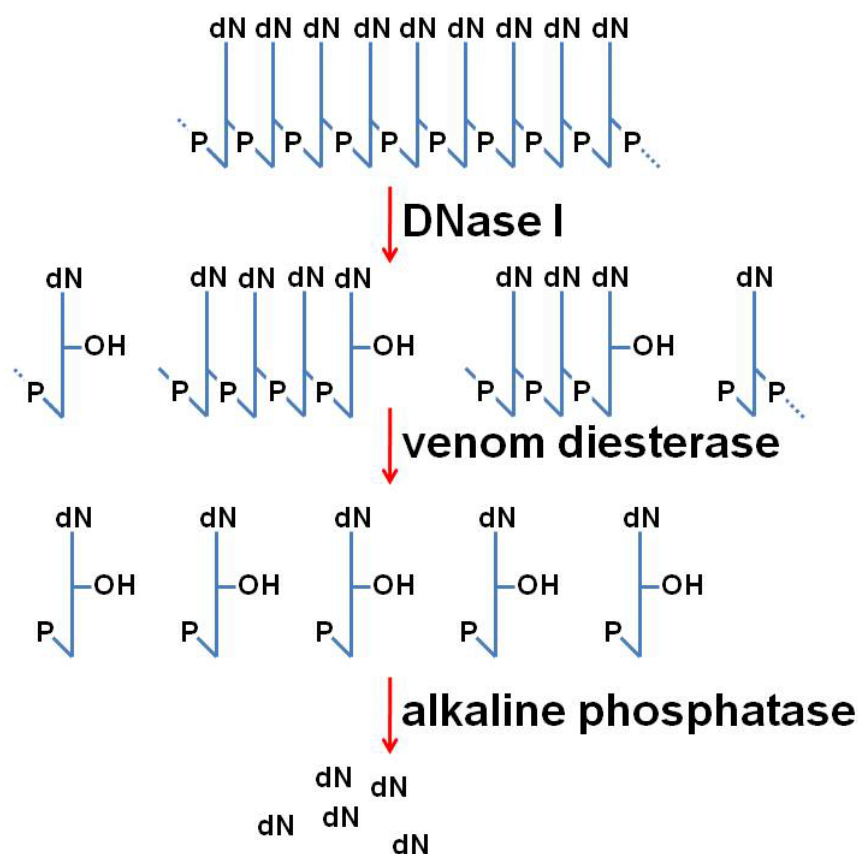
Figure 2-1: Overview of DNA hydrolysis to dNs protocol. DNase I cleaves DNA to yield single dNMPs and short oligonucleotides. SVPD breaks the oligonucleotides to single dNMPs. Alkaline phosphatase dephosphorylates dNMPs to yield dNs. Modified from Davidson [73].

2.5.2 Method for DNA hydrolysis and sample preparation

Prior to hydrolysis, purified DNA was buffered with 100 mM ammonium bicarbonate (pH 7.9) and 1 mM $MgCl_2$. Per 5 µg DNA, 0.1 µl 10 U/µl recombinant DNase I (Roche), 0.1 µl 0.002 U/µl phosphodiesterase I from *Crotalus adamanteus* venom (Sigma-Aldrich), and 0.5 µl 0.2 U/µl alkaline phosphatase (Sigma-Aldrich) were added in a final reaction volume of 30 µl, mixed by vortexing, and incubated at 37 °C overnight. Excess protein was precipitated from the solution by addition of three volumes of ice-cold methanol, mixed by vortexing, and centrifuged at 16,100 x *g* for 20 min at 4 °C. The supernatant was vacuum centrifuged at room temperature until dry. The resulting pellet was dissolved in 30 µl mobile phase buffer (5 % methanol in water).

The relative abundance of genomic dC is much higher than dU. As a result, the ^{13}C-isotope of dC with a mass spectrometrically indistinguishable molecular weight leads to an elevated background that obfuscates the dU signal despite good chromatographic separation. The nucleosides were therefore subjected to an additional separation step prior to LC/MS/MS analysis with a Zorbax C18 reverse phase chromatography column (Agilent Technologies,

Santa Clara, CA, USA) on a Hewlett Packard Series 1100 HPLC system containing a 1046A programmable fluorescence detector (Hewlett Packard, Palo Alto, CA, USA) using a 5 µl injection volume for 5 µg hydrolyzed DNA. The following program was used with a flow rate of 200 µl/min using water with 0.1 % formic acid (solution A) as the aqueous phase and methanol with 0.1% formic acid (solution B) as the organic phase: 5 % solution B until 30 sec after injection, followed by a linear gradient to 50 % solution B at 6 min, and a sharp gradient back to 5 % solution B from 7.5 to 7.6 min (figure 2-2). The nucleotides were quantitated by measuring absorption at 280 nm and analyzed with Chem Station for LC 3D software, rev. A.09.01 (Agilent Technologies). The resulting values were used to calculate the total amount of DNA per sample and compared to the previously spectrophotometrically determined concentration of purified DNA prior to hydrolysis to estimate the yield of the hydrolysis reaction. The eluate was fractionated using a Frac-920 fraction collector (Amersham Biosciences/GE Healthcare, Pittsburg, PA, USA). The elution times for dC and dU were observed to be 3 and 8 min, respectively. The dU fraction was therefore collected from minute 6 to minute 12 to ensure complete dU retrieval and maximal dC removal. Finally, the dU fraction was vacuum centrifuged at room temperature until dry and the pellet dissolved in 25 µl water containing 2 mM 2′-deoxyuridine-2-^{13}C;1,3-^{15}N$_2$ (^{13}C^{15}N$_2$-dU, C/D/N Isotopes Inc., Pointe-Claire, Quebec, Canada).
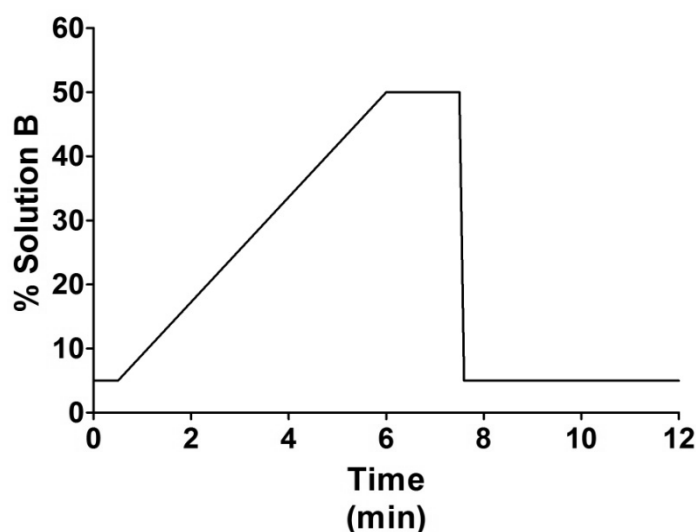

Figure 2-2: HPLC organic mobile phase gradient.

2.6 Deoxyuridine detection and quantitation

Final dU quantitation was performed using a LC-20AD HPLC system (Shimadzu Corporation, Kyoto, Japan) coupled to an API 5000 triple-quadrupole mass spectrometer (Applied Biosystems, Carlsbad, CA, USA) operating in positive electrospray ionization

mode. HPLC settings were the same as described above with an injection volume of 20 µl. Fragmentation of the dU molecular ion of 229.1 amu produced a major fragment ion of 113.1 amu, most probably corresponding to protonated U produced by breakage of the N-glycosylic bond between uracil and the deoxyribose. The collision energy used was within the range necessary to break the glycosylic bond [92]. Using identical settings, the $^{13}C^{15}N_2$-dU molecular ion of 232.1 amu produced a major 116.0 amu fragment ion. This further supported the conclusion that the glycosylic bond was cleaved. The following mass transitions were chosen to build the final MRM method: m/z 229.1 → 113.1 for dU and 232.1 → 116.0 for $^{13}C^{15}N_2$-dU (figure 2-3). A standard curve from 2 to 2000 nM dU was used for quantitation and 40 fmol $^{13}C^{15}N_2$-dU were added as an internal standard for quantitation. The results were analyzed on Analyst v.1.4.2 software (Applied Biosystems).

dU: 229.1 amu   **U: 113.1 amu**

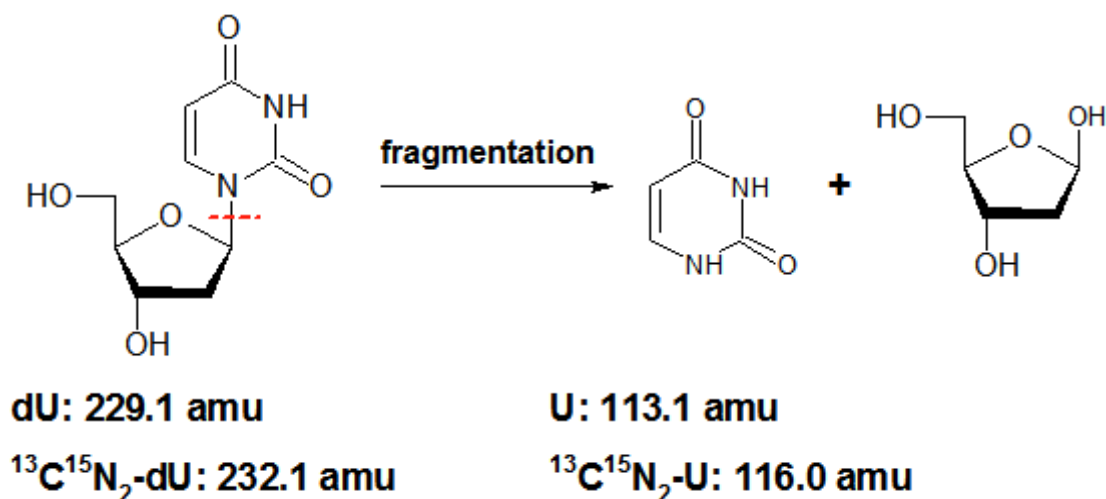**$^{13}C^{15}N_2$-dU: 232.1 amu**   **$^{13}C^{15}N_2$-U: 116.0 amu**

Figure 2-3: Overview of collision cell fragmentation. The N-glycosylic bond (red circle) between U and deoxyribose is broken, resulting in deoxyribose and U, which passes through to the mass analyzer.

3. RESULTS

3.1 Sample selection

Three sources of DNA were used in assay validation: genomic DNA from salmon sperm nuclei, genomic DNA from *UNG^{-/-}* mouse kidneys, and a 44-bp-long oligomer containing one uracil. Salmon sperm DNA was employed on the basis of its purity, widespread use in laboratories, and presumed low basal uracil levels. The oligomer provided a known quantity of dU that could be used to determine the accuracy of the assay. Finally, *UNG^{-/-}* mouse kidneys were used because of their reported increased basal uracil levels and to confirm that the assay could be applied to non-commercially-purified DNA. Both kinds of genomic DNA were sometimes heated to 95 °C for various times to induce cytosine deamination to uracil and ensure that a measurable level of dU was present in samples.

3.2 Units for genomic uracil content

Reports of genomic U content use a variety of units including grams, moles, and U molecules per diploid cell or million base pairs. Here, genomic dU was measured and will be reported in moles for experiments measuring pure dU or in dU molecules per million deoxynucleosides ($dU/10^6$ dN) for experiments measuring dU from hydrolyzed DNA. Moles dU were converted to $dU/10^6$ dN by dividing μmol dU measured by LC/MS/MS by the average of mol dG and dT measured by spectrophotometry (see section 3.4) during the precursory HPLC step. The average quantity of dG and dT represent the amount of dNs present per sample. Thus the conversion reflects one mol dU per 10^6 mol dN from hydrolyzed DNA.

3.3 Overcoming contamination and ^{13}C-deoxycytidine interference

As previously discussed, the relative abundance of dC in DNA is much greater than dU and the monoisotopic mass of ^{13}C-dC is isobaric with dU. As a result, the "tail" from the chromatographic peak of ^{13}C-dC obscured the dU peak. We therefore employed a precursory HPLC step to remove most of the dC prior to LC/MS/MS analysis. Using a solution of 10 μM pure dNs, the retention times of dC and dU were determined to be 3.6 and 8.4 min, respectively (figure 3-1). The dU fraction was subsequently collected from 6 to 12 min.
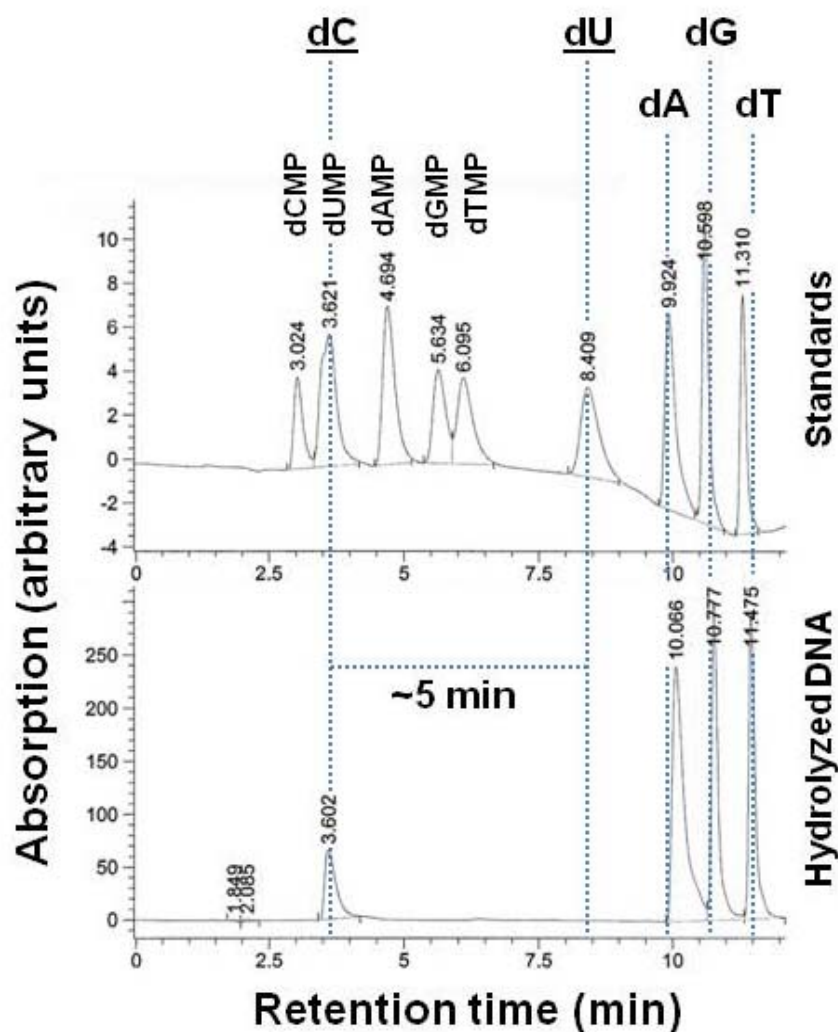
Figure 3-1: Chromatographic separation of dC and dU. The top chromatogram shows a mixture of pure dNMPs and dNs used to elucidate retention times. dC was observed to elute ~5 min before dU. Note that the dC and dUMP peaks overlap. The bottom chromatogram shows successfully hydrolyzed DNA, the only measurable peaks for which were dC, dA, dG, and dT.

A sample of 5 μg DNA from *UNG*[-/-] mouse kidney was employed to validate the effect of fractionation upon LC/MS/MS analysis. Purified dNs were used to determine the retention times of dC and dU (data not shown). Figure 3-2 shows two experiments using 5 μg of DNA with and without precursory HPLC fractionation prior to analysis. The ^{13}C-dC peak area was 83-fold greater in the unfractionated sample than in the fractionated sample; however, the amount ratio is most likely higher because the ion detector was saturated by the ^{13}C-dC peak in the unfractionated sample. Notably, the dU peak (corresponding to 250 fmol dU or 14.5 dU molecules per 10^6 dN) was obscured by the ^{13}C-dC peak tail in the unfractionated sample, underlining the importance of the precursory HPLC step.

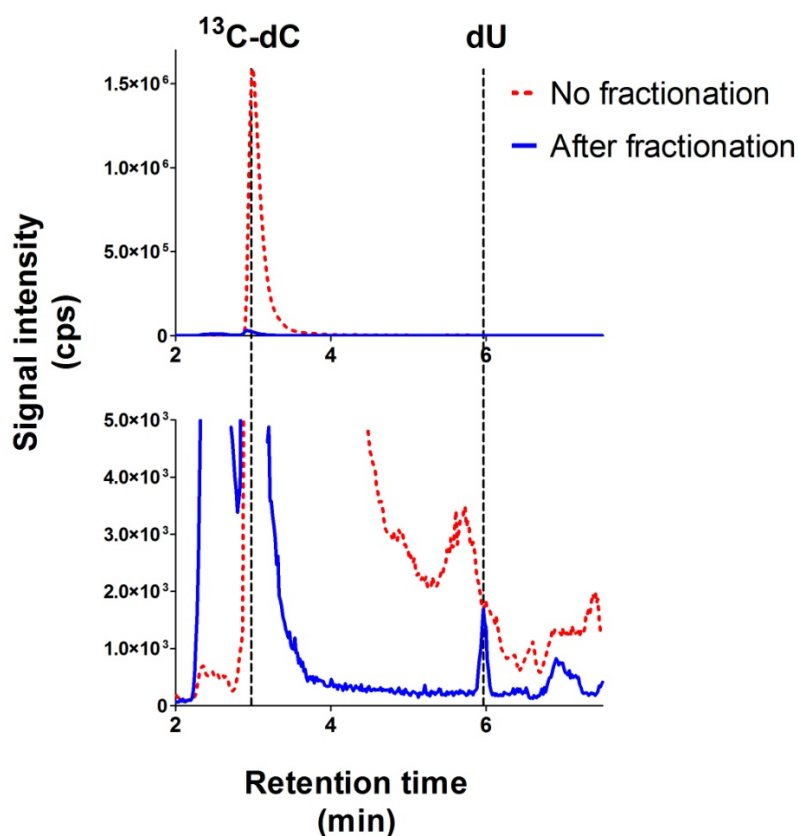**Figure 3-2: Effect of precursory HPLC step on LC/MS/MS chromatograms. Both chromatograms show the same data displayed with different y-axis scales. The dU peak is obscured by the ^{13}C-dC peak tail in the sample with no precursory HPLC step, whereas the peaks are distinguishable from one another in dU fraction that underwent precursory HPLC.**

It was suspected that intracellular dUMPs were retained during DNA isolation and dephosphorylated to dU by alkaline phosphatase, thereby introducing bias by increasing dU signal. Treating DNA samples with only alkaline phosphatase following isolation should dephosphorylate possible dUMP contaminations to dU without hydrolyzing DNA. To remove dUMP prior to DNA hydrolysis, DNA was treated with alkaline phosphatase followed by a second purification step in which only DNA and not free dU from dUMP would bind to the DNA isolation column. By this approach, it was found that up to 25% of measured dU was a result dUMP dephosphorylation, corresponding to 8.25 ± 1 fmol dU (0.478 ± 0.06 dU/10^6 dN) for 10 µg DNA. When pre-treated with alkaline phosphatase, no dU was measured in DNA from samples given only alkaline phosphatase without nucleases (figure 3-3A). Furthermore, the dUMP contamination signal combined with the signal from a hydrolyzed DNA sample pre-treated with alkaline phosphatase equaled the signal from a hydrolyzed DNA sample without pre-treatment, suggesting that intracellular dUMP was the culprit responsible for this contamination (figure 3-3B).
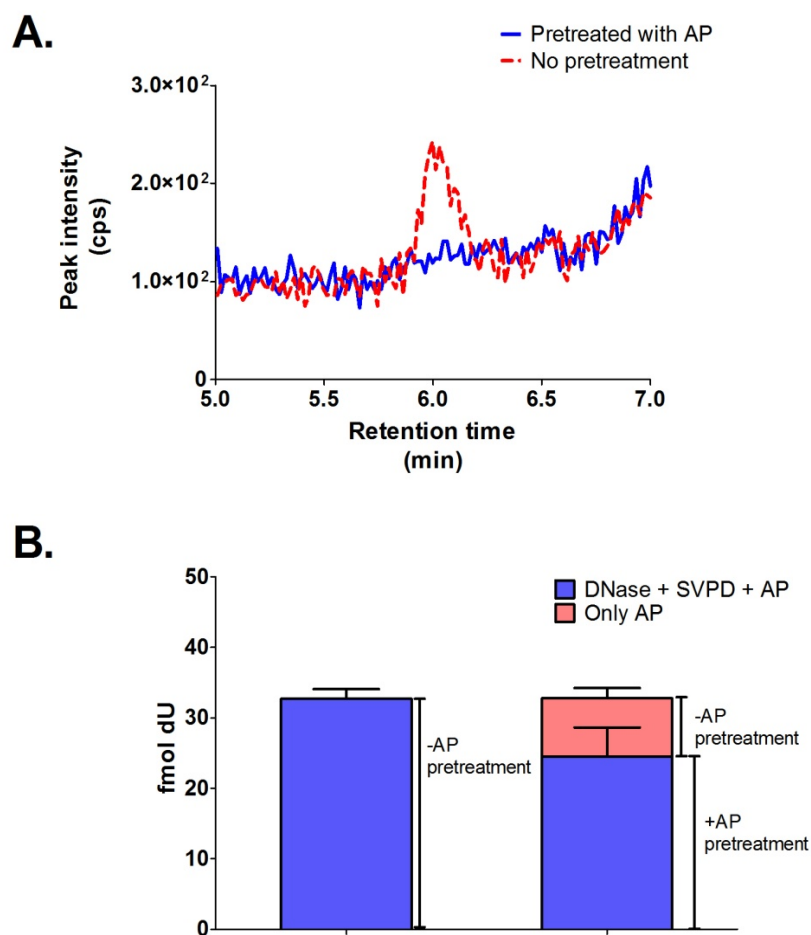
Figure 3-3: Pre-treament with alkaline phosphatase eliminates dU signal originating from intracellular dUMP contaminants in DNA samples. A. Chromatograms of hydrolyzed DNA with and without alkaline phosphatase treatment. B. dU calculated from hydrolyzed DNA without pre-treament equals that of the same DNA with pre-treatment plus that of the DNA with no pre-treatment and no nucleases in hydrolysis. Data was collected by Anastasia Galashevskaya and represents the means ±SD of triplicate LC/MS/MS runs.

Unidentified contaminations were sometimes encountered during the sample preparation steps following DNA hydrolysis. Most notably, the vacuum centrifugation step occasionally introduced dU contamination. Up to 10 fmol dU was observed in a sample containing buffer that was vacuum centrifuged in the same centrifuge as hydrolyzed DNA samples containing up to 200 fmol dU. This was attributed to aerosolized dU in the vacuum centrifuge from repeated centrifugation of hydrolyzed DNA samples. To neutralize this contamination, the following steps were taken. First, the vacuum centrifuge was regularly cleaned with water, isopropanol, and methanol. Second, samples were not centrifuged in proximity to one another. Third, control samples expected to contain no dU, such as those treated with only alkaline phosphatase, were centrifuged separately from those expected to contain dU. Finally,

31

a control vial containing the same amount of buffer as the normal samples was used every time samples were vacuum centrifuged. Figure 3-4 shows the chromatograms of a typical sample from 5 μg hydrolyzed mouse kidney DNA as well as the following controls: DNA treated with only alkaline phosphatase to control for dUMP, the HPLC fraction following that of dU to make certain all dU was collected, a vacuum centrifuge control, and a sample of the buffer in which the hydrolyzed DNA samples were dissolved.
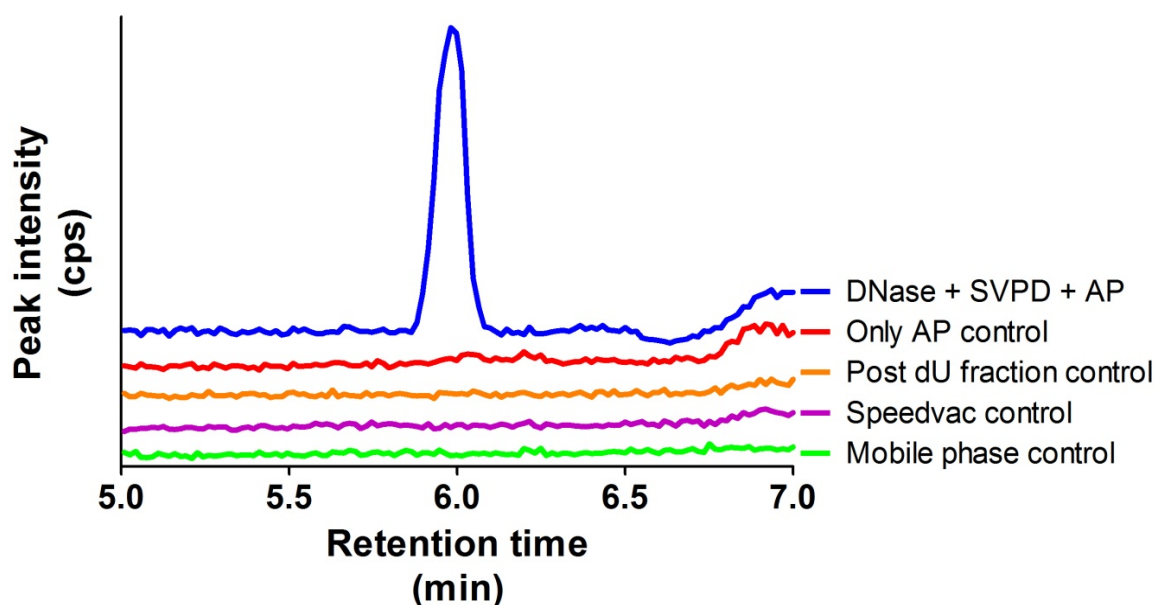

Figure 3-4: Sample LC/MS/MS chromatograms of successfully hydrolyzed DNA with no contaminations. 5 μg of UNG KO mouse kidney DNA heated to 95 °C for 45 min was used. The blue chromatogram represents a sample of hydrolyzed DNA and thus a positive signal. The controls, shown in red, orange, violet, and green, represent controls from DNA treated with only alkaline phosphatase, the fraction collected after the dU fraction during the precursory HPLC, empty buffer that was vacuum centrifuged with the samples, and pure mobile phase, respectively. Note that the chromatogram baselines overlapped, so each chromatogram was nudged above the previous one to visually distinguish them from one another.

3.4 Sample recovery

It was plausible that a substantial amount of sample could be lost during the precursory HPLC step of the assay. To test for this, known amounts of dU were subjected to precursory HPLC, processed, and analyzed with LC/MS/MS. Mean sample recovery was 97.1 % with a coefficient of variance (CV) of 8.85 %.

The yield of the hydrolysis reaction was determined by comparing the spectrophotometrically-determined DNA concentration prior to hydrolysis with the dT and dG signals measured during the precursory HPLC step. Standard curves with 0.1 to 1 mM dG and dT were analyzed by HPLC coupled to a spectrophotometer. These were used to

determine the molar extinction coefficients (ε) of dG and dT at 280 nm (data not shown). ε-values were found to be 10944 and 8245.2 mM/peak area absorption for 5 µl injections of dG and dT, respectively, and were employed to calculate the concentration of dG and dT in hydrolyzed DNA samples from the peak area absorptions. Chargaff's first rule of parity states that the global percentages of G and C as well as T and A are equal, and this has been observed as applicable to eukaryotic genomes [93,94]. Thus, the molarities of dG and dT sample were used together with their ε-values to calculate the amount of DNA in each sample. This calculation is summarized in the following equation:

$$g\ DNA = [(Abs]_G \times \varepsilon_G \times V \times MW_{dGMP+dCMP}) + [(Abs]_T \times \varepsilon_T \times V \times MW_{dTMP+dAMP})$$

where Abs is peak area absorption (in peak area absorption units), ε is the molar extinction coefficient of the nucleoside at 280 nm (in mol/l per peak area absorption units), V is injection volume (in l), and MW is the molecular weights of the nucleotide monophospate plus its base pair complement (in g/mol). The amount of DNA calculated by this method should equal the amount determined by spectrophotometry of intact DNA prior to hydrolysis. DNA from salmon sperm was employed to test this and it was observed from three separate experiments hydrolyzing 10 µg DNA that the mean yield as determined using the above quantitation was 9.88 ± 0.383 µg (98.8% yield) with a CV of 3.88%.

3.5 Assay precision

The intra-assay precision was tested in two ways. First, the LC/MS/MS portion was analyzed by comparing five individual runs of the dU standard curve. The mean observed CV for all concentrations was 7.54%, with a minimum of 0.85 % at 200 fmol and maximum of 24.44 % at 2 fmol. The precision of the entire assay was assessed by comparing six parallel dU determinations using the same batch of 10 µg salmon sperm DNA heated to 95 °C for 60 min. A CV of 5.69 % was determined for these samples.

3.6 Assay specificity

The specificity of the assay to detect only dU was confirmed in several ways. First, two chromatography steps ensure that the analyte has the same retention time as pure dU. Second, the product ion scans of dU and the internal standard, $^{13}C^{15}N_2$-dU, revealed prominent product ions isobaric to uracil and uracil-2-^{13}C;1,3-$^{15}N_2$ (figure 3-5).
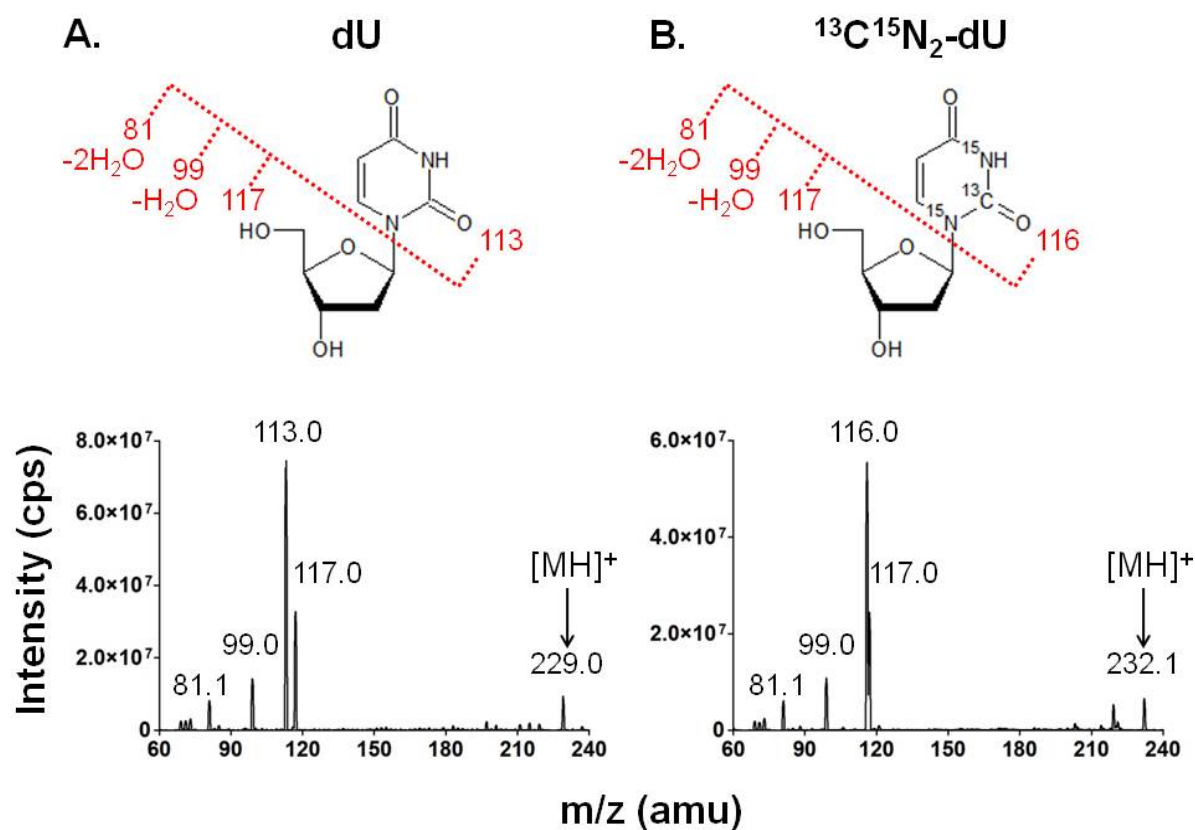
Figure 3-5: MS/MS product ion spectra of dU and $^{13}C^{15}N_2$-dU. The spectra were obtained by collision-induced dissociation of the positive molecular ions ([MH]$^+$) at m/z 229.0 (dU) and 232.1 ($^{13}C^{15}N^2$-dU). The proposed origins of key fragment ions are as indicated.

An experiment was performed with unheated salmon sperm DNA as "blank matrix" (i.e. DNA not containing U) to verify the lack of signal. The hypothesis was that salmon sperm DNA would contain no measurable dU and would therefore serve as a negative control; however, it was found that the salmon sperm DNA contained 75.4 fmol dU (4.37 dU/10^6 dN) and could therefore not be used to demonstrate that the assay does not detect dU in samples with no DNA uracil. Thus 100 and 500 fmol of an oligonucleotide containing a single U were added to the DNA to verify that the signal observed was from genomic dU (figure 3-6). The sum of the dU measured from the 100 and 500 fmol samples and the salmon sperm DNA equaled 86.96 and 95.33 % of the dU measured in the mixture of the two samples, respectively.
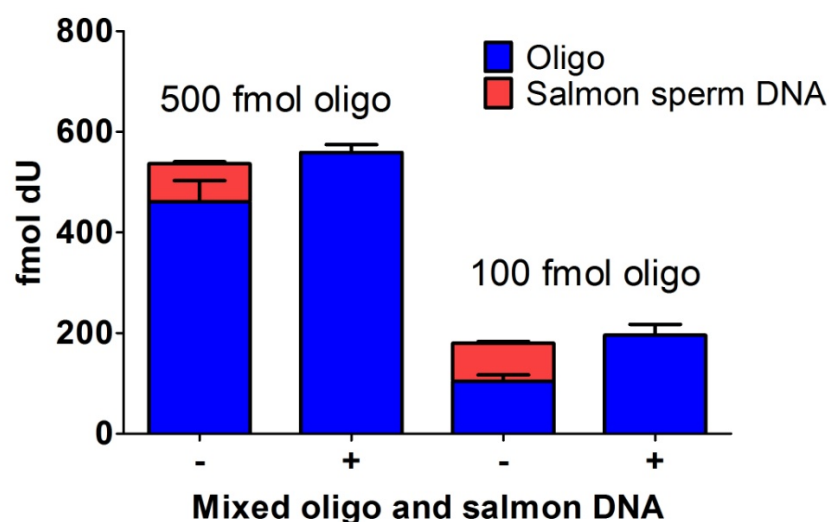
Figure 3-6: DNA from salmon sperm was spiked with known quantities of an oligo containing one U. "+" indicates that oligo and salmon sperm DNA were mixed, whereas "-" indicates separate samples of oligo and salmon sperm DNA. Signals from the oligo and salmon sperm DNA run separately equaled the signals from a mixture of the two. Data represents the means ±SD of duplicate LC/MS/MS runs.

3.7 Assay sensitivity and linearity

The limit of detection of the LC/MS/MS detection for the assay was observed to be 2 fmol dU (figure 3-7A). The signal-to-noise ratio for this amount of dU was 2.23, and the CV of five individual measurements was 24.44 %. The limit of quantitation (LOQ) was calculated to be 5 fmol dU, with a signal-to-noise ratio of 6.70 and a CV of 17.04 % (figure 3-7B).

Figure 3-7: LC/MS/MS chromatograms of limit of detection (A.) and limit of quantitation (B). The areas under the peaks between the blue lines were used for quantitation.

The linearity of the detection step was ascertained by running a standard curve of 2 to 2000 fmol dU in 40 fmol $^{13}C^{15}N_2$-dU (figure 3-8). The detection exhibited a high range and was

linear across three orders of magnitude. The R^2-value of the linear regression curve was 0.9969. The mean CV for all concentrations of the standard curve was 6.11 %.

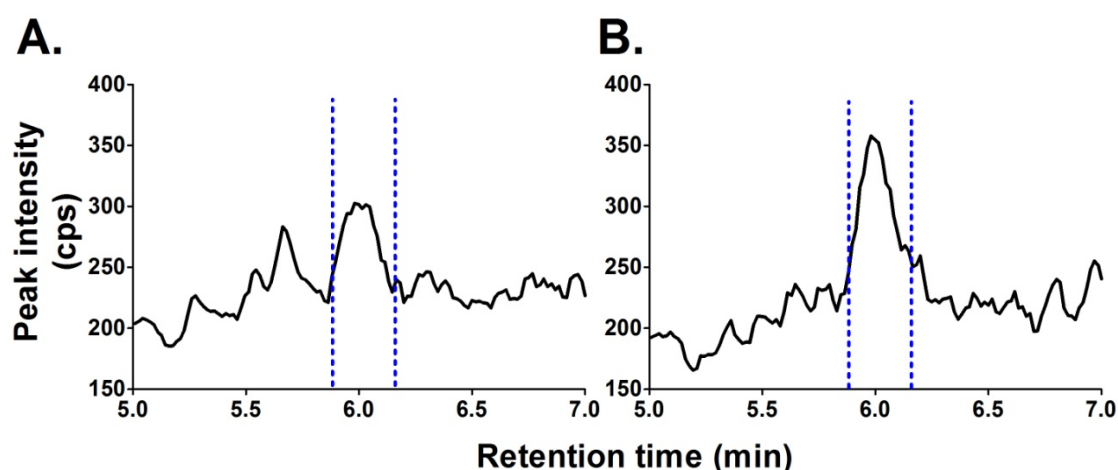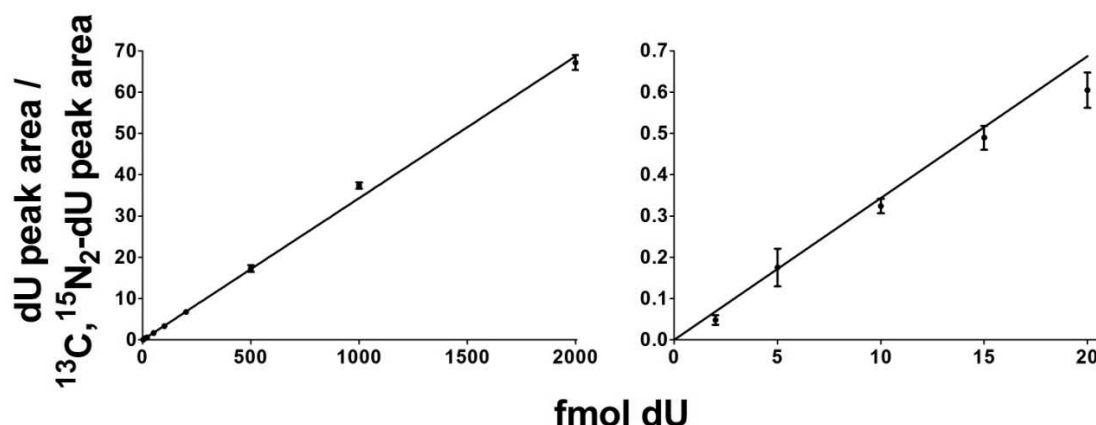**Figure 3-8: Plot of standard curve used for dU quantitation. A concentration gradient from 2 to 2000 fmol dU was used with 40 fmol $^{13}C^{15}N_2$-dU as an internal standard. The calculated R^2-value was 0.9969. Both graphs show the same data with different axes to emphasize that the assay is linear at the entire range of concentrations used. Data represents the means ±SD of triplicate LC/MS/MS runs.**

3.8 Spontaneous cytosine deamination by heating DNA to 95 °C

To monitor dU formation from cytosine deamination, 10 µg salmon sperm DNA was heated in water to 95 °C for up to 120 min (figure 3-9). The measured basal dU level of salmon sperm DNA was 4.57 ± 0.081 dU/10^6 dN. After five minutes of heating, measured dU levels rose almost three-fold to 13.0 ± 0.691 dU/10^6 dN. dU levels rose linearly to 218 ± 5.86 dU/10^6 dN at 120 min, exhibiting a slope of 1.78 ± 0.030 dU/10^6 dN per min and a R^2 of 0.9955.
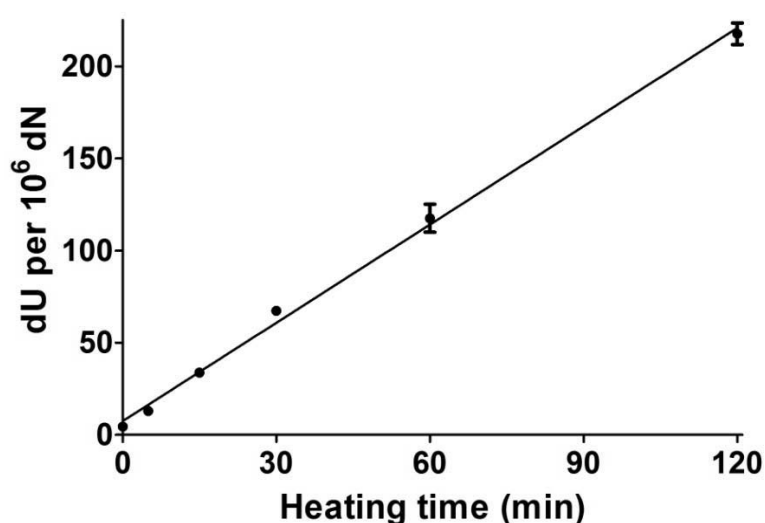
Figure 3-9: Spontaneous deamination of dC to dU by heating DNA to 95 °C for up to 120 min. Data represents the means ±SD of triplicate LC/MS/MS runs.

3.9 UNG excision of genomic uracil *in vitro*

To test whether UNG excised all uracil from DNA *in vitro*, 5 µg of preheated and non-heated DNA from *UNG$^{-/-}$* mouse kidney was treated with 1 µg of UNG and purified prior to hydrolysis (figure 3-10). The calculated activity for the batch of UNG-Δ84 used was 0.750 U/µg. UNG-treated preheated DNA samples exhibited 49.8 ± 1.95 % of dU in non-UNG-treated samples, whereas UNG-treated samples in non-heated DNA exhibited 79.1 ± 1.33 % of dU in non-UNG-treated samples. Namely, UNG excised ~50 % of uracil from the heated DNA sample, which was calculated to have an initial dU concentration of 10.75 nM, and ~20 % of uracil from the unheated DNA sample, which was calculated to have an initial dU concentration of 2.42 nM. It was therefore observed that UNG did not excise all uracil substrate from DNA *in vitro*. Notably, the preheated sample, which was higher in dU, displayed a more pronounced reduction in dU when treated with UNG. Given that the reported K_m-value for recombinant UNG-Δ84 is 4.5 µM U, this suggests that UNG's inability to excise all uracil in the DNA may have been a consequence of low substrate concentration [68].

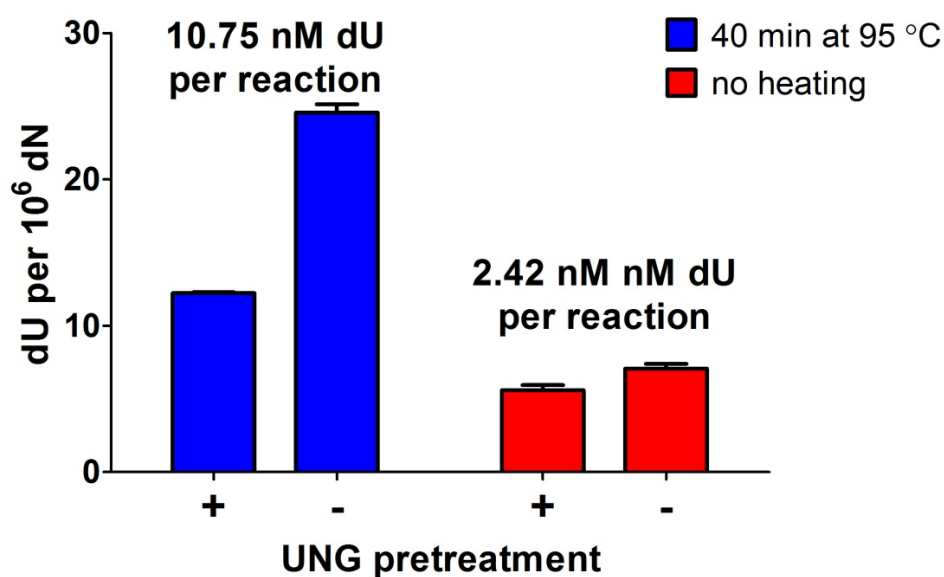**Figure 3-10: Pre-treatment with UNG did not excise all genomic uracil. Blue bars represent DNA that was artificially deaminated by heating for 40 min at 95 °C and red bars represent unheated DNA. The calculated dU concentration in the initial UNG reactions were 10.75 and 2.42 nM dU per reaction for heated and unheated DNA, respectively. Data represents the means ±SD of duplicate LC/MS/MS runs.**

4. DISCUSSION

Genomic uracil can be both beneficial and damaging. It can be mutagenic and is therefore repaired by the BER pathway. Contrastingly, AID damages DNA in the Ig-loci in B-cells by inducing C deamination to U, a crucial step in adaptive immunity. Several attempts have been made to directly measure the absolute quantity of U in DNA samples, but wide variation in reported genomic U values have been suggested as a result of technical shortcomings in the methods used [3]. The method presented in this thesis (summarized in figure 4-1) aims to address these shortcomings.
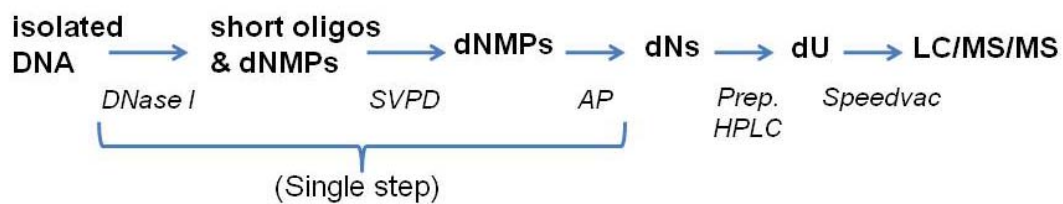
Figure 4-1: Method overview. DNA is extracted by P:C:I and treated with alkaline phosphatase to remove dUMP. Purified DNA is then treated with DNase I, SVPD, and alkaline phosphatase to yield dNs. A precursory HPLC step separates the majority of dC from dU, which is finally quantitated by LC/MS/MS.

The validation of the assay presented here was centered on the following facets of analysis: specificity, accuracy, precision, and sensitivity. The initial questions were whether dU was indeed what was being measured and whether all dU was recovered from the samples. To begin to answer the proposed questions, steps were taken to make certain that samples were not contaminated with dU. The controls used in the experiments and their rationales were as follows:

- Samples treated with only alkaline phosphatase to control for dUMP in the sample mixture. Alkaline phosphatase would dephosphorylate dUMPs to dUs in an inadequately prepared sample, resulting in a positive signal. Any dU in the sample solution prior to hydrolysis of DNA would also result in a positive signal with this control.

- The fractions that came after the dU fractions during the precursory HPLC step to confirm that no dU remained on the chromatography column.

- A vial containing only buffer during the vacuum centrifuge steps to control for dU contamination from the vacuum centrifuge.

- A sample containing only mobile phase to control for dU within the mobile phase used to dissolve dU pellets prior to LC/MS/MS analysis.

Thus it is reasonable to conclude that the dU measured originated from the sample DNA. That dU was indeed being measured was confirmed in several ways. First, the sample preparation included two separate HPLC steps. The retention time of pure dU was elucidated for both steps, so the analyte binds identically to both chromatography columns. Second, tandem mass spectrometry ensures that the analyte is isobaric with dU and fragments to yield a product ion isobaric with U. Lastly, salmon sperm DNA was spiked with an oligo containing one U per 44 dN. The measured dU signal corresponded to the expected dU concentration based on the amount oligo added to the samples, confirming that the analyte measure was indeed dU.

Spiking genomic DNA with an oligo containing a known amount of U also confirm that all dU was recovered from the DNA. The main control for this in every experiment was comparison of the DNA concentration calculated from the dN concentrations during the precursory HPLC step and the DNA concentration measured spectrophotometrically before hydrolysis. That the extrapolated DNA concentrations were equivalent to the concentration measured before hydrolysis indicates that all the DNA was successfully hydrolyzed and thus that all genomic dU was recovered.

The intra-assay precision and sensitivity were more straightforward to determine. For intra-assay precision, the dU content of one DNA sample was measured in six individual experiments, the variation for which was below 10 %. The variability introduced by the LC/MS/MS portion of the assay was analyzed independently by comparing five runs of the dU standard curve. The standard curve demonstrated a linear range over three orders of magnitude with a high degree of linearity at all ranges. A very high sensitivity was observed as well, with a limit of detection of 2 fmol dU and limit of quantitation of 5 fmol dU. Nevertheless, this is close to the limit relevant for measuring genomic uracil from biological samples and improvements upon sensitivity would extend the assay's biological applications. The results from heating DNA at 95 °C to induce deamination underscore the ease with which bias may be introduced to a direct quantitation. The assays described by Chango *et al.* and Dong *et al.* both employ a DNA denaturation step in which samples are heated to 95 °C for five minutes [13,58,60]. Here it was demonstrated that even such a brief heating step can significantly inflate the amount of dU in a sample. These results are in accordance with a similar experiment reported by Ren *et al.*, in which a similar treatment of DNA resulted in an increase of genomic uracil measured after one minute and a 40-fold increase after twenty minutes [59]. The amount increase in dU after twenty minutes calculated from the linear regression curve in the experiments presented here was about 36-fold. The variance between

the two values may result from a difference between the methods used. Ren *et al.* made no comment as to the linearity or precision of this data, but visual inspection of the figure presented suggests that their data were less linear than the data presented here. Another explanation is that Ren *et al.* used DNA from human peripheral blood leukocytes instead of from salmon sperm nuclei as presented here. The reaction conditions during heating may have also differed -- the rate of C deamination has been demonstrated to be pH-dependent as well as temperature dependent [2]. The sample buffer composition could therefore affect deamination. In addition, factors such as DNA length and DNA sequence may all play a role in heat induced C deamination, so the deamination rate cannot be assumed constant across samples.

The introduction of genomic uracil due to heating steps adds a note of caution to any assay involving the direct quantitation of genomic uracil because cytosine is usually present in much higher quantities than uracil and can be easily deaminated [1,2,7]. This issue may also affect the assay presented here. Although samples containing genomic DNA were not heated to 95 °C as a part of dU quantitation, the DNA isolation protocol involved a 1 h incubation step at 60 °C. Furthermore, the tactile temperature of the sample tubes in the vacuum centrifuge was observed to rise during prolonged centrifugation, though this was not quantified. It is important to note that genomic DNA is does not denature at 60 °C (though some denaturation may occur), so the deamination rate for heating samples at 60 °C is likely lower than heating them at 95 °C. Nevertheless, any sample heating increases the probability of genomic cytosine deamination, so future refinements to the assay should therefore include an attempt to maintain a low and constant sample temperature.

The samples treated with UNG-Δ84 contained a determined 0.75 U of the enzyme. A unit of U is defined as 1 nmol uracil excised per minute. In the reactions used, this corresponds to 75 fmol uracil excised per minute and 4500 fmol uracil excised total. The sample of heated DNA was calculated to contain 491 fmol dU. There are several possible explanations for the enzyme's inability to excise all uracil substrate from the DNA samples. As mentioned earlier, UNG exhibits intrinsic preference towards ssDNA over dsDNA and U:G pairs over U:A pairs, as well as sequential preference [32,67,68]; however, differences in UNG activity to excise uracil in certain contexts were not large enough to explain why so little uracil was excised.

The more likely explanation has to do with the kinetics of uracil excision by UNG. The K_m value for recombinant UNG under the reaction conditions used here was reported to be 4.5 μM substrate (uracil). This value represents half the enzyme's maximum possible reaction

rate. In line with the Michaelis-Menten model of enzyme kinetics, lower substrate concentrations result in lower reaction rates [1]. It was calculated that the DNA samples treated with UNG contained 10.75 and 2.42 nM dU for heated and unheated DNA samples, respectively. These values are well below the K_m for UNG under the reaction conditions used (419- and 1860-fold lower for heated and unheated DNA samples, respectively), so the reaction rates for UNG excision were likely very low. This explains why more U was excised from the heated DNA sample than the unheated DNA sample: the higher total U substrate in the heated sample resulted in a higher reaction rate and therefore more excised U. As U concentration approaches zero, the reaction rate also approaches zero and the time required for complete U excision by UNG becomes infinite.

It is also important to note that the experiments done to carry out the classification of the enzyme employed short oligomers containing a known quantity of U or genomic DNA with a high abundance of U [59,61,62,85]. The mechanism of DNA lesion search by UNG is not yet fully elucidated, but structural, thermodynamic, and kinetic data suggests a mechanism whereby the enzyme binds randomly to DNA and scans it one-dimensionally, sometimes "hopping" to other locations of the DNA proximal to where it was previously bound [95]. With this in mind, estimation of the rate of U excision by UNG gains complexity because the duration of DNA scanning and U recognition are DNA strand length-dependent [95,96]. That is, given two samples with short or long DNA strands and equimolar amounts of U, more U would theoretically be excised from the sample with shorter DNA because the UNG does not have to scan through as much DNA as the sample containing longer DNA. Thus it cannot be assumed that UNG excises all uracil in a genomic DNA context *in vitro*. Au contraire, it is unlikely that even a highly-concentrated amount of UNG can excise the low amounts of U found in normal genomic DNA.

The two assays by Mashiyama *et al.* and Ren *et al.* mentioned earlier that utilized UNG to excise uracil therefore may not have measured the actual levels of genomic U in their samples [59,61]. This does not necessarily mean that the conclusions they drew about the biology were inaccurate. The relative (though not the absolute) genomic U levels of the measure samples may have been accurate because of the similarity between samples. The results presented here suggest only that the absolute genomic U levels may have been inaccurate; however, the reaction conditions for U excision by UNG as well as the enzyme purification method were different than those employed here. While a more thorough comparison of the above assays may reveal that a much greater catalytic activity of UNG in

an *in vitro* genomic DNA context for the two assays above, the ability for UNG to excise all uracil in such a context nevertheless remains a source of bias for assays that employ it. The sensitivities of and some of the difference between the four genomic uracil quantitation assays discussed are summarized in table 4-1. Both Mashiyama *et al.* and Ren *et al.* present assays with high sensitivities, but they suffer the drawback of employing UNG to excise uracil. Furthermore, they derivatize their samples, which adds more bias to their results because the extent of derivatization was not controlled.

Table 4-1: Comparison of the different analytes, detection methods, and sensitivities of the genomic uracil quantitation assays discussed versus the assay presented here.

Reference	Analyte	Derivatized samples	DNA heating	Detection method	DNA / assay (μg)	LOQ (fmol)
Mashiyama *et al.* [61]	U	yes	no	GC/MS	5	17.8
Ren *et al.* [59]	U	yes	no	LC/MS/MS	5	2
Chango *et al.* [60]	dU	no	yes	LC/MS	1	20000
Dong *et al.* [13,58]	dU*	no	yes	LC/MS/MS	50	100
Current assay	dU*	no	no	LC/MS/MS	5	2

*These methods employed a precursory HPLC step.

Chango *et al.* claim to measure genomic dU from only 1 μg DNA from human liver carcinoma cells, though they report a limit of quantitation two orders of magnitude less sensitive than the other assays discussed. Furthermore, they do not address the problem that monoisotopic dU and ^{13}C-dC are isobaric, despite omitting a precursory HPLC step and reporting a chromatographic separation of only two minutes. Their method description also mentions that dU was detected at a mass of 113 amu, the molecular weight for free uracil. A single mass spectrometer was used, so the mass stated could not have been that of a dU product ion. Finally, despite reporting an absolute dU concentration within the range of reported literature values (about 9 dU/10^6 dN), their experiments claim dU/dT ratios of 0.022 to 0.500, corresponding to about 42,500 and 125,000 dU/10^6 dN, respectively. These amounts of genomic uracil are much higher than anything reported in the literature. It may therefore be possible that Chango *et al.* may have mistakenly measured something besides dU or introduced dU into their samples by e.g. contamination or heat-induced deamination. In contrast, Dong *et al.* employed a precursory HPLC separation of dU from dC prior to LC/MS/MS analysis, however, the sensitivity of their assay is at least one order of magnitude less than the assays employing UNG. Additionally, the assay subjects DNA to 95 °C, thereby deaminating dC to dU and introducing bias to the results.

The method for genomic uracil quantitation reported here circumvents the problems described above. The use of UNG to excise uracil is circumvented by hydrolyzing DNA to deoxynucleotides. Unlike other assays, the DNA is not heated prior to hydrolysis, so heat-induced cytosine deamination is reduced. Furthermore, the sensitivity of this assay is equal or greater than the assays discussed. In conclusion, the assay presented here is a sensitive and highly-specific method for the quantitation of absolute genomic uracil that ameliorates shortcomings of similar methods.

5. REFERENCES

1.   Voet D, Voet JG. Biochemistry. New York: J. Wiley & Sons; 2004.
2.   Lindahl T, Nyberg B. Heat-induced deamination of cytosine residues in deoxyribonucleic acid. Biochemistry 1974;13:3405-10.
3.   Olinski R, Jurgowiak M, Zaremba T. Uracil in DNA-Its biological significance. Mutat Res 2010;705:239-45.
4.   Hagen L, Pena-Diaz J, Kavli B, Otterlei M, Slupphaug G, Krokan HE. Genomic uracil and human disease. Exp Cell Res 2006;312:2666-72.
5.   Visnes T, Doseth B, Pettersen HS, Hagen L, Sousa MM, Akbari M, Otterlei M, Kavli B, Slupphaug G, Krokan HE. Uracil in DNA and its processing by different DNA glycosylases. Philos Trans R Soc Lond B Biol Sci 2009;364:563-8.
6.   Lindahl T. Instability and decay of the primary structure of DNA. Nature 1993;362:709-15.
7.   Hickson I. Base excision repair of DNA damage. Chapman & Hall; 1997.
8.   Impellizzeri KJ, Anderson B, Burgers PM. The spectrum of spontaneous mutations in a Saccharomyces cerevisiae uracil-DNA-glycosylase mutant limits the function of this enzyme to cytosine deamination repair. J Bacteriol 1991;173:6807-10.
9.   Seeberg E, Kleppe K, European Molecular Biology Organization., North Atlantic Treaty Organization. Scientific Affairs Division. Chromosome damage and repair. New York: Plenum Press; 1981. xiv, 623 p. p.
10.   Sousa MM, Krokan HE, Slupphaug G. DNA-uracil and human pathology. Mol Aspects Med 2007;28:276-306.
11.   Sono M, Wataya Y, Hayatsu H. Role of bisulfite in the deamination and the hydrogen isotope exchange of cytidylic acid. J Am Chem Soc 1973;95:4745-9.
12.   Morris SM, Jr., Billiar TR. New insights into the regulation of inducible nitric oxide synthesis. Am J Physiol 1994;266:E829-39.
13.   Dong M, Dedon PC. Relatively small increases in the steady-state levels of nucleobase deamination products in DNA from human TK6 cells exposed to toxic levels of nitric oxide. Chem Res Toxicol 2006;19:50-7.
14.   Holmquist GP, Gao S. Somatic mutation theory, DNA repair rates, and the molecular epidemiology of p53 mutations. Mutat Res 1997;386:69-101.
15.   Tessman I, Liu SK, Kennedy MA. Mechanism of SOS mutagenesis of UV-irradiated DNA: mostly error-free processing of deaminated cytosine. Proc Natl Acad Sci U S A 1992;89:1159-63.
16.   Slotkin RK, Martienssen R. Transposable elements and the epigenetic regulation of the genome. Nat Rev Genet 2007;8:272-85.
17.   Bronchud M. Principles of molecular oncology. Humana Press; 2008.
18.   Yebra MJ, Bhagwat AS. A cytosine methyltransferase converts 5-methylcytosine in DNA to thymine. Biochemistry 1995;34:14752-7.
19.   Shen JC, Rideout WM, 3rd, Jones PA. High frequency mutagenesis by a DNA methyltransferase. Cell 1992;71:1073-80.
20.   Neuberger MS, Harris RS, Di Noia J, Petersen-Mahrt SK. Immunity through DNA deamination. Trends Biochem Sci 2003;28:305-12.
21.   Shen JC, Zingg JM, Yang AS, Schmutte C, Jones PA. A mutant HpaII methyltransferase functions as a mutator enzyme. Nucleic Acids Res 1995;23:4275-82.
22.   Muramatsu M, Sankaranand VS, Anant S, Sugai M, Kinoshita K, Davidson NO, Honjo T. Specific expression of activation-induced cytidine deaminase (AID), a novel

member of the RNA-editing deaminase family in germinal center B cells. J Biol Chem 1999;274:18470-6.

23. Maul RW, Gearhart PJ. AID and somatic hypermutation. Adv Immunol 2010;105:159-91.

24. Liu M, Schatz DG. Balancing AID and DNA repair during somatic hypermutation. Trends Immunol 2009;30:173-81.

25. Friedberg EC, Walker G, Siede W. DNA repair and mutagenesis. ASM Press; 1995.

26. Traut TW. Physiological concentrations of purines and pyrimidines. Mol Cell Biochem 1994;140:1-22.

27. Bekesi A, Zagyva I, Hunyadi-Gulyas E, Pongracz V, Kovari J, Nagy AO, Erdei A, Medzihradszky KF, Vertessy BG. Developmental regulation of dUTPase in Drosophila melanogaster. J Biol Chem 2004;279:22362-70.

28. Carreras CW, Santi DV. The catalytic mechanism and structure of thymidylate synthase. Annu Rev Biochem 1995;64:721-62.

29. Mosbaugh DW, Bennett SE. Uracil-excision DNA repair. Prog Nucleic Acid Res Mol Biol 1994;48:315-70.

30. Nilsen H, Krokan HE. Base excision repair in a network of defence and tolerance. Carcinogenesis 2001;22:987-98.

31. Akbari M, Pena-Diaz J, Andersen S, Liabakk NB, Otterlei M, Krokan HE. Extracts of proliferating and non-proliferating human cells display different base excision pathways and repair fidelity. DNA Repair (Amst) 2009;8:834-43.

32. Hagen L, Kavli B, Sousa MM, Torseth K, Liabakk NB, Sundheim O, Pena-Diaz J, Otterlei M, Horning O, Jensen ON and others. Cell cycle-specific UNG2 phosphorylations regulate protein turnover, activity and association with RPA. EMBO J 2008;27:51-61.

33. Visnes T, Akbari M, Hagen L, Slupphaug G, Krokan HE. The rate of base excision repair of uracil is controlled by the initiating glycosylase. DNA Repair (Amst) 2008;7:1869-81.

34. Andersen S, Heine T, Sneve R, Konig I, Krokan HE, Epe B, Nilsen H. 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 2005;26:547-55.

35. Garinis GA, Jans J, van der Horst GT. Photolyases: capturing the light to battle skin cancer. Future Oncol 2006;2:191-9.

36. An Q, Robins P, Lindahl T, Barnes DE. C --> T mutagenesis and gamma-radiation sensitivity due to deficiency in the Smug1 and Ung DNA glycosylases. EMBO J 2005;24:2205-13.

37. McCullough AK, Dodson ML, Lloyd RS. Initiation of base excision repair: glycosylase mechanisms and structures. Annu Rev Biochem 1999;68:255-85.

38. Osheroff WP, Jung HK, Beard WA, Wilson SH, Kunkel TA. The fidelity of DNA polymerase beta during distributive and processive DNA synthesis. J Biol Chem 1999;274:3642-50.

39. Roitt IM, Brostoff J, Male DK. Immunology. Edinburgh ; New York: Mosby; 2001. 480 p. p.

40. Peled JU, Kuang FL, Iglesias-Ussel MD, Roa S, Kalis SL, Goodman MF, Scharff MD. The biochemistry of somatic hypermutation. Annu Rev Immunol 2008;26:481-511.

41. Kavli B, Otterlei M, Slupphaug G, Krokan HE. Uracil in DNA--general mutagen, but normal intermediate in acquired immunity. DNA Repair (Amst) 2007;6:505-16.

42. Petersen-Mahrt SK, Coker HA, Pauklin S. DNA deaminases: AIDing hormones in immunity and cancer. J Mol Med 2009;87:893-7.

43. Schrader CE, Guikema JE, Linehan EK, Selsing E, Stavnezer J. Activation-induced cytidine deaminase-dependent DNA breaks in class switch recombination occur during G1 phase of the cell cycle and depend upon mismatch repair. J Immunol 2007;179:6064-71.

44. Davies EG, Thrasher AJ. Update on the hyper immunoglobulin M syndromes. Br J Haematol 2010;149:167-80.

45. Lee WI, Torgerson TR, Schumacher MJ, Yel L, Zhu Q, Ochs HD. Molecular analysis of a large cohort of patients with the hyper immunoglobulin M (IgM) syndrome. Blood 2005;105:1881-90.

46. Revy P, Muto T, Levy Y, Geissmann F, Plebani A, Sanal O, Catalan N, Forveille M, Dufourcq-Labelouse R, Gennery A and others. Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2). Cell 2000;102:565-75.

47. Imai K, Catalan N, Plebani A, Marodi L, Sanal O, Kumaki S, Nagendran V, Wood P, Glastre C, Sarrot-Reynauld F and others. Hyper-IgM syndrome type 4 with a B lymphocyte-intrinsic selective deficiency in Ig class-switch recombination. J Clin Invest 2003;112:136-42.

48. Rada C, Williams GT, Nilsen H, Barnes DE, Lindahl T, Neuberger MS. Immunoglobulin isotype switching is inhibited and somatic hypermutation perturbed in UNG-deficient mice. Curr Biol 2002;12:1748-55.

49. Andersen S, Ericsson M, Dai HY, Pena-Diaz J, Slupphaug G, Nilsen H, Aarset H, Krokan HE. Monoclonal B-cell hyperplasia and leukocyte imbalance precede development of B-cell malignancies in uracil-DNA glycosylase deficient mice. DNA Repair (Amst) 2005;4:1432-41.

50. Marusawa H. Aberrant AID expression and human cancer development. Int J Biochem Cell Biol 2008;40:1399-402.

51. Pham P, Bransteitter R, Goodman MF. Reward versus risk: DNA cytidine deaminases triggering immunity and disease. Biochemistry 2005;44:2703-15.

52. Okazaki IM, Hiai H, Kakazu N, Yamada S, Muramatsu M, Kinoshita K, Honjo T. Constitutive expression of AID leads to tumorigenesis. J Exp Med 2003;197:1173-81.

53. Endo Y, Marusawa H, Kinoshita K, Morisawa T, Sakurai T, Okazaki IM, Watashi K, Shimotohno K, Honjo T, Chiba T. Expression of activation-induced cytidine deaminase in human hepatocytes via NF-kappaB signaling. Oncogene 2007;26:5587-95.

54. Matsumoto Y, Marusawa H, Kinoshita K, Endo Y, Kou T, Morisawa T, Azuma T, Okazaki IM, Honjo T, Chiba T. Helicobacter pylori infection triggers aberrant expression of activation-induced cytidine deaminase in gastric epithelium. Nat Med 2007;13:470-6.

55. Bhattacharya N, Sarno A, Idler IS, Fuhrer M, Zenz T, Dohner H, Stilgenbauer S, Mertens D. High-throughput detection of nuclear factor-kappaB activity using a sensitive oligo-based chemiluminescent enzyme-linked immunosorbent assay. Int J Cancer 2010;127:404-11.

56. Escarcega RO, Fuentes-Alexandro S, Garcia-Carrasco M, Gatica A, Zamora A. The transcription factor nuclear factor-kappa B and cancer. Clin Oncol (R Coll Radiol) 2007;19:154-61.

57. Cassell GH. Infectious causes of chronic inflammatory diseases and cancer. Emerg Infect Dis 1998;4:475-87.

58. Dong M, Wang C, Deen WM, Dedon PC. Absence of 2'-deoxyoxanosine and presence of abasic sites in DNA exposed to nitric oxide at controlled physiological concentrations. Chem Res Toxicol 2003;16:1044-55.

59. Ren J, Ulvik A, Refsum H, Ueland PM. Uracil in human DNA from subjects with normal and impaired folate status as determined by high-performance liquid chromatography-tandem mass spectrometry. Anal Chem 2002;74:295-9.

60. Chango A, Abdel Nour AM, Niquet C, Tessier FJ. Simultaneous determination of genomic DNA methylation and uracil misincorporation. Med Princ Pract 2009;18:81-4.

61. Mashiyama ST, Hansen CM, Roitman E, Sarmiento S, Leklem JE, Shultz TD, Ames BN. An assay for uracil in human DNA at baseline: effect of marginal vitamin B6 deficiency. Anal Biochem 2008;372:21-31.

62. Mashiyama ST, Courtemanche C, Elson-Schwab I, Crott J, Lee BL, Ong CN, Fenech M, Ames BN. Uracil in DNA, determined by an improved assay, is increased when deoxynucleosides are added to folate-deficient cultured human lymphocytes. Anal Biochem 2004;330:58-69.

63. Kohn KW, Erickson LC, Ewig RA, Friedman CA. Fractionation of DNA from mammalian cells by alkaline elution. Biochemistry 1976;15:4629-37.

64. Karpen GH. Preparation of high-molecular-weight DNA from Drosophila embryos. Cold Spring Harb Protoc 2009;2009:pdb prot5254.

65. Skoog DA. Fundamentals of analytical chemistry. Belmont, CA: Thomson-Brooks/Cole; 2004.

66. Erickson BE. HPLC and the ever popular LC/MS. Anal Chem 2000;72:711A-716A.

67. Bellamy SR, Baldwin GS. A kinetic analysis of substrate recognition by uracil-DNA glycosylase from herpes simplex virus type 1. Nucleic Acids Res 2001;29:3857-63.

68. Slupphaug G, Eftedal I, Kavli B, Bharati S, Helle NM, Haug T, Levine DW, Krokan HE. Properties of a recombinant human uracil-DNA glycosylase from the UNG gene and evidence that UNG encodes the major uracil-DNA glycosylase. Biochemistry 1995;34:128-38.

69. Snyder L, Kirkland J, Dolan J. Introduction to Modern Liquid Chromatography. John Wiley & Sons; 2009.

70. Crain PF. Preparation and enzymatic hydrolysis of DNA and RNA for mass spectrometry. Methods Enzymol 1990;193:782-90.

71. Furuichi Y, Miura K. A blocked structure at the 5' terminus of mRNA from cytoplasmic polyhedrosis virus. Nature 1975;253:374-5.

72. Wang Y, Taylor JS, Gross ML. Nuclease P1 digestion combined with tandem mass spectrometry for the structure determination of DNA photoproducts. Chem Res Toxicol 1999;12:1077-82.

73. Davidson J. The biochemistry of the nucleic acids. Academic Press; 1972.

74. Sambrook J, Russell DW. Molecular cloning : a laboratory manual. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press; 2001.

75. Lee M. Integrated strategies for drug discovery using mass spectrometry. Wiley-Interscience; 2005.

76. Silverstein RM, Webster FX. Spectrometric identification of organic compounds. New York: Wiley; 1998. xiv, 482 p. p.

77. Zumdahl SS. Chemical principles. Lexington, Mass.: D.C. Heath; 1995. xxi, 1043, 94 p. p.

78. Choi SW, Friso S, Dolnikowski GG, Bagley PJ, Edmondson AN, Smith DE, Mason JB. Biochemical and molecular aberrations in the rat colon due to folate depletion are age-specific. J Nutr 2003;133:1206-12.

REFERENCES

79.   Crott JW, Mashiyama ST, Ames BN, Fenech MF. Methylenetetrahydrofolate reductase C677T polymorphism does not alter folic acid deficiency-induced uracil incorporation into primary human lymphocyte DNA in vitro. Carcinogenesis 2001;22:1019-25.

80.   Koury MJ, Horne DW, Brown ZA, Pietenpol JA, Blount BC, Ames BN, Hard R, Koury ST. Apoptosis of late-stage erythroblasts in megaloblastic anemia: association with DNA damage and macrocyte production. Blood 1997;89:4617-23.

81.   Blount BC, Mack MM, Wehr CM, MacGregor JT, Hiatt RA, Wang G, Wickramasinghe SN, Everson RB, Ames BN. Folate deficiency causes uracil misincorporation into human DNA and chromosome breakage: implications for cancer and neuronal damage. Proc Natl Acad Sci U S A 1997;94:3290-5.

82.   Ramsahoye BH, Burnett AK, Taylor C. Nucleic acid composition of bone marrow mononuclear cells in cobalamin deficiency. Blood 1996;87:2065-70.

83.   Blount BC, Ames BN. Analysis of uracil in DNA by gas chromatography-mass spectrometry. Anal Biochem 1994;219:195-200.

84.   Nilsen H, Rosewell I, Robins P, Skjelbred CF, Andersen S, Slupphaug G, Daly G, Krokan HE, Lindahl T, Barnes DE. Uracil-DNA glycosylase (UNG)-deficient mice reveal a primary role of the enzyme during DNA replication. Mol Cell 2000;5:1059-65.

85.   Kavli B, Sundheim O, Akbari M, Otterlei M, Nilsen H, Skorpen F, Aas PA, Hagen L, Krokan HE, Slupphaug G. 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 2002;277:39926-36.

86.   Ross H, Noakes J, Spaulding J. Liquid scintillation counting and organic scintillators. Lewis; 1991.

87.   Chomczynski P, Sacchi N. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal Biochem 1987;162:156-9.

88.   Robyt J, White BJ. Biochemical techniques: theory and practice. Waveland Press; 1990.

89.   Green E. Genome analysis: a laboratory manual. Cold Spring Harbor Laboratory Press; 1997.

90.   Boom R, Sol CJ, Salimans MM, Jansen CL, Wertheim-van Dillen PM, van der Noordaa J. Rapid and simple method for purification of nucleic acids. J Clin Microbiol 1990;28:495-503.

91.   Melzak KA, Sherwood CS, Turner RFB, Haynes CA. Driving Forces for DNA Adsorption to Silica in Perchlorate Solutions. Journal of Colloid and Interface Science 1996;181:635-644.

92.   Zheng Y, Cloutier P, Hunting DJ, Wagner JR, Sanche L. Phosphodiester and N-glycosidic bond cleavage in DNA induced by 4-15 eV electrons. J Chem Phys 2006;124:64710.

93.   Chargaff E. Chemical specificity of nucleic acids and mechanism of their enzymatic degradation. Experientia 1950;6:201-9.

94.   Mitchell D. GC content and genome length in Chargaff compliant genomes. Biochem Biophys Res Commun 2007;353:207-10.

95.   Zharkov DO, Mechetin GV, Nevinsky GA. Uracil-DNA glycosylase: Structural, thermodynamic and kinetic aspects of lesion search and recognition. Mutat Res 2010;685:11-20.

96.   Zharkov DO, Grollman AP. The DNA trackwalkers: principles of lesion search and recognition by DNA glycosylases. Mutat Res 2005;577:24-54.