University of Stavanger

Faculty of Science and Technology

MASTER'S THESIS

Study program/Specialization: Biological Chemistry	Spring semester, 2017 Open / Restricted access
Writer: Aysha Arshad	…… (Writer's signature)
Faculty supervisor: Svein Bjelland External supervisor(s):	
Thesis title: Base excision repair of *N*4,5-dimethylcytosine in DNA	
Credits (ECTS): 60	
Key words: Base excision repair N^4,5-dimethylcytosine DNA glycosylases	Pages: 37 + enclosure: 5 Stavanger 13/07/2017 Date/year

ACKNOWLEDGEMENTS

All praises to **"ALLAH"**, the Almighty, Most Gracious, the Most Merciful and the Sustainer of the worlds**,** who gave me the health and opportunity to complete this work. I bow my compassionate endowments to Holy Prophet Muhammad (PBUH) who is ever an ember of guidance and knowledge for humanity.

I have great pleasure to express my gratitude to my distinguished supervisor Prof. Svein Bjelland for his consistent support and stimulating guidance throughout the research work. His guidance and valuable suggestions while writing this thesis are very important lesson for me. I appreciate all his contributions of time and ideas to make my Masters experience productive and stimulating.

I cannot forget to extend my gratefulness to my Co-Supervisors PhD students Marina Alexeeva and Almaz Tesfahun, for their valuable piece of advices and expertly support whenever I needed it. This thesis was not possible to complete without their help and guidance in day to day laboratory work. I would also like to thanks to my fellow master student Prashanna Guragain and our technician Xiang Ming Xu and all other members at Centre for Organelle Research (CORE), Stavanger for their practical helps during work in the lab.

Last but not least, I would like to thanks my family and parents: Muhammad Arshad and Rukhsana and my in-laws. It may not be enough to certain the words of thanks giving as you are a true symbol of love, generosity and benevolence for me. Special thanks to my beloved husband for his patience and concern. Thanks to my friends who have always been there for my moral support and with whom I share my tears and laughs.

"It always seems impossible until it is done"

-Nelson Mandela

ABSTRACT

5-Methylcytosine (m^5C) replaces cytosine (C) in DNA in prokaryotic and eukaryotic cells to execute a number of important cellular functions, but damage to m^5C have received little attention. For instance, almost no studies exist on erroneous methylation of m^5C by alkylating agents to double and triple methylated bases. Due to chemical evidence, and because many prokaryotes express methylases able to convert m^5C into the double methylated N^4,5-dimethylcytosine ($m^{N4,5}C$) in DNA, this base lesion is likely to be present in cellular DNA.

In this study, we used DNA with one $m^{N4,5}C$ residue incorporated at a specific site where $m^{N4,5}C$ was placed opposite G, A, C or T (called $m^{N4,5}C$:G-DNA, $m^{N4,5}C$:A-DNA, $m^{N4,5}C$:C-DNA and $m^{N4,5}C$:T-DNA, respectively. We set up several experiments to find putative repair activity for $m^{N4,5}C$ using different *E. coli* and human glycosylases, major initiators of the base excision repair (BER) pathway. We found the highest repair activity for $m^{N4,5}C$:C-DNA and $m^{N4,5}C$:G-DNA by the two *Escherichia coli* DNA glycosylases Fpg and Nei, respectively. To confirm this repair activity, endonuclease IV and T4-polynucleotide kinase were employed to define and process the 3´-end products following Fpg and Nei-mediated incision of $m^{N4,5}C$:C-DNA and $m^{N4,5}C$:G-DNA, respectively. In contrast, several other *E. coli* DNA glycosylases like Ung and Mug and the human enzyme hUNG did not show repair activity at employed conditions. To our knowledge, this is the first report describing a repair activity for a further methylated m^5C in DNA as well as the first alkylated base allocated to Fpg and Nei as substrate.

CONTENTS

LIST OF PAPERS

Paper 1

Bjelland, S., Alexeeva, M., Tesfahun, A., Tomkuviene, M., **Arshad, A**., Guragain, P., Klimasauskas, S., Jørgensen, K.B., Klungland, A. and Robertson, A.B. Spontaneous and enzymatic modifications of the epigenetic DNA base 5-methylcytosine as targets for repair and mutagenesis. *Manuscript*.

Paper 2

Tesfahun, A., Guragain, P., Alexeeva, M., **Arshad, A**., Tomkuviene, M., Laerdahl, J.K., Klungland, A., Klimasauskas, S. and Bjelland, S. Excision of the double methylated base N^4,5-dimethylcytosine from DNA by *Escherichia coli* Fpg protein. *Manuscript*.

Paper 3

Alexeeva, M., Guragain, P., Tesfahun, A., **Arshad, A**., Tomkuviene, M., Laerdahl, J.K., Klungland, A., Klimasauskas, S. and Bjelland, S. *Escherichia coli* Nei protein initiate repair of N^4,5-dimethylcytosine in DNA. *Manuscript*.

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

SAM	*S*-adenosylmethionine	oxo^{8}G	8-oxo-7,8-dihydroguanine
BER	base excision repair	DNMT	DNA methyltransferases
m^{N4}C	***N***^{4}-methylcytosine	AdoMet	*S*-adenosylmethionine
m^{N4,5}C	*N*^{4},5-dimethylcytosine	Tet	ten eleven translocation
m^{5}C	5-methylcytosine	METTL4	methyltransferase-like 4
hm^{5}C	5-hydroxymethylcytosine	R-M	restriction modification
f^{5}C	5-formylcytosine	SSB	single strand break
m^{N6}A	*N*^{6}-methyladenine	dRP	deoxyribose phosphate
f^{5}U	5-formyluracil	dRPase	deoxyribosephosphodiesterase
h^{5}C	5-hydroxycytocine	Pol β	DNA polymerase β
h^{5}U	5-hydroxyuracil	3´-OH	3´-hydroxyl
hm^{5}U	5-hydroxymethyluracil	3´-PO_4	3´-phosphate
hmh	5-hydroxy-5-methylhydantoin	NEIL	endonuclease VIII-like protein
m^{3}A	3-methyladenine	SP-BER	short-patch BER
m^{3}G	3-methylguanine	LP-BER	long-patch BER
m^{7}A	7-methyladenine	Pol δ	DNA polymerase δ
m^{7}G	7-methylguanine	Pol ε	DNA polymerase ε
Tg	thymine glycol	FEN1	flap endonuclease 1
Ug	uracil glycol	LIGIIIα	DNA ligase IIIα
εA	ethenoadenine	UDG	uracil-DNA glycosylase
HX	hypoxanthine	RPA	replication protein A
X	xanthine	Nei	endonuclease VIII
oxo^{2}A	1,2-dihydro-2-oxoadenine	Nfo	endonuclease IV
oxo^{8}A	7,8-dihydro-8-oxoadenine	METTL4	methyltransferase-like 4

MMS	methyl methanesulfonate
Xth	exonuclease III
APTX	aprataxin
S_N1	nucleophilic substitution 1
S_N2	nucleophilic substitution 2
Pol λ	DNA polymerase λ
fapyG	2,6-diamino-4-hydroxy-5-formamidopyrimidine
TDP1	tyrosyl-DNA-phosphodiesterase1
HMGB1	high-mobility group box 1 protein
PARP1	poly (ADP ribose) polymerase 1
MNNG	*N*-methyl-*N*´-nitro-*N*-nitrosoguanidine
hUNG	human uracil-DNA glycosylase
Fpg	formamidopyrimidine-DNA glycosylase
APE1	apurinic/apyrimidic endonuclease 1
PCNA	proliferating cell nuclear antigen
XRCC1	X-ray cross-complementation protein 1
Th^5	5-hydroxy-5,6-dihydrothymine
ca^5C	5-carboxycytosine
PNKP	polynucleotide kinase phosphatase
3´-PUA	3´-phospho-α,β-unsaturated aldehyde
AP site	abasic/apurinic/apyrimidinic site
DAMT-1	DNA N^6-adenine methyltransferase 1
dHU	5,6-dihydrouracil
dHC	5,6-dihydrocytosine
h^2A	2-hydroxyadenine

AIM OF STUDY

5-Methylcytosine (m^5C) is the major epigenetic mark in eukaryotic including human cells and has been intensively studied in recent decade. However, despite of its importance almost no knowledge exists on the biological consequences if m^5C is further erroneously methylated in DNA, where one of the resulting products is N^4,5-dimethylcytosine ($m^{N4,5}C$). The objective of this work is to contribute to fill the void of this knowledge by investigating the putative initiation of repair of $m^{N4,5}C$ residues in DNA by different *E. coli* and human DNA glycosylases *in vitro*.

1 INTRODUCTION

It is a paradox that DNA is a very stable molecule and is faithfully maintained from generation to generation. Sadly it is not true, as it is so much exposed to endogenous and exogenous physical and chemical agents in both prokaryotes and eukaryotes (Ehrlich, Gama-Sosa et al. 1985, Klose and Bird 2006). Base lesions formed by a diverse array of reactions are obviously of vital concern for all biologists as they hinder both the efficiency (toxic lesions) and precision of the replicative DNA machinery (mutagenic lesions), causing pathological cellular responses as e.g. mutagenesis, carcinogenesis and ageing (Serre, de Jésus et al. 2002). However, most chemically modified bases are damages that all cells need to and indeed remove or repair to maintain their DNA structure and genomic integrity and a sufficiently low mutation rate. For the removal of spontaneous base damages, organisms use several different kinds of DNA repair mechanisms. The most important is the base excision repair (BER) pathway initiated by a battery of DNA glycosylases with varying specificity for different modified bases. BER has evolved from bacteria to humans so mammalian cells have a similar process as microorganisms (Lindahl, Karran et al. 1997, M. TAYLOR AR LEHMANN 1998, Lindahl and Wood 1999, Krokan, Nilsen et al. 2000, Sedgwick 2004).

1.1 DNA base methylation

DNA methylation is a vital epigenetic modification and is found in both prokaryotes and eukaryotes to regulate many cellular processes including transcription, X chromosome inactivation, embryonic development, chromatin structure, genomic imprinting, chromosome stability and protection against viral genomes. The importance of DNA methylation has been

emphasized by the increasing number of human diseases that are known to be associated with aberrant DNA methylation (Robertson 2005, Klose and Bird 2006).

1.1.1 DNA base methylation in eukaryotes

I 5-Methylcytosine is the most important epigenetic regulator

5-Methylcytosine (m^5C) constitutes 4–5% of the total cytosine content of eukaryotic DNA being an important epigenetic mark in mammalian cells (Madugundu, Cadet et al. 2014). It is formed after replication by enzymatic addition of a methyl group to the C-5 position of cytosine, predominantly in a 5´-CpG-3´/3´-GpC-5′ sequence context (Weisenberger, Campan et al. 2005), accomplished by one or more DNA methyltransferases (DNMT1 and DNMT3) using *S*-adenosylmethionine (SAM) as a cosubstrate. DNMT1 preferentially catalyzes the transfer of a methyl group to hemi-methylated CpG dinucleotides (5´-m^5CpG-3´/3´-GpC-5´) and is thus a maintenance methyltransferase. DNMT3 has a greater capacity to methylate CpG sequences that lack methylated cytosine and is consequently a de-novo methyltransferase. These two enzymes work together to maintain the existing methylation pattern and to create new methylation patterns (Klungland and Robertson 2016). Cytosine methylation is directly involved mostly in the regulation of and mostly represses transcriptional activity. DNA demethylation therefore plays important roles in transcriptional activation of silenced genes (He, Li et al. 2011).

The processes of demethylation remained enigmatic until relatively recently, when it was discovered that m^5C is removed from DNA by DNA glycosylases after being modified by oxidation. In mammals, the ten eleven translocation (Tet) enzymes are responsible for this (sequential) oxidation. There are three known Tet enzymes, Tet1, Tet2, and Tet3, which catalyze the transfer of an oxygen molecule to the *N*5 methyl group of m^5C resulting in the formation of 5-hydroxymethylcytosine (hm^5C). This enzymatic reaction is dependent upon

the presence of iron and α-ketoglutaric acid. While an N–C bond within an oxidized methylated base is quite labile spontaneously liberating formaldehyde, the C–C bond between the hydroxymethyl group and the cytosine base is stable in a biological environment. Because of this and other observations as well as the fact that hm^5C is not mutagenic have given rise to proposals of hm^5C as another epigenetic mark (Klungland and Robertson 2016). However, in addition to hm^5C the Tet proteins generate 5-formylcytosine (f^5C) and 5-carboxycytosine (ca^5C) from m^5C (Figure 1), which are removed from DNA by thymine-DNA glycosylase (TDG) followed by completion of BER by downstream enzymes (Ito, Shen et al. 2011). The result is m^5C replaced by C.

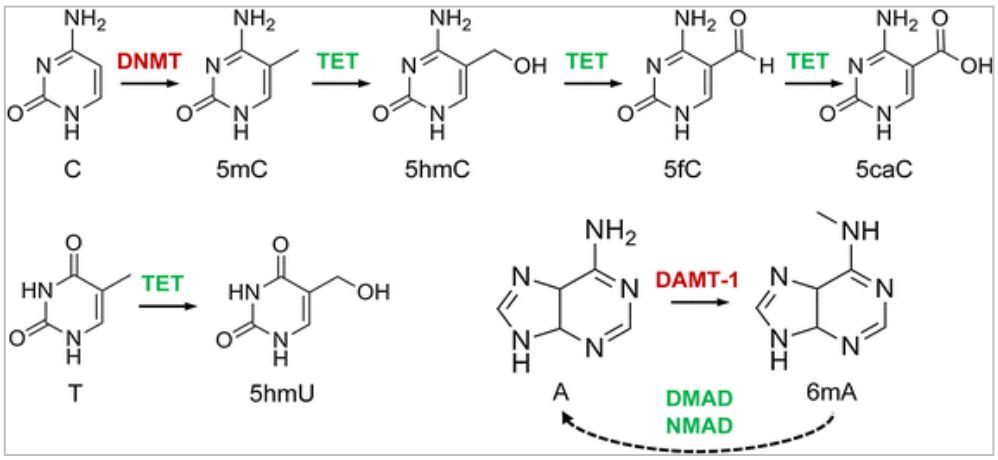


Figure 1. DNA modifications with epigenetic regulatory functions and their interdependencies. Cytosine is methylated to m^5C (5mC) by DNA methyltransferases 1 or 3 (DNMT1/3) and then further oxidized to hm^5C (5hmC), f^5C (5fC) and ca^5C (5caC) by Tet dioxygenases. 5-Hydroxymethyluracil (5hmU) is produced by Tet-catalyzed oxidation of thymine. N^6-methyladenine ($m^{N6}A$) is likely catalyzed by DNA N^6-adenine methyltransferases (DAMT-1 in *C. elegans*), even though the biochemical activity of these enzymes remains to be characterized. The Tet-like AlkB enzymes NMAD ($m^{N6}A$ demethylase 1) and DMAD (DNA $m^{N6}A$ demethylase) have been shown to be involved in $m^{N6}A$ demethylation in *C. elegans* and in *Drosophila melanogaster*, respectively, possibly by using a conserved dioxygenase mechanism (Breiling and Lyko 2015).

In spite of its established effects on gene silencing in higher eukaryotes, CpG islands are major mutational hotspots. The classical explanation is the formation of thymine by spontaneous deamination of m^5C leading to C → T transition mutations resulting in the progressive elimination of methylated CpG sites during evolution (Turek-Plewa and Jagodzinski 2005, Weisenberger, Campan et al. 2005). Many cancer-causing somatic mutations and germ line point mutations associated with human genetic disease involve loss

of CpG. It has been reported that in humans $m^5C \rightarrow TpG$ transition mutations are 10–50 fold higher than all other known transition mutations (Hendrich, Hardeland et al. 1999). The common explanation is that the mismatched T opposite G derived from deamination of m^5C is repaired less efficiently opposite G than uracil causing G:C → A:T transitions (Friedberg, Walker et al. 2006), where it is the TDG glycosylase that initiate repair of the former. Possibly the human BER protein MED1(also known as MBD4; methyl-CpG-binding domain 4), also take part in that repair (Hendrich, Hardeland et al. 1999, Kow 2002). Like TDG, MED1 acts at CpG sites as a mismatch specific T/U glycosylase for G → T and G → U mismatches raised by deamination of m^5C and C, respectively. Targeted inactivation of *Med1(Mbd4*) gene in mice resulted in enhanced mutability at CpG sequences (Cortellino, Turner et al. 2003) so, MED1 appears to act as a caretaker of genomic fidelity at CpG site. Besides that, MED1 also plays role in maintaining genomic integrity by participating in DNA mismatch repair (MMR), and the cell cycle response to DNA damage (Parsons 2003).

II N^6-methyladenine is also an epigenetic regulator

In recent years some studies hypothesized the presence of N^6-methyladenine ($m^{N6}A$) as a post-replicative DNA modification in eukaryotic genomes (Li, Delaney et al. 2012). However, compared to m^5C the levels of $m^{N6}A$ were minimal since it was only detectable by highly sensitive technologies. It has been reported $m^{N6}A$ in unicellular genomes particularly in ciliates, protozoan, chlorophyte algae and dinoflagellates (Heyn and Esteller 2015), and evidence exists for its presence in plants, insects, and mammals (Fu, Luo et al. 2015). In human, to form $m^{N6}A$ in DNA, adenine is modified by the methyltransferase-like 4 (METTL4) enzyme which is similar to DNA N^6-adenine methyltransferase 1 (DAMT-1) in *Caenorhabditis elegans* and $m^{N6}A$ is demethylated by active demethylases of the TET family (TET1–3) proteins similar to DMAD (DNA $m^{N6}A$ demethylase) in *Drosophila melanogaster*

(Heyn and Esteller 2015). Contrary to m^5C, $m^{N6}A$ correspond to an active epigenetic mark in *Chlamydomonas* reinhardtii as it is enriched at the promoter region of highly active genes, while m^5C is implicated in downstream processes of transcriptional initiation (Heyn and Esteller 2015). The research on m^6A in eukaryotic DNA is still in its infancy and detailed biological function is unclear.

1.1.2 DNA base methylation in prokaryotes

I N^6-methyladenine protects against restriction

The $m^{N6}A$ in DNA is functionally similar to N^4-methylcytosine ($m^{N4}C$) and m^5C. It is an important modification in bacteria involved in DNA replication and repair, protein−DNA interactions, host-pathogen interactions and other cellular processes. For every 200 bases, there is approximately one $m^{N6}A$ in the *Escherichia col* genome. Exemplary and solitary adenine methylases, such as Dam in *E. coli*, are involved in DNA replication as it regulates replication initiation factors. AlkB, a DNA repair enzyme in *E. coli*, most probably can demethylate $m^{N6}A$ indicating another function than in repair (Li, Delaney et al. 2012, Heyn and Esteller 2015).

II 5-methylcytosine connected with restriction-modification system

In prokaryotes, specifically in mesophiles, m^5C shares importance with $m^{N4}C$ and $m^{N6}A$ as normal DNA base modifications formed enzymatically after replication (Ehrlich, Gama-Sosa et al. 1985, Ehrlich, Wilson et al. 1987), participating in biological functions like DNA repair, defense against DNA hydrolysis and DNA restriction enzymes (Tesfahun, Alexeeva et al.). As is the case in eukaryotes, m^5C seems to be susceptible to loss by spontaneous deamination leading to C $\rightarrow$ T transition mutations in *E. coli* (Ehrlich, Norris et al. 1986).

III N^4-methyladenine replaces 5-methylcytosine in some thermophiles

Some prokaryotes especially hyperthermophiles have $m^{N4}C$ instead of m^5C in DNA due to its higher resistance to hydrolytic deamination. It is formed by the cytosine-N^4 methyltransferase; however, the evolutionary significance of this enzyme is still illusive (Trivedi, Rao et al. 2005). The $m^{N4}C$ residue is not restricted to thermophiles being a common base in many bacterial species, and has a specific role in prokaryotic restriction-modification systems. The recent discovery of an undefined cytosine derivative with similar properties to $m^{N4}C$ in trypanosome chromosomes suggests that $m^{N4}C$ might not be limited to bacterial genomes, although no evidence exists for the presence of $m^{N4}C$ in vertebrate DNA (Ehrlich, Gama-Sosa et al. 1985).

1.1.3 Enzymatically methylated DNA bases damaged by additional methylation

The possible further methylation of m^5C in cellular DNA to double and maybe triple methylated bases, either by methylating agents or by methylases, has achieved little attention. For example, some prokaryotes enzymatically modify cytosine in either the N^4- or 5-position and then a dimethylated cytosine might be feasible if the specific methyltransferases targeted the same sequence (Ehrlich, Wilson et al. 1987). Indeed, such an artificial *in vivo* condition has been made by transforming *E. coli* with a plasmid coding for a cytosine-N^4 methyltransferase, where the latter methylates m^5C formed by the *E. coli* Dcm methylase. Although N^4,5-dimethylcytosine ($m^{N4,5}C$) was hard to detect in cellular DNA, probably due to efficient repair, $m^{N4,5}C$ has been stably introduced into the same sequence in DNA *in vitro* (Figure 2) (Klimasauskas, Gerasimaite et al. 2002).

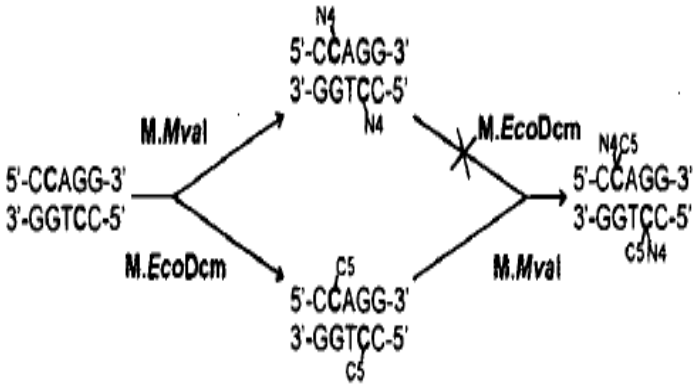


Figure 2. Enzymatic formation of N^4,5-dimethylcytosine in DNA. Only one (lower) of the two theoretical pathways leads to dimethylation of cytosine (Klimasauskas, Gerasimaite et al. 2002).

Since structural considerations on the presence of $m^{N4,5}C$ in DNA suggest possible impact on normal Watson-Crick base-pairing (Figure 3) and thus the need for repair, the above mentioned difficulty in detecting the lesion in cellular DNA appears sensible. However, no research has been undertaken to further investigate the biology of $m^{N4,5}C$ since its discovery in DNA.

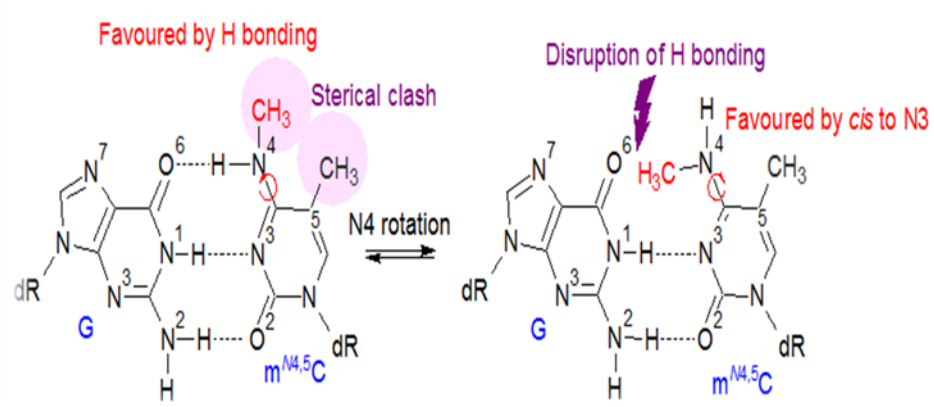


Figure 3. Destabilizing features of $m^{N4,5}C$ in DNA. Structural considerations suggest that the $m^{N4,5}C$ residue poses certain destabilizing features to DNA. The figure shows how rotation by the N4–C4 bond of $m^{N4,5}C$ relieves the putative sterical clash between opposing methyl groups (left) but at the same time disrupts the normal G:C hydrogen bonding pattern (right), to exclude $m^{N4,5}C$ from adjusting into normal Watson-Crick DNA structure. Abbreviations: dR, deoxyribose (Tesfahun, Alexeeva et al.)

1.2 DNA repair mechanisms for bases damaged by methylation

In addition to its use by methyltransferases to methylate cytosine and adenine at certain positions, SAM may also like other methylating agents, chemically methylate all DNA bases erroneously at many positions as can be seen in Figure 4.

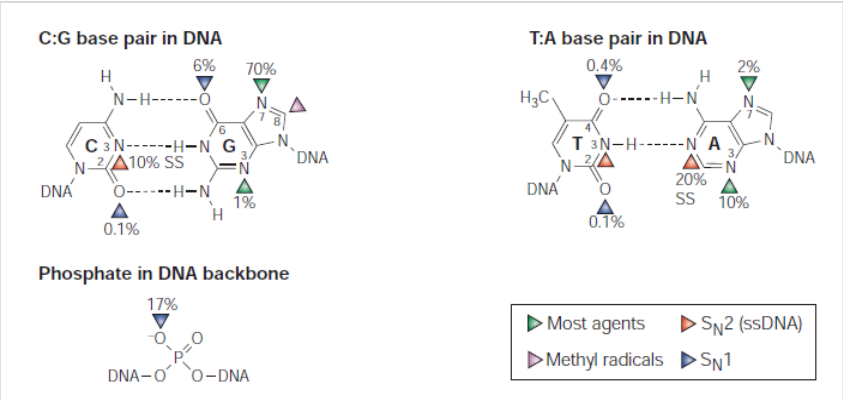


Figure 4. Sites of methylation on the bases and sugar–phosphate backbone of DNA. Blue arrows indicate oxygen atoms in DNA that are most frequently methylated by S_N1 agents, such as *N*-methyl-*N*´-nitro-*N*-nitrosoguanidine (MNNG). Red arrows indicate sites in single-stranded (ss)DNA that are methylated by S_N2 agents, such as methylmethane sulphonate (MMS). The pink arrow is a site that is methylated by methyl radicals. Green arrows indicate sites that are methylated by most agents. The percentages indicate the relative abundance of each modification (Sedgwick 2004).

To cope with such damaged bases, eukaryotes and prokaryotes have developed several repair mechanisms depending on the nature of the base lesion. The most important ones are:

i. Direct reversal of damaged base (methyltransferase and oxidative demethylation) (shown in Figure 5).

ii. Base excision repair (BER)

1.2.1 Direct reversal of damaged base (methyltransferase and oxidative demethylation)

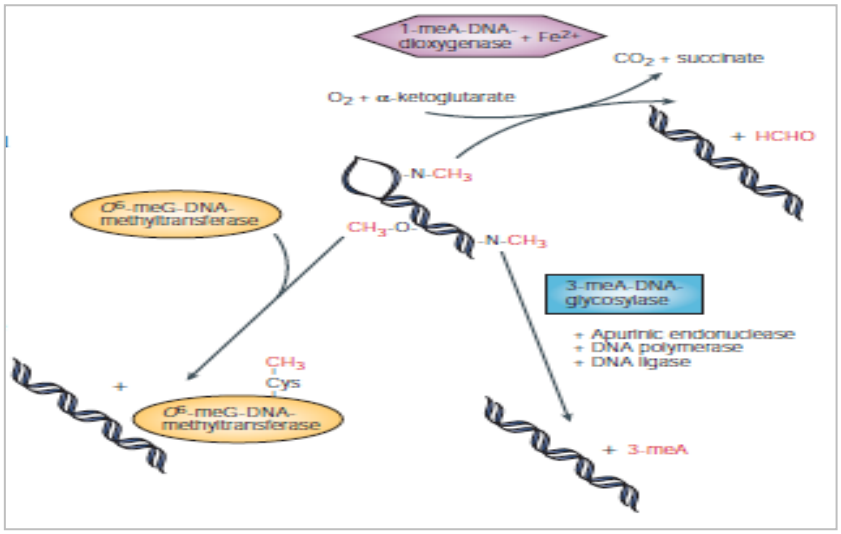


***Figure 5. Mechanisms of DNA repair of N- and O-methylated bases in DNA*: 3-Methyladenine-DNA-glycosylase:** Methylated bases with destabilized glycosyl bonds, such as 3-methyladenine (shown as –N–CH_3) are excised by 3-methyladenine (m^3A)-DNA glycosylase (blue). Several additional activities are required to complete repair of the resulting abasic sites. An apurinic endonuclease incises the baseless site, a DNA polymerase inserts a single nucleotide and a DNA ligase seals the single-stranded gap. This pathway, which is initiated by a DNA glycosylase, is known as base excision repair. **1-Methyladenine-DNA dioxygenase:** 1-Methyladenine (m^1A) and 3-methylcytosine (m^3C) in DNA are stable, modified bases that are mostly generated in single-stranded (ss) DNA (shown as –N–CH3 in a region of ssDNA). These lesions are repaired using oxidative demethylation by 1-methyladenine (m^1A)-DNA dioxygenases (purple). They require dioxygen and α-ketoglutarate as co-substrates and Fe^{2+} as a cofactor. They oxidize the methyl adduct, which results in its destabilization and release as formaldehyde (HCHO). Other products of the reaction are succinate and CO_2. The damaged base directly reverts to the unmodified form. ***O*^{6}-methylguanine-DNA methyltransferase:** The significant lesion *O*^{6}-methylguanine (*O*^{6}-meG) and the minor lesion *O*^{4}-methylthymine are methylated on their exocyclic oxygens (shown as CH3–O–), and are directly demethylated by *O*^{6}-meG-DNA methyltransferases. These proteins transfer the methyl group to an active-site Cys residue. Repair by the methyltransferases requires no cofactors or additional activities. (The methyl groups and their different fates in the different reactions as excised methylated base (for example, m^3A), released HCHO or self-methylated methyltransferase are shown in red) (Sedgwick 2004).

1.2.2 Base excision repair (BER) pathway

The BER pathway likely evolved to cope with the high level of hydrolytic decay products formed in DNA, as well as those damages created upon reactions with natural endogenous chemicals like ROS and SAM. Thus BER predominantly deals with non-bulky small nucleobase lesions, excising and replacing incorrect (*e.g.* uracil) or damaged (*e.g.* 3-methyladenine, 8-oxoG) bases derived from, as indicated above, deamination, alkylation or

oxidation. Both prokaryotes and eukaryotes have a similar BER pathway consisting of the following steps:

i. A DNA glycosylase recognizes and removes an incorrect or damaged base creating an apurinic/apyrimidinic (AP) or abasic site intermediate
ii. An AP endonuclease or AP lyase incises the AP site
iii. A lyase or phosphodiesterase removes the remaining sugar fragment
iv. Gap filling by a DNA polymerase
v. A DNA ligase seals the nick in DNA (Kim and M Wilson III 2012).

I Base recognition and removal by DNA glycosylases

DNA glycosylases are the key initiators of the BER pathway since they are responsible for detecting and specifically recognizing distinct forms of DNA damage. At least 12 genes along their splicing variants and eight ones encoding various glycosylases have been found in mammalian and *E. coli* cells, respectively, with different substrate specificities and modes of action (Krwawicz, Arczewska et al. 2007). They are generally small, monomeric proteins, which efficiently remove a broad range of damaged and inappropriately inserted bases by hydrolysis of the N–C1´ glycosylic bond between the target base and deoxyribose, releasing a free damaged base and leaving behind an AP site in DNA. DNA glycosylases fall into two categories: pure/monofunctional DNA glycosylases and bifunctional DNA glycosylases/AP lyases (Fromme and Verdine 2004). Pure glycosylases uses a water molecule as nucleophile to attack the deoxyribose C1´ to excise the damaged base, in turn activated by carboxyl side chain of an Asp (Aspartic acid) residue. Then the AP site is cleaved by an AP endonuclease leaving behind a 3´-OH and a 5´-deoxyribose phosphate (dRP) residue (see below). Bifunctional glycosylase/AP lyase enzymes perform both the first two functional steps by using an active site amine moiety of Lys (lysine) residue activated by an Asp residue

as a nucleophile to excise the damaged base and generate a covalent Schiff base protein–DNA intermediate during the catalytic process, eventually leaving behind a nicked DNA with a 3´-unsaturated or saturated aldehyde and a 5´-phosphate. (KROKAN, STANDAL et al. 1997, Parikh, Mol et al. 1997, Nilsen and Krokan 2001, Sancar, Lindsey-Boltz et al. 2004, Hitomi, Iwai et al. 2007, Krwawicz, Arczewska et al. 2007, Kim and M Wilson III 2012).

All human and *E. coli* DNA glycosylases are presented in Table 1 and 2.

II Abasic site (AP site) incision and processing

The AP sites or DNA ends produced after lesion excision or excision and incision by mono- or bifunctional glycosylases, respectively, are very lethal and mutagenic if left unrepaired (Hitomi, Iwai et al. 2007, Krwawicz, Arczewska et al. 2007). The AP site is commonly incised by an AP endonuclease, which in *E. coli* is functions of the Xth (exonuclease III) and Nfo (endonuclease IV) proteins and in human cells the APE1 (apurinic/apyrimidic endonuclease 1) protein (also called HAP1, APEX and REF1) which cleaves the phosphodiester bond 5´ to the AP site by a hydrolytic reaction. The resultant single strand break (SSB) in DNA contains a 3´-OH and a 5´-dRP (Erzberger, Barsky et al. 1998, Dianov and Hübscher 2013). In human cells, almost 90% of the AP sites are processed by hAPE1 protein (Campalans, Marsin et al. 2005). Cleavage of the phosphodiester bond by a bifunctional glycosylase occurs 3´ to the AP site resulting in a 3´-α,β-unsaturated or saturated aldehyde and a 5´-phosphate (*e.g.*, by Nth in *E. coli*) or alternatively by β/δ-elimination resulting in a both 3´- and 5´-phosphate (*e.g.*, by Fpg and Nei in *E. coli*), dependent on the type of glycosylase/lyase. While AP endonucleases create a 3´-OH ready for replication, the 5´-dRP residue is blocking to ligation and needs to be removed/converted to a 5´-phosphate. In *E. coli*, that can be performed by the Fpg and Nei proteins and DNA polymerase I (Pol I), and in mammalian cells by the lyase function of DNA polymerase β (Pol β). Bifunctional

glycosylases, in contrast, creates a 5´-phosphate ready for ligation, while the 3´-residue is blocking and needs to be removed/converted to a 3´-OH to make replication possible. Regarding 3´-aldehydes, these are removed by the 3´-phosphodiesterase function of an AP endonuclease, which in *E. coli* is the Xth and Nfo proteins and in mammalian cells APE1. APE1 can also remove 3´-phosphate groups, yet *via* a relatively weak 3´-phosphatase activity. However, polynucleotide kinase phosphatase (PNKP) is the primary enzyme for removing 3´-phosphates *i.e.* 3´-PO_4 or 3´-PUA, produced by endonuclease VIII-like protein (NEIL) during repair of oxidized bases. PNKP is a bifunctional DNA repair enzyme, fusing both a DNA kinase and phosphatase domain, and is responsible for preparing nicked DNA for ligation. In particular, the 3´-phosphatase activity hydrolyzes a 3´-PO_4 to generate a 3´-OH end, and the 5´-kinase activity phosphorylates a 5´-OH end to generate a 5´-PO_4 (Chen, Herman et al. 1991, Suh, Wilson et al. 1997, Bernstein, Williams et al. 2005, Alonso, Terrados et al. 2006, Almeida and Sobol 2007, Kim and M Wilson III 2012). Likewise in E.coli, XthA and Nfo serves as major 5´-AP endonucleases as they both have the ability to process the blocking termini by generating free 3´-OH and 5´-dRP ends. They also possess 3´-phosphodiesterase activity to remove 3´-PO_4, 3´-PUA and 3´-phosphoglycolates (Rogers and Weiss 1980).

III DNA repair synthesis and ligation of the nick by DNA ligases

After damaged base removal and AP site incision and processing, the next steps in the BER pathway is to insert the proper base by a (repair) DNA polymerase, which in *E. coli* is DNA polymerase I (Pol I) and in mammalian cells DNA polymerase β (Pol β), followed by final strand sealing by DNA ligase. In mammalian cells, other proteins assist in these processes. For example, Pol β interacts with non-catalytic X-ray cross-complementation protein 1 (XRCC1) subunit of the XRCC1–DNA ligase III heterodimer along with the nick sensor poly (ADP ribose) polymerase 1 (PARP1). Consequently, XRCC1 acts as a scaffold

protein by bringing the polymerase and ligase together at the site of repair. For the final catalysis of phosphodiester bond formation, DNA ligases utilize either NAD^+ as in *E. coli* or ATP as in mammalian cells (Lindahl and Wood 1999, Kim and M Wilson III 2012). The pathway outlined above is the major mode of BER and is called short-patch repair (SP-BER) (Hitomi, Iwai et al. 2007). In *E. coli*, Fpg lyase activity removes the 5´-dRP residue and hence favors SP-BER. If 5´-dRP is not processed before DNA synthesis then Pol I will continue replication beyond one nucleotide resulting in long-patch repair (LP-BER). In mammalian cells glycosylase products are also alternatively processed by LP-BER involving DNA polymerase δ (Pol δ) and ε (Pol ε) (Hitomi, Iwai et al. 2007), which is recruited to the BER site by the accessory 'replication clamp' protein called proliferating cell nuclear antigen (PCNA). These polymerases then perform 'strand displacement synthesis', where the downstream 5´-DNA end is 'displaced' to form a flap intermediate. The displaced strand is then removed by a structure-specific nuclease, primarily flap endonuclease 1 (FEN1), to create a ligatable substrate. Pol β can also carry out strand displacement synthesis, and therefore can potentially participate in either SP-BER or LP-BER (Mol, Hosfield et al. 2000) (Figure 6).

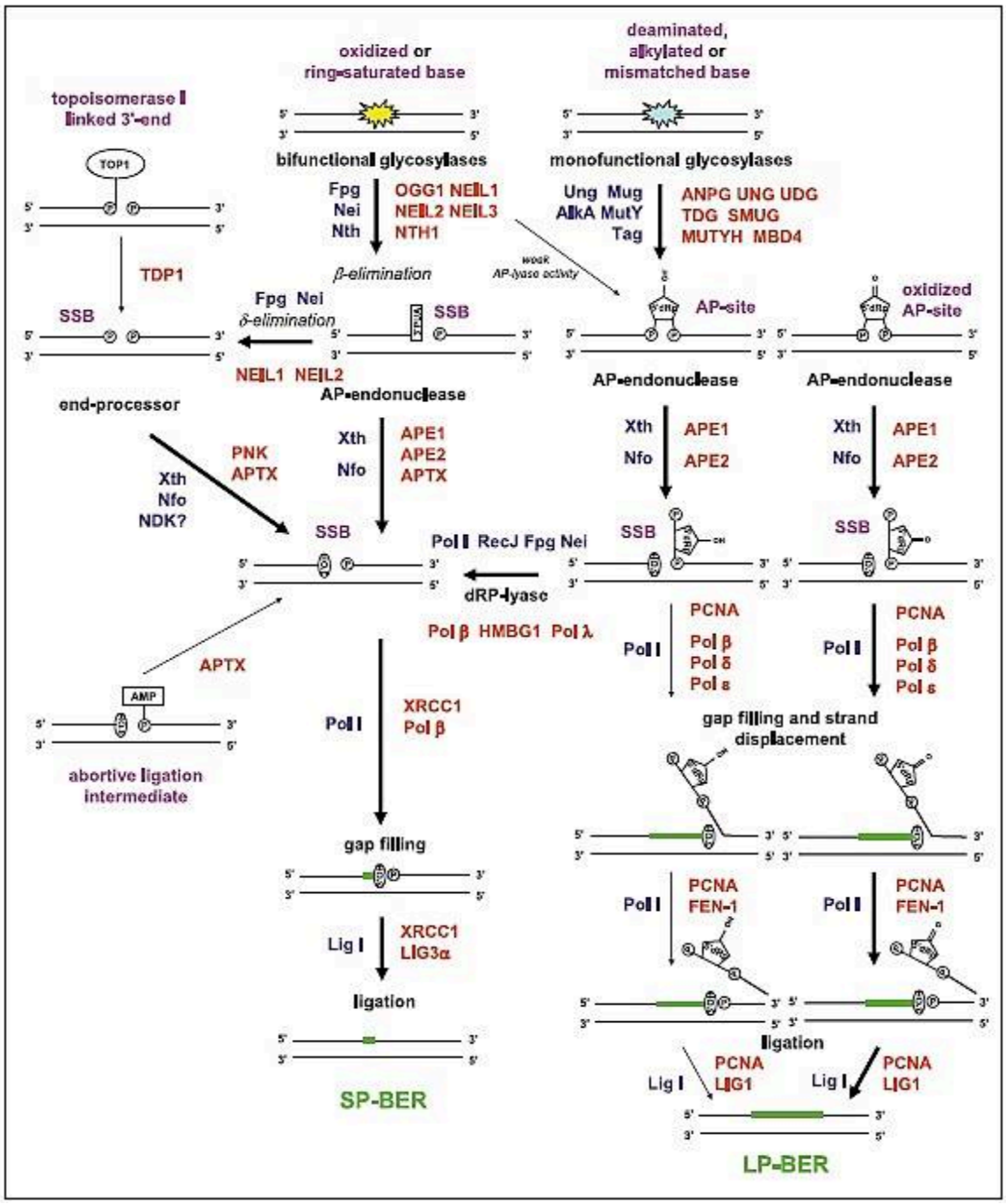


Figure 6. Model for the BER and single strand breaks repair sub-pathways. The types of DNA lesions repaired by common sub-pathways of single strand breaks repair and base excision repair are marked in purpure. *E. coli* enzymes are on left, and are in blue, human enzymes shown on right, are in red (Krwawicz, Arczewska et al. 2007).

Abbreviations: P, phosphate; OH, hydroxyl group; 3´-PUA, 3´-phospho-α,β-unsaturated aldehyde; 5´-dRP, 5´-deoxyribose phosphate; AMP, adenylate group; TOP1, topoisomerase I-linked 3´-end, SSB, single strand breaks; APTX, aprataxin; XRCC1, X-ray cross-complementation protein 1; Pol I, DNA polymerase I; Pol β, DNA polymerase β; Pol δ, DNA polymerase δ; Pol ε, DNA polymerase ε; Pol λ, DNA polymerase λ; PCNA, proliferating cell nuclear antigen; PNK, polynucleotide kinase; APE1, AP endonuclease 1; APE2, AP endonuclease 2; FEN-1, flap endonuclease 1; LIG 1, DNA ligase 1; TDP1, tyrosyl-DNA phosphodiesterase 1; HMGB1, high-mobility group box 1 protein; LIG3α, DNA ligase 3α.

Table 1. E. coli monofunctional/bifunctional DNA glycosylases involved in the BER pathway (Serre, de Jésus et al. 2002, Takao, Kanno et al. 2002, Bjelland and Seeberg 2003, Krwawicz, Arczewska et al. 2007, Kim and M Wilson III 2012)

Protein	Common Name	Substrates	Size (aa)	Additional AP lyase activity
Ung	Uracil-DNA glycosylase	U; h^5U:G; h^5U:A	229	No
Mug	Mismatch-specific uracil-DNA glycosylase	U; T; hm^5U:A; hm^5U:G; f^5U:A; f^5U:G; h^5C:G; h^5U:G	168	No
AlkA	3-Methyladenine-DNA glycosylase II	ε^3G; εA; hm^5U:A; m^3A; m^3G; f^5U:A; f^5U:G; m^7A; me^7G; m^8G; HX; X	282	No
Fgp/ MutM	Formamidopyrimidine-DNA glycosylase	oxo^8G; oxo^8A:C; oxo^8A:T; fapyG; fapyG:A; fapyG:C; h^5C; Tg; Th^5; hmh; hm^5U; f^5U; Ug; ring opened εA	269	AP lyase via β/δ-elimination, dRP-lyase
Tag	3-Methyladenine glycosylase I	m^3A >> m^3G; m^7G	187	No
MutY	Adenine-DNA glycosylase	A:oxo^8G, A:oxo^8A oxo^2A:G; A:fapyG; h^2A:G	350	No
Nth	Endonuclease III	Ug; h^5C; h^5U; dHU; Tg; Th^5; hmh; Th^5:A; Th^6:A; hm^5C; f^5U; dHC; fapyA; ring opened εA	211	AP lyase via β/δ-elimination, dRP-lyase
Nei	Endonuclease VIII	oxo^8G; oxo^8A; Tg; Th^5:A; hmh:A; hm^5U; f^5U:A; Ug:A; Ug:G; h^5U; h^5C	263	AP lyase via β/δ-elimination, dRP-lyase

Table 1. Human monofunctional/bifunctional DNA glycosylases involved in the BER pathway (Serre, de Jésus et al. 2002, Takao, Kanno et al. 2002, Bjelland and Seeberg 2003, Krwawicz, Arczewska et al. 2007, Kim and M Wilson III 2012).

Protein	Common name	Substrate	Size (aa)	Additional AP lyase activity
hUNG	Uracil-DNA glycosylase UNG1 UNG2	ssU > U:G > U:A; f^5U h^5U:G	 304 313	No
TDG	Thymine-DNA glycosylase	U:G > εC:G > T:G > T:C > T:T; hm^5U:G; f^5U	410	No
hSMUG1	Single strand specific-monofunctional uracil-DNA glycosylase	ssU > U:G > hm^5U > hm^5U:G > U:A > hm^5U:A > εC:G > f^5U; h^5U	270	No
MPG	3-Methyladenine-DNA glycosylase	m^3A; m^7A; m^3G; m^7G; εA; HX; oxo^8G:C	293	No
hOGG1	8-oxoguanine-DNA glycosylase	oxo^8G; oxo^8C:A	424	AP lyase via β-elimination
NEIL1	Endonuclease VIII-like 1	oxo^8G:C > oxo^8G:G > oxo^8G:T; Tg:G; h^5C; h^5U; dHU; fapyA; fapyG	390	AP lyase via β/δ-elimination
NEIL2	Endonuclease VIII-like 2	h^5U; dHU; h^5C; Tg; Oxo^8G:C; oxo^8G:A	332	AP lyase via β/δ-elimination
NEIL3	Endonuclease VIII-like 3	fapyG	605	AP lyase via β-elimination
MBD4/ MED1	Methyl-binding domain glycosylase	hm^5U; hm^5U:G; f^5U:G; U or T in U/TpG:m^5CpG	580	No
hNTH1	Human endonuclease III	Tg:A > Tg:G; diHU; fapyG:A/G/T; f5U; Th^5:A; h^5U; h^5C:G; h^5C:A	312	AP lyase via β-elimination

2 MATERIALS AND METHODS

2.1 Buffers and solutions

The buffers and solutions used in this thesis are enlisted in Table 3.

Table 3. Buffers and solution with ther composition used for BER assays

Buffers/solutions	Composition
HEPES buffer (5×)	225 mM HEPES, pH 7.5, 10% (v/v) glycerol, 2 mM EDTA, deionized H_2O
REC buffer (5×)	5 mM EDTA buffer pH 8, 100 mM Tris pH 7.5, deionized H_2O
Tris buffer (5×)	100 mM Tris pH 7.5, 300 mM NaOH, 5 mM EDTA, deionized H_2O
Buffer (1×) in 50% (v/v) glycerol for enzyme dilution	HEPES buffer (5×), 50% (v/v) glycerol, deionized H_2O
Stop solution	20 mM EDTA, 0.5% (w/v) sodium dodecyl sulphate (SDS), deionized H_2O
Salt-TE (STE) buffer (for DNA oligomer-hybridization)	10 mM Tris pH 8.0, 50 mM NaCl, 1 mM EDTA, deionized H_2O
Denaturing loading buffer (DLB)	80% (v/v) formamide, 1 mM EDTA, 1% (w/v) blue dextran, deionized H_2O
TE buffer (1×)	10 mM Tris pH 7.5, 1 mM EDTA pH 8.0, sterile filtrated, deionized H_2O
Taurine (20×), running buffer	1.78 M Tris base, 0.58 M Taurine, $Na_2EDTA{\times}2H_2O$, deionized H_2O
Tris/borate/EDTA buffer (10× TBE), running buffer	Tris base, boric acid, 0.5 M EDTA, pH 8.0, deionized H_2O

2.2 Preparation of duplex oligonucleotides

2.2.1 Substrate oligonucleotides

The m^{N4,5}C-containing oligonucleotide substrate fluorescently labeled at the 5´-end with Cy3 with complementary strands, where G, A, C, and T have been placed against m^{N4,5}C used in this thesis are listed in Table 4. (For details see Tesfahun et al.)

Table 4. m^{N4,5}C oligonucleotide substrate with complementary strands and their nucleotide sequences

No.	Oligo	Sequence (5´–3´)
1	Fw = **(Cy3)m^{N4,5}C** Rev = **G**	**Fw:** CGGTGAAGTAC**m^{N4,5}C**AGGAAGCGATTTCGACCC **Rev:** GGGTCGAAATCGCTTCCT**G**GTACTTCACCG
2	Fw = **(Cy3)m^{N4,5}C** Rev = **A**	**Fw:** CGGTGAAGTAC**m^{N4,5}C**AGGAAGCGATTTCGACCC **Rev:** GGGTCGAAATCGCTTCCT**A**GTACTTCACCG
3	Fw = **(Cy3)m^{N4,5}C** Rev = **C**	**Fw:** CGGTGAAGTAC**m^{N4,5}C**AGGAAGCGATTTCGACCC **Rev:** GGGTCGAAATCGCTTCCT**C**GTACTTCACCG
4	Fw = **(Cy3)m^{N4,5}C** Rev = **T**	**Fw:** CGGTGAAGTAC**m^{N4,5}C**AGGAAGCGATTTCGACCC **Rev:** GGGTCGAAATCGCTTCCT**T**GTACTTCACCG

2.2.2 Control oligonucleotide substrates

Polydeoxynucleotides containing dUMP at a specific site were supplied fluorescently labeled at the 5´-end with Cy3 by Sigma-Aldrich and reverse complementary strand are provided in Table 5.

Table 5. Control oligonucleotide substrates with complementary strands and their nucleotide sequences

No.	Control substrate	Sequence (5´–3´)
1	Fw = **(Cy3) U-60** Rev = **Gcomp60** (60 nt; 11 nt incision product)	**Fw:** CCCTCGAGGTA**U**CATGGATCCGATCGATCCGATTTCGACCTCAAACCTAGAC GAATTCCG **Rev:** CGGAATTCGTCTAGGTTTGAGGTCGAAATCGGATCGATCGGATCCATG**G**TAC CTCGAGGG
2	Fw = **(Cy3) U-60** Rev = **Gcomp60** (60 nt; 20 nt incision product)	**Fw:** TAGACATTGCCCTCGAGGTA**U**CATGGATCCGATTTCGACCTCAAACCTAGAC GAATTCCG **Rev:** CGGAATTCGTCTAGGTTTGAGGTCGAAATCGGATCCATG**G**TACCTCGAGGGC AATGTCTA

2.3 Hybridization of template strand with complementary strand to form duplexes

Substrate oligonucleotide and control uracil oligonucleotides with their complementary strands (given in Tables 3 and 4) were hybridized to a final concentration of 10 pmol/µl. For hybridization, reaction mixture of total volume 10 µl was prepared by mixing 1 µl of each 100 pmol/µl forward (Fw) strand and 1 µl of 100 pmol/µl of reverse (Rev) complementary strand with 8 µl 1× STE (Sodium chloride-Tris-EDTA) buffer in PCR tubes. Tubes were kept on ice and in darkness during the assay. BIO-RAD thermocycler was used for annealing purpose. Fw and Rev strands were annealed by heating at 95 °C for 3 min and cooling down to 20 °C at a rate of 0.01 °C/s to form duplexes.

After hybridization, duplexes with concentration of 10 pmol/µl were further diluted with 9 µl 1× Tris-EDTA (TE) buffer to make final concentration of 1 pmol/µl and stored at –20 °C in the dark.

2.4 Enzymes

Concentration and dilution buffers of all the glycosylases used in this thesis are described below.

hUNG (hUNGΔ84 with His-tag; 823 pmol/µl; dissolved in 20 mM Tris-HCl, pH 7.5, 60 mM NaCl, 1 mM EDTA, 1 mM DTT) (1,2) was a gift from Bodil Kavli and Geir Slupphaug. Fpg [Cat. No. M0240S; 8000 units/ml (17 pmol/µl); lot No. 0061405; dissolved in 20 mM Tris-HCl, pH 8.0, 50 mM NaCl, 0.5 mM ethylenediaminetetraacetic acid (EDTA), 200 µg/ml bovine serum albumin (BSA), 50% glycerol], Ung [Cat. No. M0280S; 5000 units/ml (1.95 pmol/µl)], endonuclease IV [Nfo; Cat. No. M0304S; 10,000 units/ml (0.083 pmol/µl)] and T4 polynucleotide kinase [PseT; Cat. No. M0201S; 10,000 units/ml (0.289 pmol/µl)] were obtained from New England Biolabs. Mug (Cat. No. 4125-100-EB; 100 units/ml; dissolved in

20 mM Tris-Cl, pH 8.0, 2 mM EDTA, 2.5 mM 2-mercaptoethanol, 1 mM phenylmethylsulfonyl fluoride, 50% (v/v) glycerol) was obtained from Trevigen. Nei (Cat. No. M0299S; 2.08 pmol/µl).

2.5 Assay for excision/incision of $m^{N4,5}C$ from DNA by *E. coli* DNA glycosylases Ung and Mug

To verify the excision activity of glycosylases against $m^{N4,5}C$-dsDNA, different purified *E.coli* DNA glycosylases (Ung and Mug) were incubated with $m^{N4,5}C$-ssDNA, or with $m^{N4,5}C$-dsDNA where $m^{N4,5}C$ was placed opposite G, A, C or T (called $m^{N4,5}C$:G-DNA, $m^{N4,5}C$:A-DNA, $m^{N4,5}C$:C-DNA and $m^{N4,5}C$:T-DNA, respectively. The control U:G-DNA [60 nt; 1 pmol (incision product; 11nt)] was also incubated with and without DNA glycosylases. All the buffers and other reagents with their volumes and concentrations are enlisted in Table 6.

Table 6. Reagents used for Base excision/DNA incision assay with concentration and volume for Ung and Mug

Reagents	**Reaction mixture**		**Final Concentration**
	Mug	**UDG**	
Buffer (5×)	REC buffer + DTT 4 µl	UDG buffer 4 µl	1×
1 M KCl		1.4 µl	70 mM
BSA (10 mg/ml)	1 µl	1 µl	0.1 mg/ml
Labeled DNA (1 pmol/µl)	1 µl	1 µl	0.05 pmol/µl
Enzyme (1 pmol/µl)	1 µl	1 µl	0.05 pmol/µl
Deionized H_2O	12.6 µl	11.6 µl	
Total volume	20 µl	20 µl	

Reaction mixture in a suitable reaction buffer *i.e.* REC buffer for Mug as recommended by TREVIGEN® and UDG buffer provided with enzyme by New England Biolabs in a final volume of 20 µl was prepared separately for purified DNA glycosylases Mug and Ung at varying concentrations and was incubated at 37 °C for 30 min with $m^{N4,5}C$-ssDNA, or with

m^{N4,5}C-dsDNA where m^{N4,5}C was placed opposite G, A, C or T (called m^{N4,5}C:G-DNA, m^{N4,5}C:A-DNA, m^{N4,5}C:C-DNA and m^{N4,5}C:T-DNA, respectively. The same amount of purified Mug and Ung were also incubated with and without U:G-DNA (60 nt; 1 pmol) used as positive and negative control at 37 °C for 30 min. Reactions were terminated by the addition of stop solution (20 mM EDTA, 0.5% (w/v) SDS) and proteinase K (150 µg/ml) followed by incubation at 37 °C for 10 min. DNA was precipitated with ethanol/NaAc (sodium acetate) and the precipitate was solubilized in 10 µl water (for samples without NaOH) or 10 µl 0.1 M NaOH followed by heating for 10 min at 90 °C. Then mixed with 10 µl denaturing loading buffer (DLB) solution containing 80% (v/v) formamide, 1 mM EDTA and 1% (w/v) blue dextran to prepare samples (20 µl) for electrophoresis and were incubated at 95 °C for 5 min to denature DNA and were cooled down on ice.

2.6 Polyacrylamide gel electrophoresis (PAGE)

Polyacrylamide gels 20 % (w/v) with 8 M urea in 10× TBE buffer for Mug and Ung and 20 % (w/v) polyacrylamide gels with 7 M urea in 20× Taurine buffer for hUNG with 15 wells were prepared. (See Appendix for gel composition). The solution was poured into gel cast and allowed to solidify for 30 min. A volume of 5 µl of samples was loaded on the denaturing PAGE. Electrophoresis was run for 2 h 30 min at a constant voltage of 180 V or otherwise stated. Visualization and quantification was performed by using ImageQuant Software (Molecular Dynamics Inc.) for fluorescence imaging analysis.

2.7 Assay for excision/ incision of m^{N4,5}C from DNA by human DNA glycosylase hUNG

To verify the excision/ incision activity against m^{N4,5}C-dsDNA, purified human DNA glycosylase hUNG was incubated with m^{N4,5}C-ssDNA, or with m^{N4,5}C-dsDNA where m^{N4,5}C was placed opposite G, A, C or T (called m^{N4,5}C:G-DNA, m^{N4,5}C:A-DNA, m^{N4,5}C:C-DNA and

m^{N4,5}C:T-DNA, respectively. The control U:G-DNA [60 nt; 1 pmol (incision product; 11 nt)] was also incubated with and without DNA glycosylases. All the buffers and other reagents used for this assay with their volumes and concentrations are enlisted in Table 7.

Table 7. Reagents used for Base excision/DNA incision assay with concentration and volume for hUNG

Reagents	Reaction mixture	Final concentration
HEPES buffer (5×) + DTT	4 µl	1×
1 M KCl	1.4 µl	70 mM
BSA (10 mg/ml)	1 µl	0.1 mg/ml
Labeled DNA (1 pmol/µl)	1 µl	0.05 pmol/µl
Enzyme (hUNG) (1 pmol/µl)	1 µl	0.05 pmol/µl
Deionized H_2O	11.6 µl	
Total volume	20 µl	

Reaction mixture in HEPES buffer (Slupphaug, Mol et al. 1996) in a final volume of 20 µl was prepared for hUNG and was incubated at 37 °C for 30 min (Slupphaug, Mol et al. 1996) with m^{N4,5}C-ssDNA, or with m^{N4,5}C-dsDNA where m^{N4,5}C was placed opposite G, A, C or T (called m^{N4,5}C:G-DNA, m^{N4,5}C:A-DNA, m^{N4,5}C:C-DNA and m^{N4,5}C:T-DNA, respectively. The same amount of purified DNA glycosylase was also incubated with and without U:G-DNA (60 nt; 1 pmol) used as positive and negative control at 37 °C for 30 min. Protocol was further followed by as described in Sections 2.5 and 2.6.

2.8 Fpg/Nei-mediated incision of m^{N4,5}C:C-DNA

To confirm the presence of a 3´-phosphate following incision of m^{N4,5}C:C-DNA or m^{N4,5}C:G-DNA by Fpg and Nei respectively (Bailly, Verly et al. 1989), we treated Fpg or Nei-exposed DNA with T4 polynucleotide kinase (PseT), which specifically removes phosphate from 3´ ends (Cameron and Uhlenbeck 1977, Midgley and Murray 1985) as opposed to other types of residues. The second period of incubation was carried out together with *E. coli*

endonuclease IV (Nfo) (Warner, Demple et al. 1980) instead of PseT. All the buffers and other reagents with their volumes and concentrations are enlisted in Table 8.

Table 8. Reagents with concentration and volume used for DNA incision assay by Fpg and Nei

Reagents	**Fpg reacation mixture**	**Nei reacation mixture**
Buffer (10×)	NEBuffer 1 1 µl	EndoVIII Buffer 1 µl
BSA (10 mg/ml)	1 µl	1 µl
Labeled DNA (1 pmol)	1 µl	1 µl
Enzyme (varying)	1 µl (17 pmol/µl)	1 µl (2.08 pmol/µl)
Deionized H_2O	6 µl	6 µl
Total volume	10 µl	10 µl

Reaction mixture in a suitable reaction buffer specified for each enzyme (NEBuffer 1 for Fpg and Endonuclease VIII buffer for Nei provided with enzymes by New England Biolabs) in a final volume of 10 µl was prepared separately for Fpg and Nei at varying concentrations. Then, both were incubated at 37 °C for 30 min with $m^{N4,5}$C:C-DNA, or with $m^{N4,5}$C:G-dsDNA, respectively, followed by no addition or addition of endonuclease IV (Nfo) or addition of T4 polynucleotide kinase (PseT) and incubation for additional 30 min. Reactions were terminated by the addition of stop solution (20 mM EDTA, 0.5% (w/v) SDS) and proteinase K (150 µg/ml), followed by incubation at 37 °C for 20 min. DNA was precipitated with ethanol and the precipitate was solubilized in 10 µl denaturing loading buffer (DLB) solution containing 80% (v/v) formamide, 1 mM EDTA and 1 % blue dextran) and was incubated at 95 °C for 5 min to denature DNA. After cooling on ice, a volume of 5 µl was loaded to the denaturing PAGE [20% (w/v) polyacrylamide big gels prepared with 8 M urea in 10× TBE buffer; pre-run PAGE for 30 min before loading samples (see appendix; Table 9]. Electrophoresis was performed for 4 h at a constant voltage of 500 V. Visualization and

quantification was performed by using ImageQuant Software (Molecular Dynamics Inc.) for fluorescence imaging analysis.

3 RESULTS

3.1 Characterization of excision of $m^{N4,5}C$ from DNA by DNA glycosylases hUNG, Ung and Mug

To investigate possible enzymatic excision of $m^{N4,5}C$ from DNA, a 5´ fluorescently labelled polydeoxynucleotide substrate with this dimethylated base inserted at a specific position was prepared (see Materials and Methods for details). This $m^{N4,5}C$-ssDNA was alone, or annealed to a complementary strand with $m^{N4,5}C$ placed opposite either cognate G or non-cognate A, T or C, exposed to different *E. coli* and human DNA glycosylases. Enzymatic excision of the aberrant or damaged base results in an alkali-labile apurinic/apyrimidinic (AP) site, which can be monitored by the extent that *e.g.* NaOH cleaves this site by a β/δ-elimination reaction (Bailly and Verly 1987). Bifunctional DNA glycosylases will cleave the resulting AP-DNA following $m^{N4,5}C$ excision and these consecutive activities are thus monitored without NaOH/heat treatment. Cleaved is separated from un-cleaved DNA by polyacrylamide gel electrophoresis (PAGE) under denaturing conditions (Figure 1B). All the enzymes tested *i.e. E. coli* Ung, Mug (Figure 8) as well as hUNG (Figure 7) did not exhibit any detectable DNA glycosylase activity for $m^{N4,5}C$ at the concentrations and incubation conditions employed.

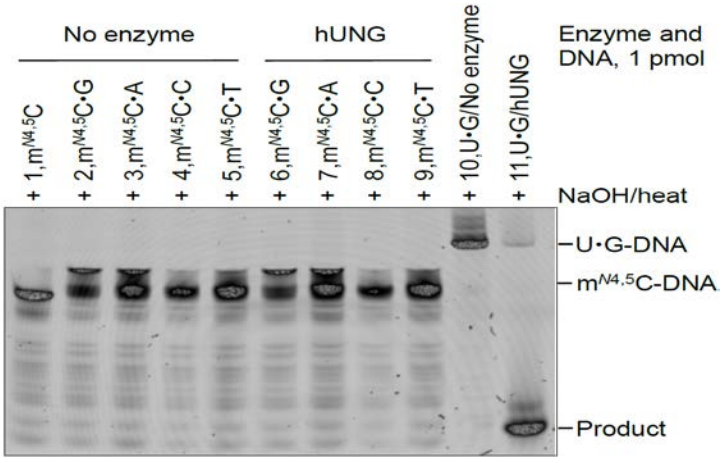


Figure 7. Human DNA glycosylases without detectable $m^{N4,5}C$-excising ability. DNA substrate alone (lanes 1–5) or with hUNG (lanes 6–9) was incubated at 37 °C in 45 mM Hepes-KOH, pH 7.5, 0.4 mM EDTA, 1 mM DTT, 70 mM KCl, 2% (v/v) glycerol, 0.1 mg/ml BSA for 30 min (final volume, 20 µl). The same amount of hUNG incubated with and without U:G-DNA [60 nt; 1 pmol (incision product; 11 nt)] was used as positive and negative control for hUNG activity, respectively. Incised was separated from un-incised DNA by denaturing PAGE at 180 V for 2 h 30 min.

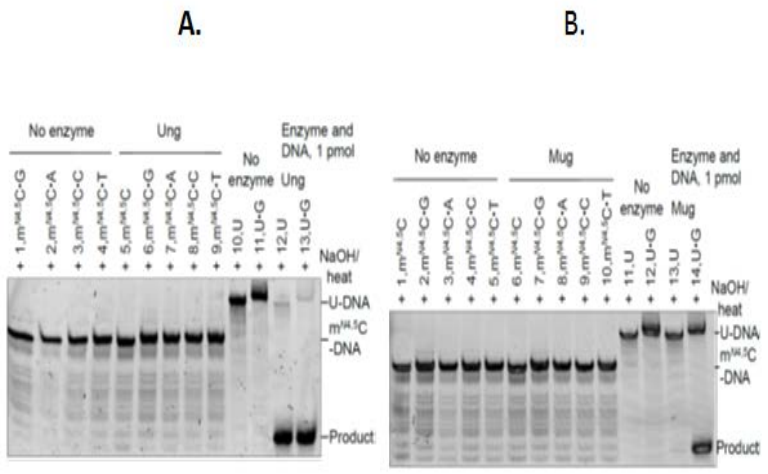


Figure 8. E. coli DNA glycosylases without detectable $m^{N4,5}C$-excising ability. DNA substrate or U-DNA or U:G-DNA was incubated at 37 °C in the reaction buffer containing the reagents added as indicated below with or without repair enzyme for 30 min (final volume, 20 µl), if not otherwise stated. (**A**) DNA substrate (1 pmol) was incubated alone (lanes 1–4) or together with Ung (1 pmol; lanes 5–9) in 100 mM Tris-HCl, 5 mM DTT, 5 mM EDTA, pH 8.0 for 30 min. Single-stranded U- or U:G-DNA [60 nt; 1 pmol (incision product; 11 nt)] incubated without (lanes 10–11) and with Ung (lanes 12–13) was used as negative and positive control for Ung activity, respectively. Incised was separated from un-incised DNA by denaturing PAGE at 180 V for 3 h. (**B**) DNA substrate (1 pmol) was incubated alone (lanes 1–5) or with Mug (1 pmol; lanes 6–10) in 20 mM Tris-HCl, pH 7.5, 1 mM EDTA, 1 mM DTT, 0.1 mg/ml BSA. Single-stranded U- or U:G-DNA (60 nt; 1 pmol) incubated without (lanes 11–12) and with Mug (lanes 13–14) was used as negative (lanes 11–13) and positive control for Mug activity (lane 14). Denaturing PAGE was performed as above at 180 V for 2 h 30 min.

3.2 Characterization of Fpg-mediated incision of $m^{N4,5}C$:C-DNA

To confirm the presence of a 3´-phosphate following incision of $m^{N4,5}C$:C-DNA by Fpg (Bailly, Verly et al. 1989) (Paper II), we treated Fpg-exposed DNA with T4 polynucleotide kinase [(PseT), which specifically removes phosphate from 3´ ends (Cameron and Uhlenbeck 1977, Midgley and Murray 1985) as opposed to other types of residues. PAGE was run using conditions favouring separation of different end products. As opposed to incubation without enzyme which formed no product (Figure 9, 1st lane) and with Fpg alone which only formed one 3´-end product (Figure 9, 2nd lane), subsequent incubations with Fpg alone and together with PseT showed that about 2/3 of the 3´-end product now appeared as a slower-migrating band corresponding to the removal of the negatively charged 3´-phosphate forming a 3´-OH end (Figure 9, 4th lane), thus confirming that Fpg-mediated incision of $m^{N4,5}C$:C-DNA forms a 3´-phosphate by β/δ-elimination. In contrast, when the second period of incubation was carried out together with *E. coli* endonuclease IV (Nfo) (Warner, Demple et al. 1980) instead of PseT, the slower-migrating band corresponding to the 3´-OH product

formed, however, in smaller amount than with PseT (Figure 9, 3rd lane). This accords with the presently accepted knowledge that Nfo is unable to remove 3´-phosphate pre-formed by Fpg alone, but is able to displace Fpg from the lesion site following base removal to incise the resultant AP site, to form the 3´-OH product necessary to prepare for repair synthesis.

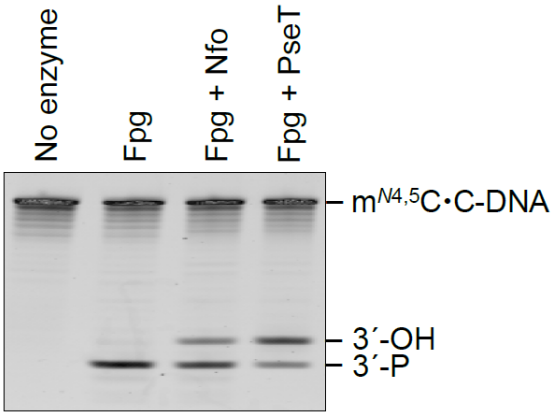


Figure 9. Definition and processing of the 3´-end following Fpg-mediated incision of $m^{N4,5}$C:C-DNA. DNA substrate ($m^{N4,5}$C:C-DNA; 1 pmol) was incubated without (1st lane) or with Fpg (17 pmol; 2nd–4th lane) at 37 °C for 30 min, followed by no addition (1st and 2nd lane), addition of endonuclease IV (Nfo; 0.083 pmol/µl; 3rd lane) or addition of T4 polynucleotide kinase (PseT; 0.289 pmol/µl; 4th lane) and incubation for additional 30 min (final volume, 10 µl). Incised was separated from un-incised DNA by denaturing PAGE at 500 V for 4 h. Abbreviation: 3´-P, 3´-phosphate.

3.3 Characterization of Nei-mediated incision of $m^{N4,5}$C:G-DNA

To confirm the presence of a 3´-phosphate following incision of $m^{N4,5}$C:G-DNA by Nei (Paper III), we treated Nei-exposed DNA with T4 polynucleotide kinase (PseT), which, as mentioned above, specifically removes phosphate from 3´ ends (Cameron and Uhlenbeck 1977, Midgley and Murray 1985) as opposed to other types of residues. Also this time PAGE was run using conditions favouring separation of different end products. As opposed to incubation without enzyme which formed no product (Figure 10, 1st lane) and with Nei alone which only formed one 3´-end product (Figure 10, 2nd lane), subsequent incubations with Nei alone and together with PseT showed that a significant amount of the 3´-end product now appeared as a slower-migrating band corresponding to the removal of the negatively charged 3´-phosphate forming a 3´-OH end (Figure 10, 4th lane), thus confirming that Nei-mediated incision of $m^{N4,5}$C:G-DNA forms a 3´-phosphate by a β/δ-elimination reaction. In contrast, when the second period of incubation was carried out together with *E. coli* endonuclease IV

(Nfo) (Warner, Demple et al. 1980) instead of PseT, the slower-migrating band corresponding to the 3´-OH product formed in a similar amount as with PseT (Figure 10, 3rd lane). This accords with the notion that Nfo displaces Nei from the lesion site following base removal to incise the resultant AP site, to form the 3´-OH product necessary for repair synthesis.

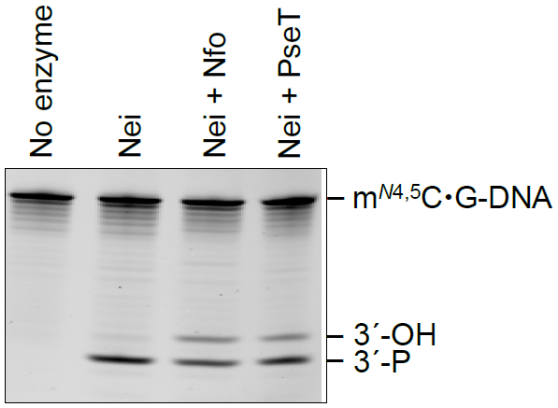


Figure 10. Definition and processing of the 3´-end following Nei-mediated incision of $m^{N4,5}$C:G-DNA. DNA substrate $m^{N4,5}$C:G-DNA (1 pmol) was incubated without (1st lane) or with Nei (2.08 pmol; 2nd–4th lane) at 37 °C for 30 min, followed by no addition (1st and 2nd lane), addition of endonuclease IV (Nfo; 0.083 pmol/µl; 3rd lane) or addition of T4 polynucleotide kinase (PseT; 0.289 pmol/µl; 4th lane) and incubation for additional 30 min (final volume, 10 µl). Incised was separated from un-incised DNA by denaturing PAGE at 500 V for 4 h. Abbreviation: 3´-P, 3´-phosphate.

4 DISCUSSION

In prokaryotic DNA m^5C shares importance with $m^{N4}C$ and $m^{N6}A$ as normal base modifications formed enzymatically after replication (Ehrlich, Wilson et al. 1987); participating in biological functions like defense against DNA hydrolysis, DNA restriction and DNA repair. In eukaryotic including mammalian DNA m^5C is even more important by being the major epigenetic mark mostly residing in CpG sequences throughout the genome (Klungland and Robertson 2016). All three base modifications are products of enzyme-catalyzed transfer of a methyl group from SAM to the cognate unmodified base.

Most human promoters contain CpG islands, which are genomic regions that have a significant overrepresentation of the CpG sequence. There is a strong correlation between promoter methylation and gene repression, also called gene silencing, and *vice versa* (Klungland and Robertson 2016). Aberrant hypo- or hyper-methylation resulting in transcriptional anomalies is connected with many disease states such as *e.g.* cancer progression. Consequently, interference with m^5C integrity may result in deregulation of genes in addition to mutagenicity and cytotoxicity.

As mentioned before, CpG sequences are hot spots for mutation induction which primarily has been ascribed to hydrolytic deamination of m^5C to thymine more readily (by a factor of three) than cytosine to uracil (Wang 1982, Ehrlich 1986). It logically follows that if a G**:**m^5C base pair is deaminated to a G**:**T base pair, an A would be inserted opposite the T during the next replication event, resulting in a G**:**C $\rightarrow$ A**:**T transition. In *E. coli* the very-short-patch repair system repairs these G**:**T mismatches, while in mammalian cells the BER pathway initiated by TDG is involved. These mechanisms are comprehensively described elsewhere (Friedberg, Walker et al. 2006, Iyer, Pluciennik et al. 2006, Hashimoto 2014). In contrast to hydrolytic damage to m^5C, few reports exists on damage inflicted by reactive oxygen species (Bjelland and Seeberg 2003, Gu, Zhang et al. 2006, Cao and Wang 2007, Cao, Jiang et al. 2009, Madugundu, Cadet et al. 2014) and almost nothing is known about damages induced by methylation exposure or arising spontaneously in cells because of erroneous

reaction with SAM (Bjelland, Tesfahun et al.). One such alkylation damage likely to arise in DNA is $m^{N4,5}C$, the subject of the present study.

The stability of the $m^{N4,5}C$–deoxyribose bond was demonstrated 50 years ago by chemical synthesis of the $m^{N4,5}C$ 2´-deoxynucleoside ($m^{N4,5}dC$) and its 5´-mono and 5´-triphosphate (Kulikowski, Żmudzka et al. 1969), indicating that the $m^{N4,5}C$ residue also would be stable if present in DNA. This was confirmed several decades later when sequential exposure of the DNA sequence 5´-C<u>C</u>AGG-3´/3´-GGT<u>C</u>C-5´ to *E. coli* Dcm and *Micrococcus varians* MvaI methyltransferases introduced $m^{N4,5}C$ at the sites underlined on the forward and complementary strand. However, such double methylation of cytosine is only achieved when the C5-methyltransferase Dcm acts first and the N^4-specific enzyme MvaI second, which also apply to other cytosine-N^4 methylases like *Bacillus centrosporus Bcn*IB methyltransferase, which methylates at the sites underlined in the sequence 5´-C<u>C</u>CGG-3´/3´-GGG<u>C</u>C-5´ (Klimasauskas, Gerasimaite et al. 2002, Vilkaitis, Lubys et al. 2002).

Although the crystal structure of $m^{N4,5}C$ in DNA is not known, it has been determined for $m^{N4,5}dC$ (Audette, Kumar et al. 1998), showing that the N^4-methyl group has a *trans* orientation with respect to the C5-methyl group (*trans* conformer; Figure 3). The N^4 amino group is sp^2-hybridized and planar (lone pair in p orbital allows overlap with the p system of the pyrimidine ring), and consequently, N^4, C4, C5 and the C5 and N^4-methyl group carbons are all in the same plane. In contrast, if the N^4-methyl group has a *cis* orientation with respect to the C5-methyl group (*cis* conformer; Figure 3) it will result in an energetically unfavourable steric clash between the two methyl groups (Figure 3, left panel). In spite of its preference in $m^{N4,5}dC$ crystal formation, the *trans* conformer does not seem to be compatible with the proper Watson-Crick base-pairing pattern. Therefore, a certain distortion in the B-helical structure is required to accommodate $m^{N4,5}C$ in DNA, and it was suggested that this distortion may promote certain DNA repair modes like nucleotide excision repair (Klimasauskas, Gerasimaite et al. 2002) (Figure 3, right panel), which has not yet been investigated.

We have used DNA with one $m^{N4,5}C$ residue incorporated at a specific site, and found that the two *E. coli* enzymes Fpg and Nei complement each other in initiating the repair of this lesion *in vitro*. As it is already known that the Fpg and Nei catalyse the AP site using N-terminal Proline for the nucleophile attack and a 3´-PO_4 terminus is generated at the DNA strand break after the removal of deoxyribose residue by β/δ elimination reaction (Hegde, Hazra et al. 2008). Since this terminus 3´-PO_4 product is not desirable as it hinders the further elongation of strand by DNA polymerase. We used PseT and Nfo to confirm and process this terninus blocking 3´-end product and Pset processed about 2/3 of the 3´-end product by removing this negatively charged 3´-phosphate and formed a 3´-OH end, favourable product for further elogation. It confirmed the fact that Fpg/Nei-mediated incision of $m^{N4,5}C$:C-DNA/$m^{N4,5}C$:G-DNA forms a 3´-phosphate by β/δ-elimination reaction. In contrast, when the second period of incubation was carried out together with *E. coli* endonuclease IV (Nfo) (Warner, Demple et al. 1980) instead of PseT, the 3´-OH product was formed, however, in smaller amount than with PseT (Figure 9 and 10). This accords with the presently accepted knowledge that Nfo is unable to remove 3´-phosphate pre-formed by Fpg/Nei alone, but is able to displace Fpg/Nei from the lesion site following base removal to incise the resultant AP site, to form the 3´-OH terminus that is amenable to elongation by DNA polymerase to precede the repair activity.

Interestingly, this is the first report describing a repair activity for a further methylated m^5C residue in DNA. In contrast, several other *E. coli* DNA glycosylases like Ung and Mug (Figure 8) and the human enzyme hUNG (Figure 7) did not show repair activity at the incubation conditions employed.

Although *E. coli* does not contain a cytosine-N^4 methyltransferase to methylate Dcm-methylated cytosines, other bacterial species or strains contain both m^5C and $m^{N4}C$ in their DNA (Ehrlich, Wilson et al. 1987) and thus both types of methylases which suggest formation

of the $m^{N4,5}C$ damage as a likely event. However, in the case of *E. coli*, we cannot totally rule out the possibility that cytosine-N^4 methyltransferase may be acquired by horizontal gene transfer or viral infection as well as $m^{N4,5}C$ may arise as a consequence of chemical insults to the cellular DNA, explaining the need for its repair. Mutated versions of methyltransferase genes and thus proteins may also be a source of erroneous methylations (Alexeeva et. al).

Indeed, this finding that two enzymes complement each other in the initiation of BER of $m^{N4,5}C$ strengthen the notion that this lesion is a challenge to the genomic integrity of *E. coli*. Likewise, the lack of activity for $m^{N4,5}C$ exhibited by the human Nei orthologues and other human glycosylases (Tesfahun, Alexeeva et al.) might indicate that $m^{N4,5}C$ is a primarily "prokaryotic" lesion. However, the much higher abundance of m^5C in eukaryotic including mammalian DNA argues against that notion and suggests a thorough search for human or mammalian repair activities for $m^{N4,5}C$ and to possibly detect the lesion in biological samples.

5 REFERENCES

Almeida, K. H. and R. W. Sobol (2007). "A unified view of base excision repair: lesion-dependent protein complexes regulated by post-translational modification." *DNA repair* **6**(6): 695-711.

Alonso, A., et al. (2006). "An intrinsic 5′-deoxyribose-5-phosphate lyase activity in DNA polymerase beta from *Leishmania infantum* supports a role in DNA repair." *DNA repair* **5**(1): 89-101.

Audette, G. F., et al. (1998). "N^4,5-dimethyl-2 '-deoxycytidine." *Acta Cryst. C* **54**: 1987–1990.

Bailly, V. and W. G. Verly (1987). "*Escherichia coli* endonuclease III is not an endonuclease but a β-elimination catalyst." *Biochem J* **242**(2): 565–572.

Bailly, V., et al. (1989). "Mechanism of DNA strand nicking at apurinic/apyrimidinic sites by *Escherichia coli* [formamidopyrimidine]DNA glycosylase." *Biochem J* **262**: 581–589.

Bernstein, N. K., et al. (2005). "The molecular architecture of the mammalian DNA repair enzyme, polynucleotide kinase." *Molecular Cell* **17**(5): 657-670.

Bjelland, S. and E. Seeberg (2003). "Mutagenicity, toxicity and repair of DNA base damage induced by oxidation." *Mutat Res* **531**: 37–80.

Bjelland, S., et al. "Spontaneous and enzymatic modifications of the epigenetic DNA base 5-methylcytosine as targets for repair and mutagenesis." *Manuscript.*

Breiling, A. and F. Lyko (2015). "Epigenetic regulatory functions of DNA modifications: 5-methylcytosine and beyond." *Epigenetics & chromatin* **8**(1): 24.

Cameron, V. and O. C. Uhlenbeck (1977). "3′-Phosphatase activity in T4 polynucleotide kinase." *Biochemistry* **16**(23): 5120–5126.

Campalans, A., et al. (2005). "XRCC1 interactions with multiple DNA glycosylases: a model for its recruitment to base excision repair." *DNA repair* **4**(7): 826-835.

Cao, H., et al. (2009). "Kinetics of deamination and Cu(II)/H_2O_2/Ascorbate-induced formation of 5-methylcytosine glycol at CpG sites in duplex DNA." *Nucleic Acids Res* **37**(19): 6635–6643.

Cao, H. and Y. Wang (2007). "Quantification of oxidative single-base and intrastrand cross-link lesions in unmethylated and CpG-methylated DNA induced by Fenton-type reagents." *Nucleic Acids Res* **35**(14): 4833–4844.

Chen, D. S., et al. (1991). "Two distinct human DNA diesterases that hydrolyze 3′-blocking deoxyribose fragments from oxidized DNA." *Nucleic Acids Res* **19**(21): 5907-5914.

Cortellino, S., et al. (2003). "The base excision repair enzyme MED1 mediates DNA damage response to antitumor drugs and is associated with mismatch repair system integrity." *Proc Natl Acad Sci USA* **100**(25): 15071-15076.

Dianov, G. L. and U. Hübscher (2013). "Mammalian base excision repair: the forgotten archangel." *Nucleic Acids Res* **41**(6): 3483-3490.

Ehrlich, M., et al. (1985). "DNA methylation in thermophilic bacteria: *N*4-methylcytosine, 5-methylcytosine, and *N*5-methyladenine." *Nucleic Acids Res* **13**(4): 1399-1412.

Ehrlich, M., Norris, K.F., Wang, R.Y.-H., Kuo, K.C., Gehrke, C.W. (1986). "DNA cytosine methylation and heat-induced deamination." *Biosci Rep* **6**: 387–393.

Ehrlich, M., et al. (1987). "*N*4-methylcytosine as a minor base in bacterial DNA." *J Bacteriol* **169**(3): 939–943.

Erzberger, J. P., et al. (1998). "Elements in abasic site recognition by the major human and *Escherichia coli* apurinic/apyrimidinic endonucleases." *Nucleic Acids Res* **26**(11): 2771-2778.

Friedberg, E. C., et al. (2006). *DNA Repair and Mutagenesis*. Washington, DC., ASM Press.

Fromme, J. C. and G. L. Verdine (2004). "Base excision repair." *Adv Protein Chem* **69**: 1-41.

Fu, Y., et al. (2015). "*N*6-methyldeoxyadenosine marks active transcription start sites in *Chlamydomonas*." *Cell* **161**(4): 879-892.

Gu, C., et al. (2006). "Recognition and incision of oxidative intrastrand cross-link lesions by UvrABC nuclease." *Biochemistry* **45**(35): 10739–10746.

Hashimoto, H. (2014). "Structural and mutation studies of two DNA demethylation related glycosylases: MBD4 and TDG." *Biophysics (Nagoya-shi)* **10**: 63–68.

He, Y.-F., et al. (2011). "Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA." *Science* **333**(6047): 1303-1307.

Hegde, M. L., et al. (2008). "Early steps in the DNA base excision/single-strand interruption repair pathway in mammalian cells." *Cell Research* **18**(1): 27-47.

Hendrich, B., et al. (1999). "The thymine glycosylase MBD4 can bind to the product of deamination at methylated CpG sites." *Nature* **401**(6750): 301-304.

Heyn, H. and M. Esteller (2015). "An adenine code for DNA: a second life for *N*6-methyladenine." *Cell* **161**(4): 710-713.

Hitomi, K., et al. (2007). "The intricate structural chemistry of base excision repair machinery: implications for DNA damage recognition, removal, and repair." *DNA repair* **6**(4): 410-428.

Ito, S., et al. (2011). "Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine." *Science* **333**(6047): 1300-1303.

Iyer, R. R., et al. (2006). "DNA mismatch repair: functions and mechanisms." *Chem Rev* **106**(2): 302–323.

Kim, Y.-J. and D. M Wilson III (2012). "Overview of base excision repair biochemistry." *Curr Mol Pharmacol* **5**(1): 3-13.

Klimasauskas, S., et al. (2002). "*N*4,5-dimethylcytosine, a novel hypermodified base in DNA." *Nucleic Acids Res Suppl*(2): 73–74.

Klose, R. J. and A. P. Bird (2006). "Genomic DNA methylation: the mark and its mediators." *Trends Biochem Sci* **31**(2): 89-97.

Klungland, A. and A. B. Robertson (2016). "Oxidized C5-methyl cytosine bases in DNA: 5-Hydroxymethylcytosine; 5-formylcytosine; and 5-carboxycytosine." *Free Radic Biol Med* **http://dx.doi.org/10.1016/j.freeradbiomed.2016.11.038**.

Klungland, A. and A. B. Robertson (2016). "Oxidized C5-methyl cytosine bases in DNA: 5-Hydroxymethylcytosine; 5-formylcytosine; and 5-carboxycytosine." *Free Radic Biol Med.*

Kow, Y. W. (2002). "Repair of deaminated bases in DNA." *Free Radic Biol Med* **33**(7): 886-893.

Krokan, H. E., et al. (2000). "Base excision repair of DNA in mammalian cells." *FEBS Letters* **476**(1-2): 73-77.

KROKAN, H. E., et al. (1997). "DNA glycosylases in the base excision repair of DNA." *Biochem J* **325**(1): 1-16.

Krwawicz, J., et al. (2007). "Bacterial DNA repair genes and their eukaryotic homologues: 1. Mutations in genes involved in base excision repair (BER) and DNA-end processors and their implication in mutagenesis and human disease." *Acta Biochim Pol* **54**(3): 413-434.

Kulikowski, T., et al. (1969). "Preparation and properties of N^4-alkyl analogues of 5-methyl-2´-deoxycytidine, their 5´-mono- and triphosphates, and N^4-alkyl derivatives of 1,5-dimethylcytosine." *Acta Biochim Pol* **16**(2): 201–217.

Li, D., et al. (2012). "Exocyclic carbons adjacent to the *N*6 of adenine are targets for oxidation by the *Escherichia coli* adaptive response protein AlkB." *J Am Chem Soc* **134**(21): 8896-8901.

Lindahl, T., et al. (1997). "DNA excision repair pathways." *Curr Opin Genet Dev* **7**(2): 158-169.

Lindahl, T. and R. D. Wood (1999). "Quality control by DNA repair." *Science* **286**(5446): 1897-1905.

M. TAYLOR AR LEHMANN, E. (1998). "Conservation of eukaryotic DNA repair mechanisms." *Int J Radiat Biol* **74**(3): 277-286.

Madugundu, G. S., et al. (2014). "Hydroxyl-radical-induced oxidation of 5-methylcytosine in isolated and cellular DNA." *Nucleic Acids Res*: gku334.

Midgley, C. A. and N. E. Murray (1985). "T4 polynucleotide kinase; cloning of the gene (pseT) and amplification of its product." *EMBO J* **4**(10): 2695–2703.

Mol, C. D., et al. (2000). "Abasic site recognition by two apurinic/apyrimidinic endonuclease families in DNA base excision repair: the 3′ ends justify the means." *Mutat Res* **460**(3): 211-229.

Nilsen, H. and H. E. Krokan (2001). "Base excision repair in a network of defence and tolerance." *Carcinogenesis* **22**(7): 987-998.

Parikh, S. S., et al. (1997). "Base excision repair enzyme family portrait: integrating the structure and chemistry of an entire DNA repair pathway." *Structure* **5**(12): 1543-1550.

Parsons, B. L. (2003). "MED1: a central molecule for maintenance of genome integrity and response to DNA damage." *Proc Natl Acad Sci* **100**(25): 14601-14602.

Robertson, K. D. (2005). "DNA methylation and human disease." *Nat Rev Genet* **6**(8): 597-610.

Rogers, S. G. and B. Weiss (1980). "[26] Exonuclease III of *Escherichia coli* K-12, an AP endonuclease." *Methods Enzymol* **65**: 201-211.

Sancar, A., et al. (2004). "Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints." *Annu Rev Biochem* **73**(1): 39-85.

Sedgwick, B. (2004). "Repairing DNA-methylation damage." *Nat Rev Mol Cell Biol* **5**(2): 148-157.

Serre, L., et al. (2002). "Crystal structure of the *Lactococcus lactis* formamidopyrimidine-DNA glycosylase bound to an abasic site analogue-containing DNA." *EMBO j* **21**(12): 2854-2865.

Slupphaug, G., et al. (1996). "A nucleotide-flipping mechanism from the structure of human uracil-DNA glycosylase bound to DNA." *Nature* **384**(6604): 87.

Suh, D., et al. (1997). "3′-Phosphodiesterase activity of human apurinic/apyrimidinic endonuclease at DNA double-strand break ends." *Nucleic Acids Res* **25**(12): 2495-2500.

Takao, M., et al. (2002). "A back-up glycosylase in Nth1 knock-out mice is a functional Nei (endonuclease VIII) homologue." *J Biol Chem* **277**(44): 42205-42213.

Tesfahun, A., et al. "Excision of the double methylated base *N*4,5-dimethylcytosine from DNA by *Escherichia coli* Fpg protein." *Manuscript*.

Trivedi, S., et al. (2005). "Nucleic acid stability in thermophilic prokaryotes: a review." *J Cell Mol Biol* **4**: 61-69.

Turek-Plewa, J. and P. P. Jagodzinski (2005). "The role of mammalian DNA methyltransferases in the regulation of gene expression." *Cell Mol Biol Lett* **10**(4): 631.

Vilkaitis, G., et al. (2002). "Circular permutation of DNA cytosine-*N*4 methyltransferases: in vivo coexistence in the BcnI system and in vitro probing by hybrid formation." *Nucleic Acids Res* **30**(7): 1547–1557.

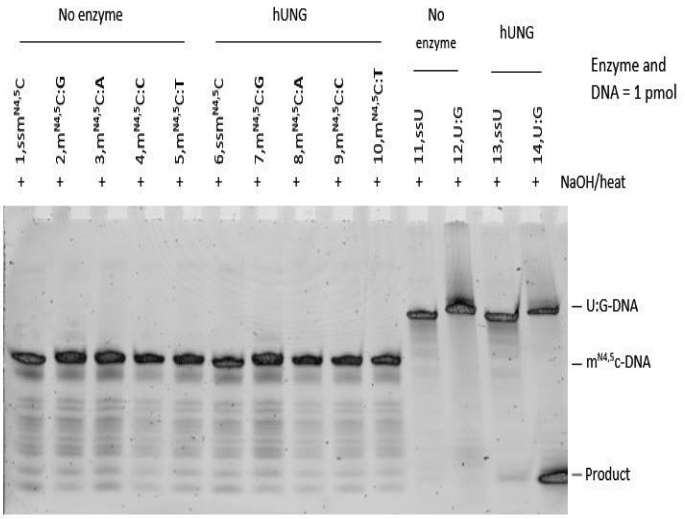
Wang, R. Y.-H., Kuo, K.C., Gehrke, C.W., Huang, L.-H., Ehrlich, M. (1982). "Heat- and alkali-induced deamination of 5-methylcytosine and cytosine residues in DNA." *Biochim Biophys Acta* **697**: 371–377.

Warner, H. R., et al. (1980). "Apurinic/apyrimidinic endonucleases in repair of pyrimidine dimers and other lesions in DNA." *Proc Natl Acad Sci USA* **77**: 4602–4606.

Weisenberger, D. J., et al. (2005). "Analysis of repetitive element DNA methylation by MethyLight." *Nucleic Acids Res* **33**(21): 6823-6836.

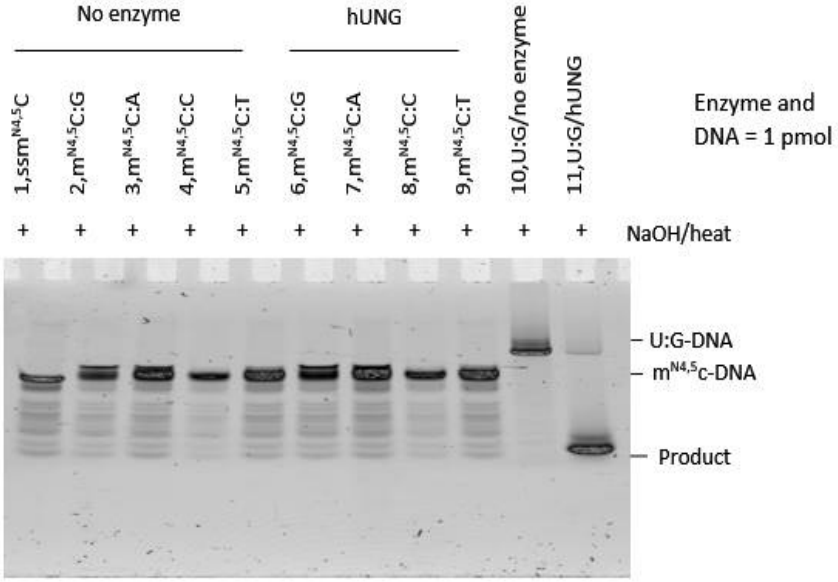
6 APPENDIX

A.



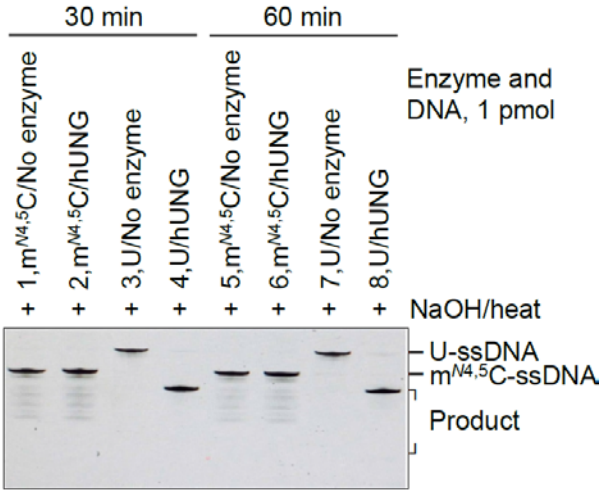
Human DNA glycosylases without detectable m^{N4,5}C-excising ability

B.



Human DNA glycosylases without detectable m^{N4,5}C-excising ability

C.



Human DNA glycosylases without detectable $m^{N4,5}C$-excising ability

D.

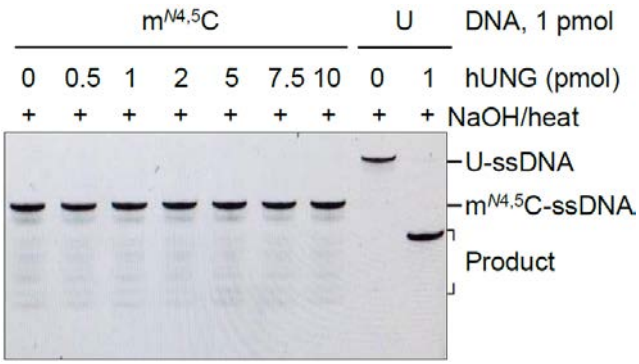
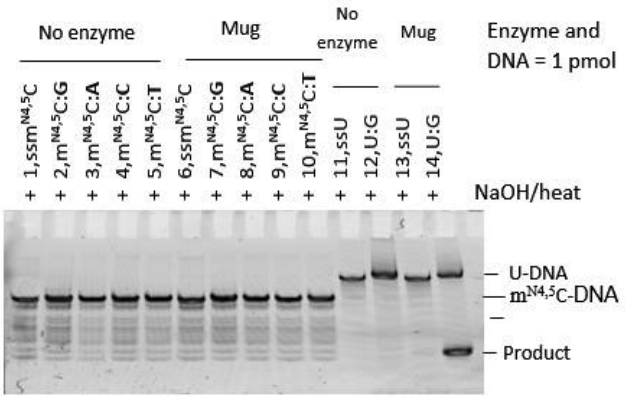


Figure 11. Human DNA glycosylase hUNG without detectable $m^{N4,5}C$-excising ability: DNA substrate or U-DNA was incubated at 37 °C in the reaction buffer containing the reagents added as indicated below with or without repair enzyme and for the time periods also indicated below or in the figures (final volume, 20 µl). Incised was separated from un-incised DNA by denaturing PAGE at 150–180 V for 2–3 h. (**A, B**): hUNG was incubated in 45 mM Hepes-KOH, pH 7.5, 0.4 mM EDTA, 1 mM DTT, 70 mM KCl, 2% (v/v) glycerol, 0.1 mg/ml BSA for 30 min. The same amount of hUNG incubated with and without U·G-DNA (60 nt; 1 pmol) was used as positive and negative control for hUNG activity, respectively. **(C):** indicates hUNG activity at different incubation time period (30 min and 1 h). **(D):** indicates the activity of hUNG (1 pmol) at different hUNG concentration (pmol/µl, 0.5, 1, 2, 5, 7.5, 10).

A.



B.

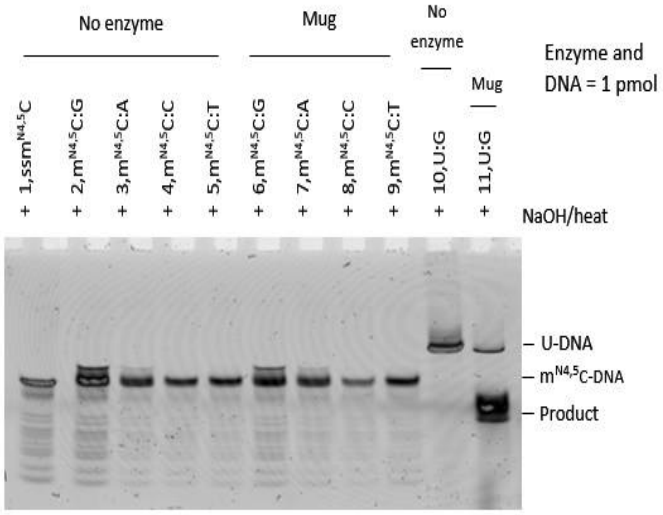


Figure 12. E. coli DNA glycosylase (Mug) without detectable m^{N4,5}C-excising ability. **(A &B):** DNA substrate (1 pmol) was incubated alone (lanes 1–5) or with Mug (1 pmol; lanes 6–10) in 20 mM Tris-HCl, pH 7.5, 1 mM EDTA, 1 mM DTT, 0.1 mg/ml BSA. Single-stranded U- or U:G-DNA (60 nt; 1 pmol) incubated without (**for A:** lanes 11–12; **for B**: lane 10) and with Mug (**for A:** lanes 13–14; **for B:** lane 11) was used as negative (lanes 11–13) and positive control for Mug activity (lane 14). Denaturing PAGE was performed as above at 180 V for 2 h 30 min.

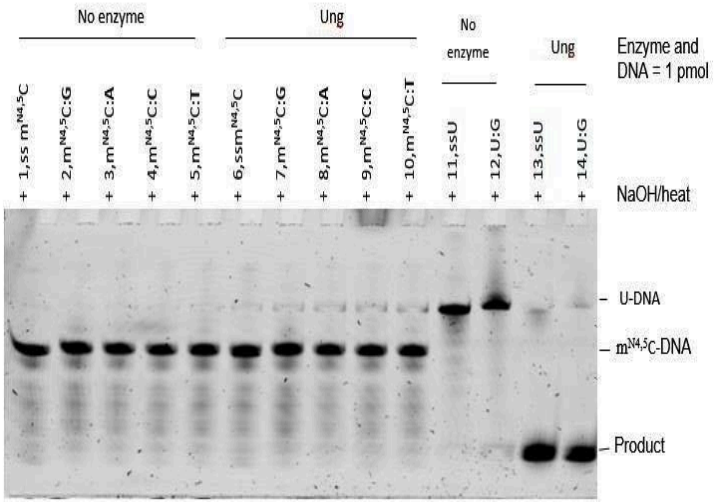


Figure 13. E. coli DNA glycosylase (Ung) without detectable $m^{N4,5}C$-excising ability. DNA substrate or U-DNA or U:G-DNA was incubated at 37 °C in the reaction buffer containing the reagents added as indicated below with or without repair enzyme for 30 min (final volume, 20 µl), if not otherwise stated. DNA substrate (1 pmol) was incubated alone (lanes 1–4) or together with Ung (1 pmol; lanes 5–9) in 100 mM Tris-HCl, 5 mM DTT, 5 mM EDTA, pH 8.0 for 30 min. Single-stranded U- or U:G-DNA (60 nt; 1 pmol) incubated without (lanes 10–11) and with Ung (lanes 12–13) was used as negative and positive control for Ung activity, respectively. Incised was separated from un-incised DNA by denaturing PAGE at 180 V for 3 h.

Table 9. Denaturing 20% (w/v) PAGE gel with 8 M urea in TAE buffer

Composition	**For small gel (5 ml)**	**For big gel (15 ml)**
Urea (final conc. 8 M)	3.36 g	10.08 g
Acrylamide 40% (w/v) (final conc. 20%)	3.5 ml	10.5 ml
TBE (Tris/borate/EDTA) buffer (10×)	0.70 ml	2.1 ml
Deionized H_2O	0.28 ml	840 µl (Mixed all the compounds together, warm in the microwave oven until urea is dissolved and cooled down at RT!)
Temed (Tetramethylethylenediamine)	3.5 µl	10.5 µl
APS (ammonium persulfate) 10%	35 µl	105 µl (APS added in the end and immediately cast the gel)

Table 10. Denaturing 20% (w/v) PAGE gel with 7 M urea in Taurine buffer

Composition	**For small gel**
Urea (final conc. 8 M)	2.625 g
Acrylamide 40% (w/v) (final conc. 20%)	3.125 ml
Taurine buffer 20×	0.313ml
Deionized H_2O	0.94 ml (Mixed all the compounds together, warm in the microwave oven until urea is dissolved and cooled down at RT!)
Temed (tetramethylethylenediamine)	6.25 µl
APS (ammonium persulfate) 10%	31.625 µl (APS added in the end and immediately cast the gel)