DNA Base Excision Repair: Mutation Induction and Novel Functions

by

Izaskun Muruzábal-Lecumberri

Thesis submitted in fulfillment of
the requirements for the degree of
PHILOSOPHIAE DOCTOR
(PhD)

Faculty of Science and Technology
Department for Mathematics and Natural Sciences
Centre for Organelle Research (CORE)
2015

University of Stavanger
N-4036 Stavanger
NORWAY
www.uis.no

ISBN: 978-82-7644-608-1
ISSN: 1890-1387
PhD thesis no. 258

Abstract

DNA is susceptible to chemical modifications corrupting its cellular information processing function, necessitating correction of such modifications. Cells are formed mostly by water – giving rise to hydrolytic reactions – and aerobic metabolism is a source of reactive oxygen species (ROS) – causing oxidation. Deamination of cytosine to uracil is an example of the former and oxidation of thymine to 5-formyluracil (f^5U) is an example of the latter.

In order to avoid the incorporation of wrong nucleosides into DNA, f^5U must be eliminated and substituted by the correct base. The main mechanism for the repair of f^5U is the base excision repair (BER) pathway, in which specific glycosylases recognize the damage and excise it from its ribose residue. Several glycosylases have been found to be involved in the repair of f^5U in *Escherichia coli*, where 3-methyladenine DNA glycosylase II (AlkA) may be the most important. However, here we present evidence to indicate that the nucleotide excision repair protein UvrA is also involved in the repair of f^5U, although the mechanism has yet to be elucidated.

Interestingly, we have found that the AlkA glycosylase, in addition to alleviating is also able to promote mutation induction by 5-formyldeoxyuridine in *E. coli*. Extrapolated to the mammalian system this observation suggests that DNA repair genes may act as oncogenes under certain cellular conditions.

Uracil lesions in DNA are repaired by the BER pathway initiated by a mono-functional uracil-DNA glycosylase (UDG), and the family 1 UDGs are the most extensively studied glycosylases. We have found that UDG-mediated repair initiation by *E. coli* Ung and hUNG also involves DNA strand incision generating a 3'-α,β-unsaturated aldehyde (UIP) and a 5'-phosphate, which demonstrates that they are indeed bi-functional enzymes. However, while the

cleavage of the N-glycosyl bond by β-elimination occurs through a covalent Schiff base intermediate between a reactive active site lysine and the deoxyribose moiety in other bi-functional glycosylases, the uracil–deoxyribose bond is suggested to be cleaved by an activated water molecule carrying out the elimination reaction.

Acknowledgements

This thesis was performed from April 2012 to April 2015 at the Centre for Organelle Research (CORE) that belongs to the Department of Mathematics and Natural Sciences, Faculty of Science and Technology, University of Stavanger, Norway.

I would like to express my gratitude to my supervisor, Professor Svein Bjelland, who gave me the opportunity to come to this awesome city and country and work in what I like the most: science. He always trusted my point of view and gave me the freedom to try, the chance to fall and the support to stand up and try again. Thanks for the three hours meetings about so many worldly things related, in one way or another, with the most important thing in life: DNA repair.

I cannot forget my co-authors, who shared with me their expertise and good advices: Marivi N. Moen, PhD, for all the material she provided and being a very useful reference for my thesis; Professor Arne Klungland, for his contributions to the papers; Meh Sameen Nawaz, PhD candidate, and Ingrun Alseth, PhD, for instructing me in radioactivity techniques in Oslo; and Professor Akira Matsuda, for kindly provide essential material for my assays.

Thanks especially to my colleagues at CORE to make the environment so cozy, for always having fresh coffee ready in the morning and being so dedicated to home baked cakes for the Happy Hour. To my lab mentor, Kristin Grøsvik, PhD, who taught me techniques both in science and Norwegian life and introduced me to lunch at 11.30 am together with Kristine M. Olsen, PhD, technician Camilla Hoel, MSc, and Ingunn W. Jolma, PhD. Thanks for welcoming me and help me with my science in so many ways. I also need to thank to my lab colleagues and friends Marina Alexeeva, MSc, and Oleg Agafonov, MSc, for the life outside CORE and the discovery that Russia and Spain are not that different; and our technician, Xiang Ming Xu, PhD, for his

experience and his capacity for knowing things about my country that I did not even suspect.

Life is hard in a foreign country, but I was lucky to find people who made me feel home. Thanks to my dear friends and also PhD candidates at UiS Charlotte D. Botter and Thanusha Naidoo, with whom I shared moments of tears and laughs, and to my good "Stavanger venner"; a piece of my heart will always belong to this city. To the UiS Doctoral Community (UiSDC) and Stipendiatorganisasjonene i Norge (SiN) that made me a proud vice-president for one year and gave me the opportunity to participate in the PhD life and meeting nice people who are now good friends.

Finally, to my family and my "Biofamily" in Spain, always so proud of my achievements, and my most important support: my parents, Jesús Mari and Maria Luisa. As far away as my dreams make me fly, they will always light up a way back home.

"In the end, we only regret the
chances we didn't take"

Nicholas Sparks, 'The Guardian'

Contents

List of papers

Paper 1

Marina Alexeeva, Marivi N. Moen, Xiang Ming Xu, Kristin Grøsvik, **Izaskun Muruzábal-Lecumberri**, Kristine M. Olsen, Ingar Leiros, Finn Kirpekar, Arne Klungland and Svein Bjelland: DNA uracil excision includes strand incision by hUNG. Manuscript.

Paper 2

Kristin Grøsvik, **Izaskun Muruzábal-Lecumberri** and Svein Bjelland: Severe repair-deficiency for oxidised DNA bases causes primarily G·C → A·T transitions in *Escherichia coli*. Manuscript.

Paper 3

Kristin Grøsvik, **Izaskun Muruzábal-Lecumberri**, Ingeborg Knævelsrud, Gyri Teien Haugland, Hilde Ånensen, Ingrun Alseth, Kousuke Sato, Akira Matsuda, Ingar Leiros, Arne Klungland and Svein Bjelland: Damage-specific mutation induction promoted by repair. Manuscript.

Paper 4

Izaskun Muruzábal-Lecumberri, Kristin Grøsvik, Kousuke Sato, Akira Matsuda and Svein Bjelland: Alleviation and promotion of damage-specific mutation induction in *Escherichia coli* are highly dependent on the *uvrA* gene. Manuscript.

List of figures

Glossary

6–4PP	6–4 photoproduct
AP	apurinic/apyrimidinic
AlkA	3-methyladenine DNA glycosylase II
BER	base excision repair
DSB	double-strand break
EcUng/ Ung	*E. coli* uracil-DNA glycosylase
fapyG	2,6-diamino-4-hydroxy-5-formamidopyrimidine
Fpg	formamidopyrimidine-DNA glycosylase
f^5dU	5-formyl-2'-deoxyuridine
f^5U	5-formyluracil
$f^5U^{\ominus}$	5-formyluracil (ionic form)
$f^5U^{C=O}$	5-formyluracil (keto form)
h^2A	2-hydroxyadenine
h^5C	5-hydroxycytosine
h^5U	5-hydroxyuracil
hAPE	human AP endonuclease
hm^5U	5-hydroxymethyluracil
hmh	5-hydroxy-5-methylhydantoin
hNTH	human endonuclease III
hOOG	human 8-oxoguanine-DNA N-glycosylase
hpm^5U	5-hydroperoxymethyluracil
hSMUG	single-strand selective monofunctional uracil-DNA glycosylase
hUNG	human uracil-DNA glycosylase
m^3A	3-methyladenine
m^3G	3-methylguanine
m^5C	5-methylcytosine
m^7A	7-methyladenine
m^7G	7-methylguanine
MMS	methyl methanesulphonate
MPG	3-methyladenine-DNA glycosylase
Nei	endonuclease VIII

NER	nucleotide excision repair
Nfo	endonuclease IV
Nth	endonuclease III
$oxo^{2}A$	1,2-dihydro-2-oxoadenine
$oxo^{8}A$	7,8-dihydro-8-oxoadenine
$oxo^{8}G$	8-oxo-7,8-dihydroguanine
rif^{R}	rifampicin resistance
ROS	reactive oxygen species
ssDNA	single-stranded DNA
dsDNA	double-stranded DNA
Tg	thymine glycol
Th^{5}	5-hydroxy-5,6-dihydrothymine
Th^{6}	6-hydroxy-5,6-dihydrothymine
UDG	uracil-DNA glycosylase
Ug	uracil glycol
Xth	exonuclease III

1 Introduction

Genetic information in all living cells is contained in their DNA, at first assumed to be a highly stable molecule. This contrasts with many viruses, which use RNA as genetic material, which leads to extremely increased mutation rates. Presently it is believed that DNA replaced RNA during early cellular evolution (Lindahl, 1993). Although the only chemical difference between DNA and RNA is an H rather than an OH group at the 3´C position of the sugar moiety, respectively (Figure 1.1), this causes a highly increased stability of the sugar-phosphate backbone (Friedberg *et al.*, 2005). However, the price paid is a more labile N-glycosyl bond in DNA compared to RNA (Figure 1.1). Due to hydrolysis of this N-glycosyl bond, apurinic/apyrimidinic (AP) sites are the most common DNA damage in all cells. Although being

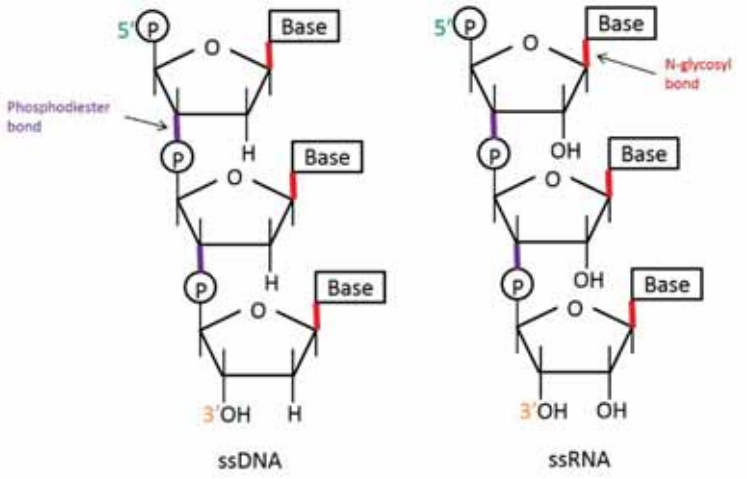

Figure 1.1 Chemical composition of ssDNA and RNA molecules. Phosphodiester bonds are marked in purple and N-glycosyl bond, in red.

mainly cytotoxic due to inhibition of replication or transcription, AP sites are believed to contribute significantly to mutagenesis when they are bypassed by certain translesion DNA polymerases, where adenine is the preferred (mis)incorporated base (the "A-rule") (Pagès *et al.*, 2008; Strauss, 2002). Because of this and because AP sites are the most common DNA lesion, cell health and survival is dependent on its almost total removal. This also applies to other lesions (Kim and Wilson, 2012).

When lesions arise in DNA bases and are not properly repaired, mutations may appear after replication due to misincorporation of a wrong base opposite the lesion. An example of C to T transition is shown in Figure 1.2.

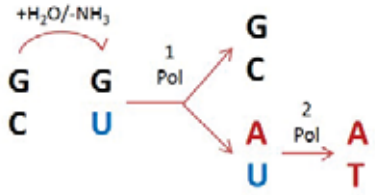

Figure 1.2 Mechanism of mutagenesis for one base substitution mutation. Hydrolytic deamination of C to U. After the first round of replication (1 Pol), U misincorporates A instead of G. In the second round of replication (2 Pol), A incorporates T. The result is a transition from GC to AT.

1.1 DNA damage formation

All cells are constantly suffering damage to their DNA caused by physical agents and chemical species coming from the outside environment or formed by the cell itself. DNA lesions can affect single nucleotides or result in changes of the helix structure. UV-B (and UV-C) light is a major producer of thymine dimers and 6–4 photoproducts (6–4PP), which both disturb the helical structure of the DNA, while ionizing radiation is particularly disruptive due to the formation of double-strand breaks (DSB). Certain aromatic compounds present in oil and tar (*e.g.* tobacco smoke) attach to DNA as bulky DNA adducts disturbing its helical structure (Henkler *et al.*, 2012). The cellular methyl donor S-adenosylmethionine (SAM) (Bauerle *et al.*, 2014) contains a reactive methyl group – like methyl methanesulfonate (MMS) (Figure 1.3) – that methylates DNA erroneously at a certain frequency. SAM is, in addition to other alkylating agents present in the environment, an important source of DNA damage. Such

damage can arise via two different chemical reactions which are called nucleophilic substitutions 1 and 2 (S_N1 and S_N2) and methylate oxygen and nitrogen residues in both single- and double-stranded DNA (ss- and dsDNA) (Sedgwick, 2004).

However, the quantitatively most important lesions formed endogenously in DNA are due to the water component of all cells and the generation of reactive oxygen species in aerobic cells described below (Friedberg *et al.*, 2005; Kim and Wilson, 2012).

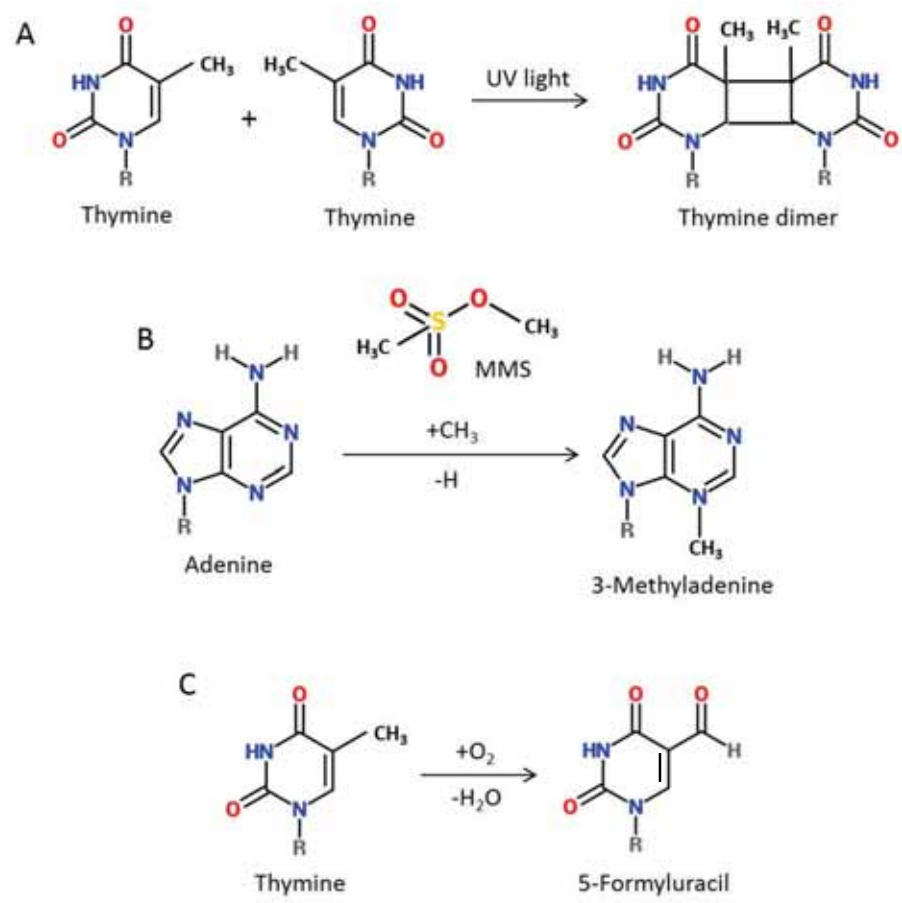

Figure 1.3 Chemical DNA base modifications. Formation of bulky adducts (A), methylation (B) and oxidation (C). Deamination is shown in Figure 1.4.

1.1.1 Hydrolytic damage

Since water is a major component of all cells, hydrolytic reactions are a major threat to the chemical integrity of their DNA and, as already mentioned, generate AP sites as the most abundant lesion. Purines are liberated from DNA at similar rates, while pyrimidines are liberated at 5% of the rate of the purines (Lindahl, 1993) due to their higher N-glycosyl bond stability (Friedberg *et al.*, 2005). The difference in depurination between ss- and dsDNA is only four-fold, suggesting a poor protection by the double helical structure (Lindahl, 1993). The measured depurination rate for duplex DNA *in vivo* translates to loss of approximately one purine per generation per *Escherichia coli* chromosome; given a doubling time of one hour. For mammalian cells, with a larger genome and a longer cell cycle, this rate is equivalent to loss of 9 000–10 000 purines during each 24 hours period (Kim and Wilson, 2012; Lindahl, 1979). In addition to being responsible for the formation of AP sites, hydrolytic deamination of cytosine, adenine, guanine and 5-methylcytosine (m^5C) resulting in uracil, hypoxanthine, xanthine and thymine, respectively (Figure 1.4), occurs extensively in a pH- and temperature-dependent manner (Friedberg *et al.*, 2005). Uracil arising from cytosine is the second most common hydrolytic DNA damage and, if not repaired prior to replication, C to T transition mutations are formed (Lindahl, 1993) (Figure 1.2).

Uracil in DNA

Although uracil is the adenine partner base in RNA it appears erroneously in the DNA as a consequence of cytosine deamination (Figure 1.3) or misincorporated from dUTP. The latter common source of genomic uracil is the principal source of abasic sites in yeast (Guillet and Boiteux, 2003) and is especially relevant in cycling cells. dUMP is a normal intermediate in the biosynthesis of dTMP and dTTP but its levels are kept very low by dUTPase (Ladner and Caradonna, 1997). Genomic uracil by deamination results in mutagenic U·G mispairs and non-mutagenic but may be genotoxic U·A pairs (Kavli *et al.*, 2007).

However, deamination of cytosine plays an important biological role in immunoglobulins diversification (Durandy and Honjo, 2001) and protection against retroviral infection (Pham *et al.*, 2005) and so it has been observed targeted cytosine deamination by specialized enzymes like APOBEC

(apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like) and AID (activation-induced cytidine) deaminases (Kavli *et al.*, 2007).

Spontaneous cytosine deamination has been estimated to take place at a rate of 60–500 events per human genome per day (Barnes and Lindahl, 2004; Krokan *et al.*, 2002). However, based on different studies on ss- and dsDNA and estimating the percentage of ssDNA in the total DNA of a cell as approximately 0.1%, this first estimation would change to 70–200 deaminations per cell per day (Kavli *et al.*, 2007).

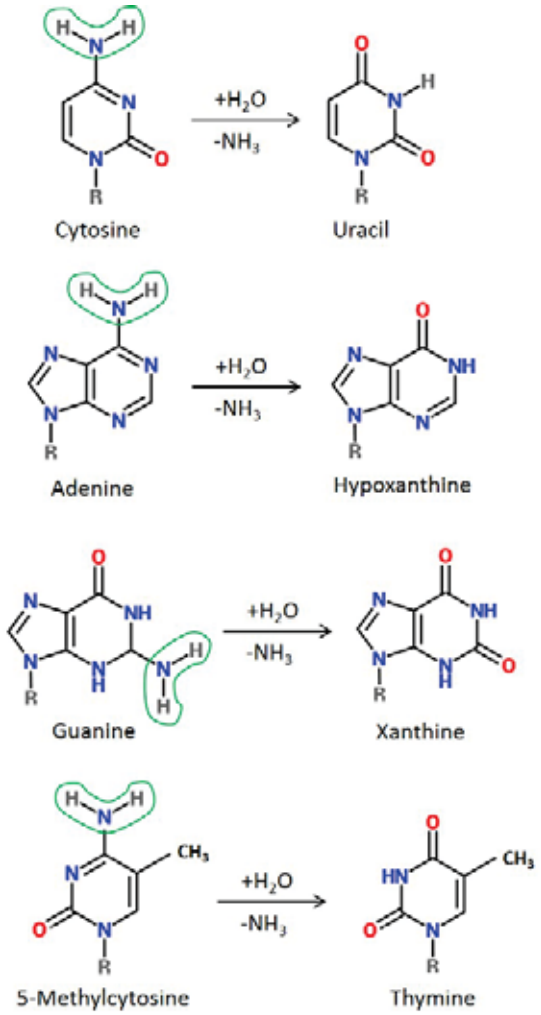

Figure 1.4 Products formed from deamination of bases in DNA. The amine group (-NH_2, circled in green) is hydrolized and replaced by the oxygen from the water molecule. Adapted from Friedberg *et al.*, 2005.

1.1.2 Oxidative damage

Due to reactive oxygen species (ROS) formed as byproducts in the electron transport chain of bacterial plasma membranes as well as mitochondria, chloroplasts and some other organelles of eukaryotes, oxidative damage to DNA is substantial in all aerobic cells. ROS are also formed by photosensitization reactions involving both ultraviolet and visible light. The most reactive ROS is the hydroxyl radical (·OH), which is formed from hydrogen peroxide (H_2O_2) and Fe^{2+} in the Fenton reaction as well as by ionizing and UVB and UVC radiations. Another important reactive but non-radical form of oxygen is singlet oxygen (Gutteridge and Halliwell, 1989).

ROS cause oxidative damage to DNA including base modifications (Figure 1.5) and strand breaks, which is believed to be of similar importance as hydrolytic damage. The most studied ROS-induced base modification is the mutagenic oxo^{8}G, which mispairs with adenine. Indeed, oxo^{8}G is easier oxidized than guanine itself giving rise to secondary oxidation products in DNA (Klungland and Bjelland, 2007). Many oxidized bases – *e.g.* the thymine lesions 5-hydroxy-5,6-dihydrothymine (Th^5), 6-hydroxy-5,6-dihydrothymine (Th^6) and 5-hydroxy-5-methylhydantoin (hmh) – are cytotoxic rather than mutagenic by inflicting an effective block to DNA replication and/or transcription (Bjelland and Seeberg, 2003).

Figure 1.5 Some important DNA base damages produced by ROS. oxo^{8}G incorporates an oxygen in position 8 of a guanine (so it is also known as 8-oxoguanine). Thymine can incorporate hydroxyl groups in *cis* (as shown in the figure) or *trans* form, resulting in cTg or tTg. hm^{5}U and f^{5}U (Figure 1.3) have a common precursor in 5-hydroperoxymethyluracil (hpm^{5}U).

5-Formyluracil

5-Formyluracil (f^5U) is formed by oxidation of the 5-methyl group of thymine (Figure 1.3C) and seems to be present in DNA at similar concentrations as oxo^8G (Bjelland *et al.*, 2001). It was reported to be formed in DNA 25 years ago (Kasai *et al.*, 1990).

The most common form of f^5U is the one with a keto group in the 5^{th} position. The enol form arises when the keto group is reduced to hydroxyl. Moreover, at basic pH the keto group adopts the anionic form. These f^5U forms (Figure 1.6) have different base-pairing abilities. The keto form ($f^5U^{C=O}$) does not obliterate the Watson-Crick pairing ability, so pairing with adenine is preferred. But the formyl group alters the charge distribution of the pyrimidine ring as well as it may directly participate in hydrogen bond formation causing *e.g.* mispairing with cytosine (Figure 1.7). The abundance of the anionic form ($f^5U^{\ominus}$) as opposed to the preferred keto form increases with pH. The amount of $f^5U^{\ominus}$ and thus the $f^5U^{\ominus}\cdot G$ mispair is significant at neutral pH. The $f^5U^{C=O}\cdot A$ and $f^5U^{\ominus}\cdot G$ pairs have been demonstrated by X-ray analyses, in contrast to the $f^5U^{C=O}\cdot C$ mispair which is tentative (Kamiya *et al.*, 2002).

The f^5U deoxynucleoside (5-formyl-2'-deoxyuridine, f^5dU) can be used to induce mutations in cells (Kamiya *et al.*, 2002; Kasai *et al.*, 1990; Yoshida *et al.*, 1997). f^5dU added to culture medium of cells is converted *in vivo* to its 5'-triphosphate derivate (f^5dUTP) which is used by cellular polymerases to incorporate fdUMP into the DNA. It has been observed in recent studies that the most common mutations induced by f^5dU in *E. coli* (Ånensen *et al.*, 2001) are different from the ones arising in mammalian cells (Kamiya *et al.*, 2002).

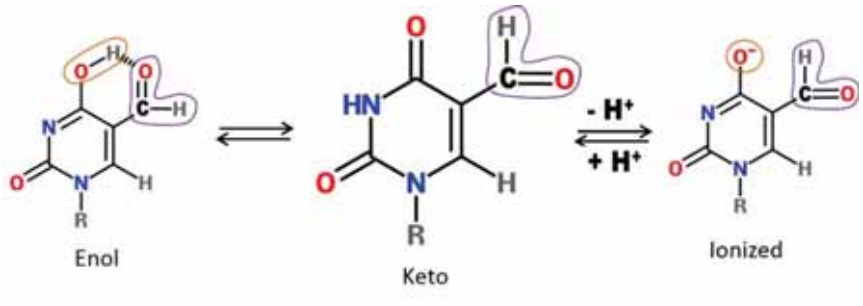

Figure 1.6 Chemistry of f^{5}U in DNA. Oxidative attack on the 5-methyl group of thymine converts it to a formyl group (circled in purple). Other forms of modifications in the 4th position are circled in orange. Adapted from Knævelsrud *et al.*, 2009.

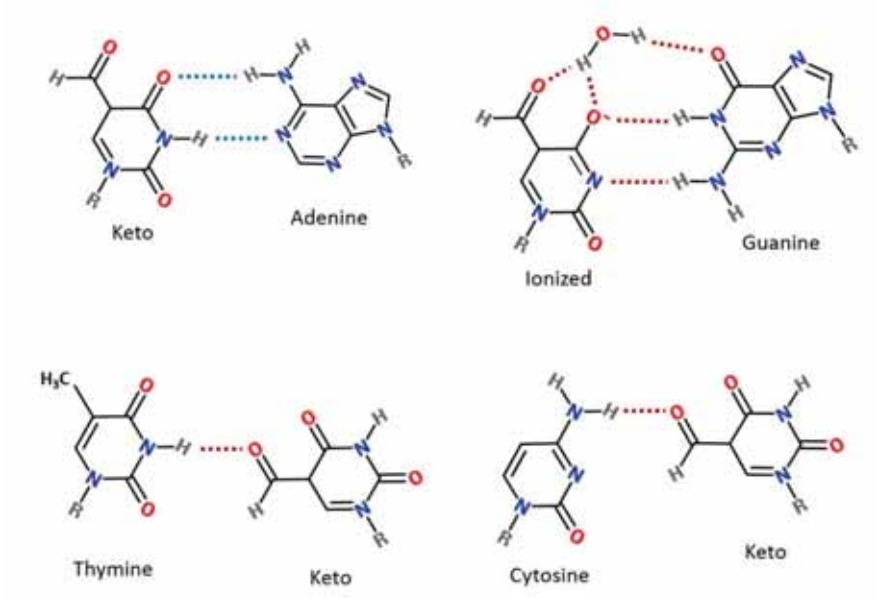

Figure 1.7 Base-pairing abilities of f^{5}U in DNA. The different forms of f^{5}U have different base-pairing abilities: cognate or Watson-Crick hydrogen bonding pattern is shown in blue dotted lines; non-cognate patterns of mispairs are shown in red dotted lines. Adapted from Kamiya *et al.*, 2002; Knævelsrud *et al.*, 2009.

1.2 DNA repair and response mechanisms

When DNA suffers a damage, this can be repaired or tolerated, or start an apoptotic response in the worst-case scenario for the cell. There are several biochemical mechanisms for DNA repair, and the enzymes involved may take part in more than one.

Sometimes the damage itself can just be removed, what restores the DNA; this is called direct reversal. When this is not possible, the damage is removed together with the damaged base/nucleotide or with a whole fragment of DNA. The cell may also show tolerance to the damage by employing specialized translesion bypass polymerases that insert a correct or an incorrect base across the lesion. Thus, cell death may be avoided at the expense of some mutation induction.

Table 1.1 lists the known biological responses to DNA damage. The main ones are explained in detail later in this text.

Table 1.1 Biological responses to DNA damage (Table 1-1 from Friedberg *et al.*, 2005)

Reversal of base damage
Excision of damaged, mispaired or incorrect bases
Base excision repair (BER)
Nucleotide excision repair (NER)
Transcription-coupled nucleotide excision repair (TC-NER)
Alternative excision repair (AER)
Mismatch repair (MMR)
Strand break repair
Single-strand break repair (SSBR)
Double-strand break repair (DSBR)
Tolerance of base damage
Translesion DNA synthesis (TLS)
Postreplicative gap filling
Replication fork progression
Cell cycle checkpoint activation
Apoptosis

The majority of the DNA damages in cells are repaired by the excision repair mechanisms: base excision repair (BER) and nucleotide excision repair (NER). These two pathways share most of the enzymes involved in insertion and ligation of the correct nucleotide, but they differ in the way they eliminate – excise – the damaged or incorrect base (Figures 1.8 and 1.10). The BER pathway employs somewhat specialized glycosylases for the excision of aberrant bases. Uracil-DNA glycosylases are known to recognize uracil in

DNA and excise it, starting a chain of downstream reactions able to convert the damaged strand to its original state. There are several enzymes reported until now as U-glycosylases (Liu *et al.*, 2003), making BER the most suitable mechanism for the repair of U and U-like damages, among other base lesions. The NER pathway is an important mechanism in the repair of "bigger" damages that distort the conformation of the DNA double helix making difficult or impossible normal copying and transcription to mRNA (Petit and Sancar, 1999). However, *in vitro* studies (Kino *et al.*, 2004) suggest the implication of NER in the repair of f^5U in DNA.

1.2.1 Base excision repair

Spontaneously arisen DNA base damages are primarily repaired by BER. Both prokaryotes and eukaryotes show similar pathways with slight differences (Figure 1.8). The mechanism starts with the recognition of the damaged base by a DNA glycosylase followed by cleavage of the N-glycosyl bond (Figure 1.1) between the base and the deoxyribose, leaving an AP site. There exist glycosylases specialized in many types of base damages or mispairings. These relatively small enzymes (~30–50 kDa) are monomeric and do not require cofactors. They seem to act by flipping-out the base 180º followed by N-glycosyl bond cleavage (Slupphaug *et al.*, 1996). In the case of bi-functional glycosylases, they exhibit in addition an AP lyase activity (O'Connor and Laval, 1989) which cleaves the phosphodiester backbone at the 3'-end of the AP site by a β- or β/δ-elimination reaction (Figure 1.9), leaving behind the AP site remnant as a (saturated/unsaturated) 3'-dRP or 3'-phosphate, respectively. The latter can be removed by a polynucleotide kinase phosphatase (PNKP) (Friedberg *et al.*, 2005).

In the short-patch BER, the DNA backbone is cleaved either by an AP endonuclease at the 5'-side or by an AP lyase – by bi-functional glycosylases – at the 3'-end of the AP site. The AP site remnant is removed afterwards by a 5'-deoxyribose phosphate (5'-dRP) lyase or a 3'-phosphodiesterase, respectively. The gap is filled by a polymerase with the correct nucleotide followed by a ligase, which seals the nick and completes the process. In the long-patch BER, the free 3'-OH group formed by incision of the AP site by an AP endonuclease is targeted by a DNA polymerase, which extends a displaced DNA strand (a flap) that is removed by a flap endonuclease (Figure 1.8). Although the molecular basis for the selection of one or the other path

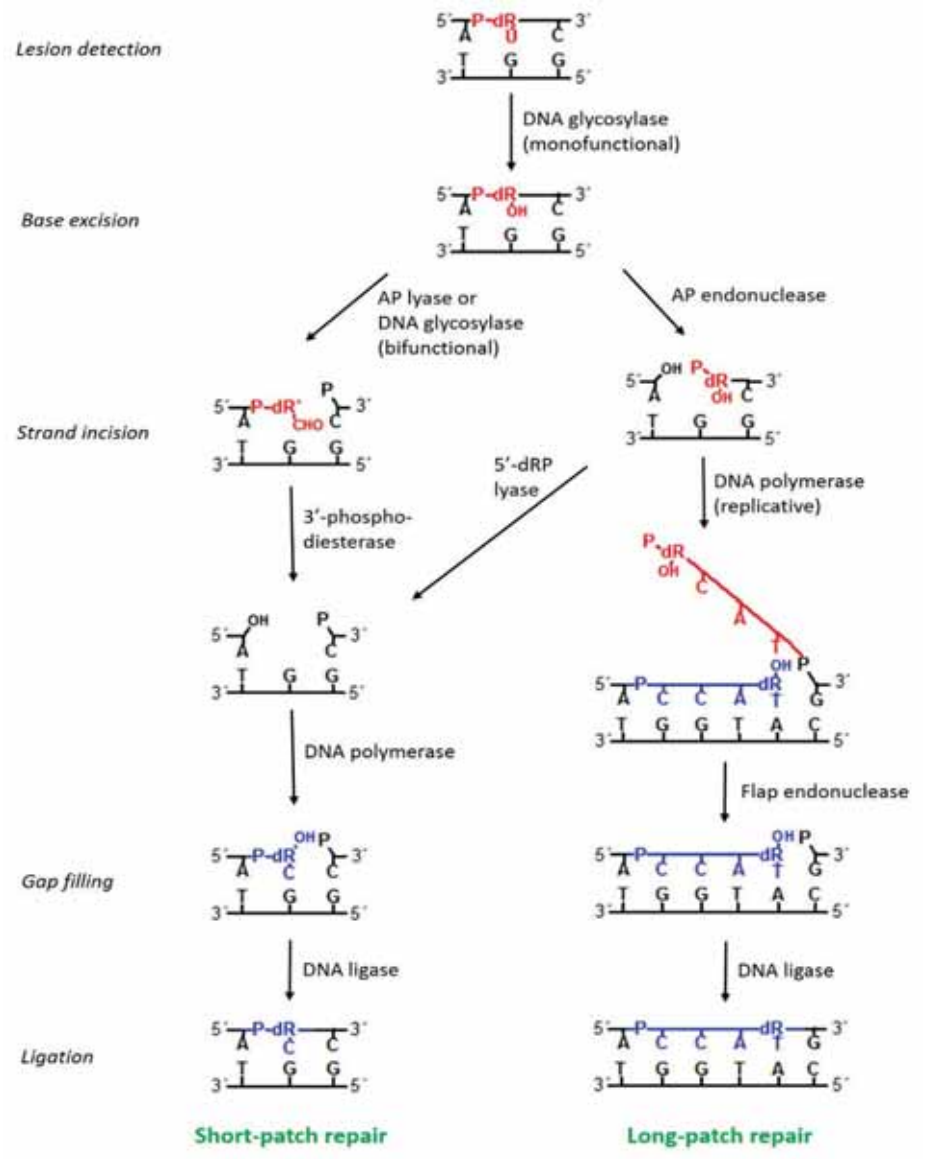

Figure 1.8 Diagram of general BER. The BER pathway follows he same steps in prokaryotes and eukaryotes, with different enzymes and cofactors involved. Adapted from Moen *et al*., 2011; Xu *et al*., 2008.

remains unclear, the presence or lack of some polymerases and cofactors seems to influence the final decision (Sander and Wilson, 2001).

Because of its abundance, the AP site is the most common substrate for BER (Kim and Wilson, 2012). In *E. coli*, the AP sites are incised by AP endonucleases such exonuclease III (Xth) and endonuclease IV (Nfo) (Hang *et al.*, 1998). In humans, this is a role for hAPE1 (human apurinic/apyrimidinic endonuclease 1) (Friedberg *et al.*, 2005).

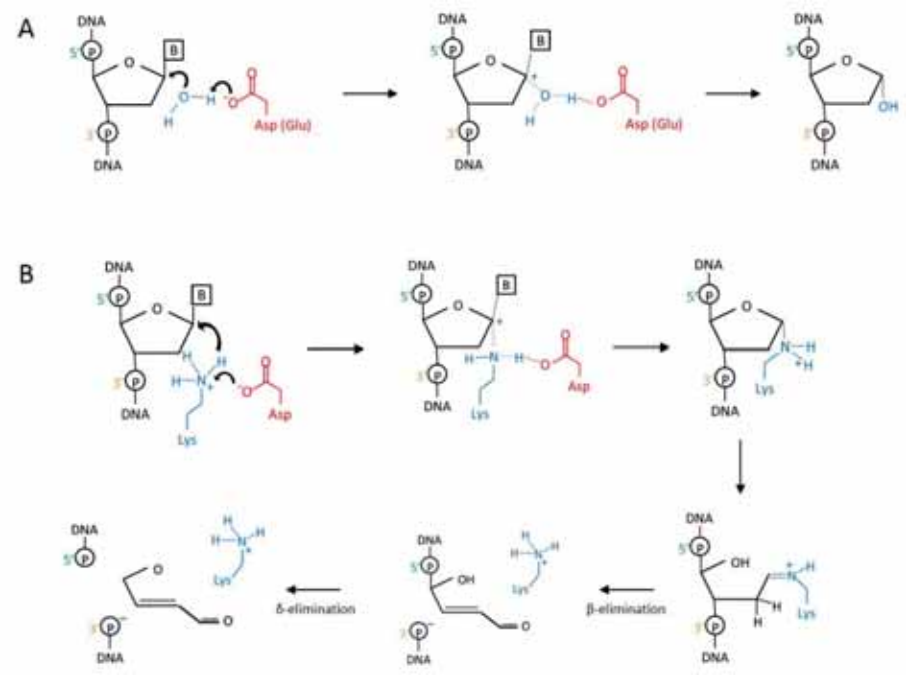

Figure 1.9 Base-damage repair mechanism by mono- and bi-functional DNA glycosylases. A) Mono-functional: a water molecule (blue) in the active site (red) is deprotonated by carboxylic acid side chain of the enzyme and positioned for nucleophilic attack at the anomeric position. The product of the reaction is an abasic site. B) Bi-functional: the nucleophile in this case is an amine group in a lysine residue of the enzyme (blue). The product of glycosidic bond cleavage is rearranged (Schiff base intermediate). A β-elimination leads to incision of the phosphodiester bond between the abasic site and the phosphate in 3' (Figure 1.1). Some enzymes also have δ-elimination function, incising the phosphate in 5' (5'-dRP lyase). Adapted from Drohat and Maiti, 2014; Schärer and Jiricny, 2001.

Uracil-DNA glycosylases

Uracil-DNA glycosylase (UDG) was the first DNA glycosylase activity to be reported 40 years ago (Lindahl, 1974). Since then, UDGs have been grouped into several families distinguished by primary structure and all members of this UDG superfamily have exclusively been described as evolutionary conserved (Schormann *et al.*, 2014) mono-functional enzymes. The most important UDGs in *E. coli* and human cells are displayed in Table 1.2.

Human uracil-DNA glycosylase (hUNG) is the human homologue for *E. coli* uracil-DNA glycosylase (EcUng) and has been the most studied UDG. It occurs in alternative sliced forms from the same gene – *ung* – designated *UNG1* for distribution in the nuclei or *UNG2* in mitochondria (Krokan *et al.*, 2001). The hUNG active site consists of a deep uracil-binding pocket with an overlying groove that binds one DNA strand, but the entire DNA-binding surface on hUNG extends considerably beyond the immediate active pocket. The non-specific DNA-binding surfaces may aid local uracil search, contribute to binding abasic DNA product and help present the DNA product to hAPE1 (Roberts *et al.*, 2012).

DNA glycosylases for lesions induced by oxidation

A multitude of different nucleic acid base modifications are induced in DNA by ROS (Bjelland and Seeberg, 2003). Many of these, including ring-saturated and ring-fragmented pyrimidines, are believed to be mainly cytotoxic lesions contributing only slightly to mutation induction. In contrast, oxo^{8}G is considered a principal lesion responsible for the mutagenic events induced by ROS, where several repair proteins counteracting its damaging effects have been extensively characterized (Friedberg *et al.*, 2005; Lindahl, 1993). Also the cytosine oxidation products 5-hydroxycytosine (h^{5}C), h^{5}U and uracil glycol (Ug) have been shown to cause mispairing contributing to mutation induction (Bjelland and Seeberg, 2003).

In *E. coli*, most oxidized base damages are removed by three enzymes: formamidopyrimidine-DNA glycosylase (Fpg) – the main enzyme for the repair of oxo^{8}G residues (Serre *et al.*, 2002) – endonuclease III (Nth) and endonuclease VIII (Nei). They exhibit somewhat overlapping specificities (Prakash *et al.*, 2012). However, the oxidized thymine residues hm^{5}U and f^{5}U are most efficiently repaired by the 3-methyladenine-DNA glycosylase II

(AlkA) (Bjelland *et al.*, 1994). Most of these glycosylases have analogues in mammalian cells (Table 1.2). The former enzymes are bi-functional and were described initially by their AP lyase activity, but they are able to recognize different oxidized substrates and, after nicking the N-glycosyl bond, they cleave the ribose by β-elimination rather than hydrolysis (Bailly and Verly, 1987).

5-Formyluracil-DNA glycosylases

f^5U has a potential mutagenic or cytotoxic effect in bacteria and may be even more mutagenic in mammalian cells (Klungland *et al.*, 2001). Several studies, both *in vitro* and *in vivo*, have demonstrated f^5U-mediated mutation induction (Ånensen *et al.*, 2001; Fujikawa *et al.*, 1998; Terato *et al.*, 1999; Yoshida *et al.*, 1997; Zhang *et al.*, 1997). It has been observed that f^5U in DNA is removed efficiently by the AlkA DNA glycosylase of *E. coli* as well as by DNA glycosylase activities present in human, murine and rat cell-free extracts (Bjelland *et al.*, 1994; Bjelland *et al.*, 1995; Zhang *et al.*, 1995). Efficient removal of f^5U from DNA was demonstrated and confirmed *in vitro* (Masaoka *et al.*, 1999).

The AlkA enzyme is a 31-kDa polypeptide containing 282 amino acids (Clarke *et al.*, 1984; Nakabeppu *et al.*, 1984). It has a broad substrate specificity, being able to recognize and catalyze the excision of several methylated bases (Table 1.2). AlkA takes part in the adaptive response to alkylation damage, since adaptation to survival in *E. coli* is associated with the induction of AlkA synthesis. The adaptive response to alkylation is regulated by the 6-oxoalkylguanine transferase I (Ada) protein. Mutations in the *ada* gene prevent the induction of AlkA (Friedberg *et al.*, 2005). However, there is no evidence of *alkA* induction by ROS exposure.

3-Methyladenine-DNA glycosylase (MPG) is the enzyme responsible for initiating repair of alkylated bases in human cells, however, exhibits no activity for f^5U. The human 8-oxoguanine-DNA N-glycosylase 1 (hOGG1) – one of the major mammalian glycosylases for oxidized bases – do not seem to recognize f^5U (Masaoka *et al.*, 1999). This contrasts with the human endonuclease III (hNTH1), which shows f^5U activity *in vitro* (Miyabe *et al.*, 2002). However, its physiological importance in f^5U repair has been questioned since it needs large amount of substrate and does not seem kinetically efficient (Matsubara *et al.*, 2003). The main mammalian f^5U-DNA glycosylase activity is believed to be a function of the single-strand selective mono-functional uracil-DNA

Table 1.2 DNA glycosylases for uracil and lesions induced by oxidation. UDGs in *E. coli* and their equivalents in human cells for the lesions mentioned in this thesis. Adapted from Bjelland and Seeberg, 2003, Friedberg *et al.*, 2005; Kim and Wilson, 2012; Krokan *et al.*, 1997; Masaoka et al., 2003; Serre *et al.*, 2002; Takao *et al.*, 2002; Wallace, 2013.

E. coli					Human cells				
Protein	**Gene**	**Common name**	**Substrates**	**AP lyase**	**Protein**	**Gene**	**Common name**	**Substrates**	**AP lyase**
EcUng	*ung*	Uracil-DNA glycosylase	U, h^5U:A, h^5U:G	No	hUNG1/ UNG2	*UNG*	Uracil-DNA glycosylase	U, U:G, U:A, fU, h^5U:G	No
					hSMUG1	*SMUG1*	Single-strand-specific mono-functional uracil-DNA glycosylase	U, U:A, U:G, hm^5U, fU, h^5U	No
Mug	*mug*	Mismatch-specific uracil-DNA glycosylase	U, T, hm^5U:A, hm^5U:G, fU:A, fU:G, h^5C:G, h^5U:G	No	TDG	*TDG*	Thymine-DNA glycosylase	U:G, T:G, hm^5U:G, hm^5U:A, fU:G, fU:A	No
					MBD4	*MBD4*	Methyl-binding domain glycosylase 4	U, T:G, hm^5U:G, fU:G, U or T in U/TpG-m^5CpG	Yes
AlkA	*alkA*	3-Methyladenine-DNA glycosylase II	m^3A, m^3G, m^7G, hm^5U:A, fU:A, fU:G	No	MPG	*MPG*	3-Methyladenine-DNA glycosylase	m^3A, m^7A, m^3G, m^7G, oxo^{8}G:C	No
MutY		Adenine-DNA glycosylase	A:oxo^{8}G, A:oxo^{8}A, h^2A:G, A:fapyG	No	MUTYH/ MYH	*MUTYH /MTH*	MutY homolog	A:G, A:oxo^{8}G: G:oxo^{8}G, h^2A	Yes/No
Fpg/ MutM	*fpg*	Formamidopyrimidine-DNA glycosylase	oxo^{8}G, oxo^{8}A:C, oxo^{8}A:T, Tg:A, Ug:G, Ug:A, h^5C:A, h^5C:G, h^5U:A, h^5U:G, fapyG:C, fapyG:A	Yes	hOGG1	*OGG1*	8-Oxoguanine glycosylase	oxo^{8}G, oxo^{8}A:C	Yes
Nth	*nth*	Endonuclease III	Tg, Th^5:A, Th^5:A, hmh:A, hm^5U:G, hm^5U:A, fU:A	Yes	hNTH1	*NTH*	Human endonuclease III	fU, Tg:A, Th^5:A, h^5U, h^5C:G, h^5C:A	Yes
Nei	*nei*	Endonuclease VIII	Tg, Th^5:A, hmh:A, hm^5U:A, hm^5U:G, fU:A, Ug:A, Ug:G, h^5C:A, h^5C:G, h^5U:A, h^5U:G, oxo^{8}G, oxo^{8}A	Yes	NEIL1/ NEIL2/ NEIL3	*NEIL*	Endonuclease VIII-like 1,2,3	Tg, h^5U, h^5C:G, oxo^{8}G, fapyG	Yes

glycosylase (hSMUG1) (Masaoka *et al.*, 2003; Matsubara *et al.*, 2003), which is a backup enzyme for hUNG and the primary repair enzyme for a subset of oxidized pyrimidines such as hm^5U and h^5U in addition to f^5U (Masaoka *et al.*, 2003).

1.2.2 Nucleotide excision repair

Some DNA lesions distort the double-helix structure significantly, which is the qualification for being a substrate for NER. Because of this, common NER substrates are cross-linked base residues caused by UV radiation, like cyclobutane pyrimidine dimers (CPDs) – including thymine dimers (Figure 1.3A) – and 6–4PPs, as well as bulky chemical species, originating from *e.g.* tar covalently attached to bases which often are guanine. An extensive review of NER substrates is described in Table 1 from Truglio *et al.* (2006).

In *E. coli* and other prokaryotes, NER is carried out by the UvrABC system, which is composed of the three proteins UvrA, UvrB and UvrC. The genes for two of these proteins – *uvrA* and *uvrB* – are controlled by the SOS response (Figure 1.11) (Janion, 2001). The *uvrC* gene, however, is not such regulated (Van Houten, 1990). The NER pathway starts with the formation of a complex consistent in one UvrB protein and two UvrA ($UvrA_2B$ complex), which docks the DNA some distance from the damage. This complex scans the DNA using its helicase activity. When the damage is encountered, $UvrA_2B$ forms an unstable complex with the DNA. Using ATP, UvrB forms a stable complex with the damaged strand and UvrA dissociates. UvrC binds to the UvrB-DNA complex and activates UvrB, which makes an incision in the damaged strand four nucleotides 3' from the lesion. UvrC is then activated and incises the damaged strand seven nucleotides 5' a few seconds after. DNA helicase II (UvrD) releases the incised fragment and UvrC. UvrB remains bound to the gapped DNA and is released prior to the new DNA synthesis performed by polymerase I. DNA ligation completes the repair (Figure 1.10) (Friedberg *et al.*, 2005; Petit and Sancar, 1999).

The endonuclease mode characteristic for NER is well conserved through evolution from prokaryotes to human cells. An important difference is, however, that the mammalian system is much more complex. Only three proteins are employed by *E. coli* in contrast with 16/17 polypeptides by human cells, among other differences (Petit and Sancar, 1999).

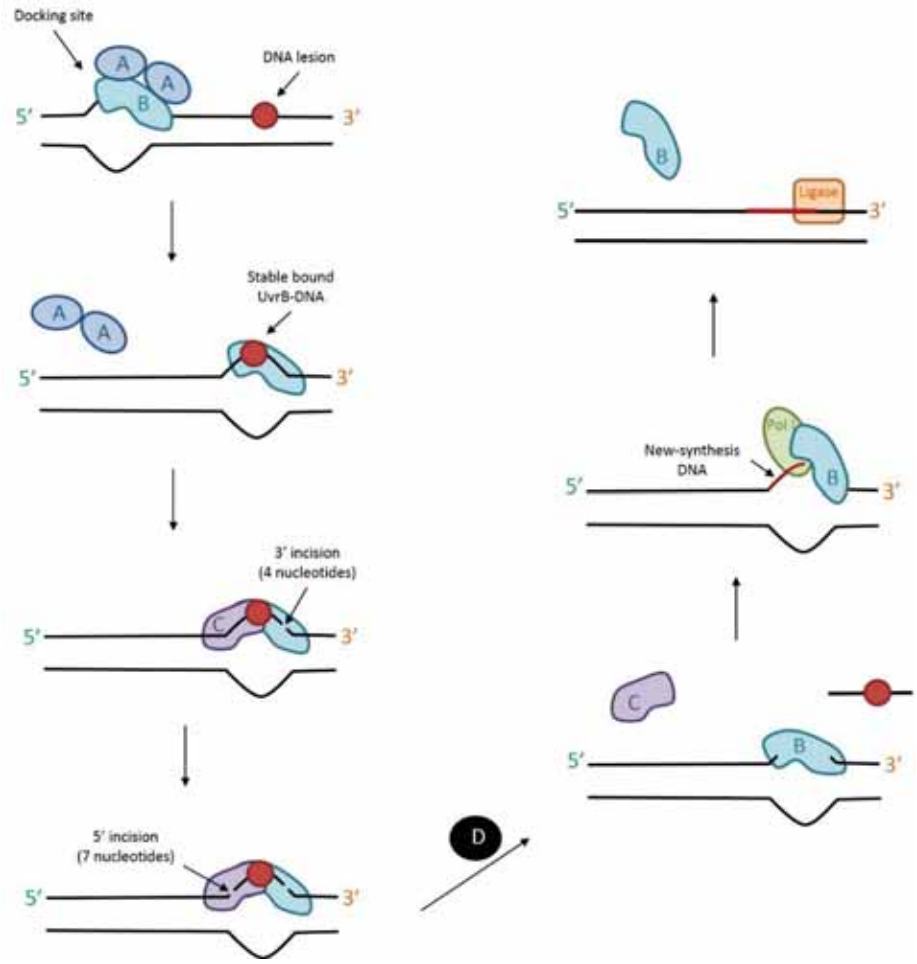

Figure 1.10 UvrABC system of nucleotide excision repair (NER). Mechanism in *E. coli*. The two molecules of UvrA bind one of UvrB in solution using ATP. ATP is also required for the stable binding between UvrB and the damaged DNA strand. UvrD (helicase II) is needed to remove UvrC and the incised fragment. Downstream enzymes – polymerase and ligase – repair the damage in the same way as in BER. Adapted from Friedberg *et al.*, 2005; Petit and Sancar, 1999.

The SOS response

The *E. coli* SOS response is a global response to DNA damage in which the cell cycle is arrested and DNA repair and translesion synthesis/mutagenesis are induced. It was described for the first time in 1975 (Radman, 1975). The repressor protein LexA binds to the operator regions of SOS genes – the SOS box – and negatively regulates their transcription. Treatment of the cells with DNA-damaging agents causes increased inhibition of DNA synthesis and thus increased ssDNA regions where RecA protein (Rad51 in eukaryotes) binds, which induces LexA to cleave itself resulting in the expression of the SOS genes (Figure 1.11). Under high load of oxidative stress, genes with a weak SOS box are fully induced, and those include *uvrA*, *uvrB* and *uvrD* (Crowley and Hanawalt, 1998), activating the NER repair pathway. The increased rate of mutation during the SOS response is caused by three low-fidelity DNA polymerases with stronger SOS boxes, induced when NER does not suffice: Pol II, Pol IV and Pol V (Saha *et al.*, 2007). Pol IV is involved in bypass replication

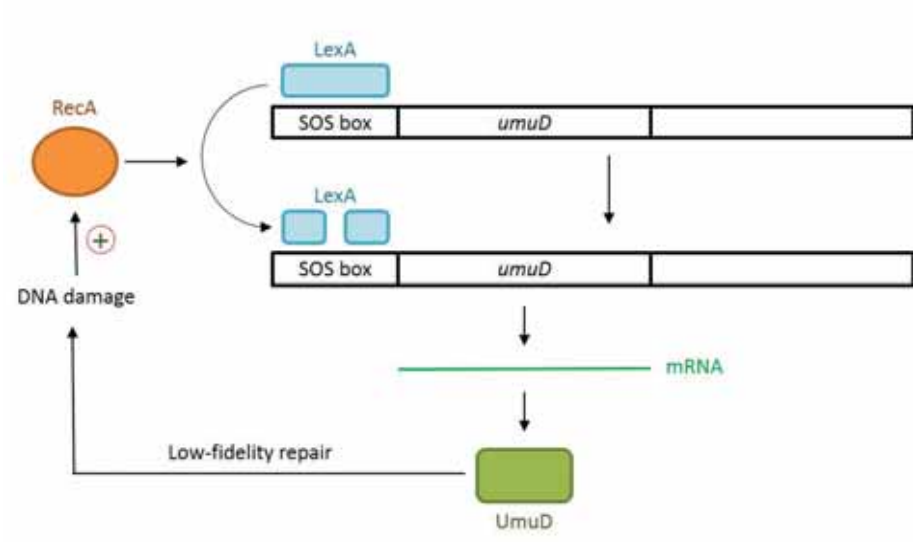

Figure 1.11 Regulation of SOS response in *E. coli* gene of the UmuD protein. DNA damage activates RecA protein, which induces LexA to degrade at its Ala84–Gly85 bond. The operon (SOS box) is released and the *umuD* gene expresses into the UmuD protein, a subunit of the DNA polymerase V, low-fidelity polymerase that repairs the damaged DNA introducing replication errors. Adapted from Friedberg *et al.*, 2005.

of damaged bases generated by ROS, and seems to play an important role in misincorporation of dCTP opposite 1,2-dihydro-2-oxoadenine (oxo^2A), leading to G·C → A·T mutations. Pol IV prefers to insert dATP opposite f^5dU, not leading to any mispairing as know until know (Hori *et al.*, 2010).

Nucleotide excision repair of 5-formyluracil

As previously mentioned, f^5U is repaired by the BER pathway, which in *E. coli* is initiated by damaged base removal by the AlkA, Fpg, Nth or Nei DNA glycosylases. In addition, the mammalian NER system has been reported to exhibit activity for repair of f^5U *in vitro* – through the XPC-HR23B complex (Kino *et al.*, 2004) – while no such evidence exists for bacterial NER (Van Houten *et al.*, 2005). Transcription coupled NER (TC-NER) do not require XPC (Le May *et al.*, 2010) indicating that f^5U seems to be interacting with the global genome NER (GG-NER).

Previous studies in *E. coli* have shown that 5-fluorouracil (F^5U) – anti-metabolite employed clinically to manage solid tumors (Wyatt and Wilson, 2009) – induces the SOS response via expression of the *umuC* gene in the wild-type and *uvrA* strains, but not in *lexA* and *recA* strains (Oda, 1987). Moreover, the same glycosylases seem to be involved in the repair of F^5U and f^5U, both in *E. coli* and humans (Liu *et al.*, 2003; Wyatt and Wilson, 2009).

2 Aim of the studies

During enzymatic characterization of UDGs using 5'-labelled oligonucleotides with uracil incorporated at a specific site, we surprisingly discovered that hUNG and hSMUG1 glycosylases might incise the DNA strand at the uracil site following base removal. Thus, the task was to verify this activity and characterize the incision products chemically (Paper 1).

The main BER glycosylases involved in the removal of virtually all the base damages formed by ROS are AlkA, Fpg, Nth and Nei. Since, to our knowledge, the consequences on mutagenesis if knocking out these enzymes have yet to be investigated, we decided to study the spontaneously arisen base substitutions in the quadruple mutant *alkA fpg nth nei* (Paper 2).

Our research group has previously shown that f^5U is efficiently removed from DNA *in vitro* by the *E. coli* AlkA DNA glycosylase, which later has been confirmed by others. To possibly provide *in vivo* evidence for this enzyme function, experiments using different concentrations of f^5dU were performed on *E. coli* wild-type and bacteria deficient in the *alkA* gene (Paper 3).

Since *in vitro* experiments suggest a role for mammalian NER in f^5U repair (Kino *et al.*, 2004), we decided to investigate whether this also was the case in *E. coli*, a much simpler system to compare the possible roles or relationships between BER and NER in f^5U repair. First, we wanted to possibly establish an *in vivo* effect regarding mutagenesis using different concentrations of f^5dU in *E. coli* by monitoring the consequences of knocking out each of the three genes coding for the UvrABC proteins system. The study was initiated by determining

the mutation rate and base substitutions induced by f^5dU in the *uvrA* mutant (Paper 4).

3 Overview of results

3.1 DNA uracil excision includes strand incision by hUNG (Paper 1)

During the performance of DNA incision assays with different uracil-DNA glycosylases (UDGs), we observed that the major UDGs in human cells (hUNG) and in *E. coli* (EcUng) – both family 1 UDGs – were able to incise the lesion site after base excision. hUNG was not able to recognize and incise AP-containing DNA, showing that it needed uracil for damage recognition.

Gel migration analyses using certain DNA glycosylases/AP lyases and AP endonucleases to define different 3'-end products indicated that the 3'-end was identical to a 3'-α,β-unsaturated aldehyde, which we decided to designate uracil-DNA incision product (UIP). The chemical nature of this new UDG-endonuclease activity was also investigated using MALDI-TOF-MS (matrix-assisted laser desorption/ionization time of flight mass spectrometry) for product analysis, confirming that the incision resulted in a 3'-α,β-unsaturated aldehyde as well as a 5'-phosphate (Paper 1, Fig. 3a). Enzyme reactions were performed in the presence $H_2{}^{16}O$ and $H_2{}^{18}O$ followed by subsequent sample preparation confirmed the unsaturated nature of the 3'-remnant by showing post-enzymatic quantitative addition of either water or ammonia to the double bond of the 3'-α,β-unsaturated aldehyde. These experiments were consistent with both a β-elimination and a hydrolytic reaction mechanism. However, when the UDG–DNA reaction intermediate could not be trapped as a stable covalent complex it points to the latter mechanism. hUNG also lacks a lysine residue in the active site to carry out a β/δ-elimination reaction, which indicates a hydrolytic mechanism. Thus, we proposed that a catalytic or activated H_2O molecule initiates a β-elimination reaction at the 2'C–3'C bond by acting as a general base and attracting a proton from the deoxyribose 2'C. The electron-

withdrawing power of the deoxyribose 1'C formyl group should greatly facilitate such a reaction, which in addition might be stabilized by His268 in the hUNG active site (Paper 1, Fig. 3a). This scheme accords well with crystal structures of hUNG in complex with damaged DNA, showing both accessibility of water molecules, formyl group stabilization (Paper 1, Fig. 3c) and reactive distance between a reactive H_2O and the 2'C atom of the flipped-out abasic nucleoside (Paper 1, Fig. 3c).

3.2 Severe repair-deficiency for oxidized DNA bases causes primarily GC to AT transitions in *E. coli* (Paper 2)

Inactivation of the main BER glycosylase genes (*alkA*, *fpg*, *nth* and *nei*) involved in repairing lesions induced by ROS in *E. coli* caused a ten-fold reduction in the bacterial growth. The mutation rate was three times higher as measured by scoring for rifampicin resistance (rif^R) (Paper 2, Table 1). This agrees with the described cytotoxicity of some ROS-induced lesions (Bjelland and Seeberg, 2003).

The abundance of the six types of base substitutions detected by sequencing the rif^R region of mutants obtained from wild-type and *alkA fpg nth nei* varied significantly between the former and latter. In wild-type, the most common mutation was the G·C → A·T transition accounting for almost 40% whilst the G·C → C·G transversion as the most infrequent one accounted for roughly 4%, of the total base substitutions recorded. Interestingly, in the quadruple *alkA fpg nth nei* mutant more than 90% of the mutations were G·C → A·T transitions while the rest were A·T → G·C transitions (Paper 2, Table 1).

Two thirds of the G·C → A·T transitions recorded in *alkA fpg nth nei* occurred in two hot spots (1576 and 1592) shared with the wild-type; most of the rest occurred in two other hot spots (1546 and 1586). Virtually all A·T → G·C transitions recorded in *alkA fpg nth nei* occurred at the site 1547, which also was a common site for such mutations in wild-type (Paper 2, Figure 2). Taken together with known mispairing abilities of oxidized G and C (Bjelland and Seeberg, 2003), our results suggest that cytosine oxidation seems more important than oxidation of other bases including guanine in spontaneous mutagenesis.

3.3 Damage-specific mutation induction promoted by repair (Paper 3)

When f^5dU at different concentrations (0.1–0.4 mM) was added to exponentially growing *E. coli* – wild-type, *alkA* and *xth nfo* – the relative growth decreased only moderately and never fell below 50%. This agrees with the character of f^5U as a mutagenic rather than cytotoxic lesion (Bjelland and Seeberg, 2003). However, since *alkA* seems slightly more affected than wild-type, a phenotypic difference may exist.

The spontaneous mutation rates (scored using the rif^R system as above) were similar for wild-type, *alkA* and *xth nfo*, although slightly lower for *xth nfo*. The addition of 0.1 mM f^5dU increased the mutation rate of all three to roughly the double, demonstrating the mutagenic capacity of f^5dU. However, higher concentrations of f^5dU (0.2 and 0.4 mM) decreased the mutation rates. In wild-type this decrease was moderate while the decrease in mutation rate for *alkA* was >50% when the f^5dU concentration added was increased from 0.2 to 0.4 mM, which fell even below the spontaneous level (Paper 3, Table I). Thus, the presence of AlkA (in wild-type) seems to prevent an abrupt decrease in f^5dU-induced mutagenesis at the highest concentration of f^5dU studied.

The A·T → G·C transition was clearly the most predominant base substitution formed in repair-proficient *E. coli* at the lowest concentration of f^5dU (0.1 mM). In addition to its large increase in abundance by f^5dU supplementation (Paper 3, Figure 3A), its connection with the f^5dU damage *per se* was indicated by its lower formation in *xth nfo* (Paper 3, Figure 3C), arguing against un-specific formation via AP sites. The increased A·T → G·C formation in *alkA* at 0.2 mM f^5dU (Paper 3, Figure 3B), which theoretically should be connected to poorer f^5U repair (Bjelland *et al.*, 1994), supports this notion. The A·T → G·C transition is challenged by G·C→ A·T at higher (0.2 and 0.4 mM) f^5dU concentrations (Paper 3, Figure 3A). Evidence that the latter is induced by f^5dU is indicated by its higher rate in *alkA* than in wild-type at 0.1 mM f^5dU (Paper 3, Figure 3B). The G·C→ A·T transition is a typical A-rule mutation that can be formed via AP sites, which is underscored by its consistent higher rates in *xth nfo* (Paper 3, Figure 3C). The formation of G·C→ T·A transversions is favored at higher f^5dU concentrations, and evidence that it is formed as a consequence of the A-rule at the higher rate in *xth nfo* compared to wild-type and *alkA* at 0.2 mM f^5dU (Paper 3, Figure 3A,C). However, the

similar rate of G·C→ T·A at 0.1 mM f^5dU under the different repair conditions (Paper 3, Figure 3) may indicate that translesion of AP sites is not the only mechanism for G·C→ T·A formation. Thus, Ånensen *et al.* (2001) reports all three mutations induced in *E. coli* by f^5dU using another mutation detection system.

Based on the above results we can conclude that the presence or absence of AlkA does not seem to be straightforward, and may have different and even opposite consequences on the f^5dU-mediated induction of the most prominent substitutions at different mutagen concentrations (Paper 3, Figure 4). The most striking result applies to the G·C→ A·T transition, which rate decreases notably in the wild-type compared to *alkA* at 0.1 mM f^5dU but opposite increases dramatically at 0.2 and 0.4 mM in the former compared to the latter (Paper 3, Figure 4B). This indicates mutation alleviation by AlkA at low f^5dU concentrations but a pronounced promotion at the highest concentrations measured.

Base substitutions were detected at 27 different sites in the rif^R region, 24 of them described previously by Garibyan *et al.* (2003). Eight sites were identified as "hot spots", harboring 8% or more of the total mutations detected in the wild-type and *alkA* strains (Paper 3, Figure 2). The majority of all A·T → G·C transitions were localized at only two sites, both exhibiting the sequence context 5'-GTC-3'. Mutation initiation requires a T to be replaced by f^5U, which would be alleviated or promoted by AlkA.

The excision of f^5U from DNA opposite A and G by AlkA at increasing enzyme concentrations was analyzed. Kinetic analysis showed that the f^5U·G substrate was cleaved much more efficiently than the f^5U·A substrate (Paper 3, Figure 5B,C). The same results have been previously reported for the human glycosylase hSMUG1 (Knævelsrud *et al.*, 2009), and accord with the flipping-out mechanism for base excision, where the enzyme must disrupt the DNA base-pairing and base-stacking interactions of the targeted base to be accommodated in the active site pocket, what determines the reaction rate (McCullough *et al.*, 1999; Wibley *et al.*, 2003).

Automated docking simulations were carried out with f^5U-DNA into the substrate-binding pocket of AlkA. While the keto form of f^5U fits well into the active site being stabilized by a hydrogen bond from the 5-formyl group to Arg22, the ionized $f^5U^{\ominus}$ form is excluded from the AlkA active site pocket (Paper 3, Figure 6A). That can be seen by the much stronger decline in

glycosylase activity for f^5U compared to the primary AlkA substrate m^3A at high pH when $f^5U^{\ominus}$ is abundant (Paper 3, Figure 6B).

3.4 Alleviation and promotion of damage-specific mutation induction in *E. coli* are highly dependent on the *uvrA* gene (Paper 4)

The growth of the *uvrA* mutant was only modestly affected by the addition of f^5dU to the culture medium (Paper 4, Figure 1), which we, as mentioned above, also observed for wild-type, *alkA* and *xth nfo* (Paper 3, Figure 1). Therefore, the results confirm the notion that f^5U is a mutagenic rather than cytotoxic lesion (Bjelland and Seeberg, 2003).

The spontaneous mutation (scored using the rif^R system as previously) rate was slightly lower in *uvrA* than in wild-type and *alkA*. The distribution of the spontaneously arisen base substitutions in *E. coli uvrA* deviated significantly from the wild-type (Paper 4, Table 2), but they were in both cases detected mostly at the same hot spots. Most spontaneous mutagenesis in *uvrA* causes G·C → A·T and G·C → T·A mutations – in contrast with a more equal distribution of the six base substitutions in wild-type (Paper 4, Figure 2) –, thus being closer to *alkA* in its behavior (Paper 4, Table 2).

Interestingly, supplement with 0.1 mM f^5dU doubled the mutation rates in wild-type and *alkA* but only increased it slightly in *uvrA* (Paper 4, Table 1) indicating that the UvrA protein or the UvrABC complex seems necessary for promoting mutations induced by f^5dU (Paper 4, Table 1). However, the abundance of A·T → G·C transitions increased ~6 times and the abundance of A·T → T·A transversions increased ~3 times above the spontaneous level (Paper 4, Table 2). It should be noticed that no increase in the A·T → T·A transversion was found by addition of f^5dU to wild-type, *alkA* and *xth nfo* (Paper 3, Figure 3). In conclusion, although the presence of the UvrA protein only affects the spontaneous mutagenesis slightly, it almost doubles the f^5dU-induced mutagenesis, which is mostly due to promotion of A·T → G·C transitions.

4 Discussion

4.1 Family 1 glycosylases are most probably bi-functional enzymes

Uracil formed by hydrolytic deamination of cytosine must be replaced in DNA prior to replication to avoid generation of C to T transition mutations (Friedberg *et al.*, 2005; Lindahl, 1993), which are the most common type of spontaneous mutation arising in cells and are frequently found in human tumors (Alexandrov *et al.*, 2013). DNA glycosylases of the BER system involved in uracil excision have been described until now as mono-functional enzymes, requiring downstream proteins to complete the BER pathway and repair the damage (Figure 1.8). The studies performed show that UDG-mediated repair initiation by EcUng and hUNG also involves DNA strand incision generating a 3'-α,β-unsaturated aldehyde (UIP) and a 5'-phosphate. UIP can be removed by hAPE1 to prime for replication whereas the 5'-end is ready for ligation, demonstrating that the U-DNA incision products are appropriate for complete repair.

Hitherto, all known bi-functional glycosylases are lyases, which incise an AP site, either formed by the enzyme itself or another enzyme, by a β- or β/δ-elimination reaction (Friedberg *et al.*, 2005). This contrasts largely with the U-DNA incision activity exhibited by the family 1 UDGs, which carries out the reaction by an activated water molecule rather than by a Schiff base intermediate and is not able to recognize and incise AP-DNA, only U-DNA, thus showing strong dependency on uracil.

The rationale for this novel enzyme activity is obscure. It may be argued that it is a consequence of the inaccuracy of the hydrolytic uracil excision activity rather than being evolved to serve a cellular function. However, U-DNA incision generates a 3' block to the initiation of repair replication, which under certain cellular stress conditions could be beneficial to the cell. Another

argument is that if U-DNA incision has been detected *in vitro* it is also probably takes place to a certain extent *in vivo*, and consequently must play a role, although minor, in cell function.

4.2 Oxidative DNA damage may initiate tumorigenesis due to C to T transitions

Since guanine is the DNA base most susceptible to oxidation, and its major oxidation product – oxo^{8}G – has a high capacity for mispairing with adenine during replication, oxo^{8}G has been regarded as the most important oxidized base generated during aerobic respiration (Fortini *et al.*, 2003; Markkanen *et al.*, 2013), leading to G·C → T·A transversions. This is indicated by an increased level of such mutations in *E. coli* when Fpg, the major enzyme for oxo^{8}G removal, and MutY, the enzyme removing mispaired A opposite oxo^{8}G, are knocked out (Garibyan *et al.*, 2003). In contrast, our study on the spontaneous mutagenesis in the *alkA fpg nth nei* mutant, where virtually all glycosylase activity for oxidized bases are knocked out, shows that ~90% of the base substitutions were G·C → A·T and the rest A·T → G·C. Specifically, only one (<1%) G·C → T·A transversion was detected. Consequently, our study on the BER deficient quadruple mutant supports previous studies on repair proficient *E. coli* concluding that the most common base substitution, G·C → A·T arising from oxidative damage to DNA, is caused by mispairing of oxidized cytosines, where presented evidence pointed to h^{5}U and Ug as important mutagenic lesions (Kreutzer and Essigmann, 1998).

However, G·C → A·T transitions also originate by deamination of cytosine to uracil, but since the *alkA fpg nth nei* mutant has complete capacity for uracil removal, virtually all the transitions found are expected to having arisen because of mispairing by unrepaired and thus persistent oxidized bases. The present and previous results underscore the importance of oxidized cytosines in the generation of C to T transitions which are frequently formed in human tumors.

4.3 Repair of 5-formyluracil in DNA by base excision repair can either alleviate or promote mutagenesis in *E. coli*

More than 20 years ago, AlkA was found to have significant activity for removal of f^5U from DNA *in vitro* (Bjelland *et al.*, 1994). At that time this was unexpected since the enzyme was known to be induced in the adaptive response to alkylation (Evensen and Seeberg, 1982; Karran *et al.*, 1982). Since then, no evidence of its role in the repair of oxidized bases *in vivo* has been demonstrated. Thus, this is the first study showing a role of AlkA in the repair of an oxidized base *in vivo*. We found that the presence of AlkA influences the distribution of base substitutions induced by f^5U and, more intriguingly, that AlkA, in addition to alleviating, promotes mutagenesis dependent on the concentration of mutagen. We have suggested a working hypothesis to explain both mutation alleviation and promotion by AlkA. This model takes into account whether f^5U removal takes place before or after replicative mispairing and considers that the ionic form of f^5U ($f^5U^{\ominus}$) is excluded from the AlkA active site which makes it accessible for mispairing (Paper 3, Figure 7).

It is also important to take into account that several back-up activities for f^5U removal like Fpg, Nth and Nei (Zhang *et al.*, 2000) are present in the absence of AlkA.

4.4 Nucleotide excision repair is involved in the repair of oxidative base lesions

Our results show that the distribution of spontaneously arisen base substitutions deviates significantly more between *E. coli uvrA* and the wild-type than between *alkA* and the wild-type (Paper 4, Table 2), which is not surprising since the UvrA protein is necessary for a functional NER pathway which is involved in the repair of several DNA lesions counteracting mutation induction. However, it was intriguing the observation that *uvrA* exhibits a more different distribution of f^5dU-induced base substitutions compared with the wild-type than *alkA* (Paper 4, Table 2), considering the believed importance of AlkA in f^5U repair. Indeed, the results indicate that the presence of the UvrA protein is necessary for mutation alleviation as well as induction (Paper 4, Figure 3B). Whether this is due to repair activity for f^5U exhibited by the UvrABC complex

per se, which has been shown for human NER *in vitro* (Kino *et al.*, 2004), or might be a consequence of UvrA or UvrABC collaborating with BER *e.g.* in the recognition of the f^5U lesion, we cannot yet conclude.

5 Future perspectives

After showing a novel bi-functionality of EcUng and hUNG suggested being able to incise U-DNA by a water-mediated elimination reaction, similar studies on hSMUG1 has been initiated, and should be extended to other mono-functional glycosylases like *E. coli* Tag and Mug, hMPG and hTDG. This might tell us whether the novel U-DNA incision activity is a specific function of some few enzymes or is coupled to the involvement of activated water in catalysis.

The spontaneous mutagenesis due to damage by ROS was measured in *E. coli* by comparing wild-type with a quadruple mutant (*alkA fpg nth nei*) deficient in the genes for the main glycosylases involved in the repair of oxidized base residues. A future goal would be to extent the comparison of the spontaneous mutagenesis using the same system for mutation detection to alkylation repair-deficient (involving the genes *tag*, *alkA*, *alkB*, *ada* and *ogt*) and deamination repair-deficient mutants (involving the genes *ung*, *mug* and *alkA*). This might tell us something about the importance of these different chemical processes in mutagenesis.

DNA repair involves many different pathways and enzymes establishing a complex network in the cell cycle. Defects in the DNA repair system have been demonstrated to have a connection with cancer development, *e.g.* deficiency in NER genes related to *xeroderma pigmetosum* and skin cancer (Pesz *et al.*, 2014). Mutations in the *OGG1* gene have been reported to be related to several types of renal cancers (Audebert *et al.*, 2000) and UNG-deficient mice have shown an increased morbidity and higher than normal incidence of B-cell lymphomas (Nilsen *et al.*, 2003). However, correlation of BER with tumorigenesis needs to be further investigated. Since the results presented in this thesis show that repair glycosylases both alleviate and promote mutagenesis in *E. coli*, BER might behave in the same way in human cells and thus be linked to cancer initiation and progression.

FUTURE PERSPECTIVES 5

Since f^5U causes a minimal perturbation in the DNA structure (Bjelland and Seeberg, 2003), it is tempting to speculate whether the large effect on f^5dU mutagenesis exhibited by the deficiency in the UvrA protein is due to collaboration with BER glycosylases rather than due to NER activity *per se*. To answer this interesting question, we have started similar mutagenesis experiments with *E. coli uvrB* and *uvrC*, from which results we hope can encourage further experiments to reach a conclusion. Molecular interaction and repair studies involving f^5U-DNA, BER glycosylases and the UvrA, UvrB and UvrC proteins should be appropriate.

6 References

Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S *et al.* (2013). Signatures of mutational processes in human cancer. *Nature* **500**, 415–421.

Ånensen H, Provan F, Lian AT, Reinertsen SH, Ueno Y, Matsuda A, Seeberg E, Bjelland S (2001). Mutations induced by 5-formyl-2'-deoxyuridine in *Escherichia coli* include base substitutions that can arise from mispairs of 5-formyluracil with guanine, cytosine and thymine. *Mutat. Res.* **476**, 99–107.

Audebert M, Chevillard S, Levalois C, Gyapay G, Vieillefond A *et al.* (2000). Alterations of the DNA repair gene *OGG1* in human clear cell carcinomas of the kidney. *Cancer Res.* **60**, 4740–4744.

Bailly V, Verly WG (1987). *Escherichia coli* endonuclease III is not an endonuclease but a beta-elimination catalyst. *Biochem. J.* **242**, 565–572.

Barnes DE, Lindahl T (2004). Repair and genetic consequences of endogenous DNA base damage in mammalian cells. *Annu. Rev. Genet.* **38**, 445–476.

Bauerle M, Schwalm E, Booker S (2014). Mechanistic diversity of radical S-adenosylmethionine (SAM)-dependent methylation. *J. Biol. Chem.* **290**, 3995–4002.

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

Bjelland S, Birkeland NK, Benneche T, Volden G, Seeberg E (1994). DNA glycosylase activities for thymine residues oxidized in the methyl group are functions of the AlkA enzyme in *Escherichia coli*. *J. Biol. Chem.* **269**, 30489–30495.

REFERENCES 6

Bjelland S, Eide L, Time RW, Stote R, Eftedal I, Volden G, Seeberg E (1995). Oxidation of thymine to 5-formyluracil in DNA: mechanisms of formation, structural implications, and base excision by human cell free extracts. *Biochemistry* **34**, 14758–14764.

Bjelland S, Ånensen H, Knævelsrud I, Seeberg E (2001). Cellular effects of 5-formyluracil in DNA. *Mutat. Res.* **486**, 147–154.

Clarke ND, Kvaal M, Seeberg E (1984). Cloning of *Escherichia coli* genes encoding 3-methyladenine DNA glycosylases I and II. *Mol. Gen. Genet.* **197**, 368–372.

Crowley DJ, Hanawalt PC (1998). Induction of the SOS response increases the efficiency of global nucleotide excision repair of cyclobutane pyrimidine dimers, but not 6–4 photoproducts, in UV-irradiated *Escherichia coli*. *J. Bacteriol.* **180**, 3345–3352.

Drohat AC, Maiti A (2014). Mechanisms for enzymatic cleavage of the N-glycosidic bond in DNA. *Org. Biomol. Chem.* **12**, 8367–8378.

Durandy A, Honjo T (2001). Human genetic defects in class-switch recombination (hyper-IgM syndromes). *Curr. Opin. Immunol.* **13**, 543–548.

Evensen G, Seeberg E (1982). Adaptation to alkylation resistance involves the induction of a DNA glycosylase. *Nature* **296**, 773–775.

Fortini P, Pascucci B, Parlanti E, D'Errico M, Simonelli V, Dogliotti E (2003). 8-Oxoguanine DNA damage: at the crossroad of alternative repair pathways. *Mutat. Res.* **531**, 127–139.

Friedberg EC, Walker GC, Siede W, Wood RD, Schultz RA, Ellenberger T (2005). DNA Repair and Mutagenesis, 2nd ed. *ASM Press*.

Fujikawa K, Kamiya H, Kasai H (1998). The mutations induced by oxidatively damaged nucleotides, 5-formyl-dUTP and 5-hydroxy-dCTP, in *Escherichia coli*. *Nucleic Acids Res.***26**, 4582–4587.

Garibyan L, Huang T, Kim M, Wolff E, Nguyen A *et al.* (2003). Use of the *rpoB* gene to determine the specificity of base substitution mutations on the *Escherichia coli* chromosome. *DNA Repair (Amst.)* **2**, 593–608.

Guillet M, Boiteux S (2003). Origin of endogenous DNA abasic sites in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **23**, 8386–8394.

Gutteridge JM, Halliwell B (1989). Iron toxicity and oxygen radicals. *Baillieres Clin. Haematol.* **2**, 195–256.

Hang B, Chenna A, Sági J, Singer B (1998). Differential cleavage of oligonucleotides containing the benzene-derived adduct, 1,N6-benzetheno-dA, by the major human AP endonuclease HAP1 and *Escherichia coli* exonuclease III and endonuclease IV. *Carcinogenesis* **19**, 1339–1343.

Henkler F, Stolpmann K, Luch A (2012). Exposure to polycyclic aromatic hydrocarbons: bulky DNA adducts and cellular responses. In Molecular, Clinical and Environmental Toxicology, Volume 3: Environmental Toxicology, 1st ed. *Springer Basel*, pp. 107–131.

Hori M, Yonekura SI, Nohmi T, Gruz P, Sugiyama H, Yonei S, Zhang-Akiyama QM (2010). Error-prone translesion DNA synthesis by *Escherichia coli* DNA polymerase IV (DinB) on templates containing 1,2-dihydro-2-oxoadenine. *J. Nucleic Acids* **2010**, 807579.

Janion C (2001). Some aspects of the SOS response system – a critical survey. *Acta Biochim. Pol.* **48**, 599–610.

Kamiya H, Murata-Kamiya N, Karino N, Ueno Y, Matsuda A, Kasai H (2002). Induction of T → G and T → A transversions by 5-formyluracil in mammalian cells. *Mutat. Res.* **513**, 213–222.

Karran P, Hjelmgren T, Lindahl T (1982). Induction of a DNA glycosylase for N-methylated purines is part of the adaptive response to alkylating agents. *Nature* **296**, 770–773.

Kasai H, Iida A, Yamaizumi Z, Nishimura S, Tanooka H (1990). 5-Formyldeoxyuridine: a new type of DNA damage induced by ionizing radiation and its mutagenicity to *Salmonella* strain TA102. *Mutat. Res.* **243**, 249–253.

Kavli B, Otterlei M, Slupphaug G, Krokan HE (2007). Uracil in DNA – General mutagen, but normal intermediate in acquired immunity. *DNA Repair (Amst.)* **6**, 505–516.

REFERENCES 6

Kim YJ, Wilson DM (2012). Overview of base excision repair biochemistry. *Curr. Mol. Pharmacol.* **5**, 3–13.

Kino K, Shimizu Y, Sugasawa K, Sugiyama H, Hanaoka F. (2004). Nucleotide excision repair of 5-formyluracil *in vitro* is enhanced by the presence of mismatched bases. *Biochemistry* **43**, 2682–2687.

Klungland A, Bjelland S (2007). Oxidative damage to purines in DNA: role of mammalian Ogg1. *DNA Repair (Amst.)* **6**, 481–488.

Klungland A, Paulsen R, Rolseth V, Yamada Y, Ueno Y, Wiik P, Matsuda A, Seeberg, E, Bjelland S (2001). 5-Formyluracil and its nucleoside derivatives confer toxicity and mutagenicity to mammalian cells by interfering with normal RNA and DNA metabolism. *Toxicol. Lett.* **119**, 71–78.

Knævelsrud I, Slupphaug G, Leiros I, Matsuda A, Ruoff P, Bjelland S (2009). Opposite-base dependent excision of 5-formyluracil from DNA by hSMUG1. *Int. J. Radiat. Biol.* **85**, 413–420.

Kreutzer DA, Essigmann JM (1998). Oxidized, deaminated cytosines are a source of C → T transitions *in vivo*. *Proc. Natl. Acad. Sci. U S A* **95**, 3578–3582.

Krokan HE, Standal R, Slupphaug G (1997). DNA glycosylases in the base excision repair of DNA. *Biochem. J.* **325**, 1–16.

Krokan HE, Otterlei M, Nilsen H, Kavli B, Skorpen F, Andersen S, Skjelbred C, Akbari M, Aas PA, Slupphaug G (2001). Properties and functions of human uracil-DNA glycosylase from the *UNG* gene. *Prog. Nucleic Acid Res. Mol. Biol.* **68**, 365–386.

Krokan HE, Drabløs F, Slupphaug G (2002). Uracil in DNA – occurrence, consequences and repair. *Oncogene* **21**, 8935–8948.

Ladner RD, Caradonna SJ (1997). The human dUTPase gene encodes both nuclear and mitochondrial isoforms. Differential expression of the isoforms and characterization of a cDNA encoding the mitochondrial species. *J. Biol. Chem.* **272**, 19072–19080.

Le May N, Egly JM, Coin F (2010). True lies: the double life of the nucleotide excision repair factors in transcription and DNA repair. *J. Nucleic Acids* **2010**.

Lindahl T (1974). An N-glycosidase from *Escherichia coli* that releases free uracil from DNA containing deaminated cytosine residues. *Proc. Natl. Acad. Sci. U S A* **71**, 3649–3653.

Lindahl T (1979). DNA glycosylases, endonucleases for apurinic/apyrimidinic sites, and base excision-repair. *Prog. Nucleic Acid Res. Mol. Biol.* **22**, 135–192.

Lindahl T (1993). Instability and decay of the primary structure of DNA. *Nature* **362**, 709–715.

Liu P, Burdzy A, Sowers LC (2003). Repair of the mutagenic DNA oxidation product, 5-formyluracil. *DNA Repair (Amst.)* **2**, 199–210.

Markkanen E, Dorn J, Hübscher U (2013). MUTYH DNA glycosylase: the rationale for removing undamaged bases from the DNA. *Front. Genet.* **4**, 18.

Masaoka A, Terato H, Kobayashi M, Honsho A, Ohyama Y, Ide, H (1999). Enzymatic repair of 5-formyluracil. I. Excision of 5-formyluracil site-specifically incorporated into oligonucleotide substrates by AlkA protein (*Escherichia coli* 3-methyladenine DNA glycosylase II). *J. Biol. Chem.* **274**, 25136–25143.

Masaoka A, Matsubara M, Hasegawa R, Tanaka T, Kurisu S, Terato H, Ohyama Y, Karino N, Matsuda A, Ide H (2003). Mammalian 5-formyluracil-DNA glycosylase. 2. Role of SMUG1 uracil-DNA glycosylase in repair of 5-formyluracil and other oxidized and deaminated base lesions. *Biochemistry* **42**, 5003–5012.

Matsubara M., Masaoka A., Tanaka T, Miyano T, Kato N, Terato H, Ohyama Y, Iwai S, Ide H (2003). Mammalian 5-formyluracil-DNA glycosylase. 1. Identification and characterization of a novel activity that releases 5-formyluracil from DNA. *Biochemistry* **42**, 4993–5002.

McCullough AK, Dodson ML, Lloyd RS (1999). Initiation of base excision repair: glycosylase mechanisms and structures. *Annu. Rev. Biochem.* **68**, 255–285.

Miyabe I, Zhang QM, Kino K, Sugiyama H, Takao M, Yasui A, Yonei S (2002). Identification of 5-formyluracil DNA glycosylase activity of human hNTH1 protein. *Nucleic Acids Res.* **30**, 3443–3448.

Moen MN, Knaevelsrud I, Haugland GT, Grøsvik K, Birkeland NK, Klungland A, Bjelland S (2011). Uracil-DNA glycosylase of *Thermoplasma acidophilum* directs

long-patch base excision repair, which is promoted by deoxynucleoside triphosphates and ATP/ADP, into short-patch repair. *J. Bacteriol.* **193**, 4495–4508.

Nakabeppu Y, Miyata T, Kondo H, Iwanaga S, Sekiguchi M (1984). Structure and expression of the *alkA* gene of *Escherichia coli* involved in adaptive response to alkylating agents. *J. Biol. Chem.* **259**, 13730–13736.

Nilsen H, Stamp G, Andersen S, Hrivnak G, Krokan HE, Lindahl T, Barnes DE (2003). Gene-targeted mice lacking the Ung uracil-DNA glycosylase develop B-cell lymphomas. *Oncogene* **22**, 5381–5386.

O'Connor TR, Laval J (1989). Physical association of the 2,6-diamino-4-hydroxy-5N-formamidopyrimidine-DNA glycosylase of *Escherichia coli* and an activity nicking DNA at apurinic/apyrimidinic sites. *Proc. Natl. Acad. Sci. U S A* **86**, 5222–5226.

Oda Y (1987). Induction of SOS responses in *Escherichia coli* by 5-fluorouracil. *Mutat. Res.* **183**, 103–108.

Pagès V, Johnson RE, Prakash L, Prakash S (2008). Mutational specificity and genetic control of replicative bypass of an abasic site in yeast. *Proc. Natl. Acad. Sci. U S A* **105**, 1170–1175.

Pesz KA, Bieniek A, Gil J, Laczmanska I, Karpinski P, Makowska I, Trusewicz A, Sasiadek MM (2014). Polymorphisms in nucleotide excision repair genes and basal cell carcinoma of the skin. *Int. J. Dermatol.* **53**, 1474–1477.

Petit C, Sancar A (1999). Nucleotide excision repair: from *E. coli* to man. *Biochimie* **81**, 15–25.

Pham P, Bransteitter R, Goodman MF (2005). Reward versus risk: DNA cytidine deaminases triggering immunity and disease. *Biochemistry* **44**, 2703–2715.

Prakash A, Doublié S, Wallace SS (2012). The Fpg/Nei family of DNA glycosylases: substrates, structures and search for damage. *Prog. Mol. Biol. Transl. Sci.* **110**, 71–91.

Radman M (1975). SOS repair hypothesis: phenomenology of an inducible DNA repair which is accompanied by mutagenesis. *Basic Life Sci.* **5A**, 355–367.

Roberts VA, Pique ME, Hsu S, Li S, Slupphaug G *et al.* (2012). Combining H/D exchange mass spectroscopy and computational docking reveals extended DNA-binding surface on uracil-DNA glycosylase. *Nucleic Acids Res.* **40**, 6070–6081.

Saha S, Jana B, Basu T (2007). The two inducible responses, SOS and heat-shock, in *Escherichia coli* act synergistically during Weigle reactivation of the bacteriophage ϕX174. *Int. J. Radiat. Biol.* **83**, 463–469.

Sander M, Wilson SH (2001). Base excision repair, AP endonucleases and DNA glycosylases. In Encyclopedia of Life Sciences. *John Wiley & Sons, Ltd.*

Schärer OD, Jiricny J. (2001). Recent progress in the biology, chemistry and structural biology of DNA glycosylases. *Bioessays* **23**, 270–281.

Schormann N, Ricciardi R, Chattopadhyay D (2014). Uracil-DNA glycosylases – Structural and functional perspectives on an essential family of DNA repair enzymes. *Protein Sci.* **23**, 1667–1685.

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

Serre L, Pereira de Jésus K, Boiteux S, Zelwer C, Castaing B (2002). Crystal structure of the *Lactococcus lactis* formamidopyrimidine-DNA glycosylase bound to an abasic site analogue-containing DNA. *EMBO J.* **21**, 2854–2865.

Slupphaug G, Mol CD, Kavli B, Arvai AS, Krokan HE, Tainer JA (1996). A nucleotide-flipping mechanism from the structure of human uracil-DNA glycosylase bound to DNA. *Nature* **384**, 87–92.

Strauss BS (2002). The "A" rule revisited: polymerases as determinants of mutational specificity. *DNA Repair (Amst.)* **1**, 125–135.

Takao M, Kanno S, Shiromoto T, Hasegawa R, Ide H, *et al.* (2002). Novel nuclear and mitochondrial glycosylases revealed by disruption of the mouse *Nth1* gene encoding an endonuclease III homolog for repair of thymine glycols. *EMBO J.* **21**, 3486–3493.

Terato H, Masaoka A, Kobayashi M, Fukushima S, Ohyama Y, Yoshida M, Ide H (1999). Enzymatic repair of 5-rormyluracil II. Mismatch formation between 5-formyluracil and guanine during DNA replication and its recognition by two proteins involved in base excision repair (AlkA) and mismatch repair (MutS). *J. Biol. Chem.* **274**, 25144–25150.

Truglio JJ, Croteau DL, Van Houten B, Kisker C (2006). Prokaryotic nucleotide excision repair: the UvrABC system. *Chem. Rev.* **106**, 233–252.

Van Houten B (1990). Nucleotide excision repair in *Escherichia coli*. *Microbiol. Rev.* **54**, 18–51.

REFERENCES 6

Van Houten B, Croteau DL, DellaVecchia MJ, Wang H, Kisker C (2005). "Close-fitting sleeves": DNA damage recognition by the UvrABC nuclease system. *Mutat. Res.* **577**, 92–117.

Wallace SS (2013). DNA glycosylases search for and remove oxidized DNA bases. *Environ. Mol. Mutagen.* **54**, 691–704.

Wibley JE, Waters TR, Haushalter K, Verdine GL, Pearl LH (2003). Structure and specificity of the vertebrate anti-mutator uracil-DNA glycosylase SMUG1. *Mol. Cell* **11**, 1647–1659.

Wyatt MD, Wilson DM (2009). Participation of DNA repair in the response to 5-fluorouracil. *Cell. Mol. Life Sci.* **66**, 788–799.

Xu G, Herzig M, Rotrekl V, Walter CA (2008). Base excision repair, aging and health span. *Mech. Ageing Dev.* **129**, 366–382.

Yoshida M, Makino K, Morita H, Terato H, Ohyama Y, Ide H (1997). Substrate and mispairing properties of 5-formyl-2'-deoxyuridine 5'-triphosphate assessed by *in vitro* DNA polymerase reactions. *Nucleic Acids Res.* **25**, 1570–1577.

Zhang QM, Fujimoto J, Yonei S (1995). Enzymatic release of 5-formyluracil by mammalian liver extracts from DNA irradiated with ionizing radiation. *Int. J. Radiat. Biol.* **68**, 603–607.

Zhang QM, Sugiyama H, Miyabe I, Matsuda S, Saito I, Yonei S (1997). Replication of DNA templates containing 5-formyluracil, a major oxidative lesion of thymine in DNA. *Nucleic Acids Res.* **25**, 3969–3973.

Zhang QM, Miyabe I, Matsumoto Y, Kino K, Sugiyama H, Yonei S (2000). Identification of repair enzymes for 5-formyluracil in DNA. Nth, Nei, and MutM proteins of *Escherichia coli*. *J. Biol. Chem.* **275**, 35471–35477.