Copyright Declaration

Investigating enzyme communication during base excision repair in *Escherichia coli*

Qiyuan Zhao

Division of Molecular Biosciences

Department of Life Sciences

Imperial College London

2013

A thesis submitted for the degree of Doctor of Philosophy and the

Diploma of Imperial College London

Abstract

Mismatch uracil DNA Glycosylase (MUG) from *Escherichia coli* is an initiating enzyme in the base excision repair (BER) pathway and is responsible for the removal of 3,N^4-ethenocytosine and uracil from DNA during the stationary phase of *E.coli* cell growth. As with other DNA glycosylases, the abasic product is potentially more harmful than the initial lesion. MUG is widely regarded as a "single turnover" enzyme because it still remains tightly bound to its abasic product after cleavage, thus impeding its catalytic turnover. This may be a general protective mechanism to protect the abasic BER intermediate, whereby coordination of enzyme activity in BER is achieved through displacement of the DNA glycosylase by the downstream apurinic-apyrimidinic (AP) endonuclease. Numerous DNA glycosylases have now been cited as having an enhanced turnover in the presence of an AP endonuclease. The aim of this project is to investigate enzyme coordination between MUG and its both downstream AP endonucleases, Exonuclease III (ExoIII) and Endonuclease IV (EndoIV), in the initial steps of BER. We show here that MUG binds its substrate, abasic DNA and non-specific DNA in the differential modes. A 2:1 cooperative binding stoichiometry with abasic DNA is demonstrated to be of functional significance in both product binding and catalysis via fluorescence anisotropy assays, band shift assays and loss-of-function site-directed mutagenesis methods. The effects of the ExoIII and EndoIV on the MUG turnover kinetics with a U•G containing substrate was investigated. Both ExoIII and EndoIV greatly enhance the turnover of MUG. Furthermore, the analysis of both ExoIII catalytic activity dependent and concentration dependent on MUG turnover demonstrate ExoIII may employ a product scavenging mechanism to enhance MUG turnover. These combined results constitute a new concept that MUG has a pre-catalytic discrimination ability to coordinate its reactivity behavior with the other enzymes.

Author Declaration

I hereby declare that this thesis, submitted in fulfilment of the requirements for the degree of Doctor of Philosophy of Imperial College London, represents my own work and has not been previously submitted to any institute for any degree, diploma or other qualification. My previous publications and any ideas, work or quotations from other people described in this thesis, are fully acknowledged in accordance with the standard referencing practices of the discipline.

Qiyuan Zhao

2013

.....

This thesis is dedicated to my parents,

Dianlin Zhao and Qiaoxia Li

Acknowledgements

I would like to express my sincere gratitude to my PhD supervisor Dr. Geoff Baldwin for his excellent guidance, critical comments and continuing support, for his extreme patience with my impatience and occasionally overbearing temper and for encouraging me to develop and pursue my own ideas along the four year's journey towards completion of this thesis. He has also been very helpful in supporting and advising me regarding my future career. Thank you for letting me work for you!

I owe extreme thanks to Dr. Jan Silhan not only for patiently teaching me various lab techniques and wisely providing me numerous useful ideas inside the lab, but also for brotherly encouraging and advising me beyond work. Even after leaving the lab he always had an open ear and ready advice.

I am also very grateful to Dr. Seden Grippon for her excellent guidance when I took my first steps in experimental science in the Baldwin lab and also for sharing her unpublished data that were crucial for the interpretation of my results.

My gratitude goes to all the current and past members of the Baldwin group with whom I had the pleasure of working with for all their time, companionship and considerable reserves of help and patience, with particular mention for Tim Wilson who always lent helping hands whenever needed, as well as Tom Adie, James Field, Marko Storch, Ben Mackrow, Lucy Rayner and Liang Xue.

I would like to thank Yu Chen for always being there for me through thick and thin.

Last and by no means least, I am deeply indebted to my grandparents and parents, for all their endless support, encouragement and love throughout. Thank you for believing in me! You are the best families in the world!

Table of Contents

List of Figures

List of Tables

Abbreviations

6-4PPs	Pyrimidine pyrimidone photoproducts
8-oxoG	8-oxo-2'deoxyguanosine
A	Adenine
AAG	Alkyladenine DNA glycosylase
ALK	Alkylpurine DNA glycosylase
AP	Apurinic/apyrimidinic
APS	Ammonium persulfate
ATP	Adenosine 5'-triphosphate
BER	Base excision repair
β-ME	β-mercaptoethanol
bp	Base pair
BSA	Bovine serum albumin
C	Cytosine
CPD	Cyclobutane pyrimidine dimer
dATP	Deoxyadenosine triphosphate
dCTP	Deoxycytidine triphosphate
DEAE	Diethylaminoethyl cellulose
dGTP	Deoxyguanosine triphosphate
Da	Dalton
DSBs	Double strand breaks
DNA	Deoxyribose nucleic acid
dRP	Deoxyribose phosphate
dRPase	Deoxyribophosphodiesterase
DTT	Dithiothreitol
dTTP	Deoxythymidine triphosphate
dUMP	Deoxyuridine monophosphate
dUTP	Deoxyuridine triphosphate
εC	Etheno Cytosine
E.coli	*Escherichia coli*
EDDS	Ethylenediamine-*N,N'*-disuccinic acid
EDTA	Ethylene diamine tetra-acetic acid
EndoIII	Endonuclease III
EndoV	Endonuclease V
EndoIV	Endonuclease IV
ExoIII	Exonuclease III
FaPy	Formamidopryrimidine
FEN1	Flap structure-specific endonuclease
Fpg	Formamidopyrimidine DNA glycosylase
FPLC	Fast performance liquid chromatography

G	Guanine
hAPE1	Human AP endonuclease 1
Hex	6-carboxy-2',4,4',5',7,7',-hexachlorofluorescein
hOGG1	Human OG glycosylase
HR	Homologous recombination
HSV-1	Herpes simplex virus Type-1
IPTG	Isopropyl-β-D-thiogalatopyranoside
kbp	Kilo base pair
kDa	Kilo dalton
LB	Lauria Bertani
M	mol L^{-1}
min	Minute
MMEJ	Microhomology-mediated end-joining
MMR	Mismatch repair
MUG	Mismatch uracil DNA glycosylase
NApe	*Neisserial* AP endonuclease
NEIL	*nei* like
NER	Nucleotide excision repair
NExo	*Neisserial* Exonuclease
NHEJ	Non-homologous end-joining
nt	nucleotide
^{o}C	Degrees Celsius
OD280	Optical density at 280 nm
OD600	Optical density at 600 nm
OH	Hydroxyl
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCNA	Proliferating cell nuclear antigen
PCR	Polymerase chain reaction
PNK	Polynucleotide kinase
Pol	Polymerase
RNA	Ribose nucleic acid
ROS	Reactive oxygen species
rpm	Revolutions per minute
s	Second
SDS	Sodium dodecyl sulfate
SMUG	Single-strand selective monofunctional uracil-DNA glycosylase
T	Thymine
Taq	*Thermus aquaticus*
TBE	Tris/Borate/EDTA
TDG	Thymine DNA glycosylase

TEMED	N, N, N', N'-etramethylethylenediamine
TIM	Triose Phosphate Isomerase
U	Uracil
UDG	Uracil DNA glycosylase
wt	Wild-type
XRCC1	X-ray cross complementing gene 1

Abbreviations for Amino Acids					
A	Ala	Alanine	M	Met	Methionine
C	Cys	Cysteine	N	Asn	Asparagine
D	Asp	Aspartic acid	P	Pro	Proline
E	Glu	Glutamic acid	Q	Gln	Glutamine
F	Phe	Phenylalanine	R	Arg	Arginine
G	Gly	Glycine	S	Ser	Serine
H	His	Histidine	T	Thr	Threonine
I	Ile	Isoleucine	V	Val	Valine
K	Lys	Lysine	W	Trp	Tryptophan
L	Leu	Leucine	Y	Tyr	Tyrosine

Chapter 1

Introduction

1.1 DNA structure

DNA, short for deoxyribonucleic acid, was first discovered by the Swiss physician Friedrich Mieschcer in 1869; however, it wasn't until 1943 that this biggest organic macromolecule was demonstrated to play a crucial role in determining genetic inheritance by Avery-MacLeod-McCarty experiment [1]. This experiment and the later Hershey-Chase experiment paved the way for James Watson and Francis Crick's discovery of the double helical structure of DNA in 1953, which marked the birth of modern genetics and molecular biology [2].

DNA is the repository of genetic information used in the development and functioning of all eukaryotes and most prokaryotes, including many viruses. The determination of the native DNA structure led to the central dogma of molecular biology, DNA is transcribed to RNA and RNA is translated into proteins. Therefore, DNA has been at the forefront of scientific research ever since 60 years ago and until now considerable advances have been made in the understanding of DNA, and the proteins that act upon it.

The amount of information in DNA is seemingly inversely related to the structure and variability of DNA. DNA comes in the form of a simple filament-like molecular arrangement that is well-suited for genetic information storage and access. Variability of DNA governs the most traits and diversity of life on earth and stems from the sequence of DNA "elementary bricks" called nucleotides, which include deoxyadenosine triphosphate (dATP), deoxyguanosine triphosphate (dGTP), deoxycytidine triphosphate (dCTP) and deoxythymidine triphosphate (dTTP). The chemical structures of nucleotides were first elucidated by Phoebus Levene in 1930 [3]. Each of four nucleotide monomeric units consists of a phosphate group, linked up with a deoxyribonucleoside. The deoxyribonucleoside is composed of one of

heterocyclic nitrogenous purine or pyrimidine bases covalently bound to a pentose sugar (lacking the 2'-hydroxyl group) via an *N*-glycosidic bond. The variation between the nucleotides lies in four base units: adenine (A), guanine (G), cytosine (C) and thymine (T). Nucleotides are covalently polymerized via a phosphodiester bond between the 3' carbon of one nucleotide and the 5' carbon to the following nucleotide, constituting a single-stranded DNA chain. The consistent orientation of the nucleotide building blocks gives the chain overall 5' → 3' polarity. The two ends of the single-stranded chain are chemically distinct. At the 3' end, the sugar of the terminal nucleotide has a free 3' hydroxyl. At the 5' end of the chain, the 5' carbon of the nucleotide isn't linked to another nucleotide but carries a phosphate group.

The crosswise pattern of X-ray diffraction images of DNA fibers was exhibited by Rosalind Franklin and Maurice Wilkins in 1951 that unveiled the double helical structure of the DNA molecule. In detail, the pitch along the axis of helix is 3.4Å, the helix undergoes one complete turn every 34 Å and each helical turn has 10 nucleotides. The helix diameter of DNA is 20 Å that determines DNA must consist of two polynucleotide chains, which are run in antiparallel directions and coupled by hydrogen bonds between complementary base pairs, A and T are paired via double hydrogen bonding but G pairs with C using three hydrogen bonds (Figure 1).

The most common conformation that DNA adopts is the right-handed double helix structure, also known as B-form, which is the structure postulated by Watson and Crick in 1953. However, many kinds of unusual non-B form DNA structures have also been found, for example, Z-form helical DNA, which spirals to the left and has a zigzag shaped backbone; moreover, A-form helical DNA, which still right-handed, but every 23 Å makes a turn and has 11 nucleotides per turn.

The nuclear chromosomes, namely coiled DNA molecules, in all eukaryotes are long, linear double helixes; however, some smaller chromosomes are circular. These include the chromosomes of bacteria, such as *Escherichia coli*, and the chromosomes of organelles such as the mitochondria and chloroplasts that exist in eukaryotic cells, and the chromosomes of some viruses, including the papovaviruses. Despite these circular double strand helixes doesn't have end, the two strands are still formed of the same basic building blocks, nucleotides, and antiparallel in polarity.

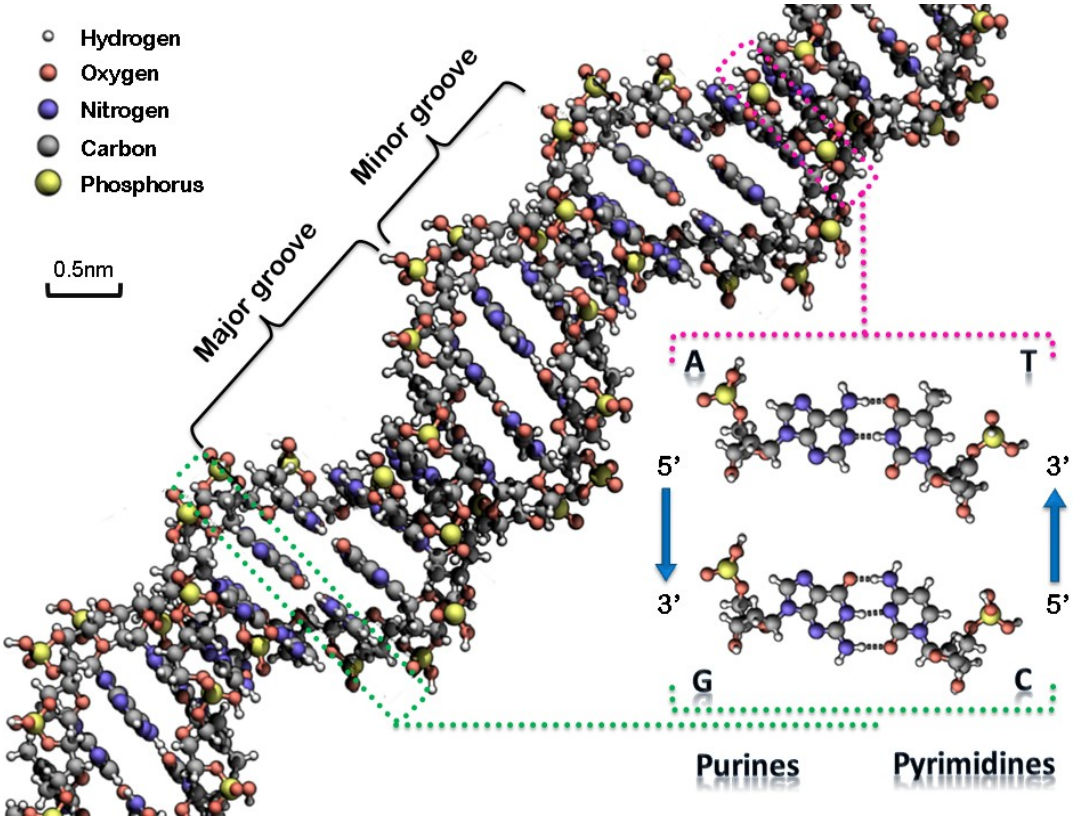


Figure 1. The structure of DNA double helix.
The DNA structure is a right handed double helix with a diameter of ~2 nm and has 10 base pairs per helical turn which are perpendicular to the helical axis. The two 5'→3' antiparallel strands of the double helix are held together by hydrogen bonds between complementary base pairs, A-T and G-C shown in the bottom right. The atoms in the structure are colour-coded by element shown in the top left (adapted from [4]).

The discovery of the double helix in 1953 immediately raised a question about how biological information is encoded in DNA. A remarkable feature of the double

helical structure is that DNA can accommodate almost any sequence of base pairs, any combination of its base units A, C, G and T, and therefore, any molecular information or instruction for life. During the following decades, it was discovered that each gene encodes a complementary RNA transcript, called messenger RNA (mRNA) [5], which is composed of A, C, G and uracil (U), instead of T. The four base units of DNA and RNA alphabets correspond to the 20 amino acids of the protein alphabet by a triplet code, every three letters or codons in a gene encodes one amino acid [6]. For instance, GCT encodes the amino acid alanine. The dictionary of DNA letters that constitute the amino acid is known as genetic code. There are 64 different triplets or codons, 61of which encode an amino acid and three of which make up the stop codon that signals the termination of the growing protein chain. Different triplets can encode the same amino acid.

Profound implications for biology are originated from the molecular complementary of the double helix. As implied by James Watson and Francis Crick in their landmark paper [2], base pairing indicates a template-copying mechanism that explains the fidelity in copying of genetic material during DNA replication. It also underpins the synthesis of mRNA from the DNA template, as well as processes of DNA damage repair [7].

1.2 DNA damage

The prime objective for every life form is to transmit its intact and unchanged genetic material to the next generation. Nature has evolved a number of biological traits to in concert protect the integrity of DNA and preserve genetic stability, for instance, the double helical nature of DNA, the condensed structure of chromosomes, and the cellular genome-maintenance enzymatic capabilities. However, DNA is still highly susceptible to damage as it continuously subject to assaults from various endogenous and exogenous mutagens [8]. In addition, the fidelity of DNA is also compromised at elevated temperature and at extremes of pH, and even due to the rare but significant fallibility of intrinsic replication, recombination and repair processing events. Any resulting damage, if not repaired or repaired incorrectly, may stall DNA replication and transcription, and give rise to mutations or wider-scale genome aberrations that could be catastrophic and threaten cell or organism viability, most prevalently implicated in aging, carcinogenesis and neurodegeneration [9, 10].

DNA damage is imputed to all kinds of modifications to the DNA native structure, ranging from large scale physical changes, such as gene transposition, to small scale chemical composition alterations of a single DNA entity [11]. DNA damage can be divided into two major classes, endogenous DNA damage and environmental DNA damage. The former class includes many hydrolytic and oxidative reactions that are a consequence of life surrounded by water and reactive oxygen. The major sites of oxidative and hydrolytic damage in DNA are summarized in Figure 2. The environmental class includes DNA damage arising from extracellularly generated physical and chemical agents. While all of the primary components of DNA, including bases, sugars and phosphodiester bonds, are subject to damage by both endogenous and environmental reactants, the damage of nitrogenous bases, which specify the genetic code, normally give rise to the most detrimental consequence.

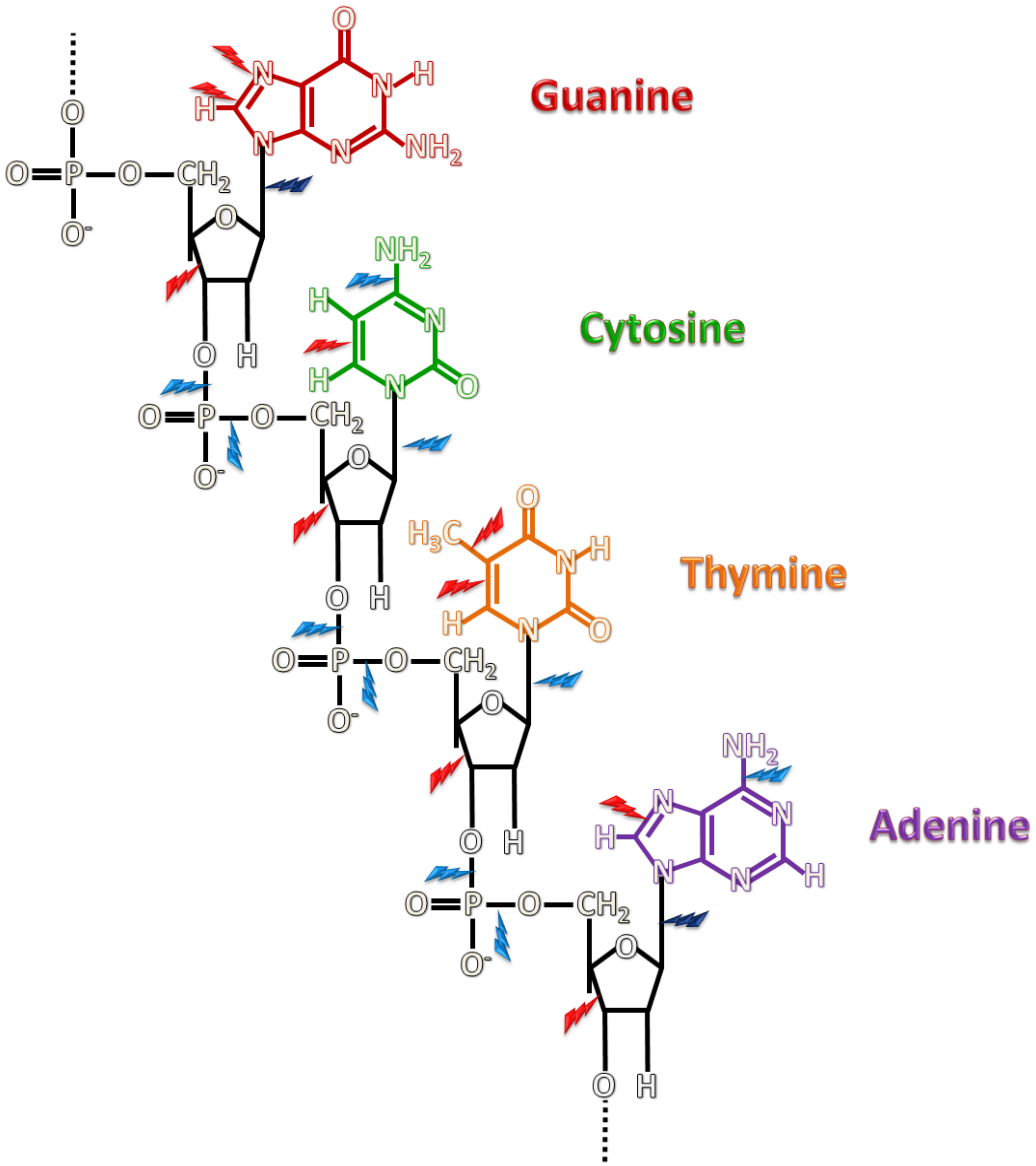


Figure 2. Major sites of hydrolytic and oxidative damage in DNA.
A short segment of one DNA strand is shown with the four canonical DNA bases. The major sites of hydrolytic depurination are shown by dark blue lightning marks. Light blue lightning marks point at other sites of hydrolytic attack. Major sites of oxidative damage are indicated by the red lightning marks.

A source of DNA aberrations arising during normal DNA metabolism is the insertion of erroneous bases during DNA synthesis, leading to nucleotide mismatches. Although the main DNA polymerases involved in DNA synthesis, such as T7 Pol, T4 Pol, Polα, Polε and *E. coli* PolIII [12], have a 3'→5' proofreading exonuclease activity to cope with misincorporation errors, mismatches can still occasionally evade the intrinsic fidelity of these enzymes and become incorporated into a DNA chain, causing distortions of helical structure, interferences of protein interactions and even mutations. Tautomeric shifts of the nitrogenous bases readily trigger mispairing. As shown in Figure 3, guanine and thymine shift between the natural keto form to the enol, and adenine and cytosine switch from the canonical amino to imino. These tautomeric shifts can result in atypical hydrogen bonds rearrangement of base-pairing, by which, during DNA replication, misincorporation arises in the daughter strand, such as, imino form of adenine or enol form of guanine in the parental strand can be mismatched with nascent cytosine or thymine respectively by DNA glycosylases, and vice versa [13].

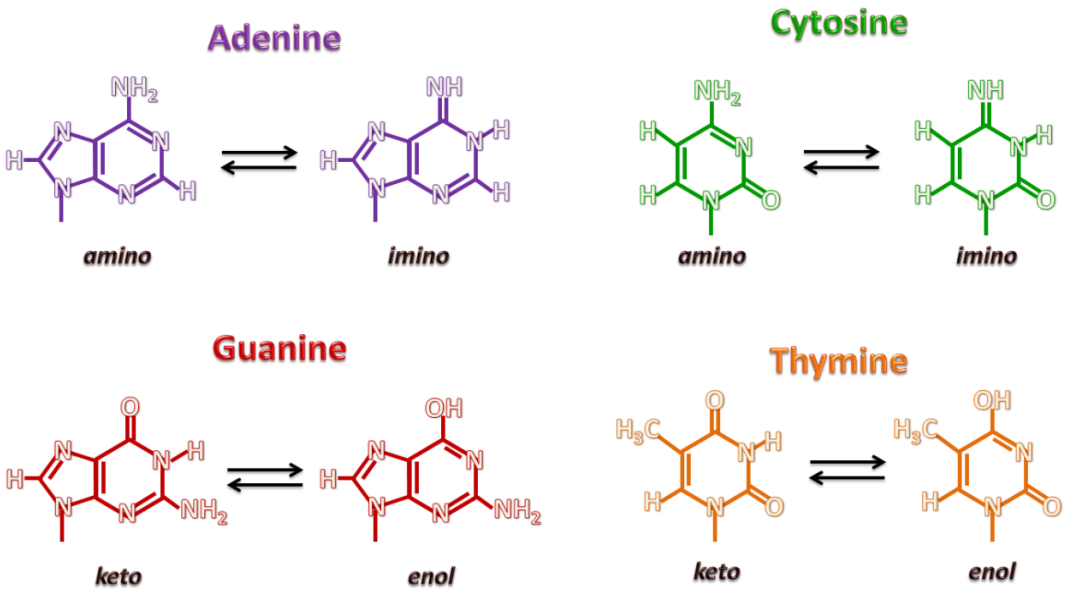


Figure 3. Nucleotide tautomers.

Spontaneous alterations in DNA base chemistry can result from deamination, depurination and depyrimidation, and oxidation. With the exception of thymine, the other four naturally occurring DNA bases, which are cytosine, 5-methylcytosine, adenine and guanine, all contain exocyclic amino groups linked up to their nitrogenous rings. Hydrolysis of these groups, in terms of deamination, occurs spontaneously in pH- and temperature-dependent reactions of DNA [8] and can result in generation of uracil, hypoxanthine, xanthine and thymine respectively (Figure 4).

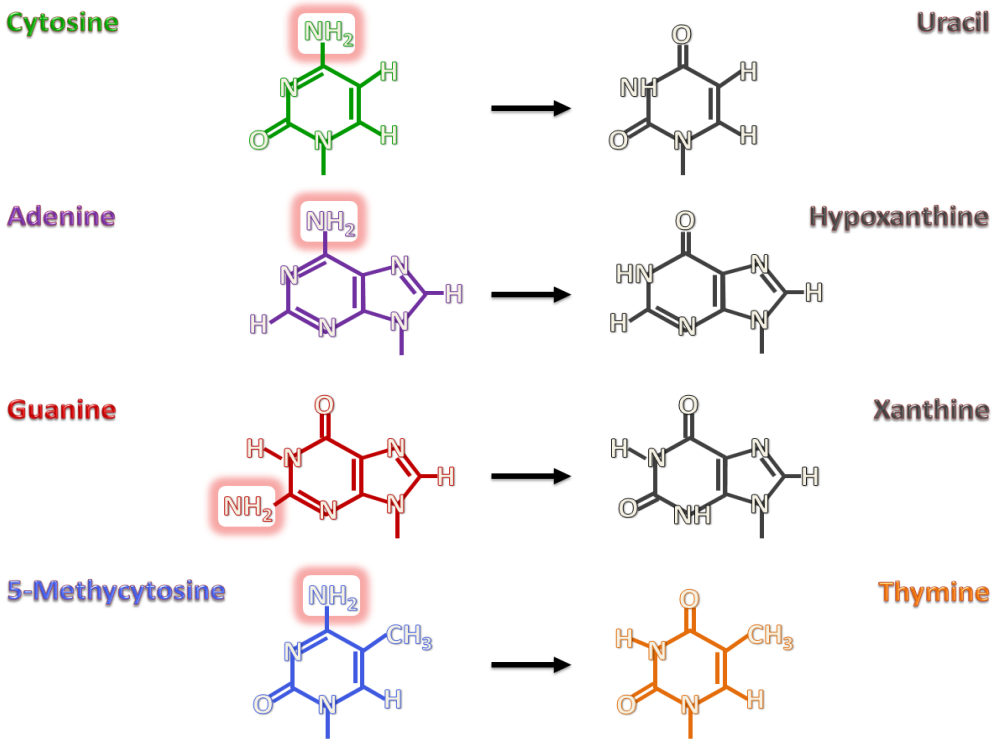


Figure 4. Products generated from the deamination of bases in DNA.

The two former deamination products are the most frequent lesions produced in a cell under normal physiological conditions [14]. It is estimated that 100-500 uracils result from cytosine deamination per cell per day and, if not repaired, can then give rise to C → T transition mutations during semiconservative DNA synthesis due to uracil's inherent complementarity for adenine [14, 15]. Deaminations of adenine and guanine occur at rates much lower than that for cytosine deamination [16]. The resulting

hypoxanthine can cause an A → G transition mutation, and xanthine cannot form stable base pairing with the natural bases and thus may arrest DNA replication and transcription [17]. Deamination of 5-methylcytosine occurs rarely and can generate thymine and hence T:G mismatches.

Base excisions from DNA continuously arise from spontaneous hydrolysis of the *N*-glycosidic bond. These reactions are also known as depurination/depyrimidination and occur at a significant frequency in cells at physiological pH and ionic strength, and can give rise to apurinic/apyrimidinic or abasic (AP) sites. It has been estimated that approximately 9,000 purines are lost per human cell genome per day via hydrolysis of the *N*-glycosidic bond [18], and pyrimidines are lost at 5% of the rate of purines [19]. AP sites are pro-mutagenic, if not repaired, they may induce base misincorporations during DNA replication; furthermore, they are unstable and may also trigger backbone cleavage via a β-elimination process [16].

Reactive oxygen species (ROS) are inevitable by-products of numerous aerobic cellular metabolic processes, such as mitochondrial respiration, in the oxygen-rich atmosphere, in addition, they can also result from water radiation and chemical agents such as paraquat. ROS constitutes the major sources of spontaneous damage to all intracellular macromolecules, including proteins, lipids, carbohydrates and DNA [20-23]. So far, more than 80 ROS-induced DNA base aberrations have been identified [24] and most of which are cytotoxic and mutagenic. For example, thymidine is very subject to the insults by the highly reactive hydroxyl (·OH), which readily adds double hydroxyl bonds to C5' and C6' of the base respectively, resulting in the generation of four isomers of thymine glycol, but the *cis* isomer is predominant (Figure 5) [25]. This unstable moiety can decompose to form various products all of which destabilize the DNA structure [26]. 8-oxo-2'deoxyguanosine (8-oxoG) is

another notable example (Figure 5), which can give rise to transversion mutations as 8-oxoG is highly competent for base pairing with adenine [27].

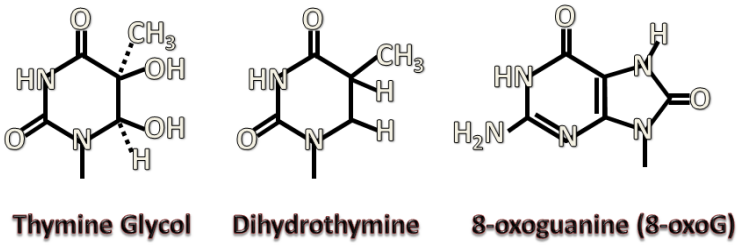


Figure 5. Possible products generated from base oxidation.

Apart from endogenous damage, DNA can also constantly incur a plethora of damage caused by environmental factors including ionizing radiation, UV light and chemical agents. X-rays and γ-rays are the most notable examples of ionizing radiation and can trigger DNA damage by either direct reaction of ionizing energy with DNA or indirectly interactions via the ionizing activation of intermediate reactive species. Direct absorption of the ionizing radiation energy by DNA can result in single and double strand breaks. The latter can cause DNA translocation, partial deletion or loss of a chromosome [28, 29]. Ionizing radiolysis of water or other surrounding molecules can generate numbers of highly reactive oxygen species such as hydroxyl and peroxide radicals and hydrogen peroxide, which can subsequently react with DNA readily, forming aberrant adducts and potentially DNA strand breaks. Hydroxyl radical is the most detrimental intermediate reactive species and responsible for approximately 65% of radiative damage [26]. It has a propensity for attacking the labile C5' and C6' double bond of pyrimidines, creating saturated derivatives, such as thymine glycol and 5, 6-dihyrothymine [30, 31]. Purine residues are less vulnerable to ionizing radiation mutagenesis of the bases, but fragmentation of the bases can occur to form formamidopryrimidine (FaPy) moieties [32] (Figure 6). In addition, ionizing radiation can trigger base cyclization of the base with the deoxyribose-phosphate

backbone [33, 34] (Figure 6). This restricts the DNA flexibility and hence can arrest many enzymatic actions.

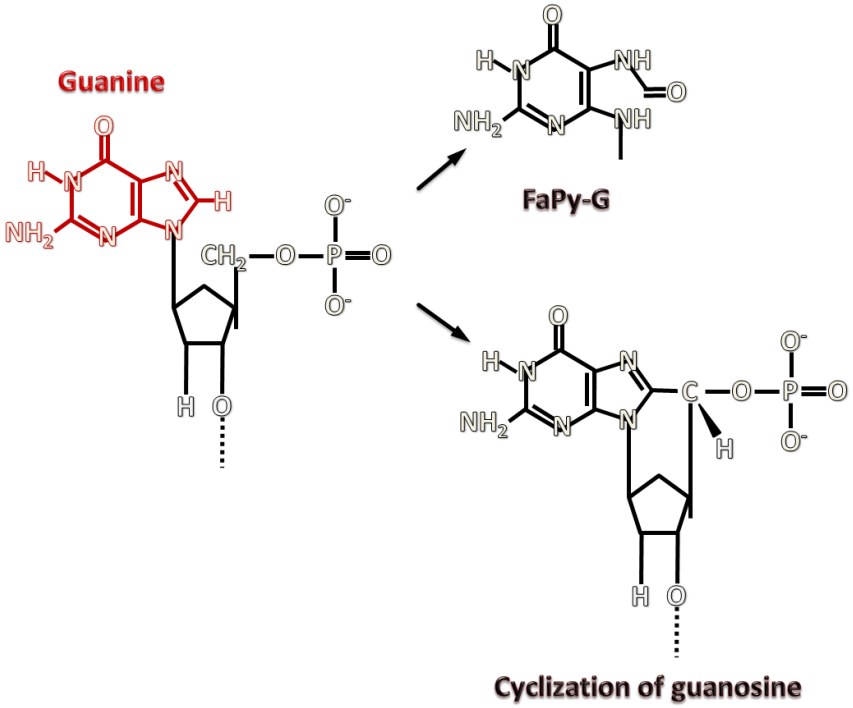


Figure 6. Products generated from ionizing radiation of guanosine.

Solar UV radiation ranges in wavelength from 100 nm to 400 nm and can be subdivided into three different wavelength bands, UV-A 315-400 nm, UV-B 280-315 nm and UV-C 100-280 nm. UV-A and UV-B can penetrate Earth's atmosphere and cause a variety of mutagenic and cytotoxic DNA lesions, especially at current period, constant depletion of the stratospheric ozone causes intensity of this radiation rises at the Earth's surface. UV radiation triggers two of the most abundant mutagenic and cytotoxic DNA lesions, cyclobutane pyrimidine dimers (CPDs) and pyrimidine pyrimidone photoproducts (6-4PPs) and their Dewar valence isomers [35-37]. CPDs occur between neighboring pyrimidines, forming a four-member saturated ring structure (Figure 7). These dimers bend the DNA backbone and further rigidify the DNA helix that can disrupt specific protein-DNA interactions and arrest DNA processing events, such as DNA recombination, replication and transcription [38]. Of these various CPD isomers, cis-syn conformation is predominant. The occurrence

frequency of CPDs depends on the identity of the nucleotides and the local DNA sequence. Thymidine dimers are the most prevalent and more likely to be found in pyrimidine rich areas. A wide variety of thymine dimers have been reported with a variety of linkages and orientations between the adjacent residues [39]. In the CPDs the nitrogenous rings are parallel, but can be either a *cis* or *trans* position, whereas, 6-4PPs are formed via perpendicular dimerization of the adjoining nitrogenous rings between C6' and C4' positions (Figure 7). UV-induced purine photoproducts have also been discovered, but they are much more susceptible to hydrolysis, breakdown and subsequent repair, so they can not persistently exist in the genome [40].

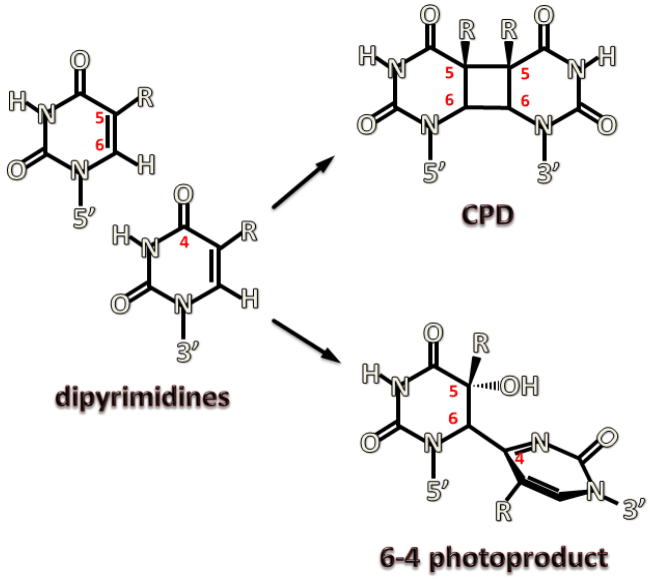


Figure 7. Two possible products generated from thymine dimerization.

Chemical agents can bring about a plethora of DNA damage via covalent chemical modification which is normally irreversible and can only be repaired by removing the erroneous lesion. Alkylating agents, such as ethylmethylsulfonate (EMS), ethylnitrosourea (ENU), *N*-methyl-*N'*-nitro-*N*-nitrosoguanidine (NMNG) and mitomycin C, are the most noteworthy DNA-damaging electrophilic chemicals and can carry out either monofunctional or bifunctional interactions with DNA. Monofunctional agents have a single reactive group and can readily add alkyl groups to on specific site on the nitrogenous bases; whereas bifunctional agents have an

additional reactive group, therefore such agents are potentially capable of interacting covalently with two base sites [41]. Many potential reaction sites for alkylation have been identified in all of four canonical DNA bases, such as N1, N3, N6 and N7 in adenine; N1, N2, N3, N7 in guanine; N3 and N4 in cytosine and N3 in thymine, but they exhibit a variety of reactivity [42, 43]. The electrophilic attacks on these sites can result in base methylations, typical examples are listed in Figure 8. Some chemical agents, such as nitrous acid, nitrogen mustard and sulfur mustards, mitomycin C, *cis*-platinum and certain psoralens, can also cause both intra and inter strand cross-links [44-48] which can inhibit DNA strand separation and hence completely block DNA replication and transcription.

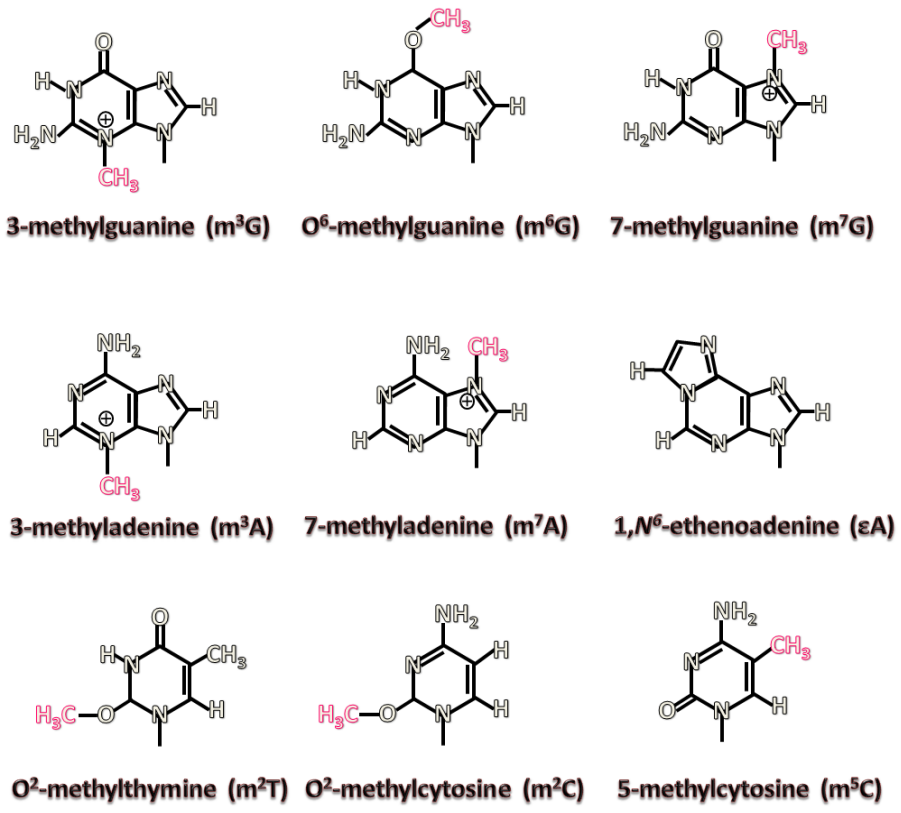


Figure 8. Typcial products generated from alkylation of four canonical bases.

1.3 Biological responses to DNA damage

"We totally missed the possible role of… [DNA] repair although… I later came to realise that DNA is so precious that probably many distinct repair mechanisms would exist [49]." This retrospect was written by Francis Crick in 1974 and manifest that DNA was prevailingly deemed to be a fundamentally stable macromolecular entity and certainly not subject to frequent insults at the early stage of DNA studies. However, upon a large number of subsequent studies on DNA damage and repair, the old perception had been tremendously overturned and DNA was ultimately acknowledged to be highly susceptible to a huge spectrum of damage that arises either spontaneously or from exposure to genotoxic environmental agents. It has been estimated that each of the ~10^{13} human cells incurs tens of thousands of DNA lesions every day [50]. In order to protect the DNA's pristine state and maintain the overall genome integrity, cells have evolved a multitude of mechanisms by which either damaged DNA is removed from the genome or the potentially lethal consequences arising from interference with general DNA metabolism are otherwise mitigated (Figure 9).

DNA repair is the most predominant biological response to combat DNA damage, by which the restoration of normal nucleotide sequence and DNA structure can be achieved. DNA repair can proceed via either the reversal of DNA damage or the excision of damaged elements. While few lesions are repaired by the direct protein-mediated reversal, such as photoreactivation by DNA photolyases and alkylation reversal by DNA alkyltransferases, the majority of DNA damage is handled by numerous stepwise catalytic events of certain enzymes via at least one type of DNA excision repair modes. On the basis of the biochemical and mechanistic modes distinctiveness, DNA excision repair is divided into three modes, which are base

excision repair (BER), nucleotide excision repair (NER) and mismatch repair (MMR).

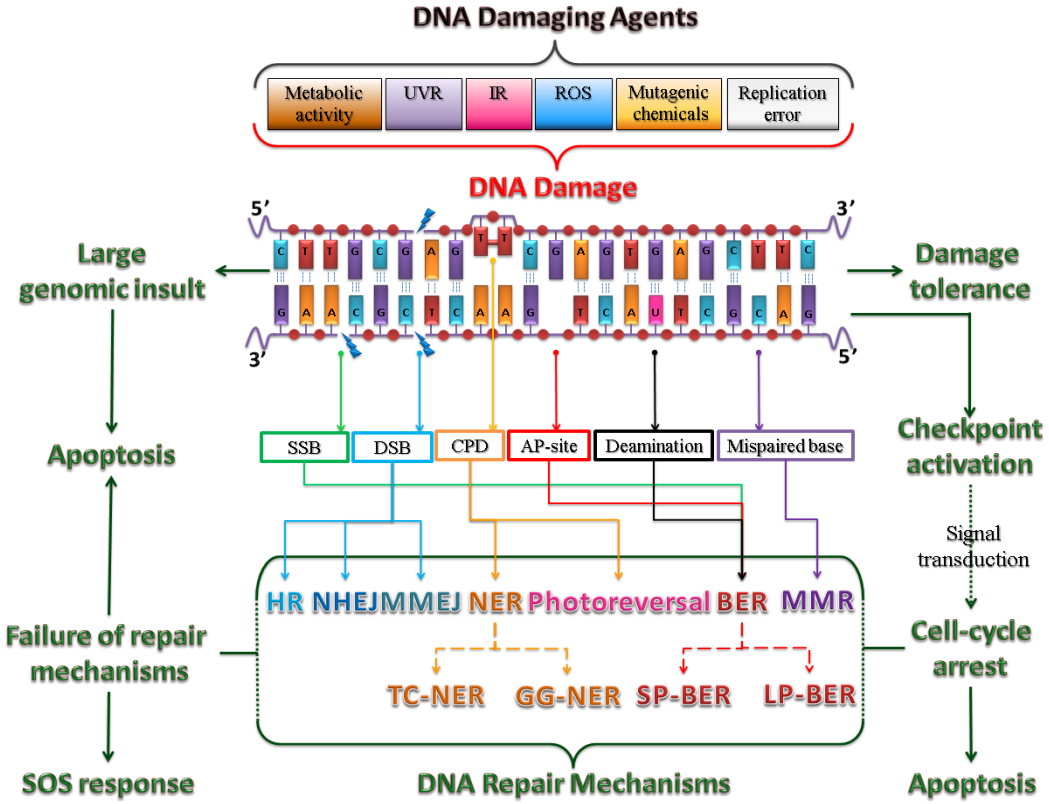


Figure 9. General scheme of biological responses to DNA damage.
Genomic lesions arise from a huge spectrum of DNA damaging agents and trigger a multitude of specific repair mechanisms to conserve the genomic integrity. In case of severe damage and/or failure of repair mechanisms, both eukaryotic and prokaryotic cells undergo apoptosis or prokaryotic cells induce a complex series of genotypic and phenotypic changes, in terms of SOS response. Sometimes the potentiality of lesions in the genome is mitigated by a biological response known as damage tolerance, during which DNA lesions are recognized by certain repair machinery, allowing the cells to undergo normal replication and gene expression without the removal of damage from DNA. The biological response to DNA damage may activate cell-cycle checkpoint by means of a network of signalling pathway that gives the cell extra time to repair the damaged DNA or may induce cell suicide response/programmed cell death (PCD).

MMR acts as a contingent for correcting rare base substitution mismatches and insertion/deletion mismatches that have evaded proof-reading during DNA replication in organisms from bacteria to mammals. MMR proteins also participate in

several other transactions that involve heteroduplex recombination intermediates, such as repair of double-strand breaks in DNA and anti-recombination.

NER removes primarily bulky, helix-distorting adducts, such as pyrimidine dimers, *cis*-platin intercalators, psoralen cross linkers and other large lesions. The NER process involves the action of about 20-30 enzymes in successive steps of damage recognition, local opening of the DNA double helix around the lesion, incision of the damaged strand on either side of the lesion, repair synthesis of the correct nucleotide sequence and ligation [51]. NER can proceed via two different pathways, transcription-coupled repair (TCR-NER) and global genome repair (GGR-NER) [52]. The former pathway specifically targets lesions that block transcription and the latter removes DNA damage from any place in the genome.

BER is responsible for removing most small base lesions by a multitude of orchestrated and sequential catalytic activity of certain enzymes, including excising the base, nicking the DNA backbone, processing the flanking termini, filling the nucleotide gap and ligating the backbone to restore the DNA duplex [53]. BER is elaborated later in this chapter due to it being the study focus of this project.

In addition to damaged nucleotides, cells also frequently sustain fracture of the sugar-phosphate backbone, resulting in either single strand or double strand breaks. Double strand breaks (DSBs) are one of the most detrimental DNA lesions. They can completely compromise the genome integrity and certainly threaten the viability of cells. In response to the threats of DSBs, two principal mechanisms have been evolved: non-homologous end-joining (NHEJ) and homologous recombination (HR) [54]. In NHEJ, DSBs are recognized by the Ku protein that subsequently binds at the damaged end of the DNA and activates the protein kinase DNA-PKcs, leading to the synapsis and activation of end-processing enzymes, polymerases and DNA ligase IV.

A less-well-known Ku-independent NHEJ pathway, called microhomology-mediated end-joining (MMEJ) or alternative end-joining, has also been found, which always causes sequence deletions [55]. Both NHEJ and MMEJ are highly error-prone because they re-join DSBs through direct ligation of the DNA ends without any requirement of sequence homology, however, both pathways function in any phase of the cell cycle in higher eukaryotes including mammals.

By contrast, HR is generally restricted to S and G2 phases of cell cycles in eukaryotes, and it repairs DSBs by using a sister-chromatid sequence as the template to retrieve genetic information and hence ensure essentially accurate repair. In *E.coli*, the DNA transaction of HR occurs in a sequential and coordinated manner, and is initiated by a single-stranded 3' DNA overhang generation, which is promoted by the RecBCD protein complex, also known as Exonuclease V. Then, RecA facilitates the single-stranded 3' DNA overhang to invade the undamaged template and, following the DNA synthesis by DNA polymerase δ, ε, branch migration by RuvA and RuvB, Holliday junctions nicking by RuvC and final substrate sealing by DNA ligase [56].

Eukaryotes have a number of enzymes homologous to RecBCD recombination proteins including Rad51, Rad52, and the breast-cancer susceptibility proteins BRCA1 and BRCA2 [57]. In addition, HR is also employed to restart stalled replication forks and to repair inter-strand DNA crosslinks, the repair of which uses the exonucleases RecJ and RecQ proteins in *E. coli* [58], and the Fanconi anaemia protein complex in eukaryotes respectively [59].

All living organisms are not only capable of removing damage from DNA by their intrinsic DNA repair pathways, but also have evolved multiple strategies for tolerating aberrant bases in DNA, all of which mitigate the potential lethal consequences of base damage arrested DNA replication and involve some perturbation of normal DNA

synthesis. The ability of cells to tolerate DNA damage is biologically as important as their ability to repair damage, and is carried out by a set of error-free and error-prone processes collectively referred to as DNA damage tolerance mechanisms.

In one type of error-free mechanism, called recombinational repair, arrested DNA synthesis is reinitiated downstream of a site of arrested replication. The resulting gap in the affected newly synthesized DNA duplex can then be filled by strand exchanges via homologous recombination, between the affected and the unaffected newly synthesized daughter DNA duplexes [60].

Alternatively, arrested replication forks can form a "chicken-foot" intermediate by folding back on themselves and annealing nascent strands, that facilitates copying information from newly replicated template strand instead of the parental damaged template strand. This template switching process is known as replication fork regression. Neither of recombination and replication fork regression mechanisms use the damaged DNA strand as a template for DNA synthesis, therefore, they are both referred to as error-free mechanisms [60].

By contrast the third DNA damage tolerance mechanism, translesion DNA synthesis, the replication machinery always bypasses DNA lesions due to the employment of specific DNA polymerases, such as Polη, Polκ, Polι and Rev1 in eukaryotes, and PolIV, DpoIV, PolII, PolIV and PolV in prokaryotes [61]. These proteins can preserve replication fork stability and hence ensure completion of DNA synthesis in the presence of DNA damage; however, their extremely relaxed fidelity can lead to promiscuous nucleotide incorporation, thereby, translesion DNA synthesis is associated with a high frequency of mutagenesis, showing an error-prone phenomenon [62].

Both DNA repair and DNA damage tolerance mechanisms represent biological responses that directly process damaged DNA in the genome, in addition to these, prokaryotic and especially eukaryotic cells have also evolved responses that help the efficiency of DNA repair and tolerance. In eukaryotes, various types of DNA damage may activate specific cell cycle checkpoints that result in cell cycle arrest, thereby giving the cell extra time to process DNA damage; or even, may finally induce programmed cell death, or apoptosis, that can rid a multicellular organism of cells that have sustained a large mutational burden or extensive genomic instability that may cause undesirable phenotypic consequences for the whole organism.

In prokaryotes, extensive sudden increases in DNA damage induce an emergency repair system, known as the SOS response. This system was the first DNA repair system described in *E.coli* induced upon treatment of bacteria with DNA damaging agents arrest DNA replication and cell division [63, 64]. It is regulated by two key proteins: LexA and RecA. During normal cell growth, the SOS genes are repressed by LexA repressor dimers that bind to a specific consensus operator sequence also known as the SOS box, nevertheless, *recA* and other SOS genes are still able to express very small amounts of the proteins they encode, therefore, there are some RecA proteins constitutively present in dividing cells. Activation of the SOS genes occurs when the cell senses highly accumulated DNA damage. RecA binds to single stranded DNA (ssDNA) at damage sites in an ATP hydrolysis driven reaction, forming RecA-ssDNA filaments. RecA-ssDNA filaments then activate LexA autoprotease activity, which leads to the LexA repressor's self-cleavage from the operator. The loss of LexA repressor induces transcription of the SOS genes but not all at the same time [65]. The first genes to be induced are *uvrA*, *uvrB*, and *uvrD*. These proteins, together with the endonuclease UvrC, catalyze NER as described above and are responsible for the repair of DNA damage induced by UV radiation, such as cyclobutane pyrimidine dimers (CPDs) [35, 66].

If, however, NER does not suffice to mend the damage, as a second defense against DNA lesions, expression of RecA and other homologous recombination (HR) functions are induced more slowly, about 10 fold. The cell division inhibitor SfiA, also known as SulA, is also induced to give the bacteria extra time to repair DNA lesions [67].

Finally, three error prone lesion bypass DNA polymerases PolII, PolIV and PolV are induced respectively [68-71]. These last-ditch proteins allow bacteria to render DNA lesions double-stranded to be replicated, but at the expense of introducing potential mutagenic effects into the genome.

Although the SOS response has only been found in prokaryotes, species from all kingdoms possess some SOS-like proteins, such as p53 tumor suppressor protein, Rad6-Rad18 complex and Rad51 in eukaryotes [72-74]. These proteins participate in DNA repair and exhibit sequence homology and enzymatic activities related to those found in *E.coli*, but are not organized in an SOS system.

In any population of cells, in both unicellular and multicellular organisms, various abovementioned biological responses to DNA damage can be orchestrated by individual cells as an integrated signaling and genome-maintenance network, to detect different classes of stochastic DNA lesions, signal their presence and promote their repair in any specific region of the genome and at any given time.

1.4 Base excision repair

Base damage is the most pervasive form of DNA damage in all living organisms, therefore, BER is quantitatively the most important mechanism of DNA repair. An estimated rate of 10^4 damaging events in each mammalian cell per day underscores the importance of the BER pathway [18, 75]. Furthermore, BER is the most versatile among excision repair pathways and is not only primarily responsible for handling ubiquitous small, non-helix distorting base lesions such as oxidized, alkylated, deaminated and even absent bases, but also encompasses the repair of DNA single strand breaks. These DNA lesions can arise from both endogenous and exogenous sources. It has been found that defects in BER predispose cellular hypersensitivity to certain cytotoxins as well as to genomic instability and heritable deleterious mutations, and may ultimately culminate in apoptosis, aberrant cellular behavior and uncontrolled cellular proliferation, a hallmark of cancer [76]. These facts notably emphasize the significance of evolutionary conservation of BER.

Over three decades ago, the basic mechanism of BER was first characterized in *E.coli* when Tomas Lindahl discovered the bacterial enzyme uracil DNA glycosylase (UDG) in the repair of uracil [14, 77]. This discovery marked a key step in the understanding of DNA damage and repair, and propelled the DNA damage specific glycosylases into the limelight. Subsequent studies in BER revealed that this repair mechanism is also conserved in the eukaryotes including mammals, and can also be reconstituted *in vitro* with cell-free extracts or using purified protein components involved in this system from bacteria to mammals [78, 79].

BER conventionally proceeds via one of three distinct pathways (Figure 10) that are initiated by recognition of different damaged bases, and which then proceed independently and are characterized by coordination between enzymes and pass DNA

substrates from enzyme to enzyme in an ordered fashion. The recognition and removal of base lesions by one of the dedicated DNA glycosylase enzymes, makes BER unique among the excision repair pathways.

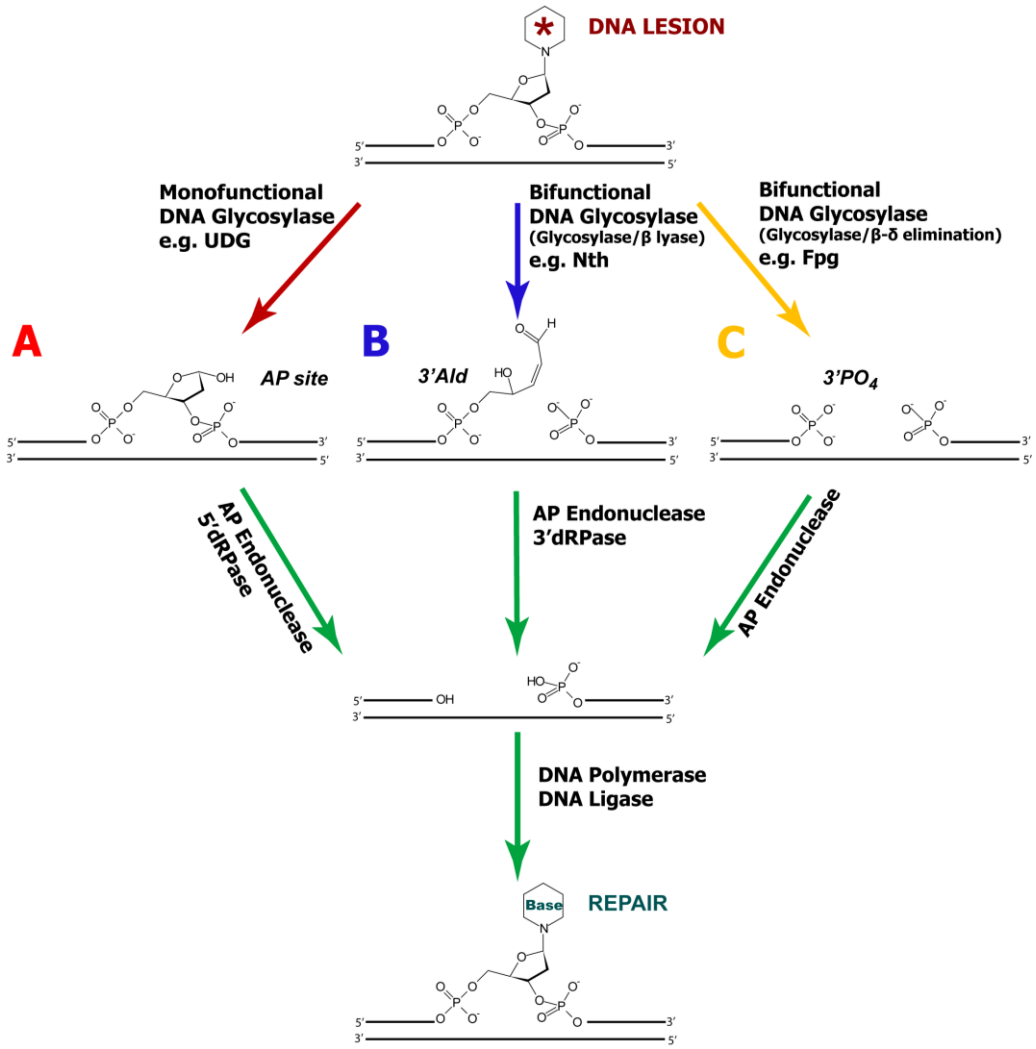


Figure 10. Schematic representaion showing three distinct pathways of BER.
Three distinct pathways of BER are divided on the basis of the functional type of lesion-specific DNA glycosylases. In the pathway (A), monofunctional DNA glycosylase hydrolyzes the *N*-glycosidic bond between the aberrant base and deoxyribose, producing an abasic/apyrimidinic (AP) site; in the pathway (B), bifunctional DNA glycosylases of the Nth family have an associated AP/β lyase activity that cleaves the DNA backbone 3' to the AP site, generating a 3'-unsaturated aldehyde (3'-Ald) via β-elimination; alternatively, bifunctional enzymes of the Fpg family initiate the pathway (C) by removal of aberrant base and then proceed with β-δ-elimination to produce a 3'-phosphate (3'-PO_4). The completion of the three distinct BER pathways is accomplished by the coordinated action of downstream enzymes, including AP endonucleases, dRPases, end-processing enzymes, DNA polymerases and DNA ligases. These downstream enzymes carry out strand incision, gap-filling and ligation.

In the first pathway, mono-functional glycosylases cleave the *N*-glycosidic bond linking the damaged base to the deoxyribose sugar, leaving an abasic or apurinic/apyrimidinic (AP) site (Figure 10A). AP sites are substrates for AP endonucleases, which are mainly classified into two families, the Xth family and the Nfo family. In human, the major AP endonuclease is APE1, which is a member of Xth family corresponding to the *E.coli* exonuclease III (ExoIII). AP endonucleases produce an incision in the duplex DNA by hydrolyzing the phosphodiester bond immediately 5' to the AP site, leaving a 3'-hydroxyl (3'-OH) and a 5'-deoxyribose phosphate (5'-dRP). Then, the 5'-dRP is converted into a 5'-phosphate by either exonucleases, such as *E.coli* exonuclease I (SBcB) and RecJ exonuclease (RecJ), or enzymes with specific DNA-deoxyribophosphodiesterase (dRpase) activity, such as DNA polymerase β (Polβ), *E. coli* 2, 6-dihydroxy-5N-formamidopyrimidine (FaPy), *E.coli* formamidopryrimidine DNA glycosylase (Fpg), endonuclease III (EndoIII) and bacteriophage T4 endonuclease V (EndoV) [80-82].

The two remaining BER pathways involve bi-functional glycosylases, which are primarily specific for oxidative base lesions and possess an additional AP lyase activity. Upon recognition of a base lesion by a bi-functional DNA glycosylase, the enzyme eliminates the damaged base in a mechanism similar to mono-functional DNA glycosylases as described above, and then incises the DNA backbone 3' to the AP site through its lyase activity, generate different products from either of two remaining pathways [82].

Bi-functional glycosylases of the Nth family yield a 3'-unsaturated aldehyde (3'-Ald) via β-elimination (Figure 10B), while members of the MutM/Fpg family catalyze two consecutive elimination steps, called β-δ-elimination, to produce a 3'-phosphate (3'-PO_4) (Figure 10C); both these 3'-terminal blocking lesions are non-ligatable and cannot be extended by DNA polymerase. Consequently they must be further

processed prior to the subsequent step in BER. Therefore, the 3'-Ald product is hydrolyzed by AP endonucleases and the 3'-PO_4 is removed by the bacterial AP endonucleases Xth or Nfo, or polynucleotide kinase (PNK) in mammals, leaving the 3'-OH required for downstream polymerase extension, respectively [83].

Once the damaged base excision step and DNA gap/end tailoring are successively achieved via either of three BER pathways, the resultant gap in the DNA duplex is then filled by a DNA polymerase. This step is referred to as repair synthesis of DNA and branches BER again into two basic modes: short-patch BER (SP-BER) and long-patch BER (LP-BER) (Figure 11). In the SP-BER, a DNA polymerase incorporates a single nucleotide and a DNA ligase seals the nick, thereby restoring the original nucleotide sequence. In eukaryotes, it has been found that XRCC1 serves as a scaffolding protein in the SP-BER and can interact and stabilize DNA ligase IIIα (LigIIIα) and form further interactions with Polβ to complete the repair in a highly coordinated fashion [79, 84]. In the LP-BER, 2-20 nascent nucleotides synthesized by DNA polymerase displace a stretch of old DNA into a 5' flap structure, that is then removed by a flap structure-specific endonuclease (FEN1) prior to ligation.

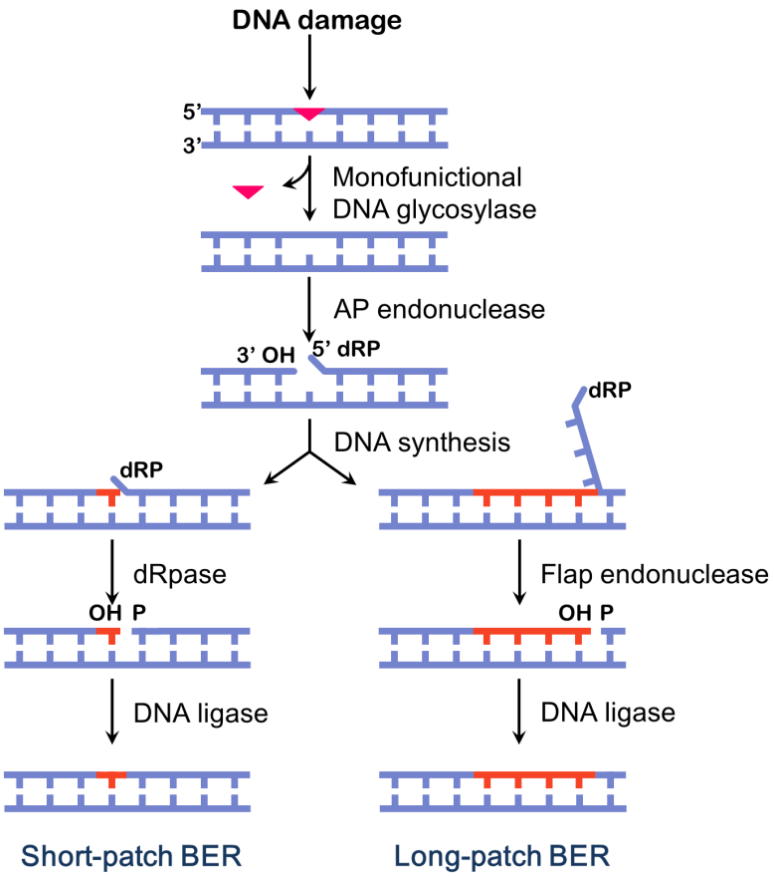


Figure 11. Basic short-pathch and long-patch BER modes.
BER carries out the repair of damaged DNA via two general modes – short patch and long patch. Initially, in the case of monofunctional DNA glycosylase, the damaged base is excised from DNA by cleavage of the *N*-glycosidic bond between the base and a deoxyribose sugar, and then the remaining AP site is removed by AP endonuclease, the following-on DNA/repair synthesis step is where BER branches into short-patch and long-patch modes. The short patch BER leads to a repair of tract of a single nucleotide and is accomplished by the coordinated action of dRpase, DNA polymerase and DNA ligase; alternatively, the long patch BER mode results in a repair tract of 2-20 nucleotides and the newly synthesized DNA strands by DNA polymerase displaces a stretch of old DNA fragment into a flap structure, which is then processed by a flap endonuclease, leaving a nick for DNA ligase.

In eukaryotes, repair mode selection between SP-BER and LP-BER depends on several factors including the lesion specificity, the local concentration of BER components, the cell cycle stage and whether the cell is terminally differentiated or actively dividing [85]. SP-BER predominates and is fully proficient while base excision is mediated by DNA glycosylase/AP lyases [86]. However, some lesions,

such as oxidized or reduced AP sites, are refractory to the dRP lyase activity of Polβ and thus must be processed via LP-BER, in which the DNA strand elongation and ligation are carried out by a multi-protein complex that consists of DNA polymerase δ (Polδ) or polymerase ε (Polε), FEN1, poly (ADP-ribose) polymerase 1 (PARP1) and DNA ligase I (LigI), and is coordinated via the proliferating cell nuclear antigen (PCNA) and Polβ [87-89]. In addition to the lesion specificity, the relative concentration of each repair enzyme also has a strong impact on the BER mode selection; for example, wild-type embryonic mouse fibroblast cell lines can process AP sites predominantly via SP-BER, whereas a mutant cell homozygous for a deletion in the Polβ gene repairs these lesions exclusively via LP-BER [90]. Moreover, it has been found that the selection between SP- and LP-BER modes in *E.coli* is strongly related to the ratio of DNA polymerase I to DNA ligase [91].

1.5 DNA glycosylases

BER is distinct from other excision repair pathways due to its unique initiation executed by DNA glycosylases. Any biological species possesses an array of DNA glycosylases (Table 1), many of which have rather broad substrate specificities, for example, mismatch uracil DNA glycosylase (MUG) from *E.coli* has wide activity against uracil, 3,N^4-ethenocytosine (εC), thymine, 5-hyrdoxymethyluracil, 8-(hydroxymethyl)-3,N^4-ethenocytosine, 1,N^2-ethenoguanine and xanthine [92-97]. Some of these activities are probably redundant in the cell whereas others may serve as potential back-up functions. Nevertheless, in contrast to the promiscuity of most DNA glycosylases for diverse substrates, some DNA glycosylases only recognize a particular class of base damage, a particular inappropriate base, or a particular mispairing. For instance, UV endonuclease from *Micrococcus luteus* is only specific for *cis-syn* isomer of cyclobutane pyrimidine dimers [98], and UDG is only capable of removing uracil from single-stranded and double-stranded DNA [99].

DNA glycosylases in *E.coli*

Gene	Protein (synonyms)	Common name	Examples of activity on DNA	Associated AP lyase
ung$^+$	UNG (UDG)	Uracil DNA glycosylase	Removes uracil from ss- and ds-DNA	No
mug$^+$	MUG	Mismatch uracil DNA glycosylase	Removes uracil, thymine, ethenocytosine or xanthine opposite guanine	No
fpg$^+$ (*mutM*$^+$)	Fpg (MutM)	FaPy DNA glycosylase	Removes oxidized and ring-opened purines including 8-oxoG and formamidopyrimidine (FaPy)	Yes
mutY$^+$	MutY	MutY DNA glycosylase	Removes adenine opposite 8-oxoG	No
nth$^+$	Nth (EndoIII)	Endonuclease III	Removes ring-saturated or fragmented pyrimidines	Yes
tagA$^+$	TagA	3-Methyladenine DNA glycosylase I	Removes 3-methyladenine and 3-ethyladenine	No
alkA$^+$	AlkA	3-Methyladenine DNA glycosylase II	Removes 3-methylpurines, 7-methylpurines, 3- and 7-ethylpurines, ethenoadenine, and O^2-methylpyrimidines	No
nei$^+$	Nei	Endonuclease VIII	Removes ring-saturated or fragmented pyrimidines	Yes

DNA glycosylases in *S.cerevisiae*

Gene	Protein (synonyms)	Common name	Examples of activity on DNA	Associated AP lyase
UNG1	UNG (UDG)	Uracil DNA glycosylase	Removes uracil from ss- and ds-DNA	No
OGG1	Ogg1	8-OxoG DNA glycosylase	Removes oxidized and ring-opened purines including 8-oxoG and FaPy	Yes
NTG1	Ntg1	Thymine glycol DNA glycosylase I	Removes ring-saturated or fragmented pyrimidines, 8-oxoG	Yes
NTG2	Ntg2	Thymine glycol DNA glycosylase II	Removes ring-saturated or fragmented pyrimidines	Yes
MAG1	Mag1	Methyladenine DNA glycosylase I	Removes 3-methylpurines, 7-methylpurines, ethenoadenine, and O^2-methylpyrimidines	No

DNA glycosylases in human cells

Gene	Protein (synonyms)	Common name	Examples of activity on DNA	Associated AP lyase
UNG	UNG (UDG)	Uracil DNA glycosylase	Removes uracil from ss- and ds-DNA	No
SMUG1	SMUG1	Single-strand selective Monofunctional uracil DNA glycosylase	Removes uracil, 5-hydroxymethyluracil	No
MBD4	MBD4 (MED1)	Methyl-binding domain glycosylase 4	Removes uracil or thymine opposite guanine at CPG sequences, thymine opposite O^6-methylguanine	Yes
TDG	TDG	Thymine DNA glycosylase	Removes uracil, thymine or ethenocytosine opposite guanine	No
OGG1	OGG1	8-OxoG DNA glycosylase	Removes oxidized and ring-opened purines including 8-oxoG and FpPy-G	Yes
MYH	MYH	MutY homolog DNA glycosylase	Removes adenine opposite 8-oxoG, 2-OH-adenine opposite guanine	No
NTHL1	NTHL1 (NTH1)	Endonuclease III	Removes ring-saturated or fragmented pyrimidines	Yes
MPG	MPG (AAG, MDG)	3-methylaadenine DNA glycosylase 1	Removes 3-methylpurines, hypoxanthine and ethenoadenine	No
NEIL1	NEIL1	Endonuclease VIII-like DNA glycosylase 1	As for NTHL1; also Fapy-A and 8-oxoG	Yes
NEIL2	NEIL2	Endonuclease VIII-like DNA glycosylase 2	Removes ring-saturated or fragmented pyrimidines	Yes
NEIL3	NEIL3	Endonuclease VIII-like DNA glycosylase 3	?	Yes

Table 1. DNA glycosylases in *E.coli*, *S.cerevisiae* and human cells (adapted from [35]).

DNA glycosylases are relatively small (~25-50 kDa) monomeric proteins that do not require cofactors for their catalysis. Upon recognizing a particular base lesion within the DNA, DNA glycosylases catalyze the excision of the aberrant nucleobases from the phosphoribose backbone by hydrolysis of the *N*-glycosidic bond, generating an AP site. Monofunctional DNA glycosylases only carry out the excision of aberrant bases (Figure 12), whereas bifunctional DNA glycosylases also successively catalyze DNA strand incision via either a β-elimination or β-δ-elimination using their intrinsic

AP lyase activity. Despite their divergent functional types, DNA glycosylases, with the exception of the alkylpurine DNA glycosylase (ALK) family [100], have evolved a common base-flipping strategy to correctly identify and optimally orient their substrates into an active site pocket for catalysis.

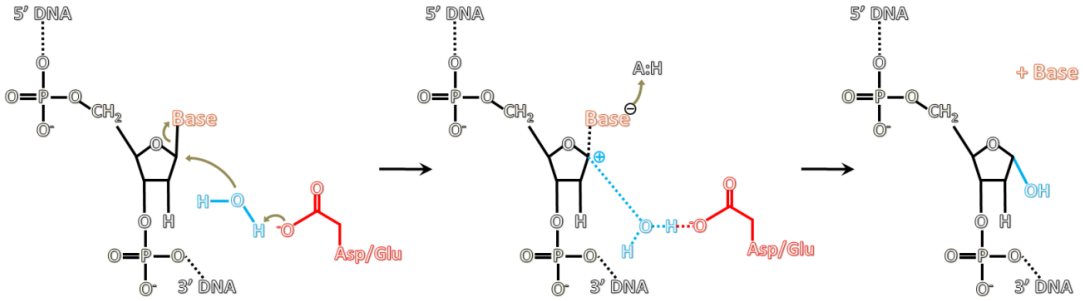


Figure 12. Basic *N*-glycosidic bond hydrolysis mechanism of DNA glycosylases using general acid/base chemistry.
In the simplest sample, an activated water molecule displaces the damaged base to generate an AP site and the free base. A general base (typically an aspartate or glutamate residue) abstracts a proton to activate the water nucleophile for attack of the anomeric C1 carbon of 2'-deoxyribose. A general acid catalyst (depicted as A:H) can accelerate the reaction by protonating the base to make it a better releasing group. (adapted from [35]).

Base flipping is a mechanism by which a target base residue from its base-stacked, hydrogen bonded, introhelical position in normal B-form DNA is swung or "flipped" completely out of the helix into a solvent-exposed, extrahelical position. It was discovered in 1994 when the first co-crystal structure was revealed for a complex of cytosine-5 DNA methyltransferase (M.*HhaI*) bound to a flipped out cytosine [101]. This structure surprisingly demonstrated that the binding to DNA by the enzyme did not distort the DNA in some crude fashion by bending or kinking, but rather the target cytosine had flipped out extrahelically on the axis of the DNA phosphodiester bonds and into the active site pocket of the enzyme.

Since then base flipping phenomenon has been observed in many systems where enzymes need access to a DNA base to exert their activity upon it. Notably, DNA

glycosylases proficiently use this mechanism to excise the aberrant bases from DNA. Based on a growing number of detailed crystal structures of various DNA glycosylases, base flipping mechanisms have been characterized in detail.

Initially, the protein probes the stability of the base pairs via processive interrogation, by either hopping or gliding, along the DNA duplex, which is followed by binding to the DNA phosphate backbone at the specific aberrant base site. Then, the enzyme kinks the DNA by compression of the backbone in the same strand as the lesion, a process that might be facilitated by weakened base pairs or base stacking interactions between the aberrant base and its base pairing partner. The aberrant base is then extruded 180^{o} out of the helix in an enzyme specific direction either through the minor groove, such as uracil DNA glycosylases (UDG) [102], or through the major groove, such as T4 endonuclease V [103]. Eventually, the flipped out base is engulfed deep in the active site pocket of the enzyme through unique energetic cost, and cleaved specifically.

It has been suggested that damaged bases might be more susceptible than normal bases to being flipped out of the DNA helix because a chemical modification could weaken the [π]-electron stacking interactions between the modified base and its adjacent normal base, making it easier to expose the modified base for DNA glycosylases-mediated excision [104]. Alternatively, the enzyme might flip bases out of the helix indiscriminately, but only specific damaged bases can fit snugly into the substrate binding site of the enzyme [105]. Most notably, a side chain of an intercalating amino acid residue of the enzyme normally protrudes into the void left by the flipped out base, thereby physically inhibiting its retrograde motion back into the duplex and promoting active site pocket ‘trapping’ of the extrahelical base. Furthermore, in some cases, the intercalating residue wedges into the widowed base stack opposite the extrahelical base and forms stable hydrogen bonds. Consequently,

the resultant enzyme-substrate complex is not only stabilized via extrahelical base contacts within the active site but also by hydrogen bonding interactions of the intercalating residues.

A steady flow of emerging crystal structures of DNA glycosylases during the past 20 years clearly illuminated that these enzymes have very similar architectural folds and employ a common base flipping mechanism, although they exhibit disparate substrate specificities due to their different active site configurations [106]. Hitherto, six structural superfamilies of DNA glycosylases have been identified on the basis of their divergent conserved architectural folds, including Endonuclease V (EndoV) superfamily, uracil DNA glycosylase (UDG) superfamily, helix-hairpin-helix (HhH) superfamily, helix-two turn-helix (H2TH) superfamily, alkyladenine DNA glycosylase (AAG) superfamily and alkylpurine DNA glycosylase (ALK) superfamily [107].

1.6 Uracil DNA glycosylase superfamily

Uracil DNA glycosylases (UDGs) (EC 3.2.2.3) comprise a prominent and extremely important DNA glycosylase superfamily, because they dominate the specific removal of uracil from DNA. Uracil can arise from spontaneous deamination of cytosine or misincorporation of deoxyuridine triphosphate (dUTP) during DNA synthesis. Among the canonical DNA bases, cytosine is extremely susceptible to deamination, by which its exocyclic amino group is hydrolyzed, resulting in a uracil, that is a fully competent base-pairing partner for adenine, therefore leading to G:C→A:T transition mutations in half of the progeny on replication (Figure 13). A range of 70 – 200 cytosine deamination events are estimated to occur in each human genome per day [108]. Furthermore, over 10,000 dUTP misincorporation issues are believed to arise every replicative cycle per human genome [109]. Therefore, uracil in DNA severely threatens the informational integrity and regulated expression of the genome, unless it is efficiently removed.

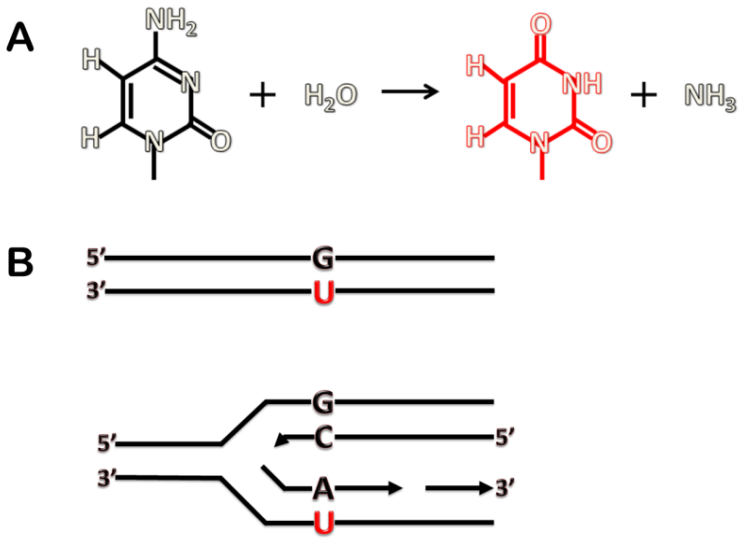


Figure 13. Spontaneous base deamination generated uracil causes G:C→A:T transition mutations.
(A) Hydrolytic deamination of cytosine generates uracil; **(B)** Unrepaired G:U mismatches results in G:C→A:T transition mutations in half the daughter duplexes on replication.

The first identified member of the UDG superfamily is *E.coli* UDG, which is encoded by the *ung* gene [14]. It has been extensively studied over the last three decades and serves as a prototypical member of UDGs. UDGs are widely distributed in all of the kingdoms of nature that emphasizes their importance in combatting the prevalence of uracil in genomic DNA. UDGs can carry out its basal activity in the presence of EDTA due to its independence of metal cofactors. They can remove uracil regardless of its base partner and from either single-stranded or double-stranded DNA, the latter case shows the fastest excision rate. Only few UDGs exhibit merely negligible activity against the canonical DNA bases, and all UDGs are 100% inactive against uracil in RNA, as discussed below.

The topology of the common core of the UDG superfamily adopts a highly conserved single-domain α/β fold, which consists of a central four-stranded parallel twisted β sheet flanked by α helices, and contains a narrow positively charged groove, whose width is approximately identical to that of a DNA duplex, therefore it is ideal for binding double-stranded DNA [110]. Loops connecting β sheet and α helices contain conserved amino acid residues that form the active site and comprise the aberrant base recognition pocket, and contribute to DNA phosphodiester backbone binding. The common α/β fold implies that UDGs have probably evolved from the same ancestral gene, however, on the basis of moderate sequence similarities and exquisite substrate specificities, UDGs can be divided into five families (Table 2).

		Eubacteria		Eukaryotes			
Family	Member	*E.coli*	Thermophiles	*S.cerevisiae*	*S.pombe*	Mammals	Archaea
Family 1	UDG	+	+	+	+	+	−
Family 2	MUG/TDG	+	+	−	+	+	−
Family 3	SMUG	−	−	−	−	+	−
Family 4	TmUDG	−	+	−	−	−	+
Family 5	UDG-b	−	+	−	−	−	+
Mj-UDG family	Mj-UDG	−	+	−	−	−	+
MIG	MIG	−	−	−	−	−	+
MBD4	MBD4	−	−	−	−	+	−

Table 2. Phyletic distribution of five families of uracil DNA glycosylases (adapted from [35]).

Family 1 is composed of the founding representative *E.coli* UDG and its close orthologs such as human UDG, yeast Ung1P and herpes simplex virus type 1 (HSV-1) UDG. Enzymes in this family are the most efficient UDGs [111] and highly conserved from bacteria to human and even viral pathogens. Apart from minor discrepancies, all characterized family 1 UDGs so far appear to primarily confer exquisite specificity toward uracil in both single-stranded and double-stranded DNA, regardless of whether the opposite base is a G or an A [112]. The tendency of the uracil-DNA binding affinity follows the order of ssU > dsU:G >> dsU:A. The repair activity of UDGs is governed by the structural conservation of the catalytic domain motifs. Five conserved motifs have been summarized through inspections of all determined crystal structures of UDGs [109, 113-115]: (i) the water-activating loop (143-GQDPYH-148, human UDG); (ii) the 5'-side backbone compression loop (165-PPPPS-169); (iii) the uracil recognition region (199-GVLLLN-204); (iv) the 3'-side backbone compression loop (246-GS-247) and (v) the leucine loop (268-HPSPLS-273) [112]. UDG binds to DNA via rigid loops in motifs ii, iv and v. The 5'- and 3'- side backbone compression loops compress the DNA backbone (pinch) and slightly bend the DNA, which becomes ~45^{o} bent and ~2^{o} kinked, facilitating the destabilization of the stacked nucleotide conformation; whilst a push from a conserved leucine residue (human UDG Leu272) in the leucine loop following its damage scanning and side chain intercalation via the minor groove, and a pull from the complementary uracil specific recognition pocket, in concert flip the uracil nearly 180^{o} from an intrahelical base stack into an extrahelical position through the major groove of DNA [116-118].

All crystal structures of the family 1 UDGs have revealed an active-site substrate-binding pocket for many selective interactions with uracil. This pocket is made of highly conserved residues in the UDGs catalytic domain, and provides both geometric and electrostatic complementarities to recognize and fit uracil explicitly in an extrahelical conformation, and prevent any other base from lodging in the active

site. The overall shape of the small binding pocket sterically blocks purine bases. The entry of 5-methylated pyrimidines such as thymine is impeded by the virtue of its 5-methyl group because the aromatic side chain of a tyrosine residue can only stack against the unmethylated C5 of uracil via van der Waals' force, becoming a steric block for thymine binding. Removal of this tyrosine residue converts UDG into a thymine DNA glycosylase (TDG) [119]. When pyrimidines move into the pocket, following-on discrimination against cytosine is achieved by a set of specific hydrogen bonding interactions provided by the bottom of the pocket. The O2 carbonyl of uracil forms a hydrogen bond to the enzyme main chain NH that joins a conserved glycine-glutamine sequence. The amide side chain of a conserved asparagine forms specific hydrogen bonds to the N3 ring nitrogen and exocyclic O4 carbonyl of uracil, whereas cytosine is excluded by repulsive interactions with its exocyclic amine N4. A water cluster buried in the base of the uracil binding pocket provides hydrogen bonding interactions that unambiguously fix the orientation of the critical amide group [105, 112].

Following target uracil extrahelical flipping and specific intrapocket selection, the catalytic cleavage of *N*-glycosidic bond is then carried out via two proposed models. A nucleophilic substitution model was proposed for human UDG on the basis of the crystallographic studies of human UDG co-crystals with substrate DNA [113]. In this model (Figure 14), the hydrogen bonds between the 2, 3 and 4 positions in uracil and Asn204, Asp145 and Gln144 in UDG, formed by enforcing stereo electronic coupling of the anomeric and σ-π_{Atom} effects [120], align the nucleotide optimally and meanwhile stabilize the anionic transition state for cleavage. Asp145 and Gln144 contribute to bulk water exclusion from the active site and their backbone amides could create an oxyanion hole for the O2 atom of uracil. The imidazole group of His268 directly attacks the C1' atom of the furanose ring and thus cleaves the *N*-glycosidic bond. Alternatively, based on crystallographic studies of HSV-1 UDG

[109], a water activation model was proposed that a nucleophilic water molecule, which is bound to and activated by Asp88 (human Asp145, *E. coli* Asp64), attacks the C1' position of dUMP using a hydroxyl nucleophile. Protonated His210 (human His268, *E. coli* His187) is hydrogen bonded to uracil O2 through another water molecule, therefore facilitating the *N*-glycosidic bond cleavage by stabilizing the developing oxyanion. The subsequent kinetic studies suggested that the water activation model is also implicated in the catalysis of *E.coli* UDG [121]. The cleavage of *N*-glycosidic bond via either of above two models ends with an abasic site in DNA and a free uracil residue.

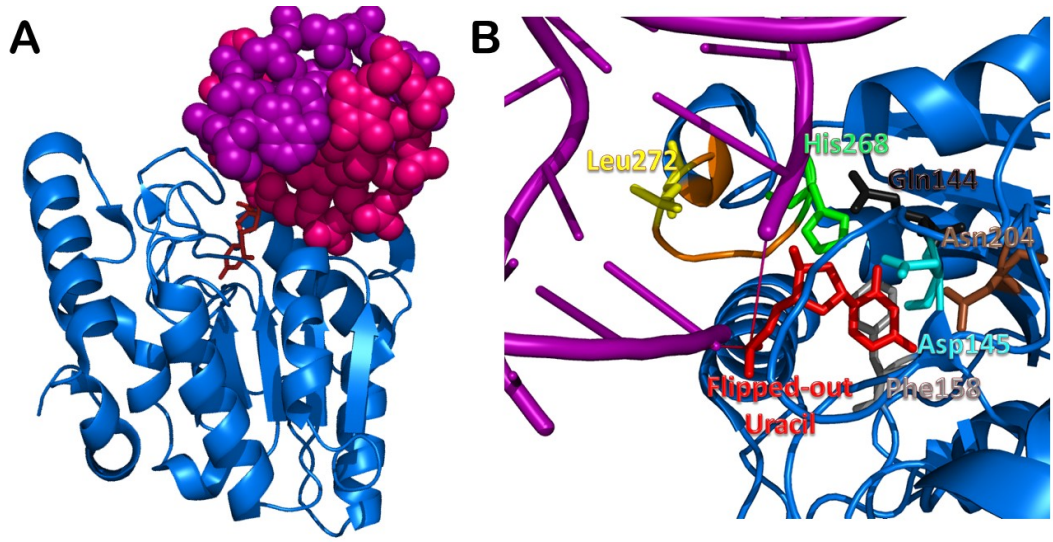


Figure 14. Crystal structure of human uracil DNA glycosylase bound to substrate DNA. (A) Human uracil DNA glycosylase (dark blue) in complex with substrate containing uracil (PDB entry 1EMH). The figure was generated using PyMOL (DeLano Scientific LLC). The two strands of duplex DNA are respectively coloured in purple and pink and viewed from the top. The uracil (red) is flipped out of the duplex DNA into the active site of hUDG. (B) Blow-up of the hUDG active site in the hUDG-DNA complex. Key residues are shown as labeled and coloured and as sticks. The side chain of Leu272 (yellow) in the conserved leucine loop motif (orange) penetrates into the helix and flips the uracil (red) out of the duplex DNA (purple and pink). In a proposed nucleophilic substitution model, the flipped uracil is oriented and stabilized via hydrogen bonding interactions with Asn204 (brown), Gln144 (black) and Asp145 (cyan), His268 attacks the *N*-glycosidic bond. Alternatively, in a proposed water activation model, Asp145 activates a nucleophilic water molecule to cleave the *N*-glycosidic bond, His268 is hydrogen bonded to flipped uracil and facilitates the cleavage. Moreover, the aromatic side chain of Phe158 (grey) can sterically clash with the ribose 2'-hydroxyl group of ribouridine in RNA, therefore, preventing the RNA molecule from binding to hUDG.

Apart from the exquisite specificity for uracil in DNA, UDGs can also remove certain closely related bases at rates some three orders of magnitude lower than that for uracil. These substrates include 5-fluorouracil, 5-hydroxyuracil and alloxan [122-125].

All UDGs do not have activity against uracil in RNA because a highly conserved phenylalanine residue (human Phe158, *E.coli* Phe77 and HSV Phe101) in UDGs' active site sterically clashes with the ribose 2'-hydroxyl group in RNA [126]. In 1995, Savva *et al.* had clearly demonstrated that the aromatic side chain of the Phe101 residue in HSV UDG protrudes into the position where the 2'-hydroxyl group of ribouridine would fit into the active site, leading to a steric hindrance and hence preventing RNA molecules from binding to UDG [109]. This hypothesis was further validated from the structure of human UDG in complex with DNA [110]. This conserved phenylalanine also plays a key role in stabilizing uracil binding in the DNA substrate via π−π interactions (Figure 14) [109].

In contrast to family 1 UDGs, which excise uracil mispaired with any canonical base, the family 2 UDGs, represented by the bacterial MUG and the eukaryotic homolog TDG, exclusively initiate BER of uracil or thymine mismatched with guanine [127, 128]. Furthermore, these enzymes are also primary DNA glycosylases for excision of 3,N^4-ethenocytosine from DNA in both bacteria and humans, but exhibit robust activity against U:G mismatches. MUG and TDG have been reported as showing a broad spectrum of activity against a variety of lesions [92, 94-97], although for many of these, including T:G mismatch, the rates of cleavage are unlikely to be of physiological significance [129]. MUG, accompanied with its homolog TDG, will be elaborated later in this chapter.

Single-stranded-selective monofunctional uracil DNA glycosylase (SMUG) comprises the third family of UDG superfamily. This enzyme has hitherto been characterized from eukaryotes only, including xenopus and human [108, 130]. SMUG is poorly sequence homologous to family 1 and 2 enzymes, but shares a conserved core fold and a mechanism of extrahelical pyrimidine recognition with the UDGs and MUG/TDG [131]. SMUG has been initially depicted as specific for uracil in single-stranded DNA, but later studies revealed the double-stranded substrates are in fact the primary target if AP endonuclease is present to relieve product inhibition of SMUG and thereby dramatically stimulate the enzyme turnover [131]. SMUG has broader specificity than UDG, primarily recognizing 5-hydroxymethyluracil but also excising uracil, 5-formyluracil and 5-hydroxyuracil from DNA [132]. It may serve as a relatively efficient backup of UDG in the repair of U:G mismatch and of single-stranded DNA [133].

Family 4 and family 5 UDGs are known only from a handful of bacterial and archaeal species [134-137]. Family 4 UDGs are different from all other members of the UDG superfamily because they are iron-sulfur (FeS) proteins with the characteristics of a 4Fe4S high potential iron protein center (HiPIP), which plays a role in substrate recognition but not catalysis [138], similar to the FeS clusters in the DNA repair enzymes of Nth/MutY family [139-142], although FeS clusters have been found in a wide range of enzymes primarily as redox active cofactors participating directly in electron-transfer catalytic mechanisms [143]. Enzymes from family 4 and 5 can remove uracil from all DNA contexts, and the latter are also responsible for excising hypoxanthine, 5-hydroxymethyluracil and 1,N^4-ethenocytosine [136].

Further to this, several enzymes in HhH superfamily also exhibit uracil excision capabilities, including the archaeal MjUDG and MIG, and the mammalian MBD4 [35].

1.6.1 *Escherichia coli* mismatch uracil DNA glycosylase and human thymine DNA glycosylase

Uracil DNA glycosylases were previously deemed to possess three characters in common: 1) activity against uracil in both single-stranded and double-stranded DNA; 2) susceptibility to the inhibition of Ugi, which is a protein encoded by *Bacillus*-infecting bacteriophages containing uracil instead of thymine in DNA; and 3) strong conservation of the essential residues implicated in exquisite specificity and catalysis. However, this perspective changed since the discovery of human thymine DNA glycosylase (TDG), which does not exhibit characteristics 1 and 2 [144].

TDG is a 410-residue and 55 kDa DNA glycosylase. It was first discovered as a G:T mismatch specific thymine DNA glycosylase for a CpG context in simian cell extracts [145], and subsequently revealed to be capable of excising thymine from G:T mismatches in other sequence contexts, as well as from C:T and T:T mispairs, but much less efficiently [146]. The later studies demonstrated that TDG also can carry out solely Ugi-insensitive excision of uracil from G:U mismatches [147]. Intriguingly, its catalytic activity for uracil excision from U:G mismatches is ~10 fold higher than that for thymine excision, whereas neither uracil nor thymine in single-stranded DNA nor U:A are substrates of TDG [148]. In addition, TDG is also a primary DNA glycosylase for removal of 3,N^4-ethenocytosine (εC) mismatched with guanine in both bacteria and humans, and it exhibits an overall more efficient catalysis for εC than for thymine [149]. The human TDG encoded gene maps on chromosome 12 from position 12q22 to q24.1, and its essential active site residues for thymine excision are located in a 304 residue N-terminal domain. The deletion mutagenesis of amino acids from the C-terminal end and N-terminal end of hTDG produced a core enzyme of 248 amino acids that had lost TDG activity but still retained double-stranded specific UDG activity [150].

Subsequent bioinformatics studies of this core enzyme identified its homologues in the Gram-negative bacteria *E.coli* and *Serrattiamarcescens*. Despite both bacterial protein sequences being significantly shorter than hTDG by approximately 120 amino acids at the N-terminus and 130 amino acids at the C-terminus, they still share greater than 30% sequence identity in their core.

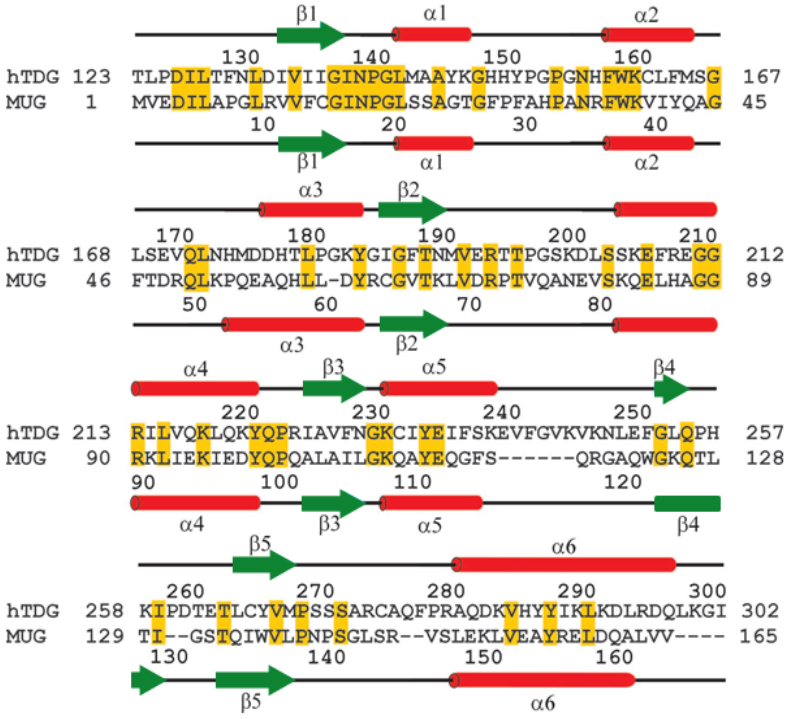


Figure 15. hTDG and MUG squence alignment.
Above both hTDG and MUG sequences, the secondary structure is shown with α-helices as red cylinders and β-sheets as green arrows. Conserved residues within both enzymes are yellow shaded.

Like hTDG, these two bacterial proteins are completely resistant to the Ugi inhibition and inactive on single-stranded substrates [127, 151]. The hTDG *E.coli* homologue was initially suggested to be subservient to the more efficient UDG to specifically excise uracil from a U:G mismatch hence it was so named mismatch uracil DNA glycosylase (MUG). However, MUG was subsequently found to be capable of processing a relatively broad range of aberrant bases mispaired with guanine *in vitro*, including uracil, 3,N^4-ethenocytosine (εC), thymine, 5-hyrdoxymethyluracil, 8-(hydroxymethyl)-3,N^4-ethenocytosine, 1,N^2-ethenoguanine and xanthine [92-97].

The crystal structures of MUG enzyme from *E.coli* offer an insight into MUG structural homologies, DNA binding and substrate recognition and active site capacity, and provide the basis for the identification of the second homologous UDG family (family 2) [128]. This new family of TDG/MUG enzymes possesses some distinct properties, such as being insensitive to Ugi inhibition, being exclusively specific to the double-stranded substrates, and lacking the conserved catalytic residues of the family 1 UDGs.

The 1.8 Å refined crystal structure of *E.coli* MUG unveiled a β-α-β topology, which is in common with many nucleotide binding proteins. It consists of a central fold of 5 stranded β-sheet flanked on both faces by α helices. The 77% score of the MUG's Sequential Structure Alignment Program (SSAP) algorithm [152] with HSV1 UDG indicated that the central fold of MUG structure has a high degree of structural homology to UDG, and suggested that MUG and UDG structures have evolved from a common ancestral fold. On the other hand, only ~10% sequence identity between MUG and UDG implies the family 1 and family 2 enzymes diverged early in evolution [128]. Like the UDGs, one face of MUG possesses excessive positive charges and is traversed by a channel connecting with a narrow substrate binding pocket, which penetrates back into the core of the enzyme. Structural alignment of the *E.coli* MUG and the HSV1 UDG revealed that the architecture of the MUG binding pocket is almost identical to the UDG active site pocket, and the highly conserved active site sequences in UDG have precise topological equivalents in MUG, suggesting that MUG's pocket may employ a similar base flipping mechanism like UDGs to process the base lesions.

A subsequent 2.85 Å refined crystal structure of MUG-substrate complex provided more detailed insights into the subtle and important roles of MUG active site pocket residues. A face and edge on one side of the MUG pocket are formed by the sequence

16-GINPG-20, which is completely conserved in the TDG/MUG enzyme sequences, and also corresponds to the sequence 86-GQDPY-90 in HSV1 UDG. The opposite face of the pocket is formed by the side chain of Phe30, which has a counterpart residue Phe101 in HSV1 UDG. Although this Phe30 residue is not conserved in all known MUG/TDG family enzyme sequences, its corresponding residue is always aromatic and therefore can make stable π-π stacking interactions with the planar pyrimidine ring of the flipped bases in the active site pocket, according with observations in UDG-base complexes [109, 110].

The entrance of the MUG active site pocket is formed by Gly20 (16-GINPG-20) (Figure 16), which corresponds to the conserved residue Tyr156 in hTDG. Tyr156 is two residues downstream of the catalytic aspartate (152-GQDPY-156) and its aromatic side chain can pack tightly against the C5 position of uracil, resulting in the steric exclusions of the 5-methyl group of thymine and the imidazole rings of adenine and guanine. Mutation of this residue to a smaller side chain confers thymine DNA glycosylase activity to human UDGs [153]. Unlike Tyr156 in hTDG, Gly20 in MUG cannot provide the same steric exclusion for the 5-methyl group of thymine. This thus explains the thymine DNA glycosylase activity of MUG and provides the first glimpse into MUG's broader substrate specificity (Figure 16).

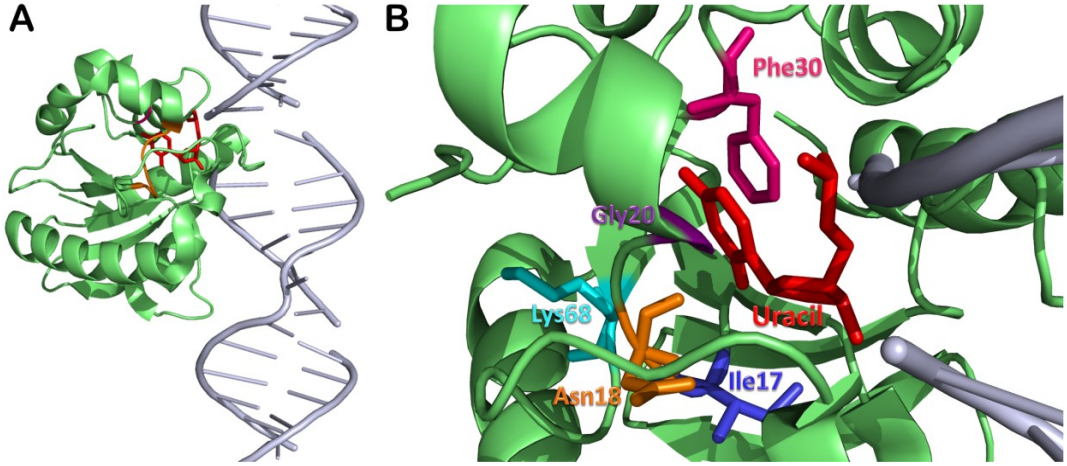


Figure 16. Crystal structure of *E.coli* mismatch uracil DNA glycosylase (MUG) in complex with uracil containing DNA (PDB entry 1MWJ).
(A) Overall structure of MUG (green) bound to a G: U substrate (grey) analog shows the

uracil nucleotide (red) is flipped deep into the active site. A face and edge on one side of the Mug active site pocket is shaped by the sequence 16-GINPG-20 (orange) and the residue Phe30 (pink) forms the opposite face. (B) Close-up view of the active site in the MUG-DNA complex. Key residues are shown as labeled and coloured and as sticks. Gly20 forms the entrance of the MUG active site pocket and unlike its equivalent residueTyr156 in UDG doesn't provide steric exclusion for the 5-methyl group of thymine. Lys68 at the bottom of the binding pocket and contributes to the discrimination against cytosine. Phe30 forms stable π-π stacking interactions with O4 of uracil flipped into the active site pocket. Ile17 forms non-specific hydrogen bonding interactions with the O2 of uracil. Asn18 is the catalytic residue of MUG. The figure was generated using PyMOL (DeLano Scientific LLC).

Tight binding of uracil and specific discrimination against cytosine in the *E.coli* MUG active site pocket result from several hydrogen-bonding interactions. MUG forms non-specific hydrogen bonding interactions between Ile17 (16-GINPG-20) and the O2 of uracil and also specific hydrogen-bonding interactions between Phe30 and O4 of uracil (Figure 16), thereby uracil is tightly bound in the active site. A non-conserved lysine residue (Lys68) forms the bottom of the binding pocket together with the main chain from residues 6 – 69 and provides specific hydrogen-bonding interactions between its ϵ-amino group and the exocyclic 4 position of a bound pyrimidine, favourably with the O4 of uracil or thymine and repulsively with the N4 of cytosine. Therefore, Lys68 contributes to the discrimination against cytosine [93]. The equivalent residue in UDG is a conserved asparagine (Asn147 in HSV1 UDG), which makes similar favourable and repulsive interactions to exclude cytosine [153].

Although structural homology between MUG and UDGs, particularly around the mouth of the base-binding pocket, suggests their catalytic mechanisms seem at least relatively conserved. However, the conserved catalytic residues aspartate and histidine in UDGs (Asp145 and His268 in human UDG) become two asparagine residues (Asn18 and Asn140) in *E.coli* MUG (Figure 16). *E.coli* MUG Asn18 and Asn140 are not charged and thus cannot perform the general acid/base catalytic mechanism like the corresponding aspartate and histidine in UDGs. The 2.85 Å refined structure of MUG-βFU-DNA complex revealed that, unlike the conserved

histidine in UDGs, Asn140 is unable to protonate the O2 of uracil as it doesn't move at any stage of the base-excision process and is > 4 Å from the O2; nor is Asn140 conserved in the MUG/TDG family enzymes, therefore, the involvement of Asn140 in catalysis is very unlikely [93]. Whereas, Asn18 in MUG is conserved in all known MUG/TDG family sequences and the 2.85 Å refined structure showed that a water molecule bound between the side-chain amide and the main-chain carbonyl of Asn18 is perfectly aligned for a nucleophilic attack on the C1 of dUMP, and in-line displacement of the base from the opposite of ring, in a manner similar to that by Asp88 in human UDG [93].

Like other nucleotide flipping proteins, the gap 'vacated' in the DNA duplex by the flipped out uracil is occupied by two *E.coli* MUG residues Gly143 and Leu144, which insert between the two bases flanking the scissile nucleotide. The peptide oxygen of Leu144 meanwhile hydrogen bonds the head group of Arg146, which inserts its side chain between the "widowed" guanine of the G:U mismatch and the preceding base, and also is hydrogen-bonded to a water molecule. The water molecule is in turn hydrogen-bonded to the deoxyribose ring oxygen of the widowed guanine. These hydrogen bonding interactions stabilize a MUG intercalating "wedge" that consists of the residues Gly143, Leu144, Ser145, Arg146 and the associated water molecule, and penetrates the base stack of the DNA duplex from the minor groove and opposite the widowed guanine once the deoxyuridine is displaced. Unique among other DNA glycosylases, the intercalating wedge in *E.coli* MUG forms strong specific hydrogen bonding interactions between the N1 imino group of the widowed guanine and the carbonyl oxygen of Gly143, and between the N2 exocyclic amino group of the widowed guanine and the carbonyl oxygens of Gly143 and Ser145 (Figure 17). These specific interactions between the intercalating wedge and the non-lesion DNA strand explains the activity preference of Mug for double stranded DNA. Furthermore, it

suggests that the intercalating wedge provides a push mechanism that facilitates the base flipping-out [93].

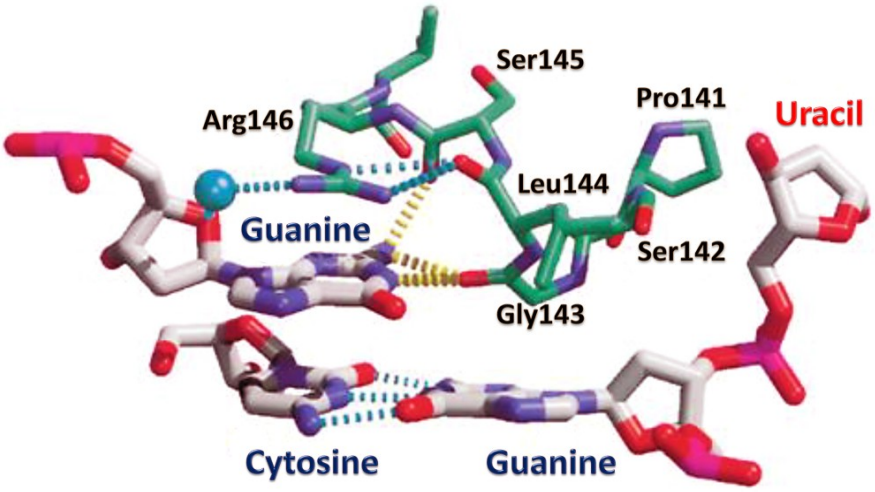


Figure 17. Hydrogen bonding interactions between the widowed guanine on the complementary strand and the MUG enzyme.
An intercalating wedge consists of Gly143, Leu144, Ser145 and Arg146 with an associated water molecule in *E.coli* MUG (green) facilitates the flipping out of uracil and makes specific hydrogen bonding interactions (yellow dash line) with the widowed guanine on the complementary strand. (adapted from [128]).

The mouth of the MUG active site pocket similar to that in UDGs is also occluded by the proximal DNA strand in both substrate analogue and product complexes, however, unlike that UDG requires prior dissociation of the abasic-DNA product to release their excised uracil base through the open mouth of the pocket [117], four surface-exposed N-terminal residues in MUG form a transient hole at the bottom of the pocket and allow the exit of an excised base, independent of the release of the abasic product DNA. This alternative 'back-door' escape route for excised bases could be linked to the observation that MUG remains tightly bound to its abasic-DNA product [93]. This post-catalytic feature impedes MUG's turnover and its activity appears to reach a plateau as the [product]/[enzyme] ratio approaches a specific stoichiometry, which still remains unsettled to date [154].

One of the notable aspects of BER is that the intermediate AP lesions produced by DNA glycosylases are frequently more unstable and thus potentially more damaging than the initial base lesion. To minimize the escape of this unstable intermediate, several DNA glycosylases, including the human thymine DNA glycosylase (hTDG), human uracil DNA glycosylase (hUDG), human OG glycosylase (hOGG1), *E.coli* mismatch uracil DNA glycosylase (MUG) and the adenine glycosylase MutY have been reported to remain tightly bound to their products, thus provoking a exceedingly rate-limiting enzymatic turnover [117, 146, 155-157]. It has been suggested that this may be a general protective mechanism, whereby coordination of enzyme activity in BER is achieved through displacement of the DNA glycosylases by the downstream AP endonucleases.

Previous studies of BER in organisms ranging from *E.coli* to human have shown that AP endonucleases can stimulate the turnover of many DNA glycosylases *in vitro* [9, 117, 158-162]. For example, human AP endonuclease/exonuclease I (hAPE1) specifically stimulates the activities of human UDG and TDG [117, 162]. These enzyme partnerships may be a biological necessity to ensure that toxic damaged DNA intermediates are never left exposed in the cell. These important findings provide a glimpse into enzyme communication in the initial steps of BER, however, a precise mechanistic understanding remains elusive.

1.7 *Escherichia coli* AP endonucleases

BER can be conceptually divided into two distinct stages. In the first and damage-specific stage of BER, as above described, DNA glycosylases excise the aberrant base and produce the central BER intermediate: an apurinic/apyrimidinic (AP) site. In the second and damage-general step of BER, the DNA backbone is incised 5' to the AP site to generate a 3' OH end for initiation of downstream DNA repair synthesis and ligation.

The hydrolytic breakage of the *N*-glycosidic bond to produce AP sites is the most frequent and significant structural abnormality that occurs in cellular DNA. Apart from being produced by the hydrolytic activity of DNA glycosylases, AP site moieties can also arise from spontaneous depurination/depyrimidination. In human lung fibroblasts, the number of AP sites in the steady state is about $0.67/10^6$ nt (~2000 AP sites/human cell) [163]. The presence of AP sites generally impedes DNA replication and transcription and may also be potentially mutagenic and cytotoxic due to lack of base coding information [164]. The first evidence regarding the existence of a specific class of enzymes capable of processing AP sites came from DNA repair studies in *E.coli*. Subsequent studies indicated that the majority of AP site repair is carried out by class II apurinic/apyrimidinic (AP) endonucleases, namely 5' AP endonucleases, which are essential components of BER [26, 165].

E.coli possesses two major families of AP endonucleases, Exonuclease III (ExoIII) family and Endonuclease IV (EndoIV) family, which are classified on the basis of amino acid sequence homology and named for their archetypal members. ExoIII and EndoIV families are structurally dissimilar, indicating that they evolved independently and not from a common ancestor. Unlike ExoIII family members present in all kingdoms of life, enzymes from EndoIV family only exist in lower

organisms but are curiously absent in plants, mammals and some other vertebrates [166]. In *E.coli*, 90% of AP endonuclease activity is provided by ExoIII, but EndoIV is capable of incising AP sites under stress conditions [167, 168].

ExoIII was the first discovered AP endonuclease. It was initially described as a 3'→5' exonuclease with an associated phosphatase activity, and the major cellular AP endonuclease activity was thought to come from endonuclease II (EndoII); however, a later study found that the same mutation of the ExoIII encoded gene *xthA* disrupts both exonuclease and endonuclease activity, and thus in 1976 Bernard Weiss attributed the two activities to one enzyme, which was ExoIII [169].

ExoIII is the major constitutively expressed 5' AP endonuclease in *E.coli* and accounts for ~90% of the total 5'AP endonuclease activity in the cell. In addition to its 3'→5' exonuclease and 5' AP endonuclease activities, the versatile ExoIII also possesses two other nucleolytic activities, which are 3' phosphodiesterase activity and ribonuclease H activity [170-172]. ExoIII is encoded by the *xthA* gene and is a 28 kDa monomeric small protein. It requires divalent cations, optimally Mg^{2+} or Mn^{2+}, to catalyze DNA backbone hydrolysis and thus can be inhibited by metal-chelating agents, such as ethylenediaminetetraacetic acid (EDTA) and ethylenediamine-*N,N'*-disuccinic acid (EDDS) [173]. ExoIII also exhibits dependency on temperature, as it can be inactivated rapidly upon incubation of the purified protein at moderately high temperatures [174].

The crystal structure of ExoIII [172] revealed it has a characteristic four-layered α,β-sandwich fold (Figure 18), which is conserved in its human homologue APE1 [175]. APE1 is also known as HAP1, APEX and REF1 [176-179] and contributes the overwhelming majority of 5' AP endonuclease activity in humans [176]. ExoIII and APE1 share 27% sequence identity. Except for the 3'→5' exonuclease activity, APE1

does possess the rest of ExoIII nucleolytic activities. Expression of human APE1 in *E.coli* ExoIII/EndoIV-deficient mutant cells restores cellular resistance to methyl methanesulfonate (MMS), thus indicating a conserved mechanism among species [176, 180].

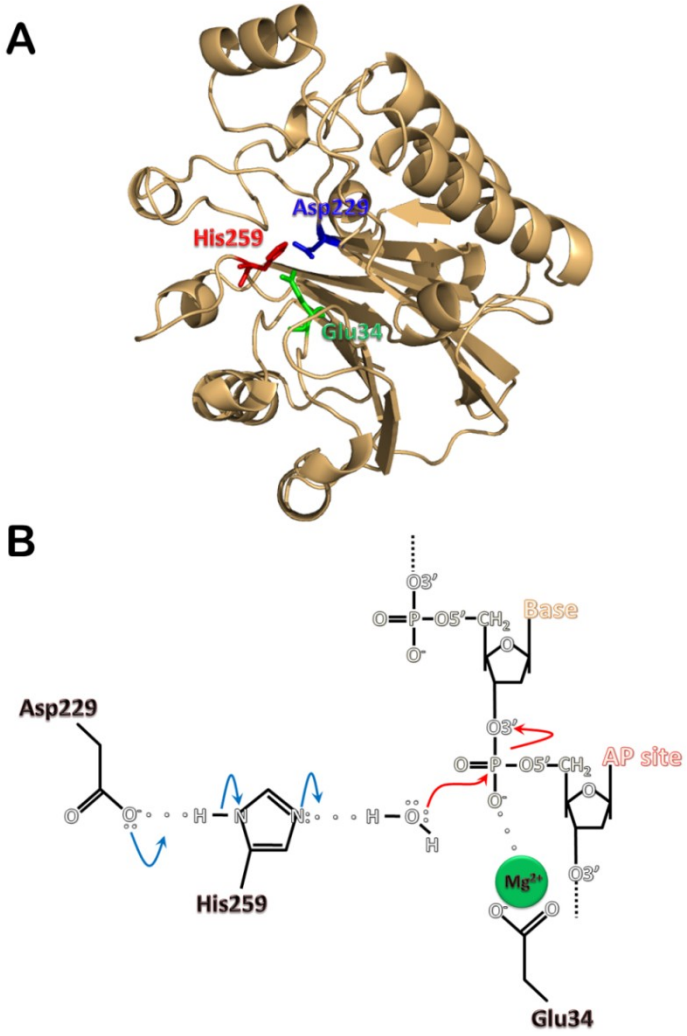


Figure 18. *E.coli* exonuclease III crystal structure and AP site cleavage mechanism.
(A) An overview of *E.coli* exonuclease III structure (PDB entry 1AKO). The figure was generated using PyMOL (DeLano Scientific LLC). The key active site residues Asp229, His259 and Glu34 are coloured in dark blue, red and green. (B) Structure based reaction mechanism for phosphodiester bond cleavage for the *E.coli* exonuclease III family of 5' AP endonucleases. Asp229 makes a hydrogen bonding interaction with His259 which in the meanwhile abstracts a proton from a water molecule. The resulting nucleophilic hydroxide ion then attacks the phosphate group. The metal ion (Mg^{2+}) bound by Glu34 interacts with the negatively charged phosphate group and aids the nucleophilic attack of the hydroxyl group (adapted from [172]).

The 2.6 Å refined crystal structure of ExoIII-Mn^{2+}-dCMP complex revealed that ExoIII cleaves O3' of phosphate bond 5' to the AP site through a nucleophilic attack aided by a single divalent metal ion [172]. ExoIII initially binds at a random non-specific site in the DNA duplex via its positively charged complementary surface, and then searches for an AP site while translocating on the DNA strand via sliding or hopping. The enzyme recognizes an AP site by a conserved tryptophan residue Trp212 in its vicinity of the catalytic center. An indole loop of the Trp212 intercalates into the AP site pocket in the DNA minor groove and slightly distorts the DNA structure via bending and kinking, facilitating the flipping out of the AP site into a sequestered enzyme catalytic pocket [181].

The divalent metal ion binding site and active site of ExoIII are located at the bottom of the groove between the two six-stand â-sheets. Polar residues within the active site form hydrogen bonds with a number of ordered water molecules, particularly within the catalytic pocket surrounded by Gln112, Asn153, Tyr109, Asn7, His259 and Trp212. Side chains of Asn153 and Gln112 hydrogen bond to the nucleotide O3' position, meanwhile the Asn153 side chain also interacts with the AP site 5'-phosphate, which are meanwhile hydrogen bonded with Asn7, Asp151, His259 and Tyr109 [172]. These hydrogen bonding interactions from the conserved ExoIII residues stabilize the extrahelical conformation of the AP site and effectively lock ExoIII on the AP-DNA, resulting in a Michaelis complex [181]. Subsequently, His259 hydrogen bonds to Asp229, as well as abstracts a proton from a water molecule, comprising an Asp-His-H_2O catalytic triad, resulting in a nucleophilic hydroxide ion which eventually attacks the phosphate group 5' to the AP site. The single metal ion Mn^{2+} bound by Glu34 functions to orient the phosphate group, stabilize the transition state and also polarize the P-O3' bond, therefore facilitating the nucleophilic attack of the hydroxyl and the release of O3' leaving group [172, 182].

Most ExoIII active site residues are conserved in APE1, therefore ExoIII and APE1 have converged upon a functionally appropriate domain, the C-terminal domain, encompassing a single divalent metal ion active site and dictating the 5'AP endonuclease activity. However, APE1 possesses another functionally specific domain, the N-terminal domain, dispensable for the AP site cleavage, to activate the DNA binding of some oxidized transcription factors *in vitro* [183].

The hunt for the EndoIV family of the AP endonucleases was initiated by the observation in *E.coli* that *xthA*-null mutant cells still have ~10% of the total cellular 5' AP endonuclease activity. EndoIV was isolated from crude extracts derived from these mutant cells [174]. EndoIV is encoded by the *nfo* gene and is a small ~30 kDa monomeric, Zn^{2+}-dependent protein. Unlike the divalent metal ion-dependent ExoIII, EndoIV has a resistance to the EDTA inhibition in normal assay conditions [184]. In common with ExoIII, EndoIV is also a multifaceted enzyme and endowed with additional 3'→5' exonuclease activity and 3' diesterase activity, however, its 3' diesterase activity always remains lower than that of ExoIII [166]. EndoIV is the only known enzyme that is capable of repairing damaged nucleotides with bases in the α configuration, such as α-deoxyadenosine, which is a major anoxic radiolysis product of adenine in DNA and results from hydroxyl radical attack at the deoxyribose C1' atom, leading to inversion of the chirality of the sugar from the normal β configuration [185]. Moreover, EndoIV can repair the DNA damage caused by the antitumor drug bleomycin [186]. Although *E.coli* EndoIV only contributes to 10% of total 5' AP endonuclease, it is notable that this enzyme expression level can be induced more than 20 fold by O_2^- or SoxRS regulon under oxidative stress conditions, therefore enhancing the capability of cells for repairing oxidative DNA damage or damage that is refractory to enzymatic processing by ExoIII [168, 187].

The most notable difference between ExoIII and EndoIV families is that EndoIV employs a trinuclear Zn^{2+}-coordinated catalytic mechanism to cleave the phosphodiester bond, whereas ExoIII uses a single Mg^{2+} or Mn^{2+} ion. The key features that governs DNA binding and catalysis of *E.coli* EndoIV have been revealed by its crystal structures, which are the only determined structures among the EndoIV family enzymes to date. The first crystal structure of *E.coli* EndoIV was determined at ultra-high 1.0 Å resolution and clearly revealed that EndoIV has a single domain αβ barrel fold, in which eight parallel β-strands are surrounded by eight peripheral α-helices (Figure 19A). This type of structure was first observed in Triose Phosphate Isomerase (TIM), so it is also known as $\alpha_8\beta_8$ TIM barrel structure.

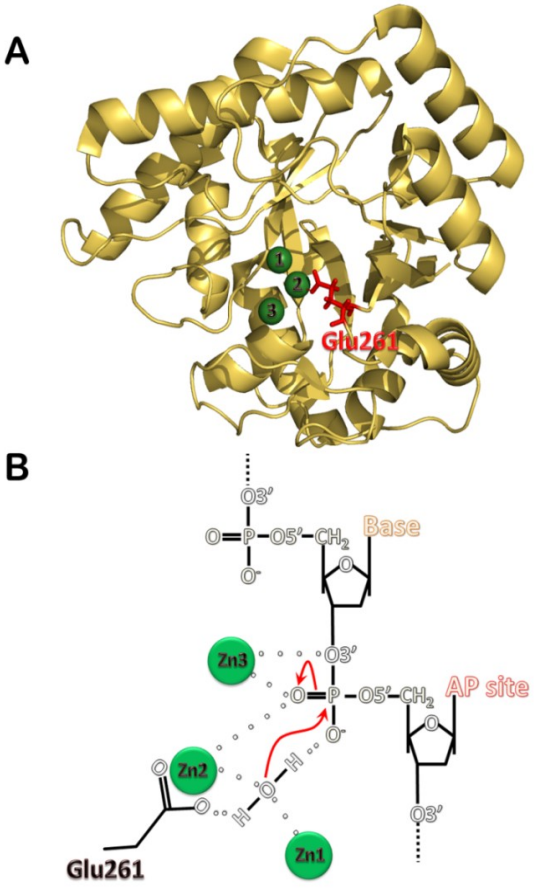


Figure 19. *E.coli* endonuclease IV crystal structure and AP site cleavage mechanism. (A) An overview of *E.coli* endonuclease IV structure (PDB entry 1QTW). The figure was generated using PyMOL (DeLano Scientific LLC). The key active residue Glu261 is coloured in red and the trinuclear metal ions are numbered and highlighted in green spheres. (B) Structure based three-metal-ion mediated reaction mechanism for phosphodiester bond

cleavage for the *E.coli* endonuclease IV family of 5' AP endonucleases. A water molecule coordinated by Zn1 and Zn2 makes the nucleophilic attack of the scissile phosphate group through a pentacoordinate transition state stabilized by all three Zn^{2+} ions. The phosphate oxygen bridging Zn2 and Zn3 remains bound to these metal ions, but Zn3 reverts its distorted coordination sphere after hydrolysis to an ideal geometry to hold cleaved DNA product and stabilize the resulting 3'hydroxyl group. Glu261 ligates Zn2 and hydrogen bonds a water molecule to activate and orient the hydroxide nucleophile for in-line attack on phosphate group.

The subsequent 2.45 Å crystal structure of EndoIV-substrate complex [188] revealed that EndoIV inserts Arg37 and Tyr72 loops from minor groove into the DNA base stack, bending the DNA at a 90^{o} angle and kinking the DNA to allow specificity for the target and facilitate the AP site flipping out. Once the AP site is flipped out, Arg37 stacks with the base pair 3' of the AP site and Tyr72 stacks with the base pair 5' of the AP site, to in concert stabilize the distorted DNA substrate. Moreover, Tyr72 plays a key role in catalysis by filling the space vacated by the flipped-out AP site sugar and phosphate moieties that thus shields the active site from bulk solvent.

Further substrate binding of EndoIV is mediated by its five R-loops (residues 10-13, 34-38, 70-73, 149-153, 224-230) that emanate from the C-terminal face of the â barrel. The extrahelical displacement of the AP site and the intricate sets of contacts of five R-loops to both DNA strands function to anchor the liberated scissile 5' phosphate into the active site metal ion cluster, which consist of three Zn^{2+} ions ligated by the side chains of the conserved aspartate, glutamate and histidine residues in a deep depression near the center of the β barrel.

To incise the phosphodiester bonds, the geometry of the EndoIV trinuclear Zn^{2+} cluster is exquisitely tuned to a pentacoordinated trigonal bipyramidal transition and all three Zn^{2+} ions participate in catalysis. Two Zn^{2+} ions (Zn_1 and Zn_2) coordinate a water molecule, which acts as the nucleophile to incise the DNA strand. The third Zn^{2+} ion (Zn_3) reverts its distorted coordination sphere after incision to an ideal

geometry to hold cleaved DNA product and stabilize the resulting 3'hydroxyl group, which is the primer terminus for downstream DNA polymerases. Glu261 is crucial for the catalysis because it coordinates one Zn^{2+} ion (Zn_2) and hydrogen bonds a water molecule to activate and orient the hydroxide nucleophile for in-line attack on the phosphate group [188].

Notably, the two different AP endonuclease families have evolved independently to process AP sites in analogous ways, including base flipping and unstable product DNA binding, these findings may therefore highlight the structural adaptations to the biological need for both creating and controlling proper 3'-OH ends to prime DNA repair synthesis. On the other hand, as evolutionary selective pressure acts at the level of biological function rather than single enzyme activity, the biological necessity to coordinate AP endonucleases with other DNA BER enzymes may guide evolution of their unrelated structures and distinct metal ion-coordinated catalysis mechanisms.

1.8 Hypothesis

There has been a great deal of interest regarding communication between enzymes during BER, namely, BER enzymes exists in an 'ordered' pathway model' in which each member implements a coordination with a specific downstream enzyme so that substrates are handed-off and repair fully proceeds in a concerted, rather than stepwise, manner. This is particularly pertinent where DNA glycosylase activity is normally inhibited by its high affinity for the AP product. It is conceivable that the *in vivo* turnover rate of DNA glycosylases should be higher than that observed *in vitro*.

Many previous studies of BER have suggested that the activity of DNA glycosylases can increase in the presence of the downstream AP endonucleases *in vitro*, yet the mechanism remains largely unknown. The project reported herein aims to address a hypothesis that enzyme communication between DNA glycosylases and AP endonucleases delineates BER pathways. Despite the fact that there is much evidence to demonstrate obvious stimulation of the DNA glycosylase activity by AP endonucleases, coordinating the initial steps of BER in *E.coli* has not been studied in detail. Along with its genetic tractability, *E.coli* has a typical product-inhibited member of UDG family 2, MUG and both prototypes of two major AP endonuclease families, ExoIII and EndoIV. Therefore, the work presented herein set out from desire to investigate these *E.coli* enzymes to specifically elucidate three major outstanding questions, aiding the validation of the project hypothesis:

1. How does MUG bind to its representative substrates and AP product?
2. Is the turnover of MUG against its representative substrates stimulated in the presence of either ExoIII or EndoIV?
3. How does ExoIII or EndoIV impact on turnover of MUG in the context of communication during BER?

Chapter 2
Differential MUG binding modes to specific and non-specific DNA

2.1 Background and objectives

In light of the hypothesized enzyme communication that delineates BER pathways, the MUG binding mode to specific and non-specific DNA is of significant interest. Despite DNA glycosylases have generally been assumed to be monomeric functional enzymes, moderately cooperative binding has been observed for human O^6-alkylguanine-DNA alkyltranferase (AGT) [189, 190]. Furthermore, TDG, the human homolog of MUG, has been crystallized in a 2:1 complex with a 22-mer DNA containing a tetrahydrofuran nucleotide (THF), which is a chemically stable mimic of the natural AP product [191]. This crystal structure revealed that one protein subunit bound at the AP site by forming numerous contacts with both target and complementary strands, meanwhile the other bound at an undamaged site through predominant interactions with the complementary strand and the contacts were less extensive than for the former subunit binding (Figure 20).

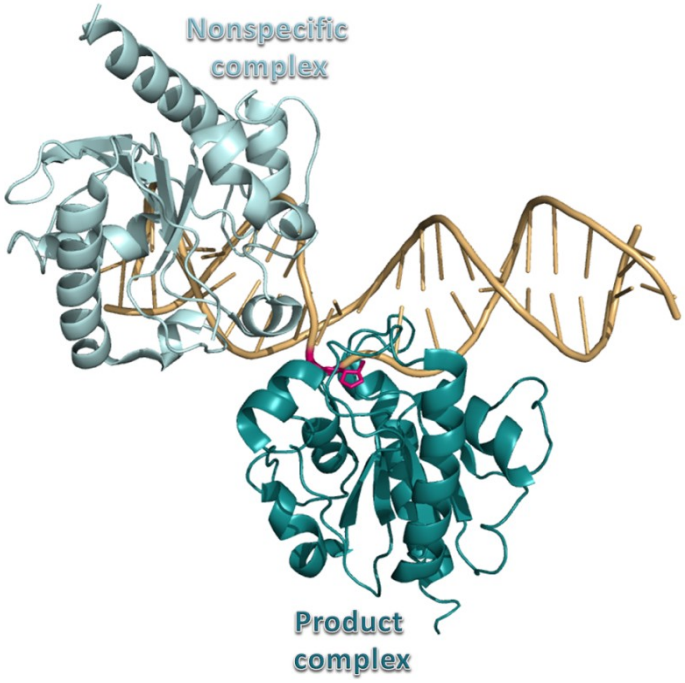


Figure 20. Overview of TDG dimeric complex with abasic product DNA.
TDG can bind the abasic DNA in a 2:1 complex, one subunit at the abasic site (magenta) to form a product complex (deep teal) and the other at an undamaged to form a non-specific complex (light teal). (PDB entry 2RBA).

A small size symmetrical dimer interface without apparent hydrogen bonds or salt bridges between the two TDG subunits was also observed in the 2:1 complex. The subsequent isothermal titration calorimetry and band shift studies indicated that TDG can bind AP-DNA with 1:1 or 2:1 stoichiometry [191]. The latest kinetics assays suggested that 2:1 binding stoichiometry for TDG repair activity is dispensable and a single subunit is fully capable of locating and processing G:U or G:T lesions [192], therefore suggesting the 2:1 stoichiometry may contribute to other critical biological roles, such as the mechanism by which APE1 simulates TDG activity, or it may just simply be an artifact because all DNA-acting enzymes will bind DNA non-specifically to some extent.

In the first reported co-crystal structures of MUG-DNA complex [128], two MUG molecules were observed to bind to the uracil containing DNA. Intriguingly, although the DNA used in these co-crystals in principle could form a blunt-ended self-complementary duplex with two central G:U mismatched base pairs, however in practice, it formed an overlapping pseudo-continuous "nicked" DNA helix via an alternative base pairing offset by six nucleotides, generating G:U mismatches (Figure 21A). In spite of the formation of this unusual DNA structure, it enabled every MUG enzyme in the co-crystals to bind to an abasic site. Interestingly, there are clear contacts between neighbouring pairs of MUG enzymes on the same extended DNA molecule. This was not reported in the original paper because the authors assumed the observed dimeric binding in these co-crystals was attributed to some interactions of each MUG enzyme with the crystal lattice while it binding to an abasic site, rather than a functional necessity. Therefore, this co-crystal structure of MUG-DNA complex was published as a 1:1 complex (Figure 21B).

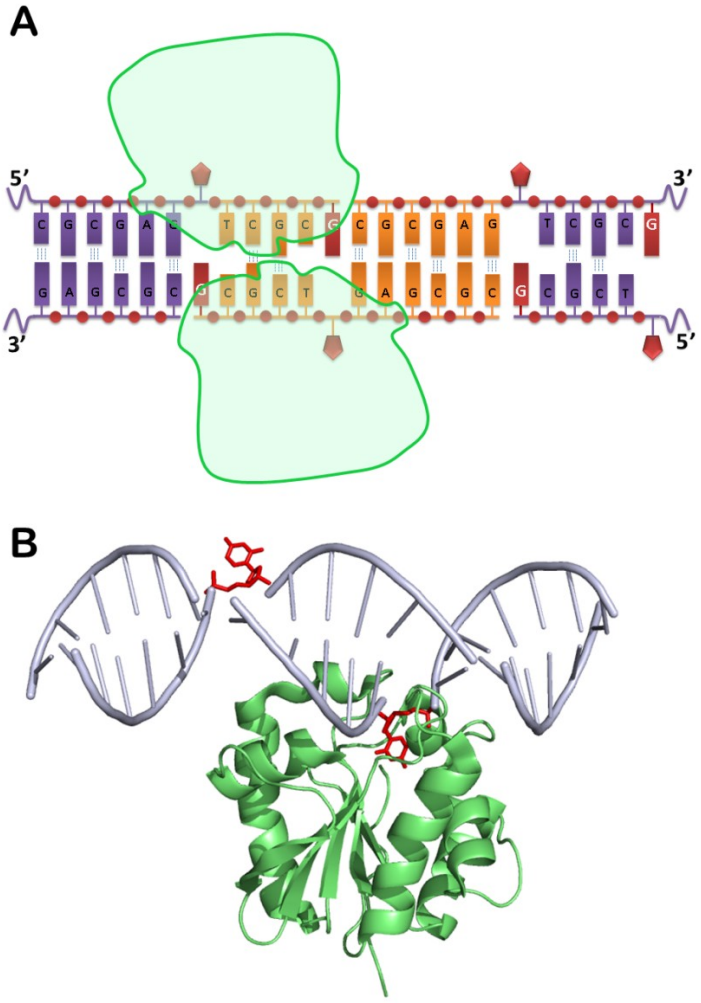


Figure 21. MUG-DNA structure.
(A) In the first reported MUG-DNA co-crystals, the 12-bp uracil containing DNA (yellow) practically formed an overlapping pseudo-continuous "nicked" DNA helix via an alternative base pairing offset by six nucleotides, generating G:U mismatches. The flipped uracil bases are red pentagons, and the opposing guanines are red. Two MUG molecules (green) had been observed to bind the DNA in a 2:1 complex in the co-crystals. (B) The published MUG-DNA crystal structure was a 1:1 complex [128], the flipped uracil bases are highlighted in red (PDB entry 1MWJ). The figure was generated using PyMOL (DeLano Scientific LLC).

Previous observations of the 2:1 binding stoichiometry of TDG and MUG enzymes in their respective co-crystals certainly gave hints about potential functions of higher order binding complexes that might not have been tested otherwise, and provided guidance for many early studies on MUG-DNA interactions in our lab. Our preliminary fluorescence anisotropy results suggested a positive cooperativity in

MUG binding to its abasic product DNA [193]. It would be desirable to further understand how MUG binds to DNA underlying considering how MUG processes lesions and how it interacts with downstream AP endonucleases in BER pathway, therefore, this chapter is devoted to depicting a detailed analysis of MUG binding mode to specific and non-specific DNA. A fluorescence anisotropy assay was firstly employed to characterize the binding mode of the MUG enzyme to fluorophore labeled oligonucleotides. Binding stoichiometry of MUG-DNA complex was then studied using band shift assays. Finally, the impact of stoichiometric binding on MUG catalytic activity was investigated.

Fluorescence-based techniques are extensively applied in the research of protein-ligand interactions due to their inherent sensitivity, and the fact that they can be implemented at true equilibrium conditions. Fluorescence is best explained with the use of a Jablonski diagram shown in Figure 22.

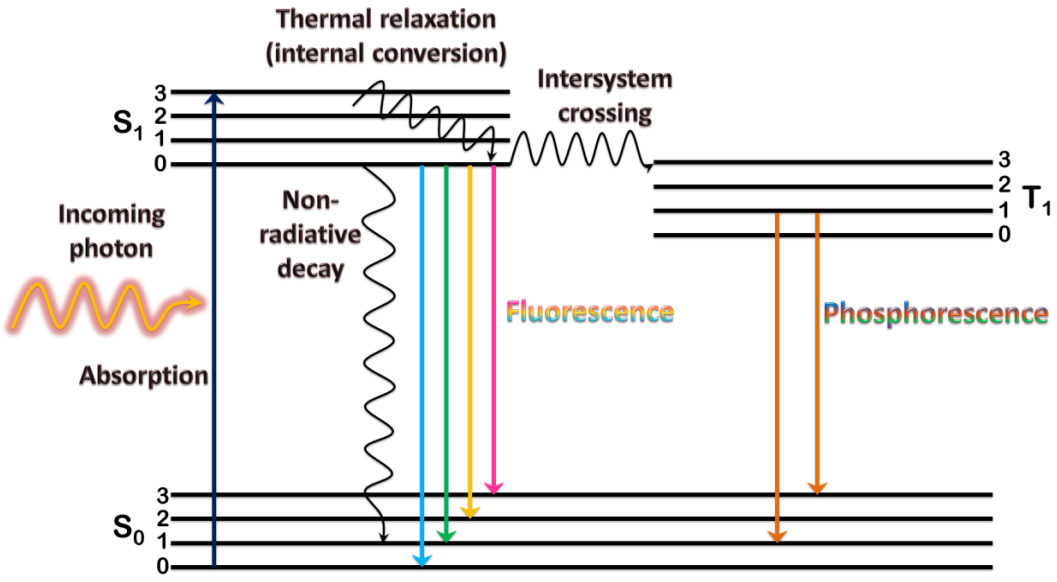


Figure 22. Jablonski diagram

S_0 and S_1 represent the singlet ground and first electronic states respectively. 0, 1, 2, and 3 are stand for the vibrational states that a fluorophore can exist in. T_1 is the first triplet state.

Once a fluorophore radiatively absorbs a photon, it undergoes an excited transition to the S_1 or higher singlet electronic states corresponding to the amount of energy transferred. Then, the dissipation of energy occurs from the excited fluorophore to its surroundings due to its thermal relaxation through internal conversion to the lowest vibrational level of S_1. The fluorophore will then return to the ground state (S_0) either radiatively with the emission of a photon, known as fluorescence, or non-radiatively. The average time an electron spends in the excited state before returning to the ground state is known as the fluorescence lifetime. Fluorescence can only occur from the first excited state because the lifetime of the fluorescence (10^{-9} s) is generally much longer than the internal conversion (10^{-13} s). At room temperature, thermal energy is not adequate to significantly populate the excited vibrational eigenstates. Absorption typically occurs from fluorophores with the lowest vibrational energy. Non-radiative intersystem crossing from the excited singlet state S_1 level to the first excited triplet state T_1 can also occur. Decay from the triplet to the ground state by emission of a photon gives rise to phosphorescence. The lifetime of phosphorescence is longer than the fluorescence lifetime.

The use of fluorescence anisotropy to monitor protein-DNA interactions has been on the rise since its first advent in 1990 [194]. When polarised light striking a fluorophore in solution, polarised emission from samples occurs along a fixed axis within the excited molecule. The polarised emission then gradually returns to unpolarised fluorescence, in terms of depolarisation, mainly depending on rotational diffusion of fluorophores among a number of other factors. Anisotropy (r) is directly related to the extent of polarisation of the emission and defined as the ratio of the polarsied-light component to the total light intensity. As shown in Equation 1, I_{II} and $I_{\perp}$ represents the emission intensities polarised parallel and perpendicular to the excitation intensity.

$$r = \frac{I_{\parallel} - I_{\perp}}{I_{\parallel} + 2I_{\perp}}$$

Equation 1

Anisotropy is correlated with an average angular displacement of the fluorophore during the lifetime of the excited state. This angular displacement directly attributes to the rate and extent of rotational diffusion in solution, which significantly rest with the size and shape of the diffusing molecule. As a solution-based, true-equilibrium, real-time, ratiometric and concentration independent method, fluorescence anisotropy can be used to probe the fluorophore's rotational mobility under highly mimic *in vivo* conditions, hence providing extremely accurate information about its molecular size and shape, and local viscosity of the fluorophore's environment, as well as offering insights into changes in molecular sizes of polymers and other macromolecules. Therefore, this method is highly desirable for studying protein-DNA interactions.

The band shift assay, which is a classic affinity electrophoresis technique, is also widely used to study the formation of protein-DNA complexes. It works on the basis of the change in migration velocity of DNA upon binding by one or more proteins during non-denaturing polyacrylamide gel electrophoresis (PAGE). This technique has the advantage that species of different molecule weight, due to formation of different stoichiometric protein-DNA complexes, can be resolved by a change in mobility and visually observed as multiple bands on the native gel. However, it can only provide a reasonable accuracy for measurement of dissociation constant values because the equilibrium between free DNA and protein-DNA complexes can be compromised by various factors during PAGE, such as local pH, temperature and salt concentration of the gel electrophoresis buffer.

2.2 Results

2.2.1 Different binding modes to product and non-specific DNA by MUG

Fluorescence anisotropy assays were performed using a Fluoromax-3 spectrofluorometer to examine the binding of the MUG enzyme to the abasic product (Hex-12AP·G) and non-specific (Hex-12C·G) oligonucleotides, which contain an abasic furanose analogue (AP) and cytosine (C) respectively, and are paired with a guanine (G) opposite the AP/C site, and labeled at the 5-terminus with 6-carboxy-2',4,4',5',7,7',-hexachlorofluorescein (Hex) via an aliphatic spacer sequence (see Materials and Methods). Previous studies have indicated Hex is well suited for characterizing MUG-DNA interactions [193]. When conjugated to DNA, this fluorophore is bright and typically exhibits fluorescence decay that is dominated by a single lifetime and relatively independent of conditions and protein binding, and its anisotropy is strongly correlated with DNA rotation, with minimal contribution from independent fluorophore mobility [192]. These properties are highly desirable for studying complex binding mechanisms as shown below for MUG enzymes.

When increasing concentrations of MUG were titrated against 100nM Hex-12AP·G product DNA, a pronounced sigmoidal response in anisotropy, characteristic of cooperative binding, was observed and the data is shown with the best fit to the Hill equation, which describes cooperative binding of multiple ligands (Figure 23A). Hill equation provides an equilibrium dissociation constant (K_d) and a Hill coefficient (n_H). The Hill coefficient is a measure of the cooperativity of the system, it gives a measure of the minimum number of interacting ligands, but it is not a direct measure of stoichiometry [195], for example, when the value of n_H is greater than 1, it means the binding mode of enzymes is so positively cooperative that once one enzyme binds to

the DNA, its affinity for other enzyme molecules increases. It also ought to be noted that the K_d is an average value of all bound protomers and thus does not represent a discrete value for binding to the abasic product site. Fitting the anisotropy data to the Hill equation gave a $K_d = 186 \pm 3.0$ nM and a $n_H = 3.4 \pm 0.2$ for MUG binding to Hex-12AP·G.

When MUG was titrated into the non-specific Hex-12C·G DNA, an obvious increase in anisotropy was observed. The data could not be properly fitted into the Hill equation (result not shown) but was best fitted into a tight binding equation, resulting in a hyperbolic binding curve ($K_d = 7 \pm 0.9$ nM) (Figure 23B). This indicates that MUG binds tightly to the non-specific DNA, but that it does so in a non-cooperative mode.

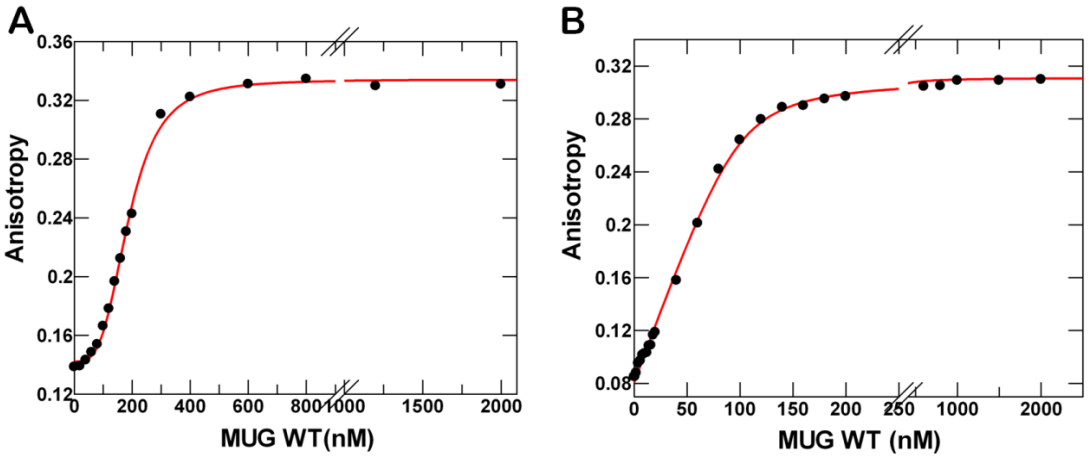


Figure 23. Product and non-specific DNA binding by wild-type MUG.
Increasing concentrations of wild-type MUG were incubated respectively with 100nM Hex labeled product DNA (Hex-12AP·G) and non-specific DNA (Hex-12C·G) in a standard reaction buffer at 25°C and the anisotropy of the fluorophore monitored as described for binding reaction (see Materials and Methods). (A) The anisotropy data for Hex-12AP·G was best fitted to the Hill equation and gave a $K_d = 186 \pm 3.0$ nM and a $n_H = 3.4 \pm 0.2$. (B) The anisotropy data for Hex-12C·G was best fitted to the tight binding equation and gave a $K_d = 7 \pm 0.9$ nM.

2.2.2 MUG binds to substrate DNA in a cooperative manner

The above results clearly indicate different binding modes for MUG to its product and non-specific DNA. It is also valuable to consider how MUG binds to its substrate DNA. To this end, a mutation in the active site of MUG was performed previously by Dr Rory O'Neil in our lab and the mutant MUG N18A enzyme was employed to examine MUG-substrate binding mode. Based on structural homology to the well-characterized UDGs, it has been predicted that the MUG residue Asn18 plays an essential role in enzyme's repair activity [93, 128]. Mutation of the equivalent residue of hTDG, Asn140, to alanine effectively abolishes catalytic activity of the enzyme, but remains proficient in binding both substrate and product [196]. The catalytic activity of N18A MUG mutant has been previously tested under single turnover condition, which means an excess of enzyme over substrate, and the activity assay showed that there was no any accumulation of product, which corresponded to the rate of *N*-glycosidic bond hydrolysis over the 2 hours reaction time course, indicating the catalytic activity of N18A MUG mutant was completely abolished by the alanine mutation and hence could not alter the nature of the substrates over the time course of either the fluorescence anisotropy or the band shift assays [129].

Initially, the binding of N18A MUG to Hex-12AP·G was examined, a similar sigmoidal response was observed to the wild type enzyme, the data was best fitted to the Hill equation with a K_d = 261 ± 1.6 nM and a n_H = 4.5 ± 0.16 (Figure 24A). When N18A MUG was titrated into the non-specific Hex-12C·G, the increased anisotropy was observed and the data was fitted better to a cooperative equation with a K_d = 126 ± 2 nM and a n_H = 1.8 ± 0.1, than the tight binding equation, gave a K_d = 17 ± 4 nM (Figure 24B), although there was clearly a less pronounced sigmoidal response than for either MUG wild-type or N181A mutant enzyme with Hex-12AP·G. With the N18A enzyme, it was notable that the sigmoidal binding curve for the abasic DNA

was exaggerated and that the non-specific binding was slightly weaker, as in contrary to wild-type MUG. It cannot be excluded that there is a possibility of the N18A mutant preparation contains a proportion of inactive enzyme, which would account for these minor differences from wild-type. However, the general curve trend/profile clearly showed that DNA binding properties are very similar for both the wild-type and mutant enzymes.

When N18A MUG was titrated into Hex-12U·G substrate DNA, an obvious anisotropy raise was observed and the data was fitted respectively into the Hill equation (K_d = 119 ± 2 nM; n_H = 1.6 ± 0.4) and the tight binding equation (K_d = 60 ± 3 nM), both obtained binding isotherms were essentially identical to those of the non-specific Hex-12C·G (Figure 24C). The degree of cooperativity for the substrate DNA was slightly less remarkable than for the non-specific DNA. It is notable that the overall change in anisotropy is comparable for both DNAs, indicating that the overall binding stoichiometry is similar.

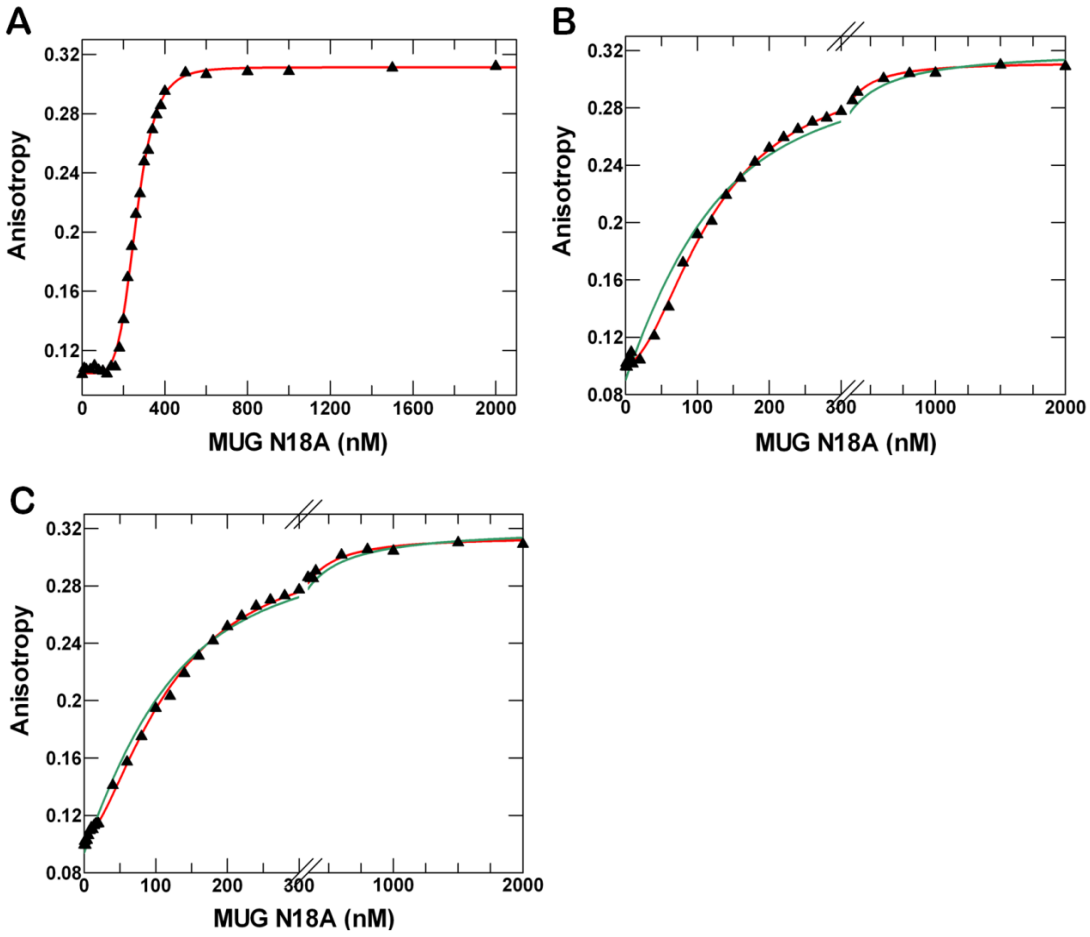


Figure 24. DNA binding by N18A MUG.

Increasing concentrations of N18A MUG were incubated separately with 100nM Hex labeled product DNA (Hex-12AP·G), non-specific DNA (Hex-12C·G) and substrate DNA (Hex-12U·G) in a standard reaction buffer at 25°C and the anisotropy of the fluorophore monitored as described for binding reaction (see Materials and Methods). (A) The anisotropy data for Hex-12AP·G was best fitted to the Hill equation and gave a K_d = 261 ± 1.6 nM and a n_H = 4.5 ± 0.2. (B) The anisotropy data for Hex-12C·G was better fitted to the Hill equation (red line) with a K_d = 126 ± 2 nM and n_H = 1.8 ± 0.1 than the tight binding equation (green line) with a K_d = 17 ± 4 nM. (C) The anisotropy data for Hex-12U·G was better fitted to the Hill equation (red line) with a K_d = 119 ± 2 nM and n_H = 1.6 ± 0.04 than the tight binding equation (green line) with a K_d = 60 ± 3nM.

2.2.3 Characterization of MUG-DNA complexes

To further our understanding of the complexes formed by MUG on binding DNA, band shift assays were performed with both abasic product DNA and non-specific DNA. The band shift on the gel is based on the different electrophoretic mobility rates at which species of different molecular weight, due to binding of one or more proteins to DNA, migrate through the gel. The shifted bands on the gel were visualized based on the Hex 5' conjugated to the DNA and then quantitated so that the binding profile could be compared to the fluorescence anisotropy data.

2.2.3.1 Optimization of band shift assay

In spite of requiring large amounts of materials, the band shift assay does not typically provide enough sensitivity for measuring equilibrium dissociation constants in the range 10^{10}-10^{12} M^{-1}, because until the various species actually enter the gel, the equilibrium distribution is quite susceptible to changes in pH and salt concentration caused by the electrophoresis running buffer once the sample is loaded onto the wells of the gel. Whereas the dissociation of protein-DNA complexes is significantly restrained once moved into the gel matrix due to the caging or crowding effects. Indeed, band shift assay and solution-based fluorescence anisotropy assay do sometimes produce quite different results under otherwise parallel conditions.

With band shift assay, the dissociation of protein-DNA complexes during PAGE can often be seen as a forward edge smearing of the complex band. In order to preserve the DNA-protein complex, prevent a subsequent underestimation of binding constants and provide a same sensitivity level as solution-based techniques, the chemical advantage of neutral osmolytes was applied to the band shift assay herein. It has been confirmed that sufficiently high concentrations of neutral osmolytes, such as betaine

glycine, triethylene glycol or methyl glucosidic, can significantly slow down the rate of protein-DNA complex dissociation by an osmotic stress effect. The osmotic stress effect can cause water hydration of the complex and thus prevents net uptake of waters during the dissociation reaction [197]. Here, 1 M betaine glycine was initially examined to see whether it could prevent the dissociation of MUG-DNA complex during the PAGE and enhance the sensitivity of the band shift assay.

Increasing concentrations of wild-type MUG were incubated with 100 nM Hex-6AP·G product DNA under standard reaction conditions (see Materials and Methods) to reach the equilibrium and pre-form complexes. 1 M betaine glycine was then added into aliquots of different reaction mixtures and incubated for further 5 minutes prior to native PAGE (Figure 25B). Equimolar amounts of equilibrated reaction mixtures without betaine glycine treatment were also analyzed under the parallel conditions as a negative control (Figure 25A). Addition of 1 M betaine glycine clearly prevented the formation of forward edge smearing bands, which were observed on the negative control gel (Figure 25A).

This result indicates that premature dissociation of the MUG-DNA complexes during loading is effectively subdued by betaine glycine. The intensities of bound complex bands within two gels were then quantitated respectively. The data of betaine glycine treated samples was better fitted into the Hill equation, resulting in a typical sigmoidal cooperative binding response with a lower dissociation constant value (K_d = 77 ± 2.1 nM) and a higher hill coefficient value (n_H = 1.9 ± 0.08) than the results of the negative control (K_d = 91 ± 1.7 nM; n_H = 1.8 ± 0.13). These findings clearly demonstrate that the sensitivity of the band shift assays is improved by incubating 1 M betaine glycine with reaction mixture samples containing MUG-DNA complexes.

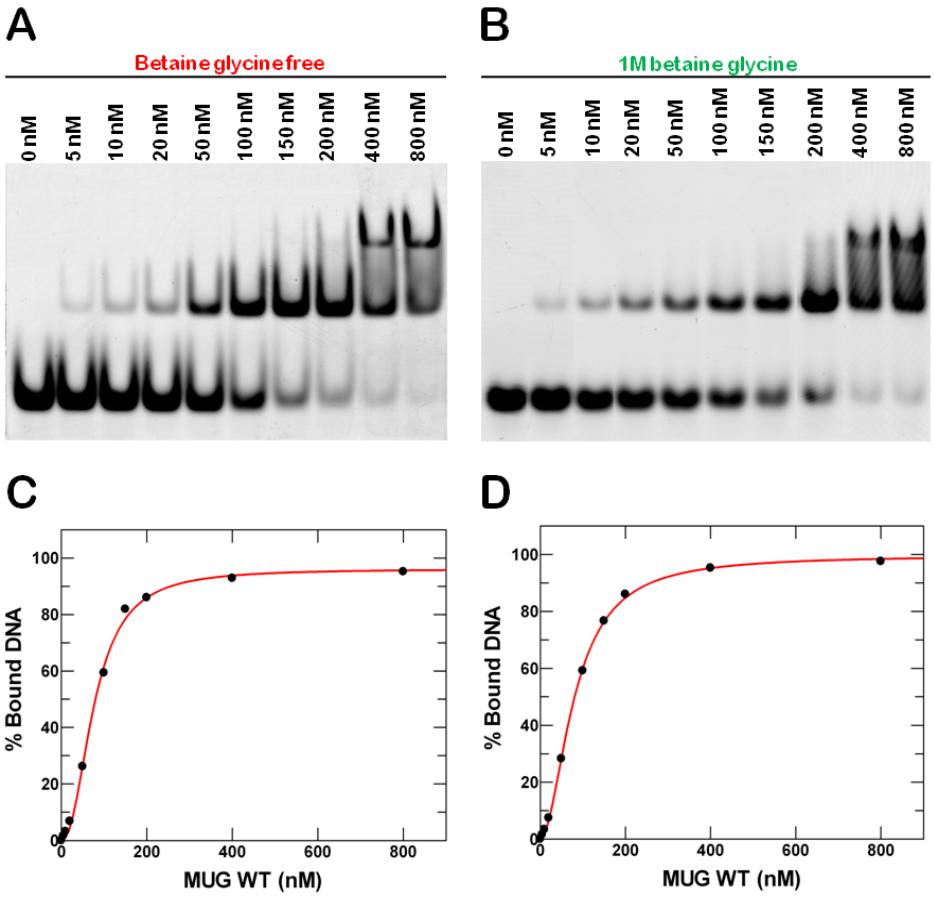


Figure 25. Optimization of band shift assay by1M betaine glycine.
Increasing concentrations of wild-type MUG were incubated with 100 nM Hex labeled product DNA (Hex-6AP·G) in a standard reaction buffer at 25°C for 30 minutes to reach the reaction equilibrium (see Materials and Methods). (A) The equilibrated reaction mixture was directly analyzed by native PAGE. (B) Native PAGE analysis after the incubation of the equilibrated reaction mixture with 1M betaine glycine. (C) The intensities of protein-DNA complex bands from gel (A) was quantitated and the data was fitted to the Hill equation with a $K_d = 91 \pm 1.7$ nM and $n_H = 1.8 \pm 0.13$. (D) The data from gel (B) quantitation was best fitted to the Hill equation with a $K_d = 77 \pm 2.1$ nM and $n_H = 1.9 \pm 0.08$.

2.2.3.2 Analysis of MUG-DNA complexes

Previous band shift assays in our lab indicated that the 25-mer Hex-12AP/U/C·G oligonucleotides used in the fluorescence anisotropy assays gave rise to multiple bands that complicated the analysis [193], further band shift assays were therefore performed with 17-mer DNA molecules, namely Hex-6AP·G and Hex-6C·G.

When MUG was titrated into Hex-6AP·G DNA (Figure 26A), a clear shifted band **b** was formed, while at higher concentrations ranging from 400 nM to 5000 nM, a second higher molecular weight band **c** appeared. Analysis of this extended titration quantified the bound complex percentage, but a tight binding equation was unable to fit the data with a defined DNA concentration of 100 nM (result not shown here), although the same data was well fitted to the Hill equation with a $K_d = 111 \pm 2$ nM and a $n_H = 1.7 \pm 0.06$ (Figure 26B). The stoichiometry of the MUG-DNA complex was determined by re-plotting this data against [MUG]/[DNA], where fitting to the tight binding equation gave a stoichiometry of 2.09 ± 0.08 and a K_d = 115 ± 0.03 nM (Figure 26C).

This analysis concurs with a further inspection of the band shift, where it can be seen that, at 1:1 MUG:DNA (100 nM MUG lane), < 50% of the DNA is bound, and the bound complex does not begin to reach saturation until it approaches a 2:1 stoichiometry at 200 nM MUG. Therefore, the best interpretation of this data is that the first bound complex (Figure 26A, band **b**) corresponds to dimeric MUG binding to the abasic DNA in a cooperative manner. The formation of the higher retarded complex (Figure 26A, band **c**) is resulted from the binding of further MUG molecules, although since this band forms a minor component it cannot be analyzed in detail, but most likely arises from additional, non-specific binding.

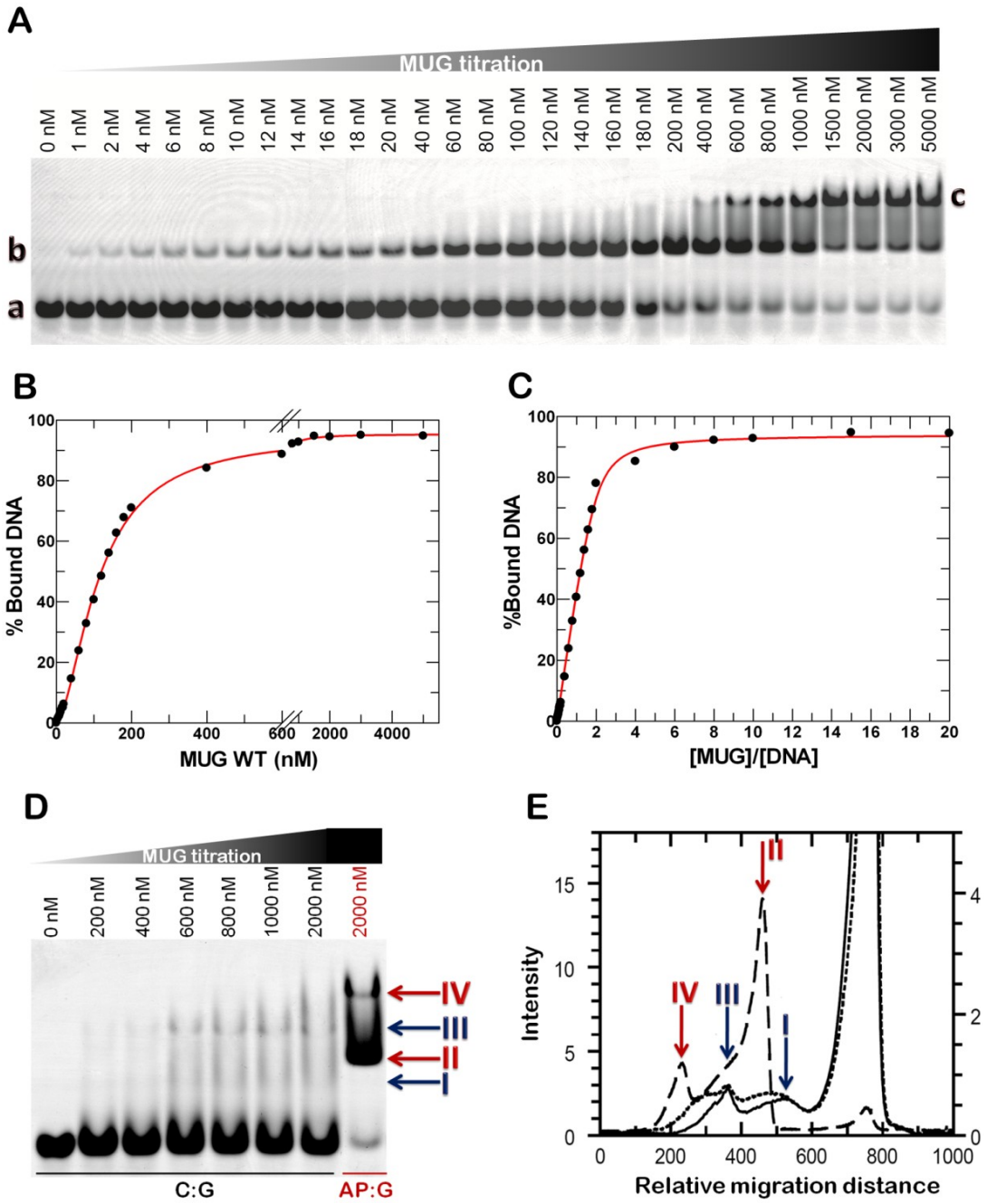


Figure 26. Band shift assays of wild-type MUG.

(A) Large-scale increasing concentrations of wild-type MUG were incubated with 100 nM Hex labeled product DNA (Hex-6AP·G) in the standard reaction buffer at 25°C for 30 minutes, and MUG-DNA complexes (bands b and c) were resolved by native PAGE (see Materials and Methods). (B) The gel (A) was quantitated to determine the percentage bound DNA using Phoretix™ 1D and the data was best fitted to the Hill equation with a $K_d = 111 \pm 2$ nM and a $n_H = 1.7 \pm 0.06$. (C) The data was re-plotted against [MUG]/[DNA] with the best fit to a tight binding equation with a stoichiometry ([MUG]/[DNA]) of 2.09 ± 0.08 and $K_d = 115 \pm 0.03$ nM. (D) Wild-type MUG was titrated into 100 nM Hex labeled non-specific DNA (Hex-6C·G)

at the concentration indicated, a control lane (red) with Hex-6AP·G was incubated for comparison; blue arrows indicate the position of the retarded bands for the non-specific DNA (I and III) and red arrows for the abasic DNA (II and IV). (E) Lane profiles were calculated using Phoretix^{TM} 1D for lane 6 (solid line, scale offset for clarity); lane 7 (dots, scale offset for clarity); lane 8 (dash line), positions of the bands in gel (D) are indicated on the profiles.

When band shift assays were carried out with the non-specific Hex-6C·G DNA, a striking increased level of smearing nature of the bands was observed (Figure 26D). Despite the fact that tight binding of MUG to non-specific DNA was indicated in the fluorescence anisotropy assays, it is clear that under the conditions of the band shift assay, the non-specific complexes are not stable; this is most likely governed by the off-rate of the complex, which must be higher for non-specific DNA. Although the bands are not distinctively separated on the gel, analysis of the lane profiles demonstrates the presence of two distinct bands, one migrates faster than the product complex, and the other, better defined complex, migrates slower (Figure 26E). These most likely corresponded to the monomeric or dimeric binding of MUG to the DNA molecule. Comparison of the abasic and non-specific complexes is revealing, as it is evident that the first shifted band with the abasic DNA is significantly more retarded than the first non-specific complex. This is consistent with the abasic complex being a dimer of MUG bound to DNA, but in a more compact enzyme-DNA complex formation, which migrates faster than two nonspecifically bound MUG molecules.

The inactive N18A MUG mutant was also used to perform band shift assays with both abasic and non-specific DNA. Identical band shift patterns to the wild-type enzyme were observed (Figure 27A). The data of N18A MUG bound to Hex-6AP·G was best fitted to the Hill equation, with a K_d of 71 ± 2.3 nM and a n_H = 2.1 ± 0.12, according with the data of the wild-type enzyme (Figure 27B) and indicating N18A MUG cooperatively binds to the abasic product DNA. Moreover, when band shift assays were performed with N18A MUG and the same sequence of DNA containing U·G mismatch in place of the abasic site, namely Hex-6U·G, the same banding pattern and

affinity as non-specific Hex-6C·G DNA was observed (Figure 27C and D). With both the wild-type and N18A enzymes and the non-specific or substrate DNAs, the higher retarded band is more pronounced suggesting some degree of cooperativity in these interactions, although the poor resolution of these complexes prevented detailed analysis.

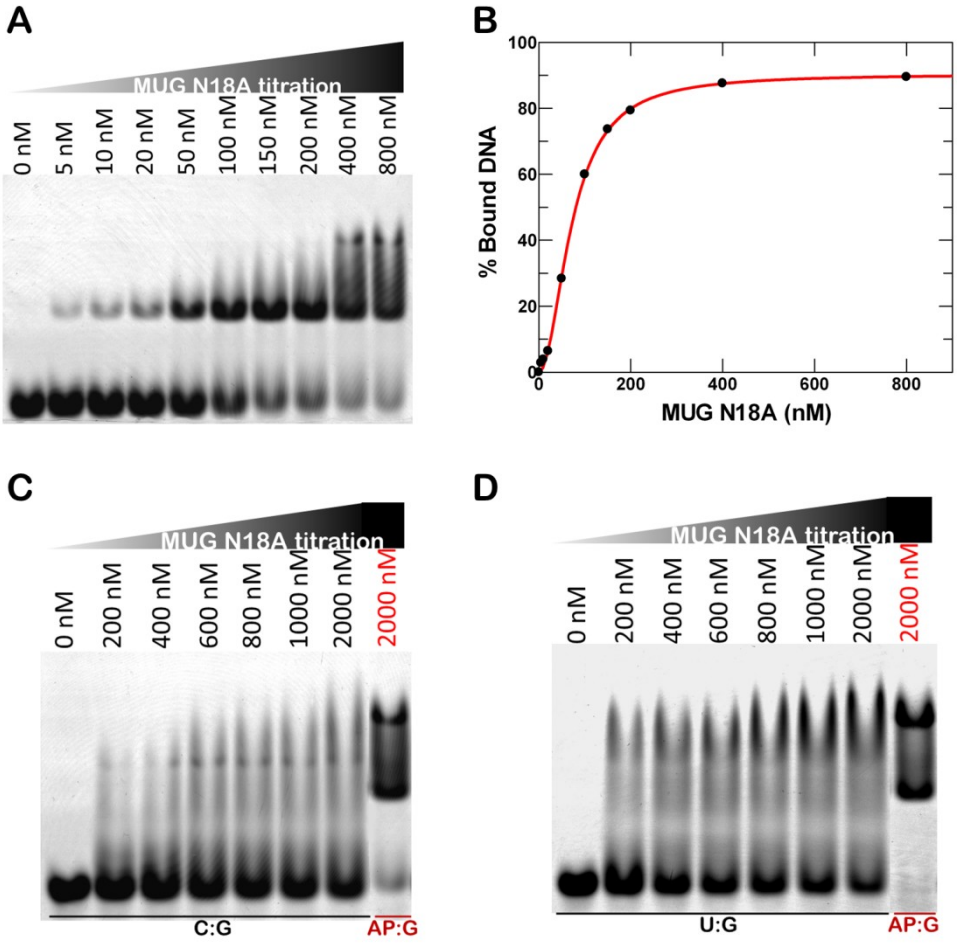


Figure 27. Band shift assays of N18A MUG.

Increasing concentrations of N18A MUG were separately incubated with 100 nM Hex labeled (A) product DNA (Hex-6AP·G), (C) non-specific DNA (Hex-6C·G) and (D) substrate DNA (Hex-6U·G) in a standard reaction buffer at 25°C for 30 minutes to reach the reaction equilibrium, and then MUG-DNA complexes were resolved by native PAGE (see Materials and Methods), control lanes (red) with Hex-6AP·G were incubated for comparisons on both gels (C) and (D). The gel (A) was quantitated to determine the percentage bound DNA using Phoretix™ 1D and (B) the data was best fitted to the Hill equation with a $K_d = 71 \pm 2.3$ nM and a $n_H = 2.1 \pm 0.12$.

2.2.3.3 Catalytic activity of MUG-DNA complexes

The data presented above clearly demonstrates higher order binding complexes of MUG with its abasic product DNA. However, it does not provide any hints that the formation of MUG-DNA complexes has any impact on the catalytic activity of the enzyme. Therefore, MUG activity assays were performed with the Hex-6U·G substrate DNA. It has been indicated that MUG turnover is severely inhibited by the rate of product dissociation (k_{off}) shown in Equation 2, there will be very few multiple turnover reactions, also known as steady state kinetics, of the enzyme within short time phase, thus k_{cat} is difficult to obtain from conventional steady-state kinetics. Alternatively, single turnover reaction, in terms of kinetics under the condition of an excess of enzyme over substrate, is preferential due to the reaction is essentially first order with respect to the substrate and therefore the formation of product will follow a single exponential. The kinetic parameter k_{cl} stands for the maximal rate of enzyme bound product formation, corresponding to the rate of *N*-glycosidic bond hydrolysis, and is regulated by the reaction steps after DNA binding and before product dissociation. Therefore, the measurement of k_{cl} is more readily by quantitating product formation under single turnover conditions.

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_{cl}} EP \xrightarrow{k_{off}} E + P \qquad (k_{cat})$$

Equation 2

To investigate the impact of different stoichiometric formation of MUG-DNA complexes on the enzyme catalytic activity, different concentrations of Hex-6U·G (200 nM, 400 nM and 600 nM) were incubated with increasing concentrations of MUG under standard reaction conditions (see Methods and Materials). All reactions were allowed to proceed for 15 minutes before being quenched with NaOH. Under

single turnover conditions, MUG is known to fully cleave its substrate in 200 seconds [129], so this reaction provides ample time for full cleavage of the substrate under single turnover conditions, but not multiple turnover. The released product bands were analyzed by quantitating the successively increased intensities and then the percentage of decrease in substrate was calculated and plotted against [MUG]/[DNA] (Figure 28). When the concentration of MUG is increased, the concentration of substrate decreases in a linear fashion until complete cleavage is achieved. In each case, complete cleavage coincides with an enzyme concentration that is double the DNA concentration (Figure 28). This is consistent with the 2:1 binding stoichiometry observed in above band shift assays (Figure 26).

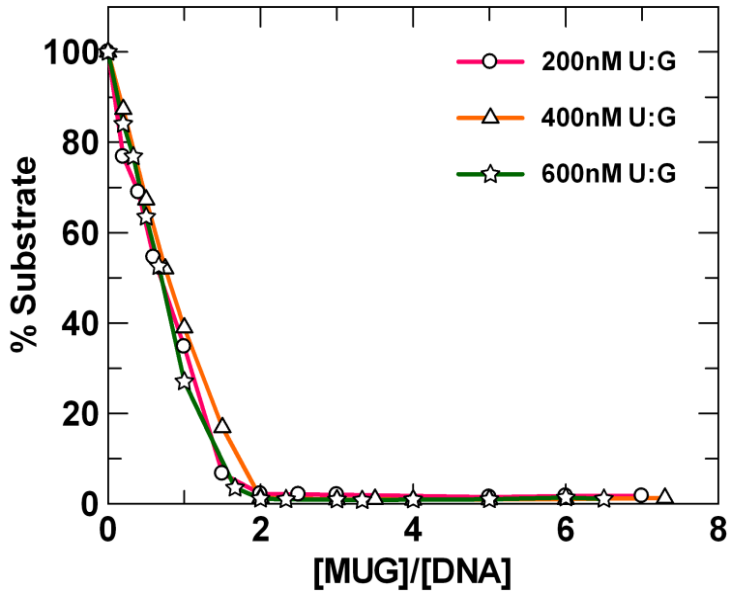


Figure 28. Stoichiometric dependence of MUG activity.
Different concentrations (200 nM, 400 nM and 600 nM) of substrate DNA (Hex-6U·G) was incubated with increasing concentrations of MUG in a standard reaction buffer at 25°C. The reaction proceeded for 15 minutes prior to quenching with NaOH and analyzing by denaturing PAGE. Product formation was quantitated and depletion of substrate was then calculated and plotted against [MUG]/[DNA].

2.3 Discussion

Since MUG is known to bind its product tightly [129], resulting in the inhibition of enzyme turnover, understanding how MUG binds DNA is of significance when considering how MUG interacts with downstream enzymes in the BER pathway. Preliminary fluorescence anisotropy data obtained in our lab had suggested that wild-type MUG enzyme binds to its abasic product DNA in a cooperative manner. To support this notion and further characterize the DNA binding mode for MUG, in this chapter, both fluorescence anisotropy and band shift assays were employed to investigate the differential binding modes of both wild-type and a catalytic mutant of MUG with substrate, abasic product and non-specific DNAs.

2.3.1 DNA binding

Initial investigation of DNA binding by MUG was carried out by monitoring the increase in fluorescence anisotropy, upon the change in rotational mobility of the diffusing fluorophore labeled DNA molecule once bound by MUG enzymes. Fitting the anisotropy data into the Hill equation gave the sigmoidal response of the DNA binding isotherms with abasic DNA, demonstrating a significant degree of cooperativity in abasic product binding by MUG. This cooperativity can only arise through interactions between molecules of MUG once bound to the DNA because MUG exists as monomer in solution [193]. The DNA binding isotherms with non-specific DNA were revealed as a hyperbolic response, with a K_d equal to the mean value for all bound ligands. The hyperbolic response corresponds to the multiple binding of non-interacting proteins. Therefore, the results establish a differential binding mode for MUG between abasic product DNA and non-specific DNA.

The N18A MUG mutant was devoted to measurement of MUG substrate binding, which was independent of following-on enzyme hydrolytic activity. N18A MUG did not have any detectable DNA glycosylase activity over a time course significant to the duration of either fluorescence anisotropy or band shift assays. The fluorescence anisotropy results indicate N18A MUG tightly binds to substrate and non-specific DNAs in exactly the same manner. Although it cannot be ruled out that mutation gave rise to perturbation on the enzyme's ability to recognize its substrate, binding of both non-specific and abasic product DNA closely resembled wild-type MUG behavior, suggesting that DNA binding mechanism was not significantly altered by the mutation. Furthermore, structural analysis of the Asn18 residue predicts that it plays a catalytic role in deprotonating a water molecule, which is required for the nucleophilic attack of the *N*-glycosidic bond, however this residue is not involved in substrate recognition [128], and mutation of the equivalent catalytic residue Asn140 in the structurally homologous hTDG does not have any impact on substrate recognition [196]. While there are some indications of cooperative binding to substrate and non-specific DNAs, this is clearly less pronounced than with the abasic product DNA. Both wild-type and mutant fluorescence data indicate that MUG forms strongly cooperative interactions when binding to abasic product DNA.

2.3.2 Stoichiometry of DNA binding and activity

MUG has been indicated above to implicate differential binding modes with specific and non-specific DNA, therefore presumably resulting in multiple species of MUG-DNA complexes. The band shift assay was employed to differentiate these MUG-DNA complexes upon the changes in gel migration rate and determine the stoichiometry of DNA binding by MUG. The band shift results with both wild-type and N18A MUG provide clear insights into the formation of a specific 2:1 complex with abasic product DNA, and additionally indicate that the specific product complex

has a significant faster migration than two MUG enzymes bound to non-specific DNA, the faster migration indicating formation of a tight complex. Data of wild-type and N18A MUG with abasic product DNA from quantitation of the bound complex percentage were both best fitted into the Hill equation, giving a sigmoidal response and a n_H of 2.1, these results further indicated that MUG cooperatively binds to its abasic product DNA, consistent with above fluorescence anisotropy results.

While the binding data indicated the formation of a specific 2:1 complex with abasic product DNA, it was not clear whether this was of importance for the catalytic activity of the enzyme. Activity assays with varying [MUG]:[DNA] ratios were performed to address this question, and clearly demonstrated that a 2:1 ratio was required for cleavage of the substrate. However, it is known that MUG exhibits a reasonable slow rate (0.04 sec^{-1}) of its single turnover substrate cleavage [129]. Therefore, there is a possibility that a single MUG molecule fully processes the lesions, but then cooperative binding to the abasic product sequesters MUG out of the reaction, resulting in the product-inhibited turnover and the observed 2:1 dependence.

Previous salt dependent assays with MUG performed in our lab also made an advance in understanding of the importance of 2:1 stoichiometry for MUG catalytic activity [154, 193]. It is known that the interactions of proteins and DNA are typically highly salt dependent due to the electrostatic interactions with the phosphodiester backbone[198]. The previous salt dependent MUG binding assays with abasic DNA in our lab demonstrated that increasing salt concentrations of the reaction buffer significantly disrupt MUG-MUG association progressively as the cooperativity of MUG enzymes still exists at 150 mM NaCl but disappears at 300 mM NaCl. The subsequent salt dependent MUG activity assays exhibited that the cleavage rate of MUG has a 4 fold decrease at 150 mM of NaCl, but a 400 fold decrease at 300 mM of NaCl. On the basis of these MUG binding and activity assays, it may be postulated

that the dimeric binding of MUG to DNA facilitates enzyme repair activity. It would be difficult to discriminate between models whereby a second MUG enzyme cooperatively binds to DNA and thus facilitates cleavage through stabilization of the enzyme-substrate complex, and the above mentioned post-catalysis product sequestration model whereby a single MUG molecule fully processes the base excision, but then cooperative binding to the abasic product sequesters MUG out of the reaction. Regardless of the fine mechanism, cooperative binding is deemed to have a functional impact on MUG repair activity.

Chapter 3
Characterization of MUG dimer interface

3.1 Background and objectives

It has been early suggested that the specific interactions made between MUG and the widowed guanine, in the form of strong hydrogen bonds between N1 and N2 groups of guanine and two MUG residues Gly143 and Ser145 greatly contribute to the slow turnover of MUG. However, the work presented in the previous chapter clearly demonstrates a cooperative formation of a specific MUG dimeric complex with its abasic product DNA. This observation provides strong evidence that the extremely poor turnover kinetics of MUG must arise from this cooperative binding, rather than specific interactions with the widowed guanine, which may well be involved in substrate recognition. Therefore, cooperative interactions between DNA glycosylases and their abasic products could be reasoned to bring on a coordinated downstream repair, by which disruption of the protein:protein interface will facilitate dissociation from the abasic product.

A likely dimer interface between two MUG molecules has been observed in the first reported MUG-DNA co-crystal structures, which reveals MUG in complex with a synthetic substrate oligonucleotide. In that study, the formation of MUG-DNA co-crystals was favored by an alternative base pairing of the substrate oligo offset by six nucleotides, yet the conformational parameters of the DNA duplex were well within the ranges observed for normal B form [128]. Through atomic coordinates manipulation of the co-crystal, as shown in Figure 29A, two opposing enzyme molecules are productively bound to intact deoxyuridine substrate analogues, and intriguingly, a fairly distinct interface between the neigbouring pair of MUG enzymes comes into sight. Upon closer inspection (Figure 29A), eight residues, including Tyr74, Val75, Gln76, Asn78, Glu79, Ser81, Lys82 and Gln83, are at the interface in either of two MUG enzymes and appear to be orientated favorably and in close proximity to specific opposing residues for forming possible hydrogen-bonding or

salt-bridging interactions, for example, hydrogen bonding interaction could occur between the peptide oxygen of one Asn78 and the side chain amino group of the other Asn78. Therefore, these eight residues were proposed to essentially contribute to the formation of MUG dimeric complex with abasic product DNA based on a true story.

The dimer story is also true for TDG but has a different scenario, in terms of dimer interface. When TDG was crystallized in a 2:1 complex with its abasic product analogue, one enzyme molecule bound at the abasic site, and the other at an undamaged site (Figure 29B), a symmetrical dimer interface between the adjacent pair of TDG enzymes was observed. The subsequent biochemical assays further identified the residues contributing to the dimer interface, which are Leu143, Met144, Tyr147, Thr196, Thr197 and Pro198 (Figure 29B). These residues are highly conserved for vertebrate TDGs but not with MUG, exceptions being Leu143 and Thr196, which have counterparts Leu21 and Thr74 in MUG [191]. The low conservation of dimer interface residues between MUG and TDG coincides with the further inspection of the MUG and TDG co-crystals alignments, as shown in the Figure 29C, the relative positions of the enzymes in TDG co-crystal are not the same as in the MUG co-crystal; the relative steric positioning of MUG and TDG dimers are also mutually exclusive, so that both complexes could not form simultaneously.

Two major questions arise from above inspections of MUG and TDG dimer crystal structures 1) do the proposed MUG dimer interface residues, Tyr74, Val75, Gln76, Asn78, Glu79, Ser81, Lys82 and Gln83 contribute to the cooperative formation of a specific MUG dimeric complex with the abasic DNA; and 2) can MUG use the dimer interface of its human homologue TDG to form new oligomeric complexes with the abasic DNA. Work presented in this chapter aimed to address above two questions and thus make further stride in understanding of the structure-function relationship in the cooperatively formed 2:1 MUG-DNA complex.

Site directed mutagenesis was initially employed to individually substitute an alanine residue for each of the proposed MUG dimer interface residues. Fast protein liquid chromatography (FPLC) was then used to purify eight alanine mutant enzymes prior to carrying out the enzyme activity assays and band shift assays, to examine the mutation effect on MUG enzyme catalytic cleavage rate, turnover kinetics and DNA binding mode. In addition, apart from the counterpart residues Leu21and Thr74 in MUG corresponding to TDG dimer interface residues Leu143 and Thr196, four MUG residues, Ser22, Gly25, Val75 and Gln76 in the homologous region of TDG dimer interface were individually replaced by their corresponding TDG dimer interface residues Met144, Tyr147, Thr197 and Pro198 via site directed mutagenesis, to let each MUG mutant have three of six homologous TDG dimer interface residues in MUG. These four mutant enzymes were subsequently purified via FPLC and investigated using band shift assays to investigate whether they can form new oligomeric complexes via the three of six homologous TDG dimer interface residues with the abasic DNA, such as a trimer.

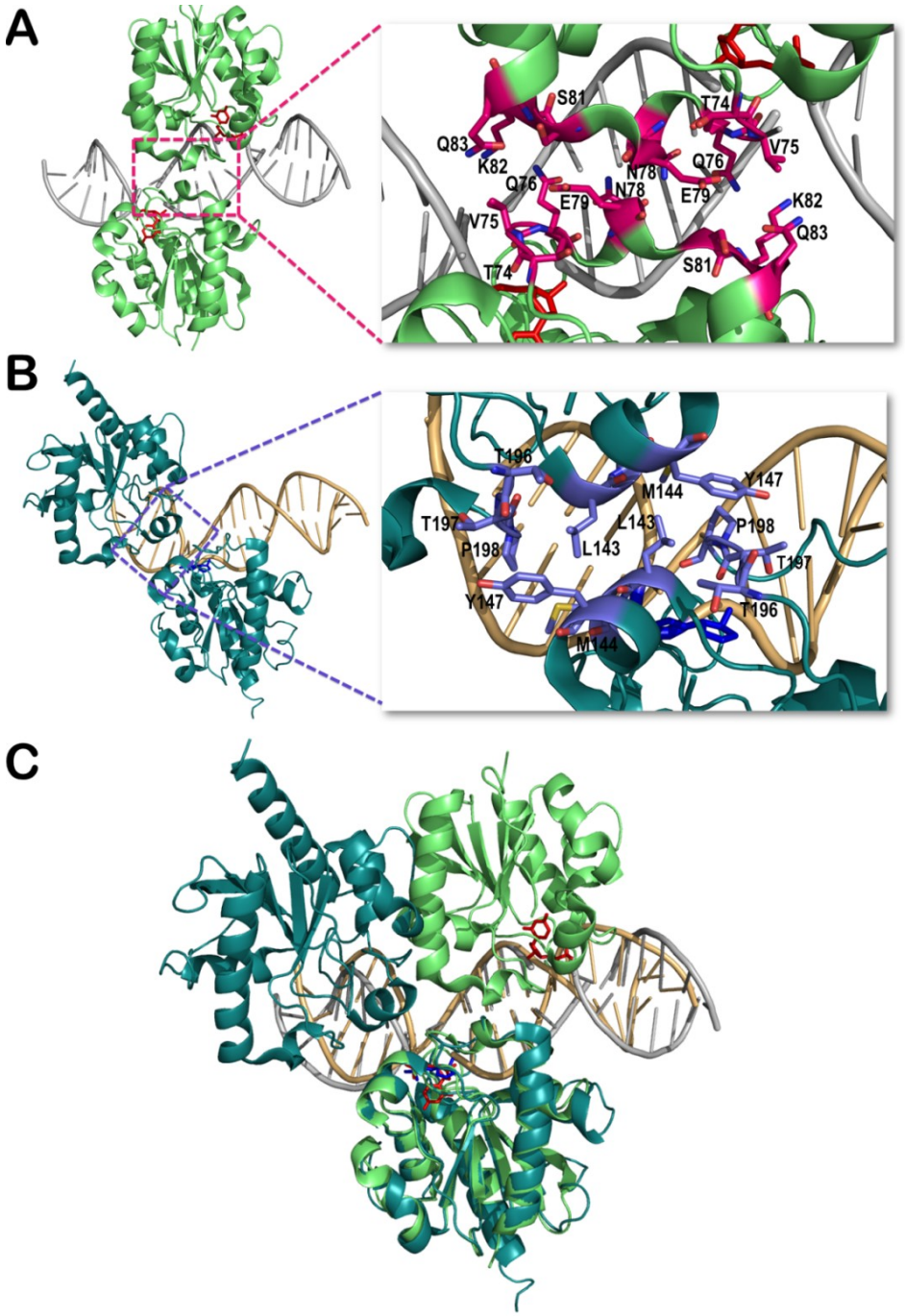


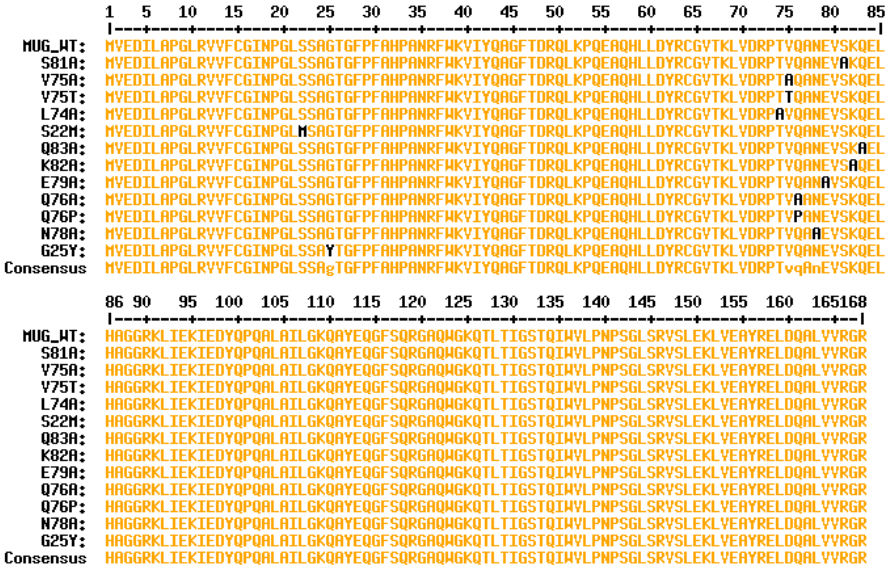
Figure 29. Comparison of crystal structures of MUG and TDG dimers bound to DNA. (A) MUG (green) binds the pseudo-continuous "nicked" DNA helix (grey) in a 2:1 complex and excises the flipped uracil (red) from the G:U mismatch (PDB entry 1MWJ). Key amino acids are shown as labeled and as sticks in the close-up view of MUG dimer interface highlighted in pink. (B) Two TDG (deep teal) molecules bind to the abasic product DNA (gold) (PDB entry 2RBA). The flipped abasic nucleotide is highlighted in dark blue. Key amino acids are shown as labeled and as sticks in the close-up view of TDG dimer interface highlighted in purple. (C) Structure overlaying of MUG and TDG dimer co-crystals in complex with DNA. This figure was generated using PyMOL (DeLano Scientific LLC).

3.2 Results

3.2.1 Construction and purification of 12 site-directed mutant MUG enzymes

Alanine substitutions of proposed MUG dimer interface residues Tyr74, Val75, Gln76, Asn78, Glu79, Ser81, Lys82 and Gln83 were performed using PCR-based site-directed mutagenesis method (see Materials and Methods), simultaneously, four MUG residues Ser22, Gly25, Val75 and Gln76 in the homologous region of TDG dimer interface were also site directed mutagenized to their corresponding TDG residues Met144, Tyr147, Thr197 and Pro198. When the 12 mutant MUG plasmids were constructed and validated by sequencing (Figure 30A), purifications of 12 mutant MUG enzymes were achieved within 3 weeks using FPLC on the basis of the ionic properties and molecular weight of the proteins (see Materials and Methods). All prepared mutant MUG enzymes displayed extremely high purities analyzed via 12% SDS-PAGE (Figure 30B) and had an average yield of 19 mg from a 1.5 liter culture calculated from the UV light absorbance at 280 nm based on an extinction coefficient of 25590 $M^{-1}cm^{-1}$.

A



B

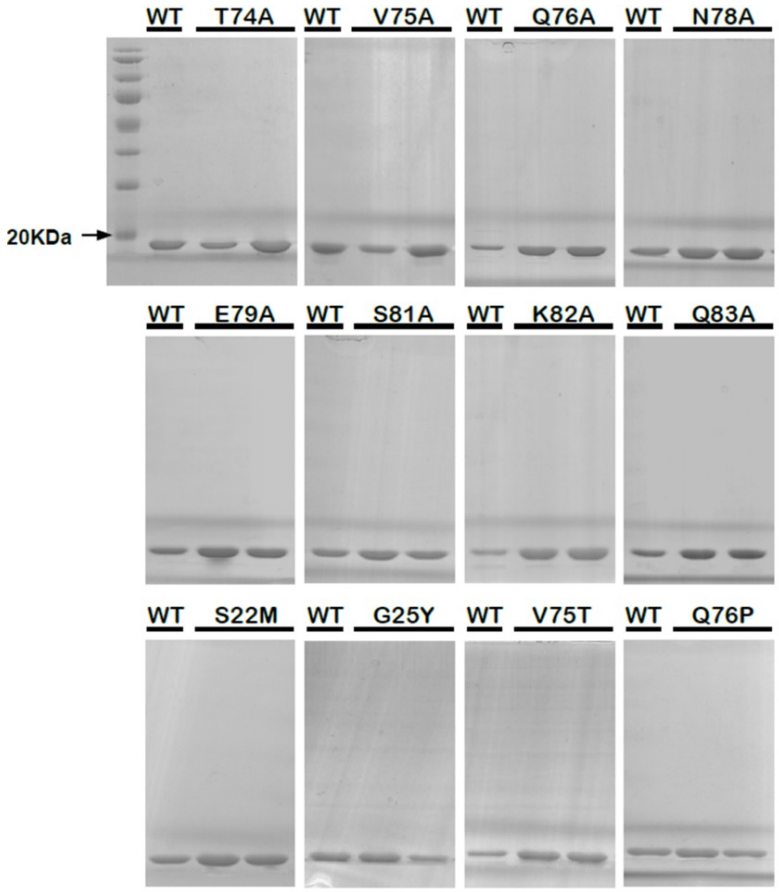


Figure 30. Mutant MUG enzymes construction and purification.

(A) The validated amino acid sequences of 12 site-directed mutant MUG enzymes were aligned using MultAlign (URL: http://multalin.toulouse.inra.fr/multalin/). The mutagenized sites of MUG variants are highlighted in red and the consensus amino acids are green. (B) Purity levels of MUG variants were analyzed by 12% SDS-PAGE (see Materials and Methods). Previously purified wild-type MUG enzyme was loaded as a marker.

3.2.2 Characterization of putative MUG dimer interface

3.2.2.1 Single turnover assays of alanine mutant MUG enzymes

Single turnover assays were conducted to determine whether the alanine substitutions of eight proposed MUG dimer interface residues have effects on the catalytic activity of the enzyme. An excess of each mutant MUG enzyme were reacted with Hex-12U·G at 25^{o}C in the standard reaction buffer for 15 minutes, resulting in 100% product obtained in the first turnover. Under saturating conditions, wild-type MUG is known to cleave this substrate fully in 200 seconds [129], so 15 minute reaction time is adequate. Samples were taken at selected time points during the reaction and quenched with NaOH, then heated at 90^{o}C to cleave at the abasic site via a β-elimination. 12-bp Hex labeled products were separated from intact substrate by denaturing PAGE. The gels of wild-type and alanine mutant MUG enzymes are shown in Figure 31A. The released product bands were analyzed by quantitating the percentages of the successively increased intensities. The data were then fitted to a first order equation to calculate the single turnover cleavage rates, which are compared in Figure 31B. Alanine substitutions of residues T74, E79 and K82 resulted in more than 90% reduction in the MUG single turnover cleavage rates; V75A and S81A MUG enzymes were impaired to 50% of wild-type level; and Q76A and N78A variants were found to perform at ~70% of wild-type rate. The only exception was Q83A, which had a wild-type like single turnover cleavage rate. These results clearly demonstrated that most proposed MUG dimer interface residues are of functional significance on the cleavage of *N*-glycosidic bond.

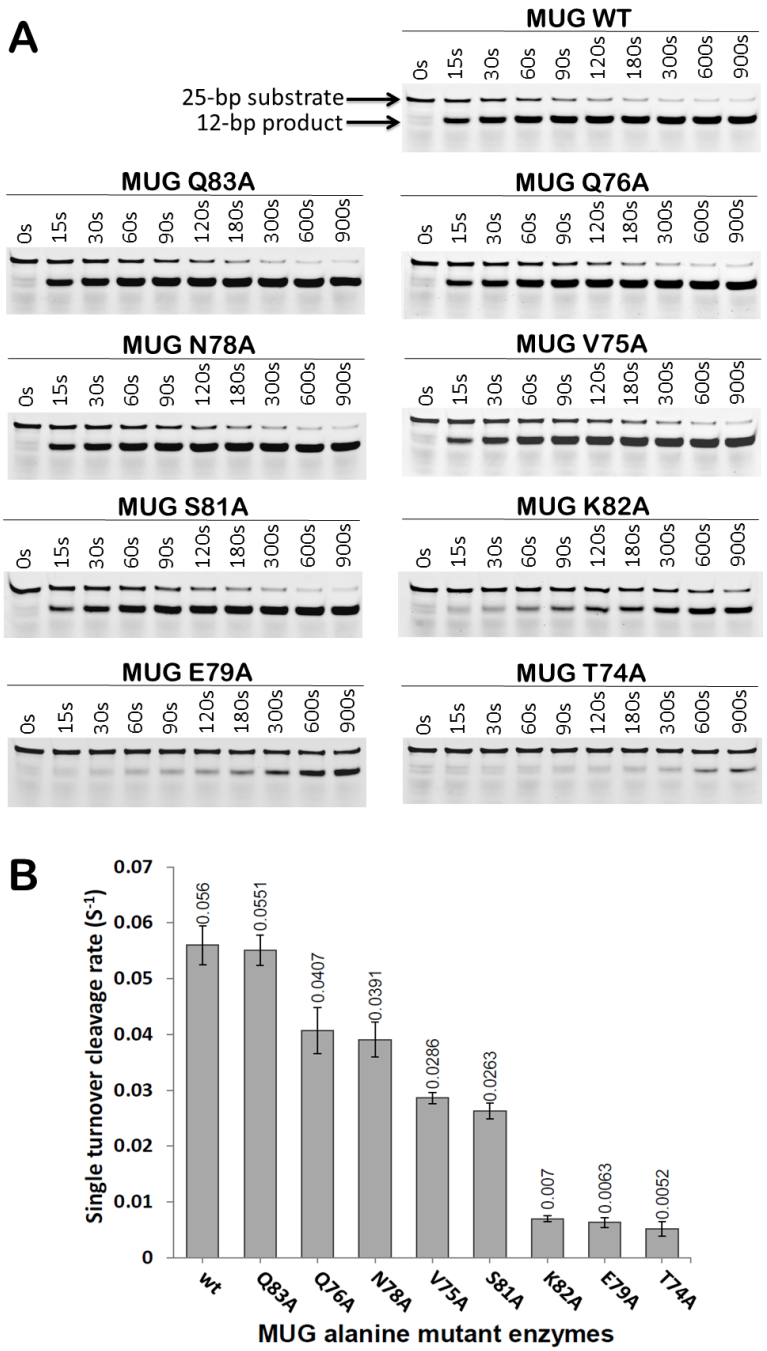


Figure 31. Single turnover assays of wild-type and alanine mutant MUG enzymes.
(A) 100 nM Hex-12U·G was incubated with 5 μM wild-type and mutant MUG enzymes separately in the standard reaction buffer at 25°C for 15 minutes prior to quenching with NaOH and analyzed by denaturing PAGE. The intensities of resolved 12-bp product bands were quantitated using the Phoretix™ 1D. (B) Single turnover cleavage rates for wild-type and alanine mutant MUG enzymes were calculated by fitting the data from product bands quantitation into a first-order rate equation and compared via bar charts. The rate values are labeled on the top of the columns.

3.2.2.2 Analysis of the complexes of alanine mutant MUG enzymes with the abasic product

Band shift assays were performed to determine whether the MUG alanine mutants lead to deficiencies of the affinity and cooperativity for their DNA binding. When each mutant enzyme was titrated into Hex-6AP·G DNA, a clear shifted dimeric complex band was seen, while at higher concentrations a second higher molecular weight band was formed (Figure 32A). The band shift patterns of alanine mutant MUG enzymes are consistent with that of wild-type enzyme described in 2.2.3.2. Data for MUG variants, by quantitating the percentage of MUG-DNA complex bands, were then well fitted to the Hill equation separately and different levels of sigmoidal response, characteristic of cooperative binding, were observed (Figure 32A). The values of K_d and n_H of wild-type and alanine mutant MUG enzymes are compared in Figure 32B and C respectively. Q76A and Q83A enzymes displayed identical binding isotherms to the wild-type enzyme and suggested they both have the same level of affinity and cooperativity for the product DNA binding as those of the wild-type MUG. V75A, N78A and S81A enzymes were found to bind to the product DNA with slightly reduced affinities and cooperativity. T74A, E79A and K82A enzymes displayed the most notable reduction in the DNA binding affinities and cooperativity. This analysis concurs with an further inspection of the band shift, where the 2:1 complex band on the most alanine mutant gels did not reach saturation at 200 nM MUG and a large amount of unbound DNA in the range of 35% - 60% was still seen (Figure 32A), these observations contradicted that of the wild-type MUG showed in 2.2.3.2 and suggested that T74A, V75A, N78A, E79A, S81A and K82A enzymes bind to product DNA in a weak cooperative manner.

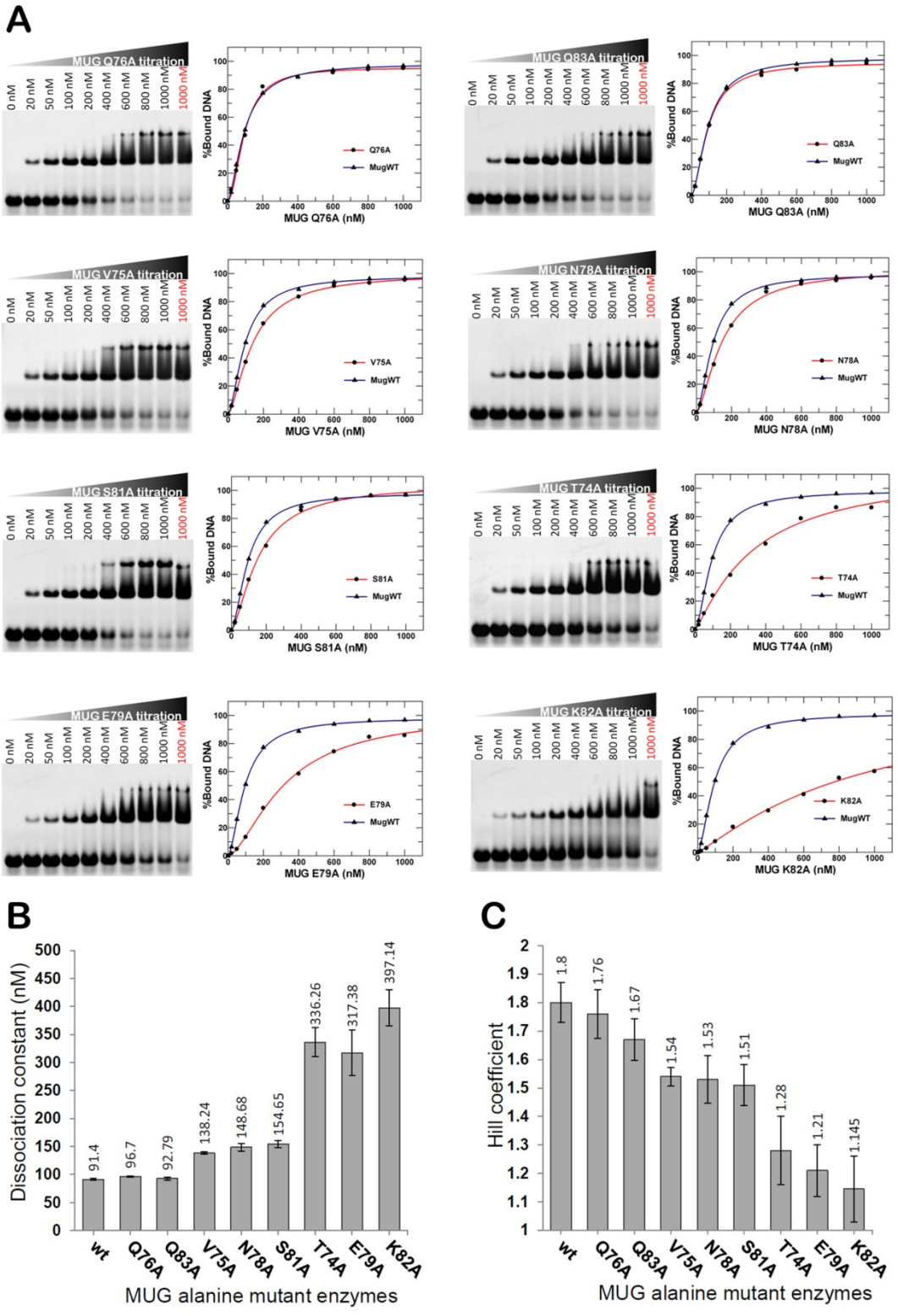


Figure 32. Band shift assays of alanine mutant MUG enzymes.

(A) Increasing concentrations of alanine mutant MUG enzymes were separately incubated with 100 nM Hex labeled product DNA (Hex-6AP·G) in a standard reaction buffer at 25°C for 30 minutes to reach the reaction equilibrium, and then MUG-DNA complexes were resolved

by native PAGE (see Materials and Methods), control lanes (red) with wild-type MUG enzyme were incubated for comparisons. The gels were quantitated to determine the percentage bound DNA using Phoretix^{TM} 1D and the data were best fitted to the Hill equation, producing sigmoidal responses (red line) which were individually compared with that of wild type (blue line), and giving the K_d and n_H values for product binding of alanine mutant MUG enzymes, which were compared with wild-type K_d and n_H using bar charts shown in (B) and (C). The values of K_d and n_H are labeled on the top of the columns.

3.2.2.3 Steady state assays of alanine mutant MUG enzymes

Wild-type MUG is known to remain tightly bound to its abasic product after the cleavage of *N*-glycosidic bond, rendering the slow turnover kinetics. The results described in 2.2.1 and 2.2.3 have clearly demonstrated the formation of MUG dimeric complex with its product and evidently suggested that the cooperative binding must result in the rate-limiting product dissociation. To investigate the effects of MUG alanine mutants on enzyme turnover kinetics, steady state assays were conducted for both wild-type and alanine mutant enzymes. This assay was done by reacting an excess of Hex-12U·G substrate DNA with wild-type and alanine mutant MUG enzymes separately at 25^{o}C in the standard reaction buffer for 12 hours. Aliquots were taken at selected time points during the reaction and quenched with NaOH, then heated at 90^{o}C to cleave at the abasic site via a β-elimination. 12-bp Hex labeled products were separated from intact substrate by denaturing PAGE. The gels of wild-type and alanine mutant MUG enzymes are shown in Figure 33. The intensities of product bands were quantitated to determine the concentrations of the successively released product at each selected time points. The increased product concentration data were plotted against the aliquot quenching time points, giving the steady state reaction profiles. Each reaction profile of alanine mutant enzymes was compared with that of wild-type (Figure 33). Wild-type MUG displayed extremely slow turnover kinetics and only 156nM product was released by 100nM enzyme, namely 1.56 times enzyme turnover was completed, in 12 hours. The steady state reaction profiles of

most alanine mutant enzymes, including V75A, Q76A, N78A, S81A, K82A and Q83A, highly matched that of wild-type MUG, indicating these mutant enzymes possess wild-type like turnover kinetics. Whereas, substituting residue T74 to alanine severely impaired the enzyme's nature and the catalytic activity of the T74A was completely abolished by 2 weeks after previous single turnover assays in 3.2.2.1, there was no any accumulation of product from T74A reaction within 12 hours. Similarly, the activity of E79A MUG was also seriously compromised by the mutation and only 0.44 time turnover was achieved, that was only 28% of wild-type level.

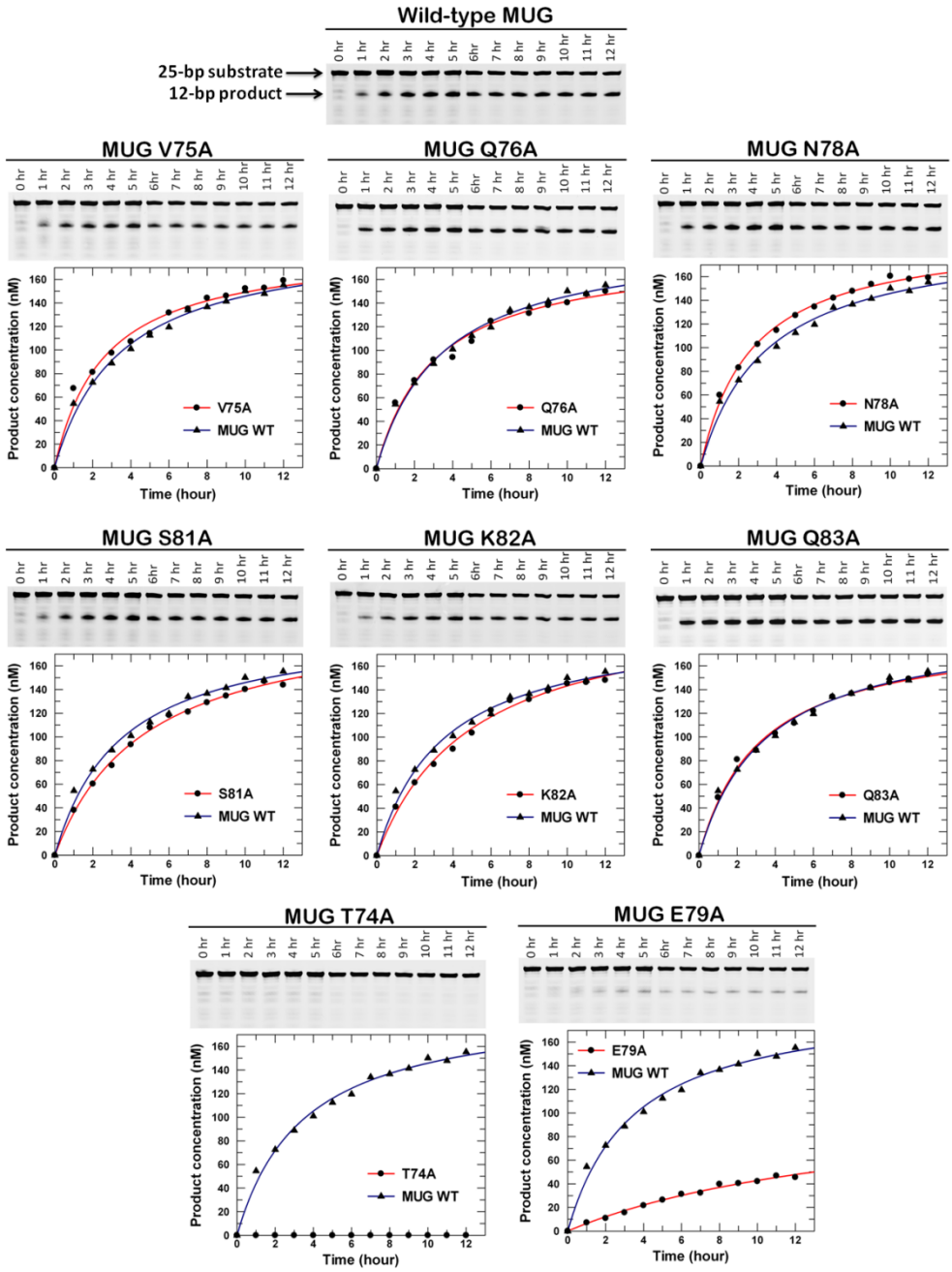


Figure 33. Reaction profiles for wild-type and alanine mutant MUG enzymes under steady state conditions.

500 nM Hex-12U·G was incubated with 100 nM wild-type and alanine mutant MUG enzymes separately in the standard reaction buffer at 25°C. Aliquots were taken at selected time points during the 12 hours course of the reaction, quenched with NaOH and analyzed by denaturing PAGE. The intensities of resolved 12-bp product bands were quantitated using the Phoretix™ 1D to determine the concentrations of released product at selected time points. The reaction profile curves for alanine mutant enzymes (red line) were generated by plotting production concentration against time and compared with that of wild type enzyme (blue line).

3.2.3 Analysis of alternative dimer interface based on TDG structural homology

Based on the structure of a dimeric complex of TDG bound to abasic DNA (Figure 29B), the alternative homologous TDG dimer interface in MUG was identified as a possible means of dimerization. Two conserved homologous TDG dimer residues Leu21 and Thr74 in MUG, four TDG dimer interface residues are different from their corresponding MUG residues, and have therefore been separately mutated in MUG, in terms of either of S22M, G25Y, V75T and Q76P MUG enzymes was created to harbor three of six homologous TDG dimer interface residues. The single turnover cleavage rates of these four MUG mutants ($k_{cl}^{S22M} = 0.056\ s^{-1}$, $k_{cl}^{G25Y} = 0.055\ s^{-1}$, $k_{cl}^{Q76P} = 0.055\ s^{-1}$, $k_{cl}^{V75T} = 0.0543\ s^{-1}$) were determined under the parallel single turnover assay conditions of MUG alanine mutants as described in 3.2.2.1.

Then, to assess whether these mutant enzymes can form new oligomeric complexes via the three homologous TDG dimer interface residues with the abasic DNA, band shift assays were performed individually with S22M, G25Y, V75T and Q76P MUG enzymes using the Hex-6AP·G product DNA. As shown in Figure 34, titrations of four mutant enzymes with Hex-6AP·G DNA exhibited the same band shift patterns as that of wild-type enzyme, the dimeric complex band was initially formed at lower enzyme concentrations and the higher order complex band then appeared at higher enzyme concentrations, apart from these, no other new bands that correspond to new species of MUG-DNA oligomeric complexes were observed. After quantitating the bound complex bands, data for all mutant MUG enzymes were well fitted into the Hill equation, giving pronounced sigmoidal response with values of $n_H>1$, indicating that four MUG variants cooperatively bound to the abasic product. S22M, G25Y and Q76P displayed the identical binding isotherms, as well as the similar product DNA binding affinities and cooperativity to those of wild-type MUG. Whereas, V75T

showed the slightly reduced affinity and cooperativity for the product binding and its sigmoidal curves didn't properly match that of wild-type enzyme, consistent with the above band shift results of V75A.

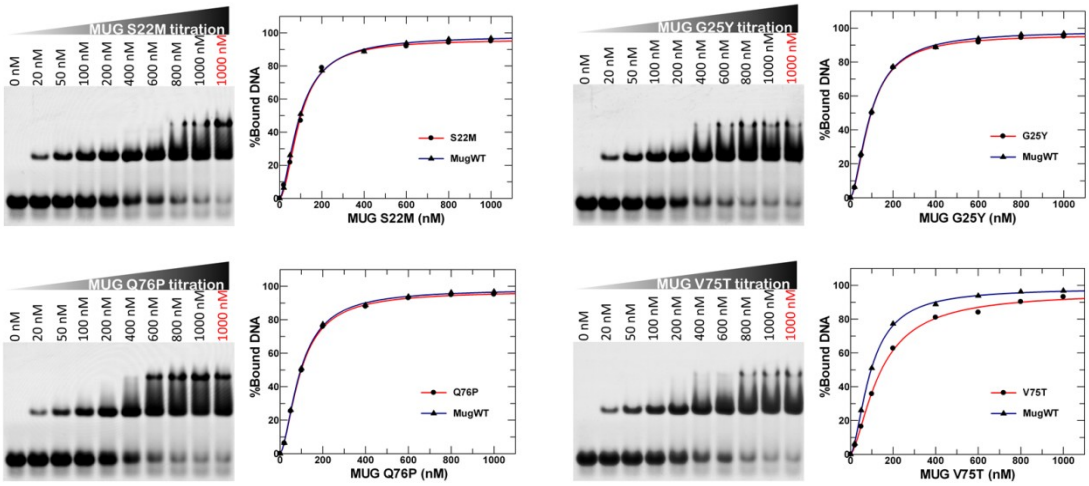


Figure 34. No new oligomeric complexes formation by MUG mutant enzymes containing alternative TDG dimer interface residues.

Increasing concentrations of MUG mutant enzymes containing three of six homologous TDG dimer interface residues were separately incubated with 100 nM Hex labeled product DNA (Hex-6AP·G) in a standard reaction buffer at 25°C for 30 minutes to reach the reaction equilibrium, and then MUG-DNA complexes were resolved by native PAGE (see Materials and Methods), control lanes (red) with wild-type MUG enzyme were incubated for comparisons. The gels were quantitated to determine the percentage bound DNA using Phoretix™ 1D and the data were best fitted to the Hill equation, producing sigmoidal responses (red line) which were individually compared with that of wild type (blue line), and giving the K_d and n_H values: S22M, K_d = 95 ± 4.9 nM and n_H = 1.9 ± 0.18; G25Y, K_d = 94 ± 1.45 nM and n_H = 1.73 ± 0.04; Q76P, K_d = 91 ± 2.75 nM and n_H = 1.81 ± 0.09; and V75T, K_d = 137 ± 6.25 nM and n_H = 1.54 ± 0.08.

3.3 Discussion

After manipulating and scrutinizing the crystal structures of MUG in complex with DNA, a likely MUG dimer interface between two enzyme molecules was identified. Residues contributing to this dimer interface were proposed to be Tyr74, Val75, Gln76, Asn78, Glu79, Ser81, Lys82 and Gln83. The side chains of these residues are orientated favorably and in close proximity to side chains of corresponding residues in the dimer interface of opposing MUG, and therefore postulated to form contacts, such as hydrogen-bonds or salt-bridges, between each other. In this chapter, site-directed mutagenesis associating with following on enzyme activity assays under both single turnover and steady state conditions and band shift assays were performed to determine the eight proposed MUG dimer interface residues whether participate in any structure-function relationships in the cooperative formation of the MUG dimeric complex.

3.3.1 The relationship between MUG single turnover cleavage rate and its cooperativity

The proposed eight MUG dimer interface residues were initially substituted individually by an alanine residue, which has a smaller and non-charged side chain, using PCR based site-directed mutagenesis. Once all alanine mutant enzymes were purified by FPLC, the single turnover assays were immediately performed and demonstrated all the alanine mutants still retained the ability of uracil excision from a G:U mismatched substrate but their single turnover cleavage rates are notably impaired by mutations, especially K82A, T74A and E79A had a significant detrimental impact on substrate cleavage by MUG; whereas Q83A was the exception which showed a wild-type like single turnover cleavage rate. It has been known that all proposed MUG dimer interface residues are on the surface of MUG, not near the

active site pocket of MUG and do not play any direct catalytic roles in the hydrolysis of *N*-glycosidic bond, taken together with the results in 2.2.3.3 that indicated the cooperative binding of MUG has a functional impact on its activity. Therefore, the findings in this chapter, that alanine mutant MUG enzymes have the impaired single turnover cleavage rates, suggest that the residues Thr74, Val75, Gln76, Asn78, Glu79, Ser81 and Lys82 are of functional significance on MUG catalytic activity, likely contributing to the cooperative binding of MUG enzymes while binding to the DNA.

Subsequent band shift assays provided clear evidence on this. The alanine mutants T74A, V75A, N78A, E79A, S81A and K82A exhibit differential degrees of declined affinity and cooperativity, indicating that the cooperative interaction between two alanine mutant enzymes is compromised by the size and charge of the alanine side chain, although all of these six mutants can still bind DNA. The E79A, K82A and T74A mutants that have poor single turnover cleavage also show significant reduced DNA binding properties with Hill coefficients that approach unity, indicating a loss of cooperativity in binding. Of these, K82A has the most impaired DNA binding, this is very likely due to Lys82 being a protein-DNA binding interface residue, as shown in Figure 35, its positively charged side chain is in very close proximity (3.68 Å) to the phosphate group of a guanine base in the DNA substrate, there may be a possible hydrogen bonding interaction mediated by water. Therefore, alanine mutation of Lys82 leads to both reduced DNA binding and catalytic activity. Excluding Lys82, the rest of residues Thr74, Val75, Asn78, Glu79 and Ser81 are very likely to make some extents of contacts via their side chains with undetermined corresponding MUG dimer interface residues in the opposing MUG molecule, consequently in concert forming a dimeric complex. In particular, due to the significantly reduced binding affinities and cooperativity of mutants T74A and E79A, Thr74 and Glu79 are suggested to make critical interactions to stabilize the dimer interface.

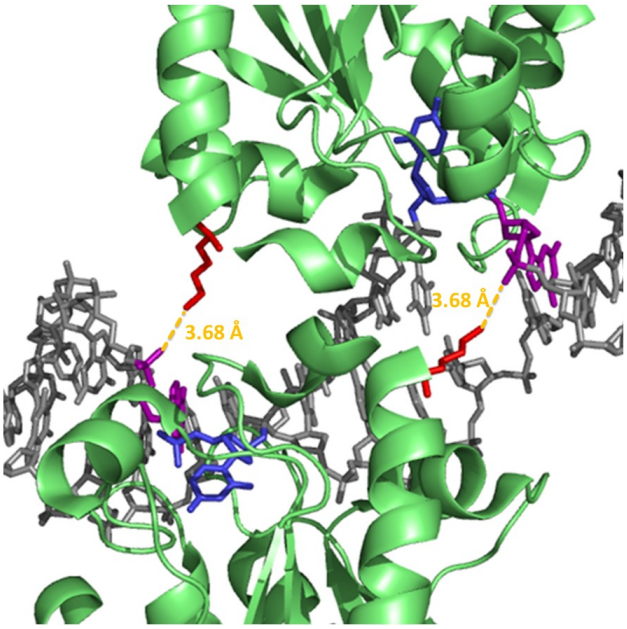


Figure 35. Structure of the MUG dimer in complex with DNA.
Two MUG molecules (green) had bound to the continuous nicked double stranded DNA molecule (grey) and cleaved uracil (blue) from the G·U mismatch (PDB entry 1MWJ). This figure was generated using PyMOL (DeLano Scientific LLC). The positively charged side chain of the lysine residues K82 (red) in each MUG molecule is in very close proximity (3.68 Å, yellow dash line) to the phosphate group of a guanine (purple) in the DNA substrate.

The alanine mutants T74A, V75A, N78A, E79A and S81A still retain some degree of DNA binding as assessed by their band shift profiles, but with differential degrees of positive binding cooperativity with Hill coefficients that are greater than one. The single alanine substitutions at these positions cause some disruption of the MUG dimer interface, but the residual binding suggests that these alanine mutant enzymes can still fold and retain some enzymatic functions, in terms of being able to bind DNA, and have been evaluated for kinetic analysis.

The declined K_d and n_H values of alanine mutant enzymes were strongly correlated with their impaired single turnover cleavage rates, as when K_d increased and n_H decreased, k_{cl} decreased. This result indicates that when the dimer is disrupted to certain extent by the mutation as explained by n_H, the cleavage of MUG is disrupted, and therefore suggests MUG has to assemble into a dimer upon substrate cleavage.

Dissimilarly, Q76A enzyme hydrolyzed the *N*-glycosidic bond at only ~70% of wild-type cleavage rate although it was able to tightly bind its abasic product in a wild-type like cooperative manner. It implies Gln76 may have some interactions with DNA rather than contributing to the MUG 2:1 dimer complex. Moreover, changing the residue Gln83 to alanine didn't affect any MUG enzyme nature on either catalytic activity or cooperativity of product binding, therefore Gln83 seems have no functional roles in the cooperative formation of the MUG 2:1 complex.

3.3.2 Residues at the MUG dimer interface affect enzyme turnover kinetics

The overall rate for the enzyme catalytic cycle is mediated by a step following base removal as evidenced by characteristic burst kinetics for product formation. It has been found that the hydrolysis of *N*-glycosidic bond by wild-type MUG proceeds 100-fold faster than the turnover rate and only 1.6 times of turnover are achieved before the reaction attaining the steady state phase, approximately within 12 hours [129]. This rate-limiting steady state phase is dominated by the rate of the enzyme dissociation from the DNA product. The steady state assay result of wild-type MUG here is consistent with the previous findings.

As above indicated in 3.2.2.2 that most alanine mutant MUG enzymes have relative weak binding affinity and mild cooperativity, therefore, steady state assays were performed to probe whether they can exhibit faster turnover kinetics, in terms of faster product dissociation rate. The steady state reaction profiles in 3.2.3.3 indicated the alanine mutant MUG enzymes don't have faster turnover kinetics, and even an expected classic initial burst phase of MUG was not observed. The kinetics of most MUG alanine mutant enzymes in single turnover state indicated that 100 nM product had been produced in the first 10 minutes (Figure 31), however, in the case of

kinetics in steady state, all the alanine mutant enzymes exhibited around 50 nM or even less product yield after 1 hour and it took around 4 hours to reach 100 nM product release for wild-type MUG (Figure 33). Although the burst phase was not explicitly examined here, these data clearly provide evidence that steady state kinetics of MUG is different from its single turnover behavior. The mismatch of MUG kinetics between single turnover state and multiple turnover state will be analyzed and discussed in detail in the next chapter.

Furthermore, except for Q76A and Q83A, the steady state reaction profiles of the other alanine mutant MUG enzymes exhibit some certain levels of variations in comparison to the profile of wild-type enzyme, suggesting that the residues Thr74, Val75, Asn78, Glu79, Ser81 and Lys82 at the MUG dimer interface have certain degree of impacts on MUG turnover kinetics.

3.3.3 No formation of new oligomeric complexes by MUG mutant enzymes containing alternative TDG dimer interface residues

Since crystal structure alignment revealed that TDG dimer interface residues correspond to Leu21, Ser22, Gly25, Thr74, Val75 and Gln76 in MUG, it is of interest to investigate whether MUG can use the dimer interface of its human homologue TDG to form new oligomeric complexes with DNA. Site-directed mutagenesis was performed to change MUG residues Ser22, Gly25, Val75 and Gln76 to the corresponding TDG dimer residues individually, taken together with two conserved homologous TDG dimer residues Leu21 and Thr74, each of four mutant MUG enzymes possesses three of six homologous TDG dimer interface residues. The subsequent band shift assays demonstrated that the band shift patterns of these mutant enzymes were identical to that of wild-type MUG and suggested no MUG-MUG interaction was formed via the partial homologous TDG dimer interface in MUG.

It is possible that two dimer interfaces in mutant MUG enzymes are partially overlapping because they share the residues Thr74 and Val75, MUG enzymes thus could not form interactions simultaneously via two dimer interfaces. Moreover, on the basis of above proposed MUG dimer interface residues, MUG dimer interface residues are inferred to have a bigger size and stronger residue side chain charge and polarity than those of homologous TDG dimer interface residues, therefore, the former is more preferable and competitive for MUG enzymes to cooperatively form routine dimeric complex, leading to the wild-type like band shift patterns and cooperativity shown in Figure 34. This result bears out the speculation from the study of MUG and TDG dimer crystal structure alignment (Figure 29C) that the steric positioning of TDG dimer and MUG dimer are mutually exclusive and thus both complexes could not form at the same time.

Chapter 4

Enzyme communication

4.1 Background and objectives

The AP site products of monofunctional DNA glycosylases are frequently more unstable and degrade autocatalytically to generate DNA strand breaks. The presence of AP sites in DNA generally impedes DNA replication and may also be potentially mutagenic due to lack of base coding information. The potent cytotoxicity of these AP lesions necessitates the repair machinery to insert the appropriate base with minimum exposure of the AP site. In keeping with this imperative, MUG has been reported previously by our lab that [129], under conditions of low MUG concentrations relative to an U:G substrate, in terms of steady state conditions, the product release step of the uracil glycosylase reaction with U:G substrate *in vitro* was exceedingly rate-limiting. This result suggests a high affinity of MUG for the AP site product, thereby leading to a reduction in MUG turnover. Moreover, many other DNA glycosylases, including the human thymine DNA glycosylase (hTDG), human OG glycosylase (hOGG1) and the adenine glycosylase MutY have also been reported to remain tightly bound to their respective AP site products [146, 155, 157].

Base excision repair involves several enzymes acting on the site of damage and it is essential that the order of their activity is precisely coordinated. Many previous studies in eukaryotes have shown that various BER enzymes assemble at the site of the DNA lesion or BER intermediate as a preformed “repairosome” and are charged with coordinated repair in each step of every sequential BER sub-pathway as summarized in Table 3, although upstream coordination of DNA glycosylase and AP endonucleases are less well studied. These enzyme interactions were originally compared to the passing of a baton, where each enzyme interacts with specific downstream enzymes so that substrates are handed off through an ordered BER pathway, as a result of that, providing an increase in specificity and efficiency to the BER pathway and facilitating the maintenance of genome integrity by preventing the

accumulation of highly cytotoxic repair intermediates. This is particularly pertinent where enzyme activity is inhibited by tight binding to its product, such as MUG, which presumably carries out a general protective mechanism, whereby coordination of enzyme activity in BER is achieved through displacement of the DNA glycosylase by the downstream AP endonuclease.

BER proteins	Interacting proteins	Citation
UNG2	PCNA, RPA, Vpr	[199-201]
SMUG	Vpr	[201]
TDG	APE1, p300, XPC/HR23B	[9, 162, 202, 203]
Aag/MPG	XRCC1, HR23A, MBD1	[204-206]
hOGG1	APE1, XRCC1, CSB	[207, 208]
MutY	ExoIII, EndoIV	[161]
MYH	APE1, PCNA, RPA	[209, 210]
NTH1	XPG	[211, 212]
NEIL1	Polβ, LigIIIα	[213]
NEIL2	Polβ, LigIIIα, PNKP, XRCC1	[213]
APE1	TDG, MYH, Polβ, LigI, FEN1, PCNA, p53, GzmA	[84, 162, 202, 214-219]
Polβ	APE1, LigI, FEN1, PCNA, PARP1,PARP2, PNKP, NEIL1, NEIL2, Aprataxin, p53m Trf2, p300, XRCC1, PRMT1, PRMT6	[79, 213, 217, 218, 220-234]
LigI	APE1, Polβ, PCNA	[216, 229, 235, 236]
LigIIIα	XRCC1, PARP1, PARP2, PNKP, NEIL1, NEIL2, Aprataxin	[79, 222, 223, 225, 237, 238]
FEN1	APE1, Polβ, PCNA, PARP1, p300, WRN, BLM	[215, 216, 220, 236, 239-243]

PCNA	APE1, Polβ, PARP1, XRCC1, p300, CSB, WRN	[199, 200, 210, 215, 232, 235, 236, 240, 244-251]
XRCC1	MPG, OGG1, APE1, Polβ, LigIIIα, PARP1, PARP2, PNKP, Tdp1, Aprataxin	[79, 84, 204, 207, 213, 222, 223, 225, 237, 252-256]
PNKP	Polβ, LigIIIα, XRCC1, NEIL1, NEIL2	[222]
PARP1	Polβ, PCNA, XRCC1, p300, CSB, WRN, Ku, LigIIIα	[220, 221, 234, 238, 244, 255, 257-261]
PARP2	Polβ, LigIIIα, PARP1, XRCC1	[225]
Aprataxin	Polβ, LigIIIα, PARP1, XRCC1	[223, 253-256]
Tdp1	XRCC1	[252]

Table 3. Protein-protein interactions among base excision repair proteins.

Previous studies of BER in organisms ranging from *E.coli* to human have shown many DNA glycosylases exhibit a stimulated turnover in the presence of their downstream AP endonucleases *in vitro* [9, 117, 155, 158, 160-162, 262-266]. The new scenario depicted by these findings indicates that some degree of enzyme coordination occurs in the initial steps of BER and AP endonucleases have been consistently observed to affect product release rather than the chemical step of the glycosylase reaction [9, 155, 161, 209]. However, the communication between DNA glycosylases and AP endonucleases has still remained elusive and may occur via distinct mechanisms. An active mechanism would require the AP endonuclease to directly interact with the DNA glycosylase and/or DNA adjacent to the AP site to disrupt the binding interface of DNA-glycosylase complex, facilitating glycosylase dissociation from its product. Coordination may also be achieved through the formation of an intermediate ternary DNA-glycosylase-AP endonuclease complex that facilitates enzyme hand off. Alternatively, a kinetic enhancement may be

observed through stochastic passive processes, whereby the AP endonuclease only acts when DNA glycosylase dissociates, the occupancy of the AP site by the AP endonucleases depletes the concentration of AP site and thus relieves the product inhibition of the DNA glycosylase and leads to increased glycosylase turnover.

If the stochastic passive mechanism prevails, both AP endonucleases in *E.coli*, ExoIII and EndoIV, would be expected to stimulate DNA glycosylases of interest in parallel. In a study of MutY, which is an *E.coli* BER DNA glycosylase that can remove misincorporated adenine residues from OG:A, G:A and C:A mispairs, Pope *et al.* found that EndoIV enhances MutY turnover rate to a higher extent than ExoIII on a G:A substrate [161], however, surprisingly, neither AP endonuclease has a stimulatory effect on the product dissociation of MutY with an OG:A substrate, which confers a higher affinity to MutY and is deemed to be an extremely important biological substrate of this DNA glycosylase [267]. These findings suggest that ExoIII and EndoIV use distinctly different substrate dependent mechanisms to stimulate MutY turnover and raise the question of the biological relevance of the stimulatory effects of the AP endonucleases.

Further analysis by band shift assay by Pope *et al*. demonstrated that ExoIII binds to the MutY-product DNA complex to generate a notable higher shifted band, in contrast, EndoIV displaces MutY from the product DNA and then forms an EndoIV-product DNA complex to give a lower shifted band. This finding suggests that the interactions between MutY-product DNA complex and the AP endonucleases are idiosyncratic, and again indicates that ExoIII and EndoIV employ clearly different coordination mechanisms to stimulate MutY turnover. Moreover, this study also indicated hAPE1, the human homologue of ExoIII, does not stimulate turnover of MutY with a G:A substrate although ExoIII does so. This indicated that the recognition of the AP site alone is not responsible for the stimulated turnover and suggested that there may be

specific recognition of the *E.coli* MutY-product DNA complex by the *E.coli* AP endonucleases.

However, the passive mechanism can not be completely ruled out because the large excess of AP endonucleases needed for the MutY turnover stimulation could readily result in the endonucleolytic depletion of spontaneously released abasic product. Prominent stimulatory effects of both ExoIII and EndoIV on the MutY turnover were observed respectively *in vitro* at high molar ratios of AP endonucleases, and until 95 fold excess of ExoIII or 17 fold excess of EndoIV was added to the MutY reactions with a G:A substrate, the product formation plots no longer retained biphasic character [161]. The stimulatory effects of AP endonucleases on many other DNA glycosylases were also ascertained but only in the presence of a large excess of AP endonucleases, for example, a 10:1 molar ratio was needed to see a stimulatory effect of APE1 on SMUG [268, 269], and 15 fold excess of APE1 was required for a two fold increase in the rate of OGG1 turnover [160].

The need for high AP endonuclease: DNA glycosylase ratios to detect the enhancement of DNA glycosylase turnover rate *in vitro* may be attributed to the absence of *in vivo* enzyme modifications, such as acetylation and sumoylation. For instance, acetylation of the N-terminal region of hTDG by CREB-binding protein (CBP) and p300 (CBP/p300) abrogates high-affinity DNA binding [270]. Acetylation of APE1 N-terminal lysine residues Lys27, Lys31, Lys32 and Lys35 by p300 plays important roles *in vivo* in deacetylating residues Lys6 and Lys7 by SIRT1, which is important in coordinating and fine-tuning the enzyme's BER activity [271, 272]. Moreover, product bound TDG is modified by small ubiquitin-like modifiers (SUMOs), which facilitates its dissociation from the product AP site [273]. Lack of these enzyme modifications *in vitro* might explain why a large excess of AP endonucleases are required necessary for some substrates.

Nevertheless, while the study by Waters *et al.* didn't show a significant increase in turnover when APE1 and hTDG were at equimolar amount, a 26 fold and 42 fold increased hTDG turnover for G:U and G:T substrates were observed by Fitzgerald and Drohat at a 1:1 molar ratio of APE1 to TDG using burst kinetics experiments. Contrary to the conventional steady state kinetics, burst kinetics has a virtue to readily determine the maximal rate constant for enzymatic turnover (k_{cat}) by conducting with a high enzyme concentration and excess substrate ([S] > [E] >> K_D) such that k_{cat} is not limited by the association of enzyme and substrate. Progress curves exhibited "burst" kinetics, with a rapid exponential phase followed by a slow linear phase, indicating that the rate of product formation (enzyme-bound) greatly exceeds that of product release [274].

Under burst kinetics conditions, Fitzgerald and Drohat subsequently analyzed the stimulatory effects of APE1 on the turnover rates of wild-type hTDG and truncated hTDG, which lacks the N- and C-terminal domains but retains the catalytic domain ($hTDG^{cat}$). Despite k_{cat} for full-length wild-type hTDG was found to be identical to that for truncated $hTDG^{cat}$, yet the APE1-stimulated turnover of full-length hTDG was lower than that of $hTDG^{cat}$. This finding suggests that the N- and/or C-terminal regions diminish the stimulatory effects of APE1 and thus is indicative of an active displacement mechanism involves interaction between hTDG and APE1. Moreover, active displacement was also indirectly ascertained by another finding that EndoIV could not stimulate the turnover of hTDG with either G:T or G:εC substrate [162, 275], because if a passive mechanism prevailed, EndoIV would be expected to stimulate hTDG turnover like APE1 does.

Since both MUG's downstream AP endonucleases, ExoIII and EndoIV, and its human homologue, hTDG, have all been observed to participate in enzyme coordination

during the initial steps of BER, it is highly likely that similar stimulatory mechanisms exist for MUG. The dissociation of MUG from its product may require the recruitment of its downstream AP endonucleases as turnover-enhancing cofactors to either actively stimulate its turnover or eliminate its AP DNA product. Therefore, in this chapter, the questions of whether ExoIII and EndoIV enhance the turnover of MUG and if the enhancements occur via an active or a passive mechanism will be addressed. Single turnover assays were firstly employed to determine if either ExoIII or EndoIV stimulates the glycosylase activity of MUG with a U:G substrate. Burst kinetics assays were then performed to ascertain if MUG turnover is facilitated in the presence of either ExoIII or EndoIV. Moreover, to determine whether the AP endonuclease activity of AP endonucleases is necessary to stimulate the glycosylase activity and turnover of MUG, a catalytically inactive ExoIII mutant D151N was constructed and used in both single turnover and burst kinetics experiments, and the results obtained will be discussed herein in comparison to those of wild-type ExoIII.

4.2 Results

4.2.1 Cloning, expression and purification of *E.coli* ExoIII

E.coli exonuclease III gene *xthA* of approximately 0.8 kbp was amplified from *E.coli* K12 genomic DNA via PCR (Figure 36A) and cloned into an expression vector pPROEX-HTb (supplied by Dr Jan Silhan, imperial college London) using NdeI and HindIII sites, which adds an N-terminal polyhistidine tag with cleavage site for TEV protease to the final protein. The constructed ExoIII over-expressing plasmid was designated as pPROEX-HTb_ExoIIIwt and transformed into $CaCl_2$ competent *E.coli* JM109 cells (see Materials and Methods).

Successfully transformed cells were selectively grown on LB plates and individual colonies were picked to inoculate starter cultures, which were incubated overnight. The pPROEX-HTb_ExoIIIwt plasmids were purified and verified for presence of the insert *xthA* gene by sequencing prior to being transformed into *E.coli* Rosetta^{TM} strain BL21 DE3 cells, which are suited for high gene expression and show limited proteolysis due to a low background of protease enzymes.

In the pPROEX-HTb_ExoIIIwt vector, ExoIII expression is under the control of the *trc* promoter, which is repressed by the upstream *lac*I gene. Addition of IPTG causes release of the repressor and switches on protein expression. 1 mM IPTG was added to a starter culture at an OD_{600} of 0.7 – 0.8 and culture aliquots taken at various time points and analyzed via 12% SDS-PAGE. The overexpression of ExoIII produced a protein with a molecular weight of 30.97 kDa (Figure 36B). Increasing the induction period increased the level of protein expression, and although overnight induction caused maximal protein expression, 6 hour induction time was utilized for ExoIII purification.

Cultures for ExoIII overexpression were typically grown in three 500 ml of LB media to an OD_{600} of 0.7-0.8 at 37°C, cooled to room temperature and then protein expression was induced by the addition of 1 mM IPTG, and growth continued at room temperature for further 6 hours. The supernatant from the cell lysis was loaded onto a HiTrap metal chelating sepharose column charged with Ni^{2+} (see Materials and Methods). His-tagged ExoIII were then eluted from the Ni-column with 400 mM imidazole. Fractions containing ExoIII were pooled, desalted and concentrated prior to being incubated with 5 mM TEV protease to cut off the His-tag from ExoIII and further purified via size exclusion chromatography using a gel filtration column (Figure 36C). Highly purified ExoIII was eluted within fractions B3 to B8 analyzed via 15% SDS-PAGE (Figure 36D). These fractions were then pooled and concentrated. A final 2 ml of 5.6 mg ml^{-1} ExoIII stock solution was obtained after adding glycerol to 20% (v/v) and the aliquots were snap frozen and stored at -80°C. The protein concentration was calculated from the OD_{280} based on an extinction coefficient of 39545 $M^{-1}cm^{-1}$.

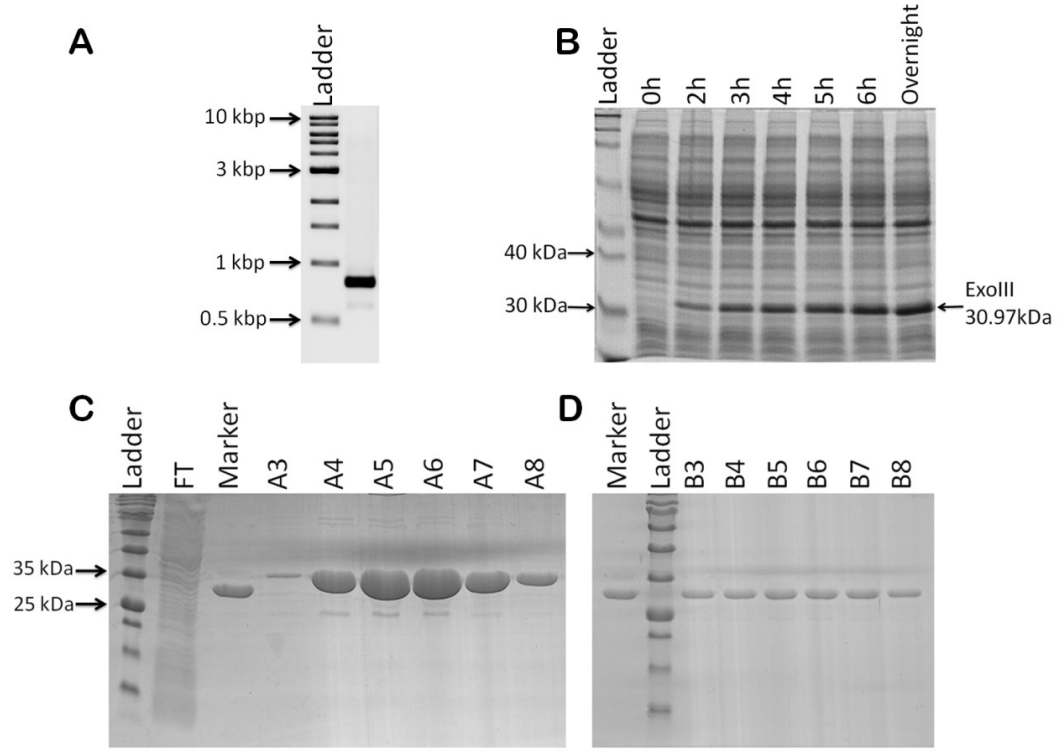


Figure 36. Cloning, expression and purification of *E.coli* ExoIII.
(A) ~0.8 kbp *E.coli* ExoIII gene *xthA* was produced via PCR-amplification of *E.coli* K12

genomic DNA and analyzed by 1% agarose gel electrophoresis. (B) 15% SDS-PAGE analysis of ExoIII expression level of 1mM IPTG-induced *E.coli* BL21 DE3 cell culture samples at OD_{600} of 0.7-0.8 within increasing induction time. (C) 15% SDS-PAGE analysis of ExoIII protein fractions collected from the Ni-column. The marker lane contains a commercial wild type ExoIII marker. (D) 15% SDS-PAGE analysis of ExoIII protein fractions eluted from the gel filtration column. The marker lane was loaded with the commercial wild type ExoIII protein.

4.2.2 Cloning, expression and purification of *E.coli* EndoIV

E.coli endonuclease IV gene *nfo* of approximately 0.85 kbp was amplified from *E.coli* K12 genomic DNA using PCR and cloned into the expression vector pPROEX-HTb using BamHI and HindIII sites (see Materials and Methods). The constructed EndoIV over-expressing plasmid, designated as pPROEX-HTb_EndoIVwt, was then purified from the *E.coli* JM109 transformants, verified by sequencing for presence of the insert *nfo* gene and transformed into *E.coli* Rosetta^{TM} strain BL21 DE3 cells.

The over-expression of EndoIV produced a protein with a molecular weight of 31.6 kDa, and 1mM IPTG addition to the cell cultures at OD_{600} of 0.7-0.8 and further 6 hour induction period at room temperature were determined to be ample for a high level of EndoIV expression (result not shown here).

The supernatant of cell lysate harvested from three 500 ml cell cultures was loaded onto a HiTrap metal chelating sepharose column charged with Zn^{2+} (see Materials and Methods). His-tagged EndoIV were then eluted from the Ni-column with 400 mM imidazole. Fractions containing EndoIV were pooled, desalted and concentrated prior to being incubated with 5 mM TEV. The pooled and TEV-digested fractions were further purified via size exclusion chromatography using a gel filtration column. Fractions B3 to B5 were detected to contain pure EndoIV protein via 15% SDS-PAGE, and then pooled and concentrated (Figure 37B). A final 1.8 ml of 1.9 mg ml^{-1} EndoIV

stock solution was obtained after adding glycerol to 20% (v/v) and the aliquots were snap frozen and stored at -80^{o}C. The protein concentration was calculated from the OD_{280} based on an extinction coefficient of 31315 $M^{-1}cm^{-1}$.

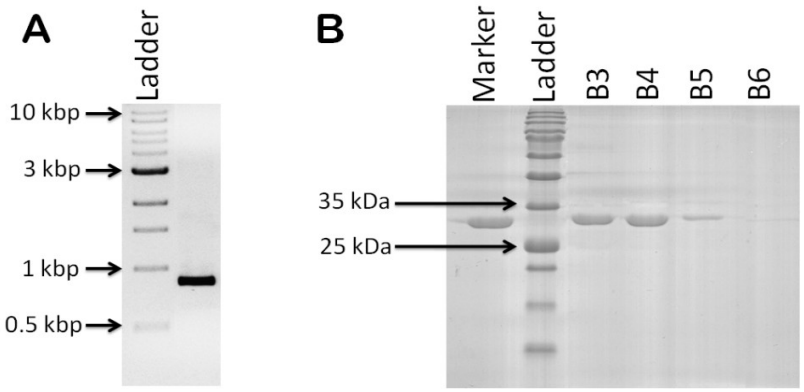


Figure 37. Cloning, expression and purification of *E.coli* EndoIV.
(A) ~0.85 kbp *E.coli* EndoIV gene *nfo* was produced via PCR-amplification of *E.coli* K12 genomic DNA and analyzed by 1% agarose gel electrophoresis. (B) 15% SDS-PAGE analysis of EndoIV protein fractions collected from the two-step chromatography using the Ni-column and the gel filtration column. The marker lane was loaded with the commercial wild type EndoIV protein.

4.2.3 Construction and purification of mutant ExoIII D151N

Neisserial AP endonuclease (NApe) is one of two Exonuclease III family enzymes in the human pathogen *Neisseria meningitidis* and exhibits a high degree of structural conservation with *E.coli* ExoIII. Previous study in our lab has shown that a conserved aspartate residue Asp149 of NApe plays a key role in dictating the enzyme's catalytic activity and its asparagine substitution brought on a catalytically inactive NApe mutant protein [182].

The Asp149 residue of NApe corresponds to the Asp151 residue of *E.coli* ExoIII. Therefore, in order to create a catalytically inactive ExoIII mutant, asparagine substitution of the Asp151 residue was performed via PCR-based site-directed

mutagenesis method to create a full length mutant ExoIII over-expressing plasmid of approximately 5.5 kbp (Figure 38A), which was designated as pPROEX-HTb_ExoIIID151N (see Materials and Methods). The mutant plasmid was then isolated from the *E.coli* JM109 transformants, verified by sequencing for presence of the mutation and then transformed into *E.coli* RosettaTM strain BL21 DE3 cells.

The expression and purification of ExoIII D151N were identical to those of wild-type ExoIII and therefore not described here. As shown in Figure 38B, fractions B3 to B5 eluted from a gel filtration column show the presence of pure ExoIII D151N protein. Fractions B3 to B5 were pooled and concentrated. A final 1.7 ml of 2.4 mg ml^{-1} ExoIII D151N stock solution was obtained after adding glycerol to 20% (v/v) and the aliquots were snap frozen and stored at -80^{o}C. The protein concentration was calculated from the OD_{280} based on an extinction coefficient of 39545 $M^{-1}cm^{-1}$.

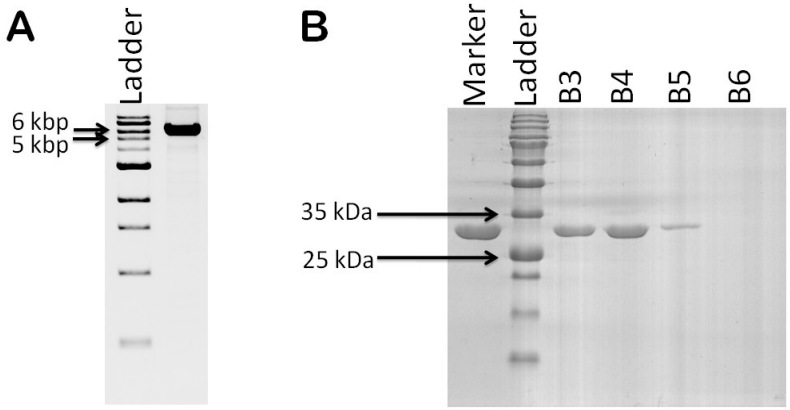


Figure 38. Site-directed mutagenesis and purification of mutant ExoIII D151N.
(A) ~5.5 kbp *E.coli* ExoIII D151N mutant plasmid was produced via PCR-based site-directed mutagenesis and analyzed by 1% agarose gel electrophoresis. (B) 15% SDS-PAGE analysis of ExoIII D151N protein fractions collected from the two-step chromatography using the Ni-column and the gel filtration column. The marker lane was loaded with the above purified wild-type ExoIII protein.

4.2.4 AP endonuclease activities of ExoIII, EndoIV and ExoIII D151N under burst kinetics conditions

The AP endonuclease activities of ExoIII, EndoIV and ExoIII D151N proteins were investigated under burst kinetics conditions in which a 5 fold excess of a 25-mer 5' end Hex-labeled 3'overhanging double stranded DNA substrate (Hex-12AP·G o/h) was reacted with enzymes respectively. Hex-12AP·G o/h possesses a AP site opposite to a guanine and 12 bases from the 5' end, and contains 4-mer 3' overhangs on both the AP and the G strand, which has previously been shown to inhibit the 3'→5' exonuclease activity of *E.coli* AP endonucleases [276].

AP endonuclease reactions were initiated by incubating 100 nM enzymes separately with 500 nM native Hex-12AP·G o/h substrates at 25^{o}C in the standard reaction buffer, and reaction aliquots were rapidly quenched in the AP endonuclease quenching buffer after different incubation times (see Materials and Methods). Products were separated from DNA substrate by denaturing PAGE (Figure 39A) and analyzed by quantitating the percentages of the successively increased intensities of product bands (Figure 39B). The AP endonuclease activity of a commercial ExoIII enzyme from NEB was also probed under the parallel conditions as a positive control in comparison to those of home-made AP endonucleases.

As shown in Figure 39B, more than 90% of the DNA substrates had been depleted within 15 seconds by ExoIII and EndoIV respectively, indicating that both enzymes exhibit extremely high AP endonuclease activities, which are consistent with that of the commercial NEB ExoIII enzyme. Conversely, over the 20 minute time course of ExoIII D151N reaction, only ~10% DNA substrates were cleaved, indicating that there is only a negligibly low residual AP endonuclease activity with the D151N mutant. Therefore, D151N was used as a catalytically inactive enzyme in the

following experiments to investigate whether the catalytic activity of ExoIII is necessary to stimulate the glycosylase activity and turnover of MUG. The burst kinetics rates of both ExoIII and EndoIV were not calculated here because the high velocity of their reactions made it impossible to collect more time point samples during the initial burst phase of product formation (within 15 seconds) to quantify and fit into a single exponential equation. In addition, as shown on both ExoIII and NEB ExoIII gels in Figure 39A, the 12-mer products generated from the cleavage at the AP site in turn acted as substrates for the 3'→5' exonuclease activity of ExoIII and were digested into a variety of exonuclease products.

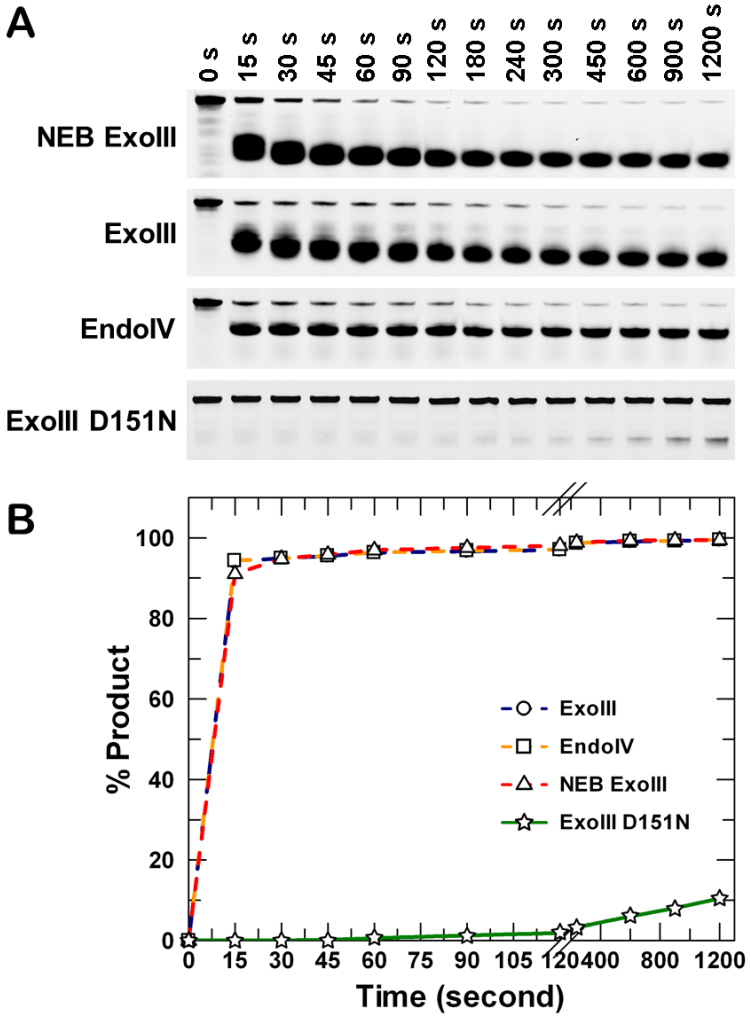


Figure 39. AP endonuclease activities of ExoIII, EndoIV and ExoIII D151N under burst kinetics conditions.

(A) Burst kinetics experiments were performed by reacting 100 nM ExoIII, EndoIV, ExoIII D151N and NEB ExoIII enzymes separately with 500 nM Hex-12AP·G o/h, reaction aliquots

were quenched after different incubation times and analyzed by denaturing PAGE (see Materials and Methods). (B) The percentages of product formation at different time points were determined by quantitating the intensities of product bands using the Phoretix 1D. Data of ExoIII (blue dash line), EndoIV (yellow dash line) and ExoIII D151N (green line) were plotted against time and compared with that of the commercial NEB ExoIII (red dash line).

4.2.5 The effect of *E.coli* AP endonucleases on MUG glycosylase activity

The rate constants of wild-type MUG glycosylase activity with a U·G containing substrate in the presence of either ExoIII or EndoIV were examined under single turnover conditions, with a 5 fold excess of enzyme (500 nM) over substrate (100 nM), and the accumulation of product corresponded to the rate of *N*-glycosidic bond hydrolysis.

Reactions were initiated by incubating equimolar amounts of MUG and either ExoIII or EndoIV enzymes in the standard reaction buffer at 25^{o}C with a 25-mer 5' end Hex-labeled, and 3'4-mer overhang and uracil containing double-stranded DNA substrate (Hex-12U·G o/h). Because ExoIII is a divalent metal ion dependent enzyme, 1 mM $MgCl_2$ was added to the standard reaction buffer to sustain its enzymatic activities. By contrast, EndoIV possesses an intrinsic trinulear zinc cluster in its active site for its enzymatic catalysis and is even active in the presence of 1-10 mM EDTA without added divalent metals [174], therefore, no further Zn^{2+} ions were added into the standard reaction buffer for the reaction with EndoIV. Aliquots of the reaction mixtures were taken at selected time points and quenched with NaOH, then products were separated from intact substrate by denaturing PAGE (Figure 40A) and analyzed by quantitating the percentages of the successively increased intensities of product bands and fitting the data to a first-order rate equation (Figure 40B) (see Materials and Methods).

The products generated from both MUG reactions with additions of ExoIII and EndoIV showed a single exponential increase in production with cleavage rates (k_{cat}) of 0.040 ± 0.003 s^{-1} and 0.039 ± 0.002 s^{-1} respectively, which coincide with that (k_{cat} = 0.038 ± 0.003 s^{-1}) of MUG reaction in the absence of AP endonucleases. These indicate neither ExoIII nor EndoIV can stimulate the glycosylase activity of MUG under single turnover conditions. In addition, the 12-mer products, which were generated from the ordered initial BER steps, in terms of *N*-glycosidic bond hydrolysis by MUG and the downstream AP site cleavage by either ExoIII or EndoIV, in turn acted as substrates for the 3'→5' exonuclease activity of both ExoIII and EndoIV and lower product bands were observed on both ExoIII and EndoIV gels. However, the lower product bands on the ExoIII gel appeared much earlier than those on EndoIV gel, indicating that ExoIII displays a higher extent of 3'→5' exonuclease activity compared to that of EndoIV.

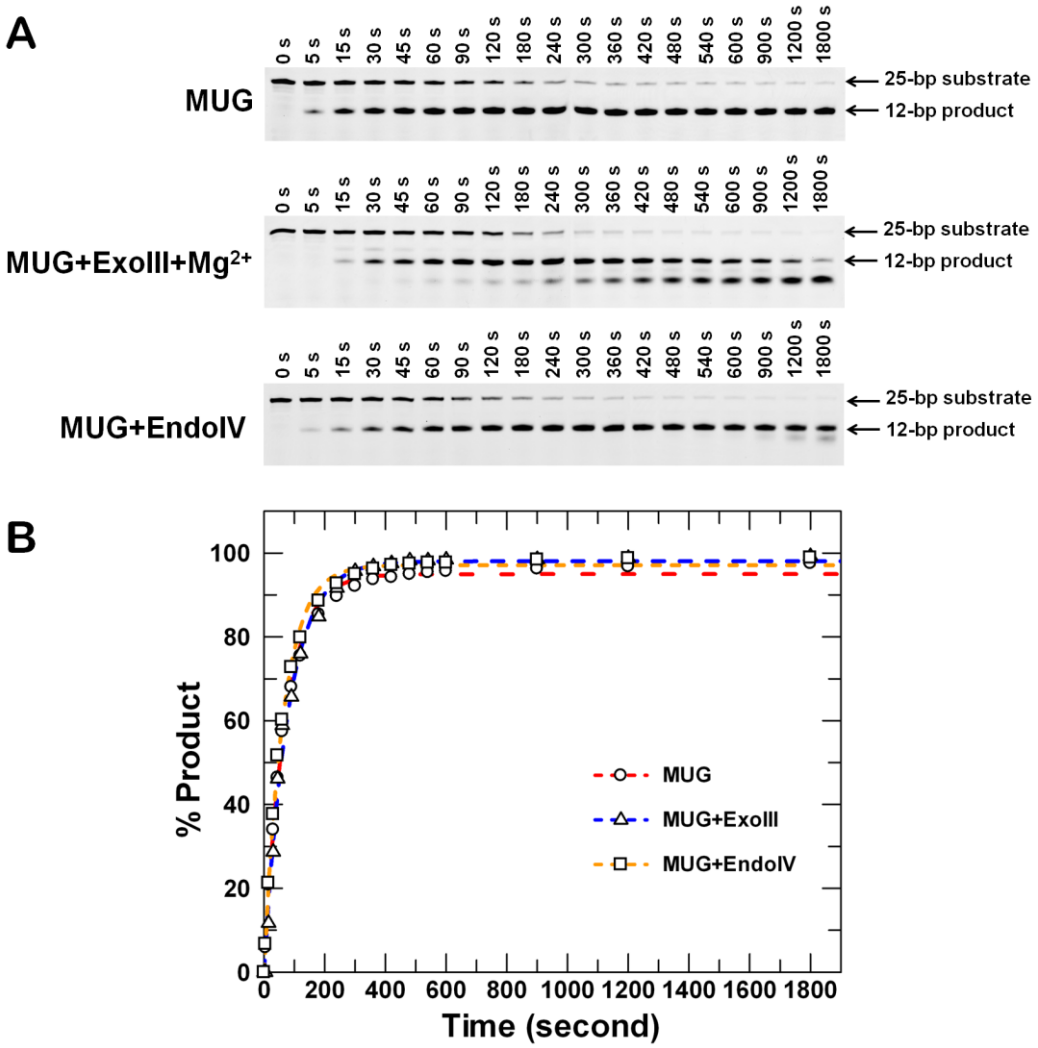


Figure 40. The effect of ExoIII and EndoIV on MUG glycosylase activity.
(A) 500 nM MUG was reacted with 100 nM Hex-12U·G o/h in the absence and presence of either 500nM ExoIII or EndoIV in the standard reaction buffer at 25°C, reaction aliquots were taken at selected time points and quenched with NaOH. The products were then separated from the Hex-12U·G o/h substrate by denaturing PAGE (see Materials and Methods). (B) The percentages of product formation at different time points were determined by quantitating the successively increased intensities of product bands using the Phoretix1D. The data of both MUG reactions in presence of ExoIII (blue dash line) and EndoIV (yellow dash line) are shown with the best fits to the first order rate equation with k_{cat} values of 0.040 ± 0.003 s^{-1} and 0.039 ± 0.002 s^{-1} respectively, which coincide with that of MUG only reaction data (red dash line) with a k_{cat} of 0.038 ± 0.003 s^{-1}.

4.2.6 The effect of *E.coli* AP endonucleases on MUG turnover

The results described above in 4.2.5 have clearly revealed that the presence of either ExoIII or EndoIV does not perturb the intrinsic chemical process of MUG associated with the uracil removal, and therefore, the possible effect of ExoIII and EndoIV is limited to the steps governing the steady state rate.

Burst kinetics assays were thus conducted herein to analyze the effect of ExoIII and EndoIV on the steady state turnover of MUG with a U·G containing substrate. Reactions were initiated by rapidly mixing a 5 fold excess of Hex-12U·G o/h substrate (2500 nM) with MUG (500 nM) in the presence or absence of the equimolar amounts of either ExoIII or EndoIV (500 nM), and incubated in the standard reaction buffer at 25^{o}C for 12 hours. Aliquots of the reaction mixture were taken at hourly intervals during the 12 hour incubation time course and then quenched with NaOH (see Materials and Methods). Products were separated from intact substrate by denaturing PAGE as shown in Figure 41A. The intensities of product bands were quantitated to determine the product concentrations, which were then plotted against the hourly quenching time points (Figure 41B).

Comparisons of the amplitudes of the exponential burst phase of MUG turnover kinetics in the absence and presence of either ExoIII or EndoIV reveal that both ExoIII and EndoIV significantly enhance the turnover of MUG with the Hex-12U·G o/h substrate, and the stimulatory effect of ExoIII on MUG turnover is much stronger than that of EndoIV. Moreover, after cleavage at the AP site, both ExoIII and EndoIV show an immediate further 3'→5' exonuclease progression along the DNA product and as shown in Figure 41A, a variety of exonuclease products bands were observed on the MUG + ExoIII + Mg^{2+} gel and a single lower exonuclease product band appeared on the MUG + EndoIV gel.

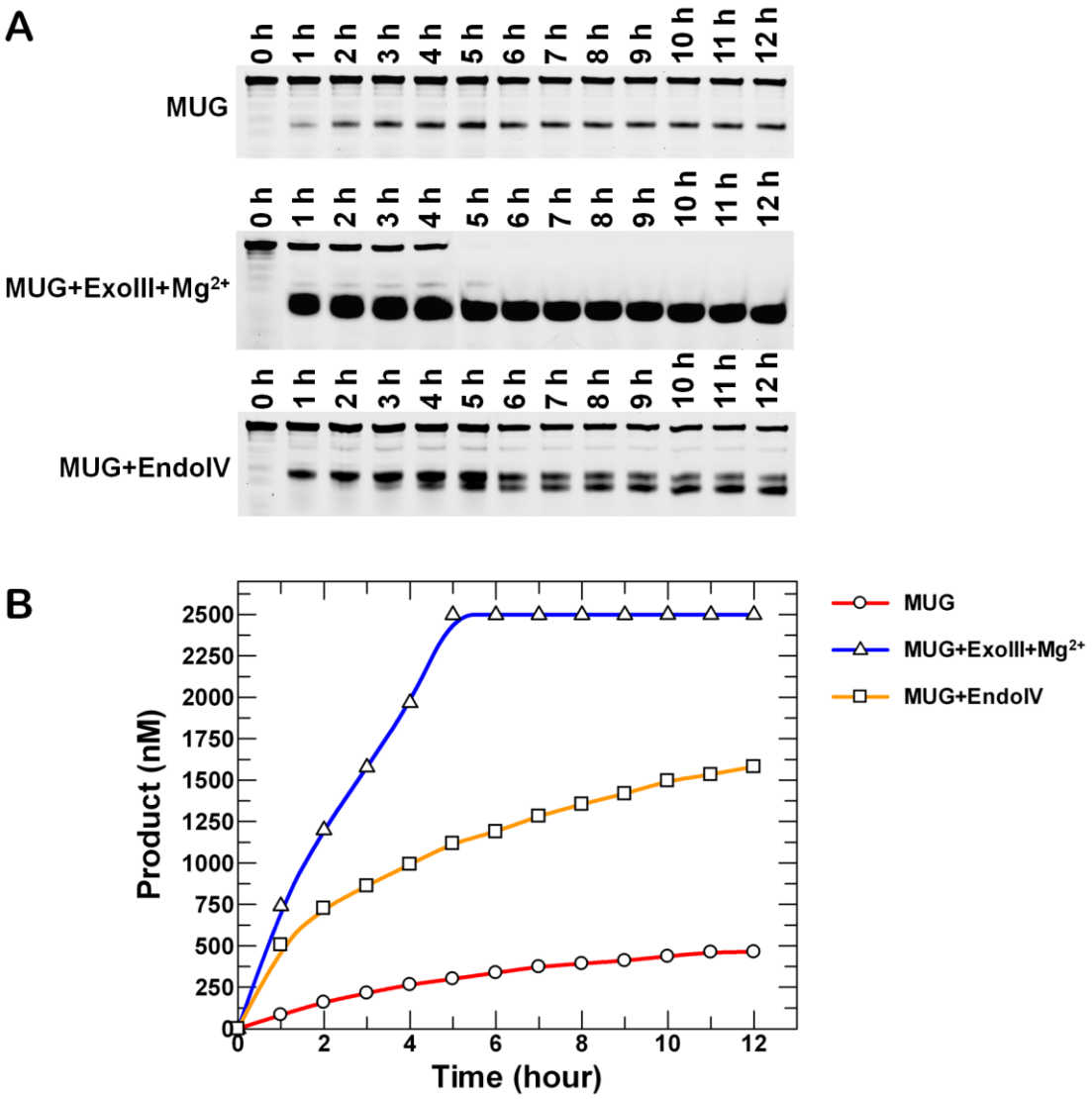


Figure 41. The effect of ExoIII and EndoIV on MUG turnover.

(A) 500 nM MUG was reacted with 2500 nM Hex-12U·G o/h in the absence and presence of either 500 nM ExoIII or EndoIV in the standard reaction buffer at 25°C, 1 mM $MgCl_2$ was added to the reaction mixture with ExoIII. Reaction aliquots were taken at selected time points and quenched with NaOH. The products were then separated from the Hex-12U·G o/h substrate by denaturing PAGE (see Materials and Methods). (B) The product concentration at different quenching time points were determined by quantitating the successively increased intensities of product bands using the Phoretix1D. The data of MUG reactions in the absence (red line) and presence of either ExoIII (blue line) or EndoIV (yellow line) are plotted against quenching time points.

4.2.7 ExoIII catalytic activity dependence of MUG turnover enhancement

To gain insight into the role of catalytic activity of *E.coli* AP endonucleases in enhancing MUG turnover, 10 mM EDTA and an inactive mutant ExoIII D151N protein were respectively employed in the burst kinetics assays of MUG with the Hex-12U·G o/h substrate.

ExoIII is a divalent metal ion dependent enzyme and it has been previously reported that 10 mM EDTA present in the reaction buffer could completely inhibit the catalytic activity of ExoIII by sequestering Mg^{2+} ions from ExoIII active site [161], but our laboratory has shown previously that ExoIII is still capable of binding to the abasic DNA in the absence of Mg^{2+} ions [193]. Moreover, mutant ExoIII D151N protein has been confirmed in 4.2.4 to be an almost catalytically inactive protein after the mutation and only retains a negligible residual activity. EndoIV was not used for the analysis of its catalytic activity dependence of MUG turnover enhancement, because 1) it is resistant to the inactivation by the metal chelating agents, such as EDTA; 2) since it was found to has a relatively weak stimulatory effect on MUG turnover compared to ExoIII, EndoIV has not been the main study focus of this project at present.

As shown in Figure 42, when 10 mM EDTA was present in the MUG reaction with ExoIII, the exponential MUG burst phase kinetics displayed a same amplitude as that of MUG alone reaction, indicating that the EDTA-treated ExoIII completely lost the stimulatory effect on MUG turnover and thus suggesting the catalytic activity of ExoIII is essential on the MUG turnover enhancement. While the mutant D151N was introduced to the reaction in the presence of 1 mM $MgCl_2$, D151N didn't show a wild type like stimulatory effect and only a very limited extent of MUG turnover

enhancement was observed, this was deemed to be due to the residual catalytic activity of D151N. This result again suggests that the MUG turnover enhancement is dependent on the catalytic activity of ExoIII.

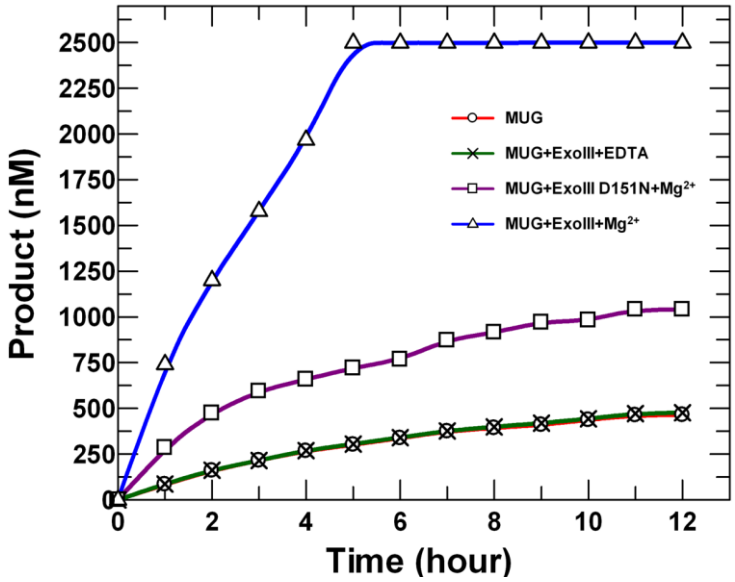


Figure 42. ExoIII catalytic activity dependence of MUG turnover enhancement.
500 nM MUG was reacted with 2500 nM Hex-12U·G o/h in the standard reaction buffer at 25°C in the presence of either 500 nM ExoIII with 10mM EDTA or 500 nM catalytically inactive mutant ExoIII D151N with 1 mM $MgCl_2$. Reaction aliquots were taken at hourly intervals during the 12 hour reaction time course and quenched with NaOH. The products were then separated from the Hex-12U·G o/h substrate by denaturing PAGE (see Materials and Methods) (gel pictures not shown). The product concentration at different quenching time points were determined by quantitating the successively increased intensities of product bands using the Phoretix1D.The data of MUG reactions in the presence of either ExoIII with 10 mM EDTA (green line) or ExoIII D151N with 1 mM $MgCl_2$ (purple line) are plotted against quenching time points and then compared to those of MUG alone reaction (red line) and MUG reaction with 1 mM $MgCl_2$activated ExoIII (blue line).

4.2.8 The effect of ExoIII concentration on MUG turnover enhancement

The requirement of catalytically active ExoIII for MUG turnover enhancement gives a hint that ExoIII may employ a product scavenging mechanism, by which ExoIII cleaves or sequesters the AP site product released by MUG and thus prevent re-association of the MUG with its product, resulting in the turnover enhancement of MUG. Therefore, if it is true, the amount of ExoIII present in MUG steady state reaction is essential to the depletion of MUG released abasic product and hence the rate of MUG turnover. Here, burst kinetics assays of MUG were carried out in the presence of increasing amounts of ExoIII with the Hex-12U·G o/h substrate. The catalytically inactive mutant ExoIII D151N was also examined under the parallel conditions as a negative control.

Reactions were initiated by rapidly adding increasing concentrations (100 nM, 500 nM, 1000 nM and 2000 nM) of either ExoIII or D151N separately into reactions of 500 nM MUG with a 5 fold excess of Hex-12U·G o/h substrate (2500 nM), and incubated in the standard reaction buffer with 1 mM $MgCl_2$ at 25^{o}C for 12 hours. Aliquots of the reaction mixture were taken at selected time points during the 12 hour incubation time course and then quenched with NaOH (see Materials and Methods). The products were separated from the Hex-12U·G o/h substrate by denaturing PAGE (gel pictures not shown). The intensities of product bands were quantitated to determine the product concentrations, which were then plotted against the quenching time points (Figure 43A). Whilst the larger amounts of ExoIII participated in the MUG reaction, the depletion of Hex-12U·G o/h substrate was achieved faster, hence representing the overall rate of MUG turnover was enhanced to a higher degree. Although all the MUG reactions in the presence of increasing concentrations of D151N were still very far from complete within 12 hours, the

amplitudes of MUG burst phase kinetics increased when D151N concentration increased in the MUG reaction. The faint extent of MUG enhancement by increasing concentrations of D151N is attributed to the residual catalytic activity of D151N. The combined results from ExoIII and D151N clearly demonstrate that the ExoIII concentration has a great impact on MUG turnover enhancement. More striking differences on MUG turnover rate were observed in the early timescale between MUG reactions with or without either 2000 nM ExoIII or EndoIV (Figure 43B). This result suggests that MUG enzyme may preferentially bind to substrate DNA as a monomer without cleavage under multiple turnover state. When ExoIII and EndoIV were separately introduced into the multiple turnover system of MUG, the initial cleavage rate or the product accumulation rate increased significantly.

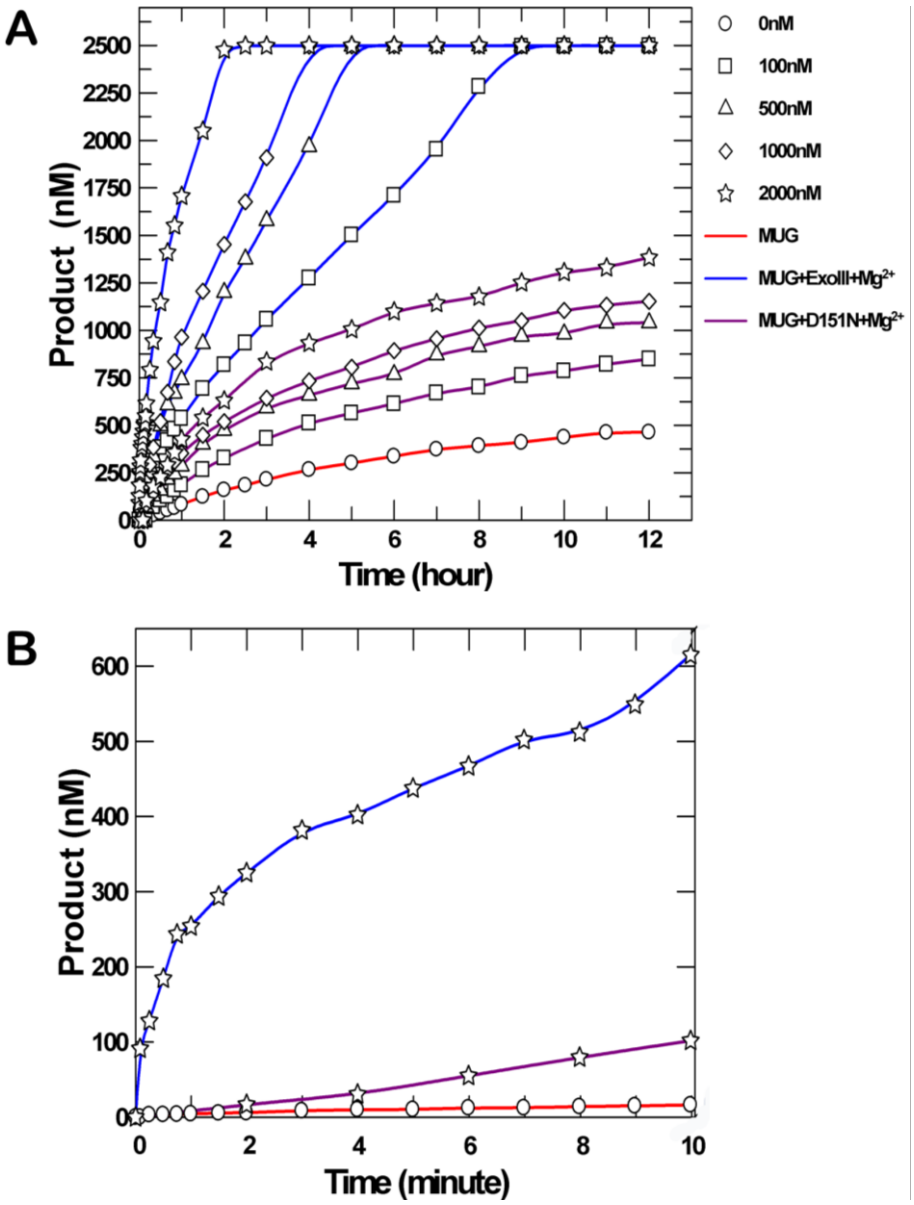


Figure 43. The Effect of ExoIII concentration on MUG turnover enhancement.
500 nM MUG was reacted with 2500 nM Hex-12U·G o/h in the standard reaction buffer at 25°C in the presence of increasing concentrations of either ExoIII or ExoIII D151N with 1 mM $MgCl_2$. Reaction aliquots were taken at selected time points during the 12 hour reaction time course and quenched with NaOH. The products were then separated from the Hex-12U·G o/h substrate by denaturing PAGE (see Materials and Methods) (gel pictures not shown). The product concentration at different quenching time points were determined by quantitating the successively increased intensities of product bands using the Phoretix1D. (A) The data of MUG alone reaction and reactions with either ExoIII (blue line) or ExoIII D151N (purple line) concentrations (red line with ○) of 100nM (□), 500nM (△), 1000nM (◇) and 2000nM (☆) in the presence of 1 mM $MgCl_2$ are plotted against quenching time points. (B) The data of MUG reactions without (red line with ○) or with either 2000nM ExoIII or EndoIV (☆) are plotted against early quenching time points.

4.3 Discussion

Like its human homologue TDG, MUG is widely regarded as a "single turnover" enzyme because it still remains tightly bound to its abasic product DNA after cleaving the uracil, thus impeding its catalytic turnover [129]. The slow turnover of MUG must arise from the enzyme's cooperative binding, which was described in chapter 2. This may reflect an imperative to protect the cell from the deleterious effect of the intermediate AP site. It is perhaps particularly important for MUG, because it is established that MUG is active in the stationary phase of the cell cycle. Under such growth conditions, enzymes normally are associated with DNA replication and repair are less abundant, so that protection of numerous unstable AP intermediates may be of greater importance.

During the past several years, the ordered pathway model of BER has been conceptually attractive as it involves defined, sequential steps, with the protection of reaction intermediates and cooperative enzyme activity. This allows sequestration of the potentially toxic intermediates, rapidly passing them along during the BER without further jeopardizing the genome stability. This model has received widespread support, with reports of enhanced DNA glycosylase turnover in the presence of AP endonucleases ranging from *E.coli* to human [9, 117, 158-162]. Therefore, it seems highly plausible that the actual turnover rate of MUG *in vivo* may be higher than that observed *in vitro.* This notion was evidenced in this chapter by the results obtained from MUG single turnover assays and burst kinetics assays in the separate presence of ExoIII and EndoIV, following the cloning, expression and purification of ExoIII and EndoIV proteins. Moreover, a catalytic ExoIII mutant, D151N, was also created and purified herein to be accompanied with wild-type ExoIII to examine the effects of catalytic activity and concentration on MUG turnover enhancement.

4.3.1 Orchestration in the initial steps of BER in *E.coli*

The initial investigation of MUG turnover enhancement by ExoIII and EndoIV were carried out using *in vitro* burst kinetics assays to monitor the change of amplitude of the exponential burst phase of MUG turnover kinetics, in terms of product accumulation/release rate, in the absence and presence of either ExoIII or EndoIV. As shown in Figure 41, in the case of the reaction of MUG (500 nM) with a 5 fold excess of 25-mer U•G substrate (2500 nM), the turnover of MUG was significantly enhanced upon addition of equimolar amounts of either ExoIII or EndoIV (500 nM). It is also noteworthy that ExoIII and EndoIV displayed significantly different degrees of stimulatory effect on MUG turnover. This indicates that recognition of the AP site alone is not responsible for the enhanced turnover and suggests that ExoIII and EndoIV may use distinctly specific mechanisms to render the orchestration with upstream MUG in the initial steps of BER in *E.coli*, which is consistent with the observation from MutY turnover kinetic studies by Pope *et al* that ExoIII and EndoIV use distinctly different substrate dependent mechanisms to stimulate MutY turnover [161].

4.3.2 Mismatch of MUG reactivity demeanor between single turnover state and multiple turnover state

The single turnover assays and burst kinetics assays collected here for MUG provide important new insight into its catalytic mechanism. As shown in Figure 40, in the case of the single turnover reaction of MUG (500 nM) with a 25-mer U•G substrate (100 nM), MUG enzymes preferentially formed a specific dimeric complex with its substrate DNA, by which 100% substrates were completely cleaved within 10 minutes with a k_{cat} value of $0.038 \pm 0.003\ s^{-1}$. When equimolar amounts of ExoIII or EndoIV (500 nM) were added into the MUG reaction, the cleavage rate constants (k_{cat}

$^{+ExoIII}$ = 0.040 ± 0.003 s^{-1} and k_{cat} $^{+EndoIV}$ = 0.039± 0.002 s^{-1}) remained unchanged. This result suggests that the presence of either ExoIII or EndoIV does not affect the intrinsic chemical process of the uracil glycosylase reaction and that the stimulatory effects are limited to the steps governing the product release, in terms of MUG enzyme turnover.

When the concentration of U•G substrate was elevated to 2500 nM but the MUG concentration remained invariant (500 nM), the reaction was therefore performed under the multiple turnover state as shown in Figure 41. In classic enzyme kinetics of a multiple turnover reaction, it would be expected to see an initial burst phase that corresponds to the single turnover activity, at least 250 nM product DNA should have been produced in the first few minutes by the dimeric cleavage of 500 nM MUG enzymes in the absence of ExoIII and EndoIV, or alternatively 500 nM product DNA if MUG is able to cleave as a monomer. However, the cleavage proceeded extremely slowly and it took 3 hours to accumulate just ~250 nM products, clearly but unexpectedly, demonstrating that the predicted burst phase was not observed (Figure 43). This result suggests that MUG enzyme may preferentially bind to substrate DNA as a monomer, but that the monomer is not capable of catalysis. When ExoIII and EndoIV were separately introduced into the multiple turnover system of MUG, the initial cleavage rate or the product accumulation rate increased significantly.

The combined results reveal that MUG won't cleave its substrate as a monomer, unless 1) it can form a stable dimeric complex with a second MUG molecule to repair the DNA lesion and then protect the intermediate AP site, or alternatively 2) when the downstream ExoIII or EndoIV is present, they can nonspecifically bind to U•G DNA somewhere and activate the cleavage of a single MUG molecule rather than two MUG molecules form a stable protein-DNA complex which directly hands off the AP product to either ExoIII or EndoIV. This pre-catalytic discrimination ability of MUG

is a completely new concept, in terms of how DNA repair enzymes may coordinate their reactivity behavior.

4.3.3 ExoIII catalytic activity dependence of MUG turnover enhancement

While the stimulatory effects of *E.coli* AP endonucleases on MUG turnover were revealed above, it was not clear whether they were dependent on the catalytic activity of ExoIII and EndoIV. To address this question, burst kinetics assays of MUG were conducted with a 25-mer U•G substrate in the presence of either 10 mM EDTA or a catalytically inactive ExoIII mutant D151N. The combined results (Figure 42) clearly demonstrated that the ExoIII catalytic activity plays a key role in enhancing MUG turnover. It is therefore implies that ExoIII may use a product scavenging mechanism to deplete the concentration of AP site, prevent MUG from rebinding to its product and recycle MUG back to carry out new uracil glycosylase reactions.

4.3.4 The effect of ExoIII concentration on MUG turnover enhancement

The stimulatory effects of *E.coli* AP endonucleases on MUG turnover described in Section 4.3.1 clearly indicates that the initial steps in BER are highly coordinated where the deficiency of one BER component can result in upstream and/or downstream repair to be rate-limited or even defective. Therefore, the relative concentration of each repair factor would be expected to have a strong impact on the efficiency and mode of repair at damaged sites. Accordingly, MUG burst kinetics assays were performed with increasing concentrations of wild-type ExoIII and mutant D151N, and demonstrated that the higher concentrations of ExoIII introduced into the

reaction were directly proportional to the higher product accumulation rates. It therefore suggests that the rate-limiting step for MUG turnover in the presence of ExoIII now becomes the cleavage of the AP DNA, and basically more ExoIII molecules in the reaction lead to more cleavages of AP site and hence more MUG molecules dissociate from the DNA to recycle into next round of uracil glycosylase reactions. This result validates the above speculation in Section 4.3.3 that ExoIII employs a product scavenging mechanism to enhance MUG turnover.

Chapter 5
General discussion & Conclusion

5.1 DNA binding modes of MUG

In this project, differential binding modes of MUG with its abasic DNA product and non-specific DNA have been demonstrated using both fluorescence anisotropy assays and band shift assays (see Chapter 2). The sigmoidal response of the DNA binding isotherms with abasic DNA demonstrates a significant degree of cooperativity in product binding by MUG, whereas the hyperbolic response obtained from either a catalytically inactive MUG mutant N18A with substrate DNA or wild-type MUG with non-specific DNA indicates there are differential binding modes between abasic product DNA and non-specific DNA. Although there are some indications of cooperative binding to non-specific DNA with MUG and substrate DNA with N18A, it is unambiguous that the cooperative binding with the abasic DNA is much more pronounced. Both mutant and wild-type anisotropy data indicate that MUG forms strongly cooperative interactions when binding to abasic product DNA.

5.1.1 DNA binding stoichiometry of MUG

A 2:1 stoichiometric MUG-abasic DNA complex has been indicated using band shift experiments in this thesis (see Chapter 2), which is consistent with the previous stoichiometry study of MUG-DNA binding in our lab using a time-resolved fluorescence anisotropy assay, as well as a MUG competition anisotropy binding assay [154].

In the former assay, the rotational correlation time, which was related to the overall rotational diffusion of the MUG-DNA complex, was determined as MUG was titrated into the Hex-labeled abasic product DNA (Hex-AP•G) (100 nM) with the magic angle position serving as an internal control. Comparison with calculated values for rotational correlation time indicated that the observed saturation values of

19 ns were in line for what was expected for MUG dimers binding to DNA monomer (18.7-22.6 ns) and were significantly higher than the calculated values of for a single MUG binding to DNA (7.3-8.8 ns). This result clearly indicates there is a 2:1 MUG binding stoichiometry with its abasic product DNA.

Then, the MUG competition anisotropy binding assay was initiated with a pre-bound mixture of increasing concentrations of MUG enzymes (500 nM, 1000 nM, 1500 nM and 2000 nM) and Hex-AP•G DNA (500 nM), and subsequently the enzyme was competed off by titrating in increasing amounts of unlabeled abasic DNA. The initial concentration of Hex-AP•G was at 500nM, to be above the K_d for the abasic product and the concentration of MUG in the initial complex was increased in stoichiometric equivalents from a 1:1 ratio with the Hex-AP•G up to 4:1 (Figure 44A). With an initial 1:1 stoichiometry of MUG- Hex-AP•G, the initial observed anisotropy was very low, and the enzyme was readily competed off with unlabeled competitor AP•G DNA. At 2:1 the initial anisotropy was much higher, and this was competed off directly with unlabeled AP•G. At higher stoichiometric equivalents of 3:1 and 4:1, the starting anisotropy was slightly higher than that at 2:1, but there was a lag before the unlabeled DNA was able to compete off the labeled MUG-DNA complex.

At low MUG concentration the initial observed anisotropy was very low, indicating that the majority of the Hex-abasic DNA was unbound and that a 1:1 stoichiometry of MUG:DNA was insufficient to fully bind the Hex-AP•G DNA. At 2:1 the observed anisotropy was much higher, consistent with a near saturated complex. The relatively small increase in anisotropy observed at 3:1 and 4:1 ratios was most likely due to additional non-specific binding of MUG. The lag observed with 3:1 and 4:1 is consistent with the unlabeled competitor first binding excess free MUG, or MUG bound in a weak non-specific manner. Only once this excess MUG has been bound,

the competitor does begin to compete for MUG bound in a tighter, specific complex.

In the competition assay, the stoichiometry was further examined by plotting the anisotropy versus the stoichiometric balance (Figure 44B), calculated from the stoichiometry of MUG-DNA minus the stoichiometric equivalents of competitor DNA added. Plotting the data in this way clearly demonstrates that the specific MUG-Hex-AP•G complex is only competed off once there is a stoichiometry of 2:1 MUG:Hex-AP•G remaining. The slight shift to the right of the higher concentration curves can be explained by additional non-specific binding of MUG to Hex-AP•G. Both time-resolved anisotropy data and competition anisotropy data are therefore consistent with the above demonstration of MUG cooperatively binding to the abasic DNA with a 2:1 stoichiometry.

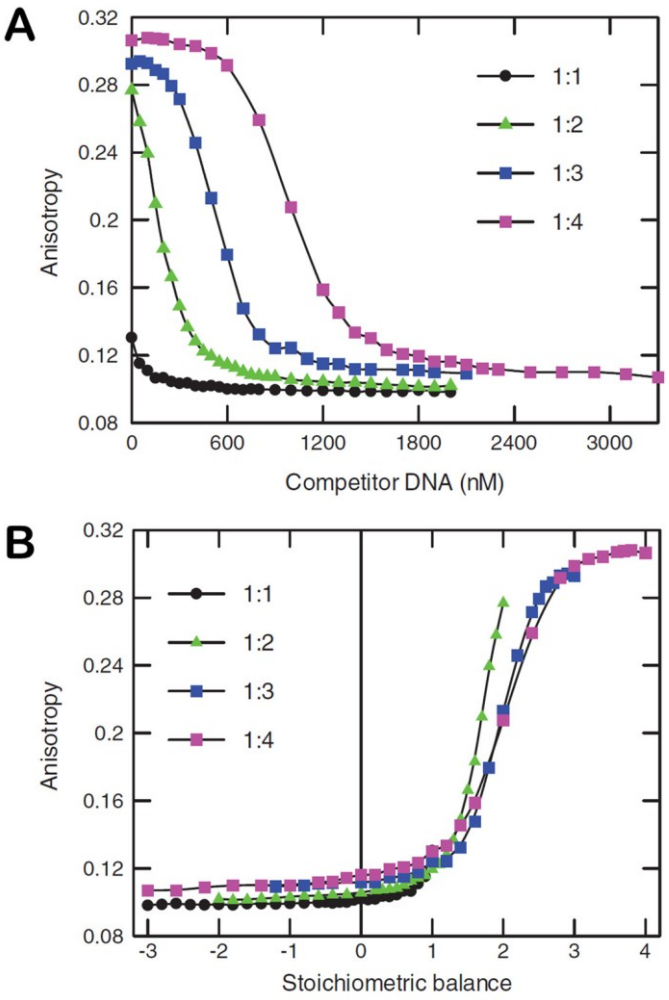


Figure 44. MUG competition anisotropy binding assays [154]**.**
500nM Hex-AP•G was incubated with 500, 1000, 1500 or 2000 nM MUG in standard MUG reaction buffer at 25°C for 15 minutes to reach equilibrium. Increasing amounts of unlabeled abasic DNA (AP•G) were titrated in and anisotropy was measured. Representative anisotropy data are shown plotted (A) versus concentration of competitor AP•G DNA and (B) versus the stoichiometric balance, calculated as the stoichiometry of MUG-Hex-AP•G [(MUG)/(Hex-AP•G)], minus the stoichiometric equivalents of unlabeled DNA added [(AP•G)/(Hex-AP•G)].

Contrary to previous time-resolved fluorescence anisotropy results and competition anisotropy binding results, the band shift results here provide a more clear perspective on that two MUG molecules bind to the abasic product DNA in a cooperative manner, and moreover, it also indicates the formation of specific MUG dimeric product complex is much tighter than that of MUG-non-specific DNA complex because the

specific product complex has a significantly higher gel migration rate than two MUG enzymes bound to non-specific DNA, the faster migration corresponding to formation of a tight complex (Figure 26).

The observation of MUG non-specific binding from the band shift assays is consistent with the non-specific binding results obtained from the fluorescence anisotropy assays. Non-specific binding is a very common phenomenon among DNA binding enzymes. DNA binding is widely deemed to initially occur at a random, non-specific site and the enzyme then translocates to its specific site either through sliding (one dimensional) or through multiple dissociation/re-association events where the protein hops along the DNA until it reaches the specific site [277, 278]. The non-specific DNA binding of MUG is aligned with other DNA binding enzymes that are capable of binding DNA non-specifically [189, 279, 280].

DNA glycosylases have generally been assumed to be monomeric functional enzymes. There has previously been no report of a member of the uracil DNA glycosylase superfamily either acting as a dimer, or binding its product in a cooperative manner, here, the new scenario, in terms of a 2:1 cooperative binding stoichiometry for abasic DNA binding by MUG, is unveiled for the first time. Although MUG's human homologue TDG has also been observed to bind DNA with a 2:1 stoichiometry, no cooperativity has been detected for the formation of its dimeric product complex.

The wealth of structural information for both MUG and TDG enzymes with DNA has been surprisingly hard to translate into mechanistic modes of binding and repair of DNA lesions. Albeit two MUG monomers have been previously observed in co-crystals in the study by Barrett *et al*. [93], dimerization was not taken into account because the observed MUG dimers were deemed to individually bind to a

U•G mismatch on each DNA strand as well as interact with the crystal lattice but not due to a functional necessity. Consequently, the first MUG-DNA co-crystal structure was published as MUG monomer bound to DNA, although the two MUG molecules had bound forming a potential protein-protein interface with each other [93].

As shown in Figure 29A, two opposing MUG molecules are productively bound to an intact U•G mismatches containing substrate through atomic coordinates manipulation of the PDB file: 1MWI. A fairly distinct MUG-MUG interface comes into sight with eight residues including Tyr74, Val75, Gln76, Asn78, Glu79, Ser81, Ly82 and Gln83 (Figure 29A). These interface residues are specifically facing one of eight opposing interface residues, as well as they are on the surface of MUG, far away from the MUG active site and don't interact with DNA.

These eight residues have been individually mutated to alanine, which has a smaller and non-charged side, and the binding characteristics of eight MUG alanine mutant enzymes have been analyzed using band shift assays as described in the Chapter 3, demonstrating that most alanine mutant enzymes bind to abasic product in a modestly weak cooperative dimeric manner, except for Q76A and Q83A, which binding isotherms, as well as the levels of affinity, cooperativity and stoichiometry for the DNA binding are completely identical to those of wild-type MUG. K82A exhibits the most impaired DNA binding, through further scrutinizing the MUG-DNA complex structure, it is very likely that Lys82 plays a key role in protein-DNA interaction because its positively charged side chain is very adjacent to the phosphate group of a guanine in the DNA substrate (Figure 35). On the other hand, the remaining five MUG interface residues Thr74, Val75, Asn78, Glu79 and Ser81 are very likely to form specific contacts with opposing interface residues on the other enzyme molecule of MUG dimer complex. Because of the significantly reduced binding affinities and cooperativity of mutants T74A and E79A, Thr74 and

Glu79 residues are particularly distinctive among these five residues and suggested to make critical interactions to stabilize the dimer interface. However, the strike of single site-directed alanine substitution of these five interface residues seems not impelling enough to disrupt the MUG interface, therefore, one of the future work is to perform the multiple site-directed mutagenesis with alanine on those interface residues simultaneously, and analyze the binding characteristics of the mutant containing multiple alanine mutation sites on the dimer interface.

It is intriguing that TDG has also been found to bind abasic DNA in a 2:1 stoichiometry [191]. As shown in Figure 20, one TDG subunit had bound at the abasic site to form a specific product complex and the other subunit had bound at an undamaged site to form a non-specific complex. A symmetrical TDG dimer interface has also been observed in this crystal structure. 6 residues in the buried surface area of TDG including Leu143, Met144, Tyr147, Thr197 and Pro198 have been determined to contribute to the TDG dimer interface (Figure 29B) via biochemical studies by Maiti *et al.* [191]. However, there are no apparent hydrogen bonds or salt bridges between these residues, therefore, no further biochemical or mutagenesis studies on these TDG dimer interface residues were carried out by Maiti *et al*.

In the Chapter 3, comparison of the TDG dimer with the relative position of the two MUG enzymes in the MUG-abasic DNA complex suggests that there is a different mode of dimeric interface interaction (Figure 29C). The alternative TDG dimer interface in MUG has been analyzed using site-directed mutagenesis and band shift assays, indicating that no other MUG-MUG interactions occur via the mutated homologous TDG dimer interface residues in MUG (Figure 34). This may be due to the relative positions of the two enzymes in the MUG crystal are not the same as in the TDG crystal, the relative positionings are also mutually exclusive (Figure 29), so

that both complexes could not form simultaneously.

Dimeric binding has been observed in studies of other DNA repair enzymes as well. MutY has been indicated to be a dimeric functional enzyme as MutY has to assemble into a dimer upon binding and then functionally processing substrate DNA, which only contains one A•8-oxoG mismatch [281]. Moreover, band shift analysis demonstrated that MutY and its murine homologue hMYH form multiple bands with an A•G bearing substrate DNA [265, 282].

5.2 Implications of 2:1 DNA binding stoichiometry for MUG catalysis

While our data demonstrate the formation of a specific dimeric complex with abasic DNA, it was not previously clear whether this had an impact on catalytic activity of MUG, since binding to non-specific and substrate DNA is either non-cooperative or only relatively weakly cooperative. A 2:1 MUG: DNA ratio has been indicated to be required for U•G bearing DNA cleavage (Figure 28). However, it is known that MUG exhibits a reasonable slow rate (0.04 s^{-1}) of its single turnover substrate cleavage [129]. There is an argument that MUG fully repairs the DNA lesion in a monomeric manner, afterwards, strong cooperative dimeric binding happens and may sequester MUG out of the further reaction, resulting in the product-inhibited turnover and the observed 2:1 dependence in Section 2.2.3.3.

TDG has been shown to be fully capable of locating and processing G•U or G•T lesions as a monomer, even though a 2:1 TDG binding stoichiometry has been observed with abasic DNA [191]. The TDG dimeric binding has hereon been thought, if adopted *in vivo*, to contribute to other critical biological roles, perhaps

binding interactions with other proteins.

Taking above findings in context of multiple turnover assays, it is apparent that under multiple turnover conditions, MUG does not exhibit a burst phase that is reflective of its single turnover behaviour. This can be reconciled if a critical role is considered for the dimeric MUG complex in catalysis: when DNA substrate is in excess, MUG is distributed among numerous molecules and the lack of cooperative binding of DNA substrate (Figure 24) indicates that MUG is limited in its ability to form dimeric enzyme substrate complexes under steady state conditions, therefore the cleavage of MUG requires a dimeric complex.

When most MUG alanine mutant enzymes have been identified to bind abasic DNA in a relatively weak cooperative manner, effects of alanine substitutions of eight proposed MUG dimer interface residues on MUG catalysis have also been characterized under both single turnover and steady state conditions in Chapter 3. It was initially predicted that if MUG is able to cleave as a monomer, an increased turnover of MUG would be observed, since disruption of the MUG:MUG interface by the loss-of-function alanine substitutions would facilitate dissociation of MUG from the abasic DNA product. In fact, unexpectedly, the alanine mutations of MUG dimer interface face residues do not only impair both affinity and cooperativity of MUG binding, but also lead to further loss of MUG activity to certain degrees. As known by scrutinizing the crystal structure of MUG (Figure 29A), the proposed MUG dimer interface residues are on the surface of MUG enzyme, far from the active site pocket of MUG and do not play any direct catalytic roles in the hydrolysis of *N*-glycosidic bond, and therefore most likely in stabilizing dimeric MUG complexes. The combined results thus suggest that MUG won't cleave its substrate as a monomer, unless it can form a stable dimeric complex with a second MUG molecule to repair the DNA lesion and then protect the AP product. This is

particularly pertinent to MUG because the single turnover cleavage rate of MUG is reasonably slow for an enzyme, at 0.04 s^{-1} for a U•G containing DNA substrate, and the half-life ($t_{1/2}$) of MUG cleavage is around 17 sec, in terms of residence time of MUG enzyme on the substrate as a monomer, the 1:1 stoichiometric MUG:DNA complex is not stable enough for this relatively long time scale (~17 sec) needed for the cleavage, therefore, the more stable MUG dimeric complex is required to be fully capable of locating long time enough on DNA and hence processing one U•G mismatch.

The implications of the cooperative formation of a specific MUG dimeric complex for catalysis is distinct from observations with other DNA glycosylases that are generally deemed to act as monomers, and there has previously been no report of a member of the uracil DNA glycosylase superfamily either catalyzing as a dimer, or binding its product in a cooperative manner [105]. One prominent contrast to MUG is its human homologue TDG. In detail, TDG can form dimeric binding with abasic DNA, however, no cooperative binding has been observed from its both anisotropy assays and band shift experiments [191]. The anisotropy data of TDG binding to a G•AP bearing DNA reveals that one TDG subunit binds to the AP site very tightly, and a second TDG subunit then binds very weakly to form a dimeric complex, which is clearly different from the observation here of pronounced strong cooperative binding of two MUG molecules to abasic DNA (Figure 23 and 26). Furthermore, it has been indicated that the 2:1 binding of TDG is dispensable for its catalytic activity and one TDG molecule is fully functional for locating and processing the DNA lesions [191]. There are striking differences between MUG and its human homologue (TDG) on the dimeric complex machinery acting to repair DNA lesions, despite the relatively high (32%) sequence identity.

5.3 The role of ExoIII in base excision repair pathway by MUG

The data obtained from MUG burst kinetics assays in the presence of either ExoIII or EndoIV clearly indicate that both *E.coli* AP endonucleases, ExoIII and EndoIV, have significant stimulatory effects on MUG turnover (Figure 41). The effect of ExoIII and EndoIV on the chemical events prior to product release has also been investigated under single turnover conditions, suggesting that the presence of either ExoIII or EndoIV does not perturb the intrinsic chemical process associated with the uracil cleavage and that the stimulatory effect is associated with the steps governing the steady-state rate, in terms of product release. Moreover, the different extents of stimulatory effects between ExoIII and EndoIV under the parallel reaction conditions give a hint that the exact mechanism of the enhanced turnover may be different for these two *E.coli* AP endonucleases. These finds here are quite consistent with the observations of MutY turnover kinetics with ExoIII and EndoIV [161], that a dramatic change in the steady state rate of the MutY reaction with a G•A containing substrate was observed upon addition of either ExoIII or EndoIV, but EndoIV shows a stronger stimulatory effects than ExoIII, and furthermore, band shift assays demonstrated that EndoIV displaces MutY from the product DNA whereas ExoIII results in a super-shifted band, in terms of forming a ExoIII-MutY complex on DNA, these observations suggest ExoIII and EndoIV employ distinct mechanisms for MutY turnover enhancement.

It has been previously demonstrated by Hang *et al.* that MUG turnover with an substrate containing 8-(hydroxymethyl)-3,N^4-ethenocytosine•guanine mismatch was enhanced by 1.5 fold in the presence of 4 fold excess of EndoIV. The stimulatory effects of EndoIV on MUG turnover with either 8-HM-εC•G substrate DNA or AP•G product DNA (Figure 41) is relatively weak, therefore, EndoIV has not been

the main study focus of this project at present.

Furthermore, the data gleaned from the single turnover assays and burst kinetics assays of MUG in the absence or presence of ExoIII in chapter 4 indicate a mismatch of MUG reactivity demeanor between single turnover state and steady state (Figure 39 and 43), and first time unveils a pre-catalytic discrimination ability of MUG, by which MUG can fully process the DNA lesion when either 1) a second MUG molecule cooperatively binds to form a stable dimer complex for catalysis, or 2) the downstream ExoIII or EndoIV nonspecifically binds to U•G bearing DNA somewhere and activate the cleavage of the MUG monomer, bypassing the step of MUG dimeric complex formation.

MUG turnover enhancement has been indicated to be dependent on both ExoIII catalytic activity and concentration as detailed in Section 4.2.7 and 4.2.8, suggesting ExoIII may enhance MUG turnover via a product scavenging mechanism, whereby the action of the ExoIII at the AP site will process the abasic lesion, thus depleting the concentration of AP site, preventing MUG from rebinding to its product and initiating new MUG catalytic cycle, therefore enhancing the turnover of MUG.

Like ExoIII, human AP endonuclease (hAPE1) has also been found to use product scavenging mechanism for OGG1 turnover stimulation [155]. In detail, hOGG1 spontaneously dissociate from the AP site intermediate generated by its DNA glycosylase activity, then hAPE1 occupies and processes the AP site to avoid the reassociation of the hOGG1 to its product, thus increasing the turnover of hOGG1.

In order to quantitatively elucidate the complexity of the resulting scenario of MUG burst kinetics in the presence of ExoIII, one of the main future work is to establish a

pathway model for enzyme coordination between MUG and ExoIII, and then fit the data obtained from MUG burst kinetics assays conducted with or without ExoIII to quantitatively determine the maximal rate constant for MUG turnover and discriminate between the distinct enzyme coordination mechanisms in the initial steps of base excision repair. As shown in Figure 45, the complex repair scheme has been preliminary mapped as shown in Figure 45 and extensive data modeling work is under way by collaborating with Dr Juliane Liepe (Imperial College London) using the approximate bayesian computation, which is a powerful method to enable exquisite tuning of the multitude of parameters within the reaction topology and give a unique solution to fit the data and explain the whole enzyme communication system between MUG and its downstream AP endonucleases during BER.

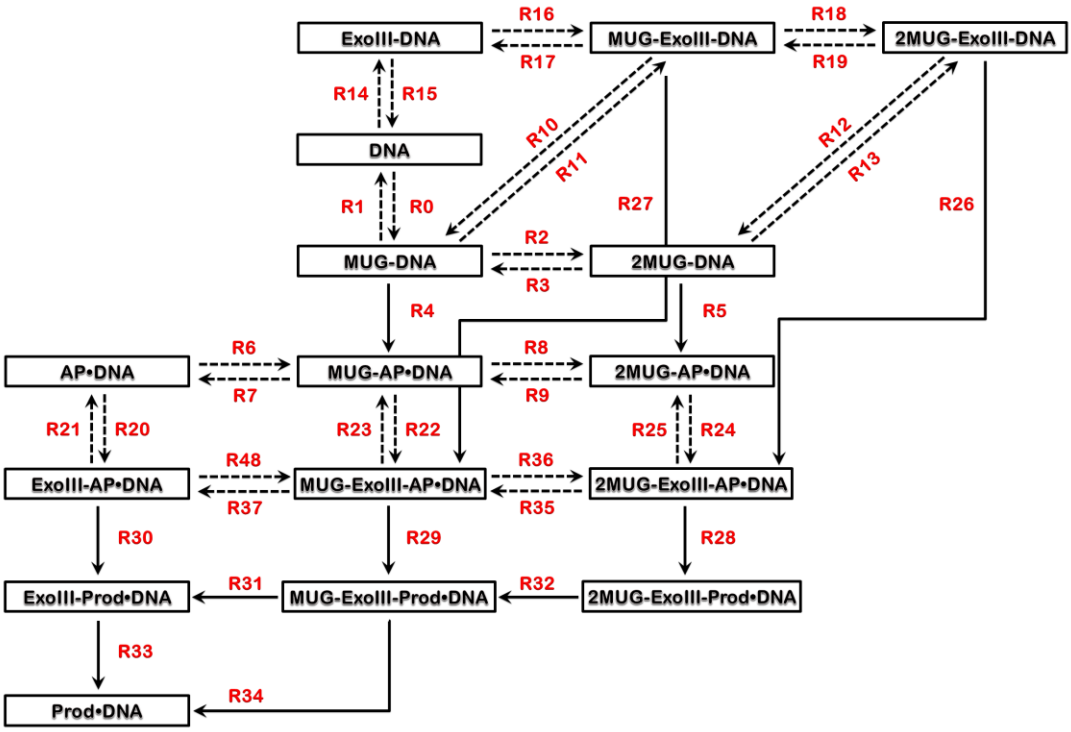


Figure 45. Preliminary model of enzyme coordination between MUG and ExoIII.

Moreover, as abovementioned, it is very likely MUG is capable of cleaving its substrate as a monomer when its downstream AP endonuclease, ExoIII or EndoIV

non-specifically binds to MUG-substrate DNA complex somewhere to activate MUG monomer catalysis. Therefore, in order to investigate whether there is a specific interface between MUG and either ExoIII or EndoIV, which is required for the burst phase rate enhancement of MUG monomer by ExoIII or EndoIV. It will be interesting to conduct the burst kinetics assays with the alanine mutant MUG enzymes (see Chapter3, Section 3.2.1) in the presence of either ExoIII or EndoIV, or even hAPE1, to see if the enhancement of MUG burst phase rate is generic or idiosyncratic.

5.4 Biological functions of MUG *in vivo*

Two possible mechanisms, including active MUG-AP endonuclease interaction mechanism and passive abasic product scavenging mechanism, have been considered above in Section 4.1 for the resolution of stable MUG complexes bound to DNA within the context of BER pathway, however, the role of MUG in DNA repair *in vivo* remains quite speculative. It has been widely argued and clearly indicated that neither εC, U nor T may be the biologically relevant substrate for MUG. Although εC is suggested to be the best substrate for MUG *in vitro* [129], it has not been detected in *E.coli* DNA yet, and lipid peroxidation, the principal source of εC in mammalian DNA, is not known to occur in *E.coli* [283]. Uracil is known to occur in DNA because of its incorporation during DNA replication as well as cytosine deamination. The latter event may cause C to T transitions. However, UDG is a very efficient enzyme to eliminate a vast majority of uracils in DNA. It has been indicated that the ability of *ung E.coli* cell extract to excise uracils from a U:G containing DNA is at least 400-fold lower than that found in ung^+ cell extract, indicating MUG is a not an effective back up enzyme for UDG [284]. Additionally, T:G containing DNA can be more efficiently repaired by a separate repair pathway in *E.coli* called very short patch repair, rather than by MUG, although it is primarily directed to sites of Dcm

methylation [285, 286]. Therefore, it is possible that, *in vivo*, MUG may have different biologically important substrates selectivity, or act using other mechanisms for the resolution of complexes as a part of other DNA processing pathways, such as replication and transcription.

MUG is expressed very poorly in exponentially growing cells, in contrast, it is fairly abundant in stationary-phase cells and the approximate number of MUG molecules per stationary-phase cell is 500 [287]. Sanath *et al.* have indicated that MUG has an important anti-mutagenic role in *E.coli* cells maintained in stationary phase using both genetic reversion assays and forward mutation assays [287]. MUG thus joins a small coterie of DNA repair enzymes that protect stationary-phase cells against mutations, including a structure protein, Dps, that reduces oxidative damage to DNA [288], an enzyme that breaks down hydrogen peroxide, catalase HPII [289], and two alkyltransferases Ada and Ogt [290].

At the onset of stationary phase, *E.coli* DNA has been observed to become more susceptible to oxidative damage and hence induces multiple protective mechanisms against oxidative damage [291]. MUG has been previously indicated in our lab to have modest activity against some products of oxidative damage, such as 5-hydroxycytosine [129]. It is thus possible that MUG may excise the oxidative damage products in non-dividing *E.coli* cells. Additionally, in stationary phase, the *xthA* gene, which codes for ExoIII, is also expressed by RpoS upregulation. ExoIII can also participate in the repair of oxidative damage [292, 293]. Therefore, it is possible that ExoIII may coordinate with MUG in the process of oxidative damage, similar to their coordination in the BER pathway as indicated above.

During late exponential and early stationary phase, the initiation of *E.coli* chromosome replication is inhibited by an uncharacterized extracellular factor [294].

The inhibition of the initiation of DNA replication is modulated by direct interactions with the replication machinery. This inhibition does not stall ongoing replication to proceed to termination before complete exhaustion of nutrients. MUG may play a repair role in associating with the ongoing DNA replication machinery to clear or bypass the abasic site, which is an obstacle to stall the replication fork. Furthermore, it can not be eliminated that MUG may interact with helicases, such as RecQ and PriA, to unwind replication fork blockage by gaps on both leading and lagging DNA strands [60].

The potential role of MUG coupled in transcription will not be discussed here because the transcription in late stationary phase bacteria is still not well studied and the late stationary phase is very likely to be highly heterogeneous in terms of transcription profiles within different bacterial cells [295, 296].

5.5 Conclusion

Differential modes of DNA binding by mismatch uracil DNA glycosylase (MUG) from *Escherichia coli* with its substrate, abasic DNA product and non-specific DNA have been demonstrated using both band shift and fluorescence anisotropy assays. MUG binds tightly to its substrate and non-specific DNA, but that it does so in a relatively weak cooperative manner, whereas the binding cooperativity of MUG with its abasic DNA product is very pronounced.

A 2:1 binding stoichiometry of MUG was demonstrated using band shift assays and the specific product complex has a significant faster migration than two MUG enzymes bound to non-specific DNA, indicating the formation of MUG dimeric complex with abasic DNA is tighter than its non-specific complex. Alanine mutations of the most MUG dimer interface residues do not only impair both affinity

and cooperativity of MUG binding, but also lead to further loss of MUG activity, therefore suggesting MUG doesn't cleave its substrate as a monomer, unless it can form a stable dimeric complex with a second MUG molecule to repair the DNA lesion and then protect the AP product.

Both *E.coli* AP endonucleases, Exonuclease III (ExoIII) and Endonuclease IV (EndoIV) greatly enhance the turnover of MUG with a U•G containing substrate. Both ExoIII catalytic activity and concentration have a big impact on MUG turnover enhancement, suggesting that ExoIII may use a product scavenging mechanism to deplete the concentration of AP site, avoid reassociation of MUG to its product and initiate new catalytic cycle of MUG. Finally, a pre-catalytic discrimination ability of MUG is unveiled for the first time, by which MUG won't cleave its substrate as a monomer, unless it can form a stable dimeric complex with a second MUG molecule to repair the DNA lesion and then protect the intermediate AP site, or alternatively when the downstream ExoIII or EndoIV is present, they can non-specifically bind to U•G DNA somewhere and activate the cleavage of a single MUG molecule, bypassing the step of MUG dimeric complex formation.

Chapter 6
Materials and Methods

6.1 General materials

6.1.1 Chemicals

All commercial enzymes were purchased from New England Biolabs (NEB) (Hitchin, Herts). QIAprep Spin Miniprep Kit and QIAquick Gel Extraction Kit were from QIAGEN, Crawley, UK. All remaining chemicals and materials were from Sigma-Aldrich (Dorset, UK) unless stated otherwise.

6.1.2 DNA substrates

Oligonucleotides used in all protein kinetic experiments in this thesis were synthesized by Integrated DNA Technologies, Inc. (IDT) (Coralville, USA) and are summarized in Table 4. Either 17-mer or 25-mer target oligonucleotide strands have the identical sequences, except for the single modified nucleotide target sites, which is either a tetrahydrofuron nucleotide, a chemically stable mimic of the natural abasic product (AP) or deoxycytidine (C) or deoxyuridine (U). Target oligonucleotide strands designated with "HEX" were labeled with 6-carboxy-2', 4, 4', 5', 7, 7' – hexachlorofluorexcein, succinimidyl ester (HEX) on their 5' terminus. Either 17-mer or 25-mer complimentary strands contain a guanine (G) opposite the AP/C/U sites in the target strands. All target and complimentary oligonucleotides were further purified by reverse-phase high pressure liquid chromatography (HPLC) using a Surveyor HPLC system, with UV1000 detector sets at 260 nm and 554 nm, and a Hichrom Z509 25 cm x 4.6 mm column thermostatted at 55^{o}C. Gradients were developed using buffer A (0.1 M triethylammonium acetate, 5% acetonitrile (v/v), pH 6.5) and buffer B (0.1 M triethylammonium acetate, 65% acetonitrile (v/v), pH 7.0) at a flow rate of 1 ml minute^{-1}. The concentrations of purified oligonucleotides were separately determined at OD_{260} with their specific extinction coefficients (ε260)

calculated from the sum of the individual extinction coefficient of the base constituents (Equation 1) [297]. Double-stranded substrates were made by mixing equimolar amounts of the target strands respectively with the complementary strands, heating to 90°C and cooling slowly to room temperature.

Oligo ID	Oligo Sequence 5' → 3'	Tm (°C)	Bases
Hex-12AP	GCT ATG GAC TAA APAA TGA CTG GGT G	54.6	25
Hex-12C	GCT ATG GAC TAA CAA TGA CTG GGT G	57.4	25
Hex-12U	GCT ATG GAC TAA UAA TGA CTG GGT G	55.5	25
12compG	CAC GCA GTC ATT GTT AGT CCA TAG C	57.8	25
Hex-6AP	GAC TAA APAA TGA CTG CG	44.3	17
Hex-6C	GAC TAA CAA TGA CTG CG	47.7	17
Hex-6U	GAC TAA UAA TGA CTG CG	44.2	17
6compG	CGC AGT CAT TGT TAG TC	47.7	17
$12compG_{o/h}$	CAG TCA TTG TTA GTC CAT AGC GTG T	56.8	25

Table 4. Summary of oligonucleotides used in all protein kinetic experiments.
Bases of interest including abasic, uracil and non-specific are highlighted in red, and their complementary bases are highlighted in dark blue.

6.1.3 Enzymes

The pTrc99A vector containing the *E.coli* mismatch uracil glycosylase (MUG) was a gift from Dr R. Savva, School of Crystallography, Birkbeck College, University of London, UK. The cloning, expression and purification of *E.coli* Exonuclease III (ExoIII) and Endonuclease IV (EndoIV) are separately detailed in Sections 6.3.1.2 and 6.3.1.3. The mutagenesis, expression and purification of 12 MUG mutant enzymes and ExoIII D151N are respectively described in Sections 6.3.1 and 6.4. Stocks of home-made enzymes were stored in 20% (v/v) glycerol and 20 mM Tris buffer pH8.0 at -80°C. Enzyme dilutions were prepared in the appropriate reaction buffers and not further refrozen and reused.

6.1.4 Growth media and antibiotics

The Lauria Bertani (LB) broth and plate were prepared as described in Table 5 and sterilized by autoclave.

Lauria Bertani (LB) broth / plate*	
Composition	**Final concentration**
NaCl	2mg/ml
Yeast Extract	1mg/ml
Trypton	2mg/ml
Agar *	6mg/ml

Table 5. Compositions of Lauria Bertani (LB) broth/plate*.

The sterilized LB media were supplemented with the appropriate antibiotics to allow for selection of the bacterial strain or plasmid of interest. The antibiotics used are summarized in Table 6, these were filter-sterilized before use.

Antibiotic	Stock concentration	Final concentration	Storage
ampicillin	10mg/ml in dH_2O	100µg/ml	-20°C
chloramphenicol	25mg/ml in ethanol	10µg/ml	-20°C
kanamycin	10mg/ml in dH_2O	25µg/ml	-20°C

Table 6. Summary of antibiotics' stock and working concentrations, and storage temperatures.

6.1.5 Competent cell lines

E.coli K12 strain JM109, which genotype is (F^- *traD36 proAB+ lacI^q, Δ(lacZ)M15I* Δ*(lac-proAB) gyrA96 recA1 relA1 endA1 thi1 hsdR17 mcrA supE44*), was bought from New England BioLabs. *E.coli* BL21 Rosetta (DE3) pRARE strain, which genotype is (F^- *ompT- Ion- dcm λ(DE3)[lacI, lacUV5-T7gene 1, ind1, sam7, nin5] pRARE [camr proL leuW argW glyT argU ilex])* was supplied from Novagen (Nottingham, UK).

6.1.5.1 Preparation of chemically competent cells

5 ml Lauria Bertaini (LB) broth containing 100 μg ml^{-1} ampicillin was inoculated with 1 colony of the relevant bacterial strain and grown overnight at 37°C. The saturated culture was inoculated into 100 ml of sterile LB medium and incubated at 37°C until OD_{600} reached 0.4-0.5. The culture was then transferred to a sterile 500 ml centrifuge bottle and cooled on ice for 20 minutes, prior to harvesting cells by centrifugation at 4,000 rpm at 4°C. All subsequent steps were carried out below 4°C. The cell pellet was resuspended in 100 ml ice-cold sterile 80 mM $MgCl_2$, 20 mM $CaCl_2$, and then re-centrifuged. The cell pellet was resuspended in 4 ml ice-chilled sterile 0.1 M $CaCl_2$ and incubated on ice for 2 hours. 0.25 volumes of ice-cold sterile 80% (v/v) glycerol were added to the cell suspension. Fresh-made competent cells were transferred to sterile 1.5 ml Eppendorf tubes in 200 ml aliquots, snap frozen and stored at -80°C.

6.2 Molecular and cellular biology methods

6.2.1 Isolation of plasmid DNA

Plasmid DNAs were obtained using the QIAGEN QIAprep spin miniprep kits. Typically, 2.5 ml cell culture was grown overnight at 37^{o}C and pelleted. Plasmid DNA was extracted and purified from the pellet using QIAGEN anion exchange columns according to the manufacturer's instructions. In brief, the bacterial cells were subjected to alkaline lysis using a series of three solutions including resuspension buffer, lysis buffer and neutralization buffer. The cell debris generated was then pellets by centrifugation at 13,000 rpm for 10 minutes at room temperature. The supernatant was then transferred to the small anion exchange column and spinned at 13,000 rpm for 1 minute. After a wash step with salt solution to remove impurities, 25-50 µl TE buffer or mQH2O water, depending on the expected yield, was added to the column to elute DNA from the column by centrifugation at 13,000 rpm for 1 minute.

6.2.2 Polymerase chain reactions (PCR)

1 µl of a 10 fold plasmid dilution (containing typically ~100 ng DNA) and 2.5 units Pfu turbo DNA polymerase, which was home-made and kindly provided by Dr Jan Silhan in our lab, was added to 0.2 µM of each of the two specific primers in a total reaction volume of 50 µl in the 1 x Pfu turbo reaction buffer (20 mM Tris-HCl, pH8.5, 150 µg/ml BSA, 16 mM $(NH_4)_2SO_4$, and 3.5 mM $MgCl_2$) with 200 µM each dNTP and 3% (v/v) DMSO in 0.2 ml STARLAB PCR tubes (STARLAB international GmbH, Hamburg, Germany). Routinely, the following cycle in Table 7 was used for PCR reactions, whereby the annealing temperature was individually adjusted based on the different specific primers used. The PCR fragments were

extracted by agarose gel electrophoresis (Section 6.2.4) and gel purification (Section 6.2.5).

Step	Temperature (°C)	Time length (second)	Cycle
Initialization	94	300	1x
Denaturation	94	30	25x
Annealing	45-57 (individually adjusted)	60	
Elongation	72	120	
Final elongation	72	600	1x
End	4	∞	

Table 7. Polymerase chain reaction (PCR).

6.2.3 Restriction endonuclease digestion

For analytical endonuclease digestions, 5 µl (typically containing ~5 µg of DNA) of plasmid DNA was digested in a total volume of 10 µl. For preparative purposes the total volume was increased to 100 µl and 10 µl of plasmid DNA was digested. All digestions were executed in the appropriate 1 x reaction buffer in the presence or absence of BSA, depending on the enzymes which were employed. Routinely, 2-10 units (analytical digests) or 6-30 units (preparative digests) of the enzyme was used and the reactions were incubated at 37°C for 1-2 hours. The cleaved fragments were separated from the host vector backbone by agarose gel electrophoresis (see Section 6.2.4) and extracted by gel purification (see Section 6.2.5).

6.2.4 Sticky end ligation

3’ or 5’ overhanging ends of double-stranded DNA can be generated by many restriction enzymes when cutting DNA as described above. The nucleic acid bases of these overhanging ends can engage in base pairing with the corresponding

overhanging of linearized vector DNA with the same enzyme. Routinely, 0.5 µl insert of interest was mixed with 0.5 µl linearized vector and 0.5 µl T4 DNA ligase (0.5-1.5 units) in 1 x Rapid Ligation Buffer (Promega, UK). The reaction mixture was incubated at room temperature for at least 1 hour, then analyzed via agarose gel electrophoresis (see Section 6.2.5), purified (see Section 6.2.6) and finally transformed into JM109 cells (see Section 6.2.7).

6.2.5 DNA analysis by agarose gel electrophoresis

Large linear double-stranded DNA fragments, which sizes range from 0.8 kbp to 2 kbp, were analysed on 0.7-2% (w/v, concentration dependent on the size of the DNA fragment) agarose gels prepared in 1 x TBE buffer (Invitrogen™, UK). The agarose gel solution was heated to allow the agarose to dissolve and poured into a gel-casting tray, where it was allowed to cast prior to use. Gels were run in 1 x TBE buffer at 70 V for 45-90 minutes at room temperature. DNA samples were stained by 0.5 x SYBER Green I (Molecular Probes®, Eugene, USA) and prepared in 1 x DNA loading buffer (0.025% (w/v) bromophenol blue, 0.05% (w/v) xylene cyanol, 0.02 M EDTA, and 6% (v/v) glycerol in H_2O). The DNA was visualized using a Fuji FLA-5000 fluorescent image analyzer (Fujifilm Corporation, Japan). Fragment size was compared to "Quick-load" 1kb DNA ladder (New England Biolabs® Inc, UK).

6.2.6 Gel purification of DNA fragments

The target DNA fragments' bands were excised to extract the DNA from the agarose gels. Gel purifications were carried out using the QIAGEN QIAquick gel extraction kits according to the manufacturer's instructions. In summary, the gel slices were dissolved in the yellow QG buffer by heating at 55°C for 10 minutes and isopropanol added to facilitate DNA precipitation. DNA was then bound to an ion

exchange column by centrifugation at 13,000 rpm for 1 minute. Impurities were washed off by adding salt solution into the column and then centrifuged at 13,000 rpm for 1 minute. Finally, purified DNA was eluted from the column in volumes of 20-50 µl EB buffer or mQH_2O depending on the expected yield.

6.2.7 Transformation of chemically competent cells

Aliquots of competent cells were deforested on ice. 1-5% (v/v) of plasmid DNA was added and cells incubated on ice for 15 minutes. Afterwards, they were heat-shocked at 42^{o}C for 45 seconds and further incubated on ice for 10 minutes. Cells harboring plasmid with ampicillin resistance were mixed with sterilized antibiotic-free LB medium and incubated at 37^{o}C for 45 minutes prior to plating the cell suspensions onto LB agar plates containing 100 µg/ml ampicillin. Plates were incubated overnight at 37^{o}C for 16-20 hours.

6.2.8 Glycerol stocks

2.5 ml LB medium containing 100 µg ml^{-1} ampicillin was inoculated with a single bacterial colony and incubated overnight at 37^{o}C. 500 µl overnight grown cell culture was mixed thoroughly with 125 µl 100% glycerol to make a 20% (v/v) cell glycerol stock, then snap frozen on dry ice and stored at -80^{o}C. To inoculate expression cultures, a few cells were scraped from the frozen stock with sterile pipette tip and resuspended in LB medium.

6.2.9 Gene cloning

6.2.9.1 ExoIII cloning strategy

E.coli ExoIII gene *xthA* of approximately 0.8 kbp was generated by PCR from *E.coli* K12 genomic DNA using primers in Table 8, which created two underlined NdeI and HindIII flanking restriction sites for the directional cloning. PCR was run under the abovementioned conditions in Section 6.2.2 except for the specific annealing temperature of 45^{o}C. PCR-amplified fragments were digested with NdeI and HindIII (see Section 6.2.3) and gel purified (see Section 6.2.6) and then subcloned into an expression vector pPROEX-HTb (see Section 6.2.4), which was kindly supplied by Dr Jan Silhan, Imperial College London. The pPROEX-HTb vector contained an N-terminal polyhistidine tag with cleavage site for TEV protease to the final protein construct. The *E.coli* Rosetta^{TM} strain BL21 DE3 cells were transformed with the ligate (see Section 6.2.7) and the introduced *xthA* insert was validated by DNA sequencing following the Cogenics' sequencing instructions (Beckman Coulter Genomics, UK).

Primer ID	Sequence 5'→3'	Restriction site	Tm (^{o}C)
ExoIII_F	GAG ATA TAC ATA TGA AAT TTG TCT CTT TTA ATA TC	NdeI	52
ExoIII_R	CAT AAG CTT AGC GGC GGA AGG	HindIII	58

Table 8. ExoIII PCR primers.

6.2.9.2 EndoIV cloning strategy

E.coli EndoIV gene *nfo* of approximately 0.85 kbp was amplified by PCR from *E.coli* K12 genomic DNA using primers in Table 9, which created two underlined BamHI and HindIII flanking restriction sites for the directional cloning. PCR was run under the abovementioned conditions in Section 6.2.2 except for the specific

annealing temperature of 51°C. PCR-amplified fragments were digested with BamHI and HindIII (see Section 6.2.3), gel purified (see Section 6.2.6) and then ligated into the compatible sites of the expression vector pPROEX-HTb (see Section 6.2.4). The *E.coli* Rosetta™ strain BL21 DE3 cells were transformed with the ligate (see Section 6.2.7) and the introduced *nfo* insert was validated by DNA sequencing following the Cogenics' sequencing instructions (Beckman Coulter Genomics, UK).

Primer ID	Sequence 5'→3'	Restriction site	Tm (°C)
EndoIV_F	GAG ATG GAT CCA TGA AAT ATA TTG GCG	BamHI	56
EndoIV_R	CAT AAG CTT ACG CCA CCG CTT TTT C	HindIII	59

Table 9. EndoIV PCR primers.

6.2.10 DNA analysis by denaturing urea PAGE

Small single-stranded DNA fragments between 1bp and 500 bp were analyzed by 15% or 20% denaturing (7 M urea) polyacrylamide (19:1 Acrylamide/Bis) gel electrophoresis (PAGE) (Severn Biotech Ltd, UK) in 1 x TBE buffer. Polymerisation of large volume gels (100 ml) was initiated by adding 800 µl 10% (w/v) ammonium persulfate solution (APS) and 80 µl N,N,N',N'-tetramethylenediamine (TEMED). Small volume gels (5 ml) were polymerised with 30 µl 10% (w/v) APS and 5 µl TEMED. Gels were run at either 60 W (large gels) or 200 V (small gels). Native gel were prepared in exactly the same manner but in the absence of urea. Gels employed to separate fluorescently labelled DNA were scanned using the Fuji FLA-5000 fluorescent image analyzer. HEX labelled DNA substrates (λ_{Ex} = 535 nm, λ_{Em} = 556 nm) were identified using a Cy3 excitation laser at 532 nm with a 570 nm (±20 nm) band pass filter. Gel band intensity quantifications were carried out from the captured images using Multi Gauge software (Fujifilm, Japan) and Phoretix 1D Gel Analysis software (TotalLab, UK).

6.3 Biochemical and biophysical protein methods

6.3.1 Protein preparation

6.3.1.1 Expression and purification of wild type and mutant MUG proteins

The mismatch uracil DNA glycosylase from *E.coli* (MUG) was expressed from the vector pTrc99A (kindly provided from Dr R Savva, Birkbeck College, University of London) in *E.coli* Rosetta^{TM} strain BL21 DE3 cells. 5 µl inoculation of the strain carrying the relevant MUG construct was grown overnight at 37^{o}C in 2.5 ml LB broth containing 100 µg ml^{-1} ampicillin. 1ml fresh overnight cultures were diluted 1,000 times into 1L LB broth containing 100 µg ml^{-1} ampicillin and grown for a further 6 hours at 37^{o}C to OD_{600} of 0.6. Expression of MUG protein was then induced by 1 mM Isopropyl-1-thio-β-D-galactopyranoside (IPTG). The cultures were further grown overnight at 37^{o}C prior to cell harvesting by centrifugation at 5,000 rpm at 4^{o}C for 10 minutes. The cell pellets were resuspended and lysed by sonication in 25 ml buffer A (20 mM Tris-HCl, pH 8.3, and 1 mM EDTA) with "Complete" protease inhibitors (Roche Applied Science). To precipitate nucleic acids, the lysates were incubated with 1/10 volume of 10% (w/v) streptomycin sulfate for 1 hour on ice prior to centrifugation at 15,000 rpm for 1 hour at 4^{o}C.

The supernatant was subjected to tandem anion/cation exchange chromatography on columns filled with DE52 cellulose resin (Whatman®, GE Healthcare) and SP Sepharose resin (Whatman®, GE Healthcare) respectively on an ÄKTA purifier (GE Healthcare). The duo-column system was equilibrated with 500 ml buffer A. MUG was thereby congregated onto the SP-Sepharose column and the DEAE-cellulose column was disconnected. Protein was eluted from the SP-Sepharose column using a gradient of 0 – 100% buffer B (20 mM Tris-HCl, pH 8.3, 1 mM EDTA, and 1 M

NaCl) over 400 ml. Fractions of the eluate were analyzed by 12% SDS-PAGE (see Section 6.3.2). The fractions containing the target protein was pooled and the volume reduced by ultrafiltration using Millipore ultrafiltration membranes (Millipore Corporation, Billerica, MA) in a stirred cell on ice. MUG was further purified by size-exclusion chromatography using a HiLoad 26/60 Sephadex75 gel filtration column (GE Healthcare) pre-equilibrated with buffer C (20 mM Tris-HCl, pH8.3, 1 mM EDTA, and 200 mM NaCl). Fractions of the eluate were again analyzed by 12% SDS-PAGE (see Section 6.3.2) and the ones containing target proteins were pooled and concentrated as above.

The protein concentration was calculated from the OD_{280} based on an extinction coefficient of 25,590 $M^{-1}cm^{-1}$ using the ultrafiltration eluent for background correction. This method for determining protein concentration was also verified using a Bio-Rad protein assay (see Section 6.3.3). Glycerol was added to 20% (v/v) based on the mass of glycerol added and its specific gravity (1.129 at 25^{o}C), and aliquots of the enzyme were snap frozen and stored at -80^{o}C. Concentrations were re-measured by spectrometry after addition of glycerol to both enzyme and the eluent used for background correction.

All MUG mutant proteins used in this project were made by site directed mutagenesis (see Section 6.4) and were expressed, purified and quantified as above.

6.3.1.2 Expression and purification of wild-type and D151N ExoIII proteins

The *E.coli* Exonuclease III (ExoIII) was expressed from the vector pPROEX-HTb (a gift from Dr Jan Silhan, Imperial College London) in *E.coli* Rosetta^{TM} strain BL21 DE3 cells. A single colony of *E.coli* BL21 DE3 transformants harboring

pPROEX-HTb_ExoIIIwt plasmid was cultured in 5 ml LB broth containing 100 μg ml^{-1} ampicillin overnight at 37°C. 3 x 1 ml of overnight cultures were diluted 500 times separately into 500 ml LB broth containing 100 μg ml^{-1} ampicillin and grown for a further 5-6 hours at 37°C to OD_{600} of 0.7-0.8. The cultures were then cooled to room temperature and protein expression was induced by the addition of 1 mM IPTG. The cultures were further incubated at 25°C (room temperature) for 6 hours before harvesting via centrifugation at 5,000 rpm at 4°C for 10 minutes. Pellets containing ExoIII were resuspended in buffer W (10 mM Phosphate, 2.7 mM KCl, pH 8.0, 500 mM NaCl, 50 mM imidazole, and 2 mM β-mercaptoethanol (β-ME)) and sonicated prior to ultracentrifugation at 15,000 rpm for 1 hour at 4°C.

The supernatant was loaded onto a metal affinity chromatography column, which was filled with HiTrap Chelating Sepharose Fast Flow resin (GE healthcare), pre-charged with Ni^{2+} and equilibrated with buffer W. His-tagged ExoIII protein was eluted with buffer E (10 mM Phosphate, 2.7 mM KCl, pH 8.0, 500 mM NaCl, 500 mM imidazole, and 2 mM β-ME). Fractions of the eluate were analyzed by 15% SDS-PAGE (see Section 6.3.2). Fractions containing the target protein were pooled and concentrated by ultrafiltration as above (see Section 6.3.1.1) prior to a 2 hour incubation with 5 mM TEV protease (kindly provided by Dr Jan Silhan, Imperial College London) at room temperature to remove the poly-histidine tag of ExoIII. Undigested ExoIII was removed on a freshly Ni^{2+} charged chelating Sepharose column, while the de-tagged protein was eluted through the column with the buffer, pooled and concentrated as above. De-tagged ExoIII was further purified by size-exclusion chromatography using a HiLoad 26/60 Sephadex75 gel filtration column (GE Healthcare) pre-equilibrated with buffer G (20 mM Tris-HCl, pH 8.0, 200 mM NaCl, 1 mM EDTA, and 1 mM Dithiothreitol (DTT)). Fractions of the eluate were again analyzed by 15% SDS-PAGE (see Section 6.3.2) and the ones containing target proteins were pooled and concentrated as above.

The protein concentration was calculated from the OD_{280} based on an extinction coefficient of 39545 $M^{-1}cm^{-1}$ using the ultrafiltration eluent for background correction. Glycerol was added to 20% (v/v) as above, and aliquots of the enzyme were snap frozen and stored at -80^{o}C. Concentrations were re-measured by spectrometry after addition of glycerol to both enzyme and the eluent used for background correction. The ExoIII D151N catalytic mutant was made by site directed mutagenesis (see Section 6.4) and was expressed, purified and quantified as above.

6.3.1.3 Expression and purification of EndoIV

The *E.coli* Endonuclease IV (EndoIV) was expressed from the vector pPROEX-HTb in *E.coli* Rosetta^{TM} strain BL21 DE3 cells. A single colony of *E.coli* Rosetta^{TM} strain BL21 DE3 transformants harboring pPROEX-HTb_EndoIVwt plasmid was inoculated in 5 ml LB broth containing 100 μg ml^{-1} ampicillin and 0.1 mM $ZnSO_4$, and grown overnight at 37^{o}C. 4 x 1 ml of overnight culture was diluted 500 times into LB broth containing 0.1 mM $ZnSO_4$ and 100 μg ml^{-1} ampicillin, and incubated for further 6 hours at 37^{o}C to OD_{600} of 0.7-0.8. The cultures were then cooled to room temperature. EndoIV protein expression was induced using 1 mM IPTG and proceeded for 6 hours at 25^{o}C (room temperature). The cells were harvested by centrifugation as above and then resuspended in 25ml of buffer F (20 mM Tris-HCl, pH 7.9, 500 mM NaCl, 20 mM imidazole, and 10% (v/v) glycerol) with “Complete” protease inhibitors. A cell lysates was prepared by sonication and then clarified by centrifugation at 15,000 rpm for 1 hour at 4^{o}C.

The supernatant was applied to a metal affinity chromatography column, which was filled with HiTrap Chelating Sepharose Fast Flow resin (GE healthcare),

pre-charged with Zn^{2+} and washed with 50ml of buffer F. His-tagged EndoIV was then eluted with 6 ml buffer P (20 mM Tris-HCl, pH 7.9, 500 mM NaCl, 400 mM imidazole, 1 mM β-ME, and 10% (v/v) glycerol). The eluted fractions were immediately diluted 5 times with buffer D (20 mM Tris-HCl, pH 7.9, 500 mM NaCl, 1 mM β-ME, and 10% (v/v) glycerol) and analyzed by 15% SDS-PAGE (see Section 6.3.2). Fractions containing EndoIV protein were pooled, and concentrated by ultrafiltration as above before 2 hour incubation with 5 mM TEV protease at room temperature to cleave the poly-histidine tag on EndoIV. Undigested EndoIV was kept on a freshly Zn^{2+} charged chelating Sepharose column, while the de-tagged protein was eluted through the column with the buffer P. Fractions containing EndoIV proteins were pooled and concentrated as above. Afterwards, EndoIV was further purified by size-exclusion chromatography using a HiLoad 26/60 Sephadex75 gel filtration column (GE Healthcare) pre-equilibrated with buffer S (20 mM Tris-HCl, pH7.9, 200 mM NaCl, and 10% (v/v) glycerol). Fractions of the eluate were again analyzed by 15% SDS-PAGE (see Section 6.3.2) and the ones containing target proteins were pooled and concentrated as above.

The protein concentration was calculated from the OD_{280} based on an extinction coefficient of 31315 $M^{-1}cm^{-1}$. Glycerol was added to 20% (v/v), and aliquots of the enzyme were snap frozen and stored at -80^{o}C. Concentrations were re-measured by spectrometry after addition of glycerol to both enzyme and the eluent used for background correction.

6.3.2 SDS PAGE

5 ml resolving gel solution comprised of 12% or 15% acrylamide (29:1 Acrylamide/Bis) 0.1% (w/v) sodium dodecyl sulphate (SDS) in 375 mM Tris-HCl, pH 8.8, and was casted with 65 µl 10% (v/v) APS and 10 µl TEMED. 100%

isopropanol was floated on top of the resolving gel and washed thoroughly with dH_2O after polymerisation. A 5% acrylamide (29:1 Acrylamide/Bis) stacking gel with 0.1% (w/v) SDS in 125 mM Tris-HCl pH 6.8 polymerised with 10µl 10% (v/v) APS and 2.5 µl of TEMED per 1.5 ml of gel was layered on top of the resolving gel.

Samples were mixed with an equal volume of 2 x SDS loading buffer (20% (v/v) glycerol, 4% (w/v) SDS, 100 mM Tris-HCl, pH 6.8, 0.2% (w/v) bromophenol blue, 200 mM DTT) and heated at 95ºC for 10 minutes prior to loading on to the gel. Gels were run at 200 V in 1 x SDS PAGE running buffer (25 mM Tris-HCl, pH 8.3, 250 mM glycine, 0.1% (w/v) SDS) at room temperature for 1 hour.

Gels were stained in Coomassie blue (45% (v/v) methanol, 10% (v/v) acetic acid, 3 mM Coomassie Brilliant Blue R-250) overnight on a rocking platform. Gels were destained in a solution of 30% (v/v） methanol, 10% (v/v) acetic acid over 3 hours and gel images were captured using a Fujifilm LAS-3000 Luminescent Image Analyzer (Fujifilm Corporation, Japan).

6.3.3 Bradford protein assay

Arginine and aromatic residues within proteins can be bound by Coomassie Blue, resulting in an absorbance shift that serves as a colorimetric test. In its unbound cationic form, Coomassie Blue has a maximum absorbance of 470 nm and appears brown. In its bound anionic form, the maximum absorbance is 595 nm and the complex appears blue colour. The Bradford protein assay is thus based on the number of arginine and aromatic residues.

The Bradford protein assay is linear within a range of 0-1 mg ml^{-1} protein and therefore is suitable for determination of lower protein concentrations. Higher

concentrations can be determined by diluting the sample with specific buffers prior to mixing it with the Bradford reagent. Routinely, 10 µl (or 5 µl for highly concentrated samples) protein sample was mixed 190 µl of Bradford reagent (5 time diluted in mQH_2O) in a 96-microwell plate (Thermo Scientific Nunc). As the colour changes immediately, this assay can be applied to determine protein concentrations very quickly. The changed blue colour is not stable and measurements were therefore taken as soon as possible by scanning the plate using a UV/Vis Scanning Spectrophotometer (Beckman Coulter).

6.3.4 *In vitro* enzyme assays

6.3.4.1 Single turnover assays of MUG proteins

Single turnover assays were carried out by reacting 100 nM Hex-12U•G with 5 µM wild-type or mutant MUG proteins at 25^{o}C in a standard reaction buffer (50 mM Tris-HCl, pH 8.0, 50 mM NaCl, 1 mM EDTA and 0.1 mg ml^{-1} bovine serum albumin (BSA)). At specified increments, over a time course up to 15 minutes, 20 µl reaction aliquots were removed and quenched with an equal volume of 0.1 M NaOH. The quenched samples were then incubated at 90^{o}C for 30 minutes. An equal volume of 2 x formamide loading buffer (95% (v/v) formamide, 0.04% (w/v) bromphenol blue, 0.04% (w/v) xylene cyanol and 20 mM EDTA) was added prior to 20% urea denaturing PAGE (see Section 6.2.10). The gels were visualized using the Fuji FLA-5000 fluorescent image analyzer and quantified using the PhoretixTM 1D software. Data were fitted to the first-order rate equation (Equation 3) using GraFit 6 software (Erithacus Software). In the Equation 3, A_t is product at time t, A_∞ is product at saturation, *t* is time and *k* is the rate constant.

$$A_t = A_\infty (1 - e^{-kt}) \qquad \textbf{Equation 3}$$

6.3.4.2 Steady state assays of MUG proteins

Steady state assays were conducted by reacting 500 nM Hex-12U•G with 100 nM wild-type or one of mutant MUG proteins at 25^{o}C in the standard reaction buffer. At specified increments, over a time course up to 12 hours, 20 µl reaction mixtures were taken and quenched with an equal volume of 0.1 M NaOH. The quenched samples were then heated at 90^{o}C for 30 minutes. An equal volume of 2 x formamide loading buffer was added prior to 20% urea denaturing PAGE (see Section 6.2.10). The gels were visualized using the Fuji FLA-5000 fluorescent image analyzer and quantified using the PhoretixTM 1D software. Data were analyzed using GraFit 6 software (Erithacus Software).

6.3.4.3 Equilibrium DNA binding assays

Equilibrium DNA binding assays were performed to examine the binding of MUG glycosylase to the 25-mer Hex-labelled double-stranded oligonucleotides (Hex-12C•G, Hex-12U•G or Hex-12AP•G). In order to measure fluorescence anisotropy, a Jobin-Yvon SPEX Fluoromax-3 spectrofluorometer (HORIBA Scientific) fitted with automated polarization filters was utilized. Data was recorded using an excitation wavelength of 535 nm and an emission wavelength of 556 nm. The binding assays were conducted by preparing 100 nM 25-mer Hex-labeled oligonucleotide duplexes in a reaction volume of 400 µl in the standard reaction buffer at 25°C, and then wild type or one of mutant MUG proteins was added by increasing increments of concentration up to 5,000 nM. The excitation was at 535 nm and emission was detected through a 556 nm cut-off filter. Five measurements of anisotropy of each specific protein concentration increment were taken and averaged. Each protein titration repeated at least in duplicate. The observed anisotropy was plotted against MUG concentration and the data were fitted to the Hill equation

(Equation 4) using GraFit 6 software (Erithacus Software). In the Equation 4, [L] is the concentration of the free legend, *K* is equilibrium dissociation constant, *n* is the cooperativity constant, Cap is the capacity and background is the anisotropy of free DNA.

$$y = \frac{[L]^n \times Cap}{[L]^n + K_d{}^n} + background$$

Equation 4

6.3.4.4 Band shift assays

Band shift assays were performed by mixing 100 nM 17-mer Hex-labelled oligonucleotide duplexes (Hex-6C•G, Hex-6U•G or Hex-6AP•G) with increasing concentrations of either wild-type or one of mutant MUG proteins in the standard reaction buffer at 25°C for 30 minutes. Betaine glycine was then added to a final concentration of 1 M and further incubated at 25°C for 5 minutes. 11.1 µl reaction mixtures were then mixed with 2.2 µl 6 x native gel loading buffer (0.042% (w/v) bromphenol blue) and were loaded immediately onto an 8% polyacrylamide gel (see Section 6.2.10), which has been pre-run at constant 40 V in 1 x TBE buffer for 60 minutes at 4^{o}C. PAGE was performed at constant 40 V in 1 x TBE buffer for 220 minutes at 4°C. Bands on the gels were visualized using the Fuji FLA-5000 fluorescent image analyzer and quantified using the PhoretixTM 1D software. The equilibrium binding constants were determined by fitting the data to the Hill equation using GraFit 6 software (Erithacus Software).

6.3.4.5 MUG uracil glycosylase activity assay

In order to determine the optimal MUG:DNA ratio for maximum uracil glycosylase activity, 200 nM, 400 nM and 600 nM Hex-6U•G were separately incubated with

increasing concentrations of MUG at 25^{o}C in the standard reaction buffer for 15 minutes. At selected time points, 20 µl reaction mixtures were removed and quenched with 10 µl 0.1 M NaOH. The quenched samples were then heated at 90^{o}C for 30 minutes. An equal volume of 2 x formamide loading buffer was added prior to 20% urea denaturing PAGE (see Section 6.2.10). Imaging and analysis were performed as above.

6.3.4.6 AP endonuclease activity assays

The apurinic/apyrimidinic endonuclease activity of wild-type ExoIII, mutant ExoIII D151N and EndoIV were investigated by incubating 100 nM enzymes separately in the standard reaction buffer (50 mM Tris-HCl, pH 8.0, 50 mM NaCl, 1 mM EDTA and 0.1 mg ml^{-1} BSA) at 25^{o}C with 500 nM native Hex-12AP•G o/h substrates, which were created by reacting 100 nM Hex-12U•G o/h substrates with 1 nM UDG (a gift from Dr Hellen Thomson, Imperial College London) overnight at 25^{o}C in the standard reaction buffer. The reactions with ExoIII and D151N proteins were supplemented with 1 mM $MgCl_2$ because ExoIII is a divalent metal ion dependent enzyme. Aliquots were removed at specified time points as detailed in the figure legends, and quenched in the AP endonuclease quenching buffer (0.01% (w/v) xylene cyanol, 0.01% (w/v) bromophenol blue, 40 mM EDTA in formamide) [83] prior to being analyzed by 20% urea denaturing PAGE. Bands on the gels were visualized using the Fuji FLA-5000 fluorescent image analyzer and quantified using the PhoretixTM 1D software. Data were plotted against time using GraFit 6 software (Erithacus Software). The AP endonuclease activity of commercial NEB wild type ExoIII was also tested under the parallel conditions as a positive control in comparison to those of home-made AP endonucleases.

6.3.4.7 Effects of ExoIII and EndoIV on MUG single turnover kinetics

In order to examine the effects of ExoIII and EndoIV on MUG single turnover kinetics, 500 nM wild type MUG was reacted with 100 nM Hex-12U•G o/h substrates in the presence of either 500 nM ExoIII or EndoIV in the standard reaction buffer at 25^{o}C. The reaction with ExoIII was supplemented with 1 mM $MgCl_2$. At specified time points as detailed in the figure legends, over a time course up to 15 minutes, 20 µl reaction aliquots were removed and quenched with an equal volume of 0.1 M NaOH. The quenched samples were then heated at 90^{o}C for 30 minutes. 20 µl 2 x formamide loading buffer was added prior to 20% urea denaturing PAGE (see Section 6.2.10). Gel imaging and analysis were carried out as above. Data were fitted to the first-order rate equation (Equation 3) using GraFit 6.

6.3.4.8 Effects of ExoIII and EndoIV on MUG multiple turnover kinetics

In order to examine the effects of ExoIII and EndoIV on MUG multiple turnover kinetics, burst kinetics assays were performed by rapidly mixing 2500 nM Hex-12U•G o/h substrates with 500 nM MUG in the presence of either 500 nM ExoIII or EndoIV. Reaction mixtures were incubated at 25^{o}C in the standard reaction buffer. 20 µl reaction aliquots were taken at selected time points as detailed in the figures over a 12 hour time course and quenched with 20 µl 0.1M NaOH. After heating at 90^{o}C for 30 minutes, the denatured reaction aliquots were mixed with 2 x formamide loading buffer and then 16 µl of those was analyzed via 20% urea denaturing PAGE (see Section 6.2.10). Gel imaging and analysis were carried out as above. Data were plotted against specified time points using GraFit 6.

6.3.4.9 Analysis of ExoIII catalytic activity dependence on MUG turnover enhancement

In order to determine the effect of ExoIII catalytic activity on MUG turnover enhancement, MUG multiple turnover reactions were performed by rapidly mixing 2500 nM Hex-12U•G o/h substrates with 500 nM MUG at 25^{o}C in the standard reaction buffer in the presence of either 500 nM wild-type ExoIII with 10 mM EDTA or 500 nM mutant ExoIII D151N. The reaction mixtures were then aliquoted in 20 μl volumes at specified time points during an incubation time course up to 12 hours and quenched by the equal volume of 0.1 M NaOH. The reaction aliquots were then denatured at 90^{o}C for 30 minutes and then mixed with 2 x formamide loading buffer prior to being analyzing via 20% urea denaturing PAGE (see Section 6.2.10). Gel imaging and analysis were carried out as above. Data were plotted against specified time points using GraFit 6.

6.3.4.10 Effects of ExoIII concentration on MUG turnover enhancement

In order to test whether the observed ExoIII stimulatory effect on MUG turnover was sensitive to enzyme concentration, a series of MUG multiple turnover reactions were performed with 2500 nM Hex-12U•G o/h substrates with 500 nM MUG in the presence of increasing concentrations of ExoIII (100 nM, 500 nM, 1000 nM and 2000 nM) at 25^{o}C in the standard reaction buffer supplemented with 1 mM $MgCl_2$. At specified time points as described in the figure legends, 20 μl reaction mixtures were removed and quenched with 20 μl 0.1 M NaOH. The quenched samples were then heated at 90^{o}C for 30 minutes. An equal volume of 2 x formamide loading buffer was added prior to 20% urea denaturing PAGE (see Section 6.2.10). A series of parallel reactions with mutant ExoIII D151N were performed with the same

buffer and conditions used in the assays for the wild type enzyme. Imaging and analysis were performed as above.

6.4 Site-directed mutagenesis

Site-directed mutagenesis for constructing 12 mutant MUG plasmids and ExoIII D151N plasmid were performed following the protocol of a Stratagene QuikChange® II Site-Directed Mutagenesis Kit. Separate primers (Table 10) were designed to be complimentary to the target region of the open reading frame encoding either MUG or ExoIII, except for the mutation codons highlighted in bond in table 10 in places of the native codons. The primers were employed in separate PCR reactions, which were conducted following the routine PCR protocol as described in Section 6.2.2 with the specific annealing temperatures listed in Table XXX, as well as either the pTrc99A plasmid containing wild-type MUG gene or the pPROEX-HTb_ExoIII plasmid harbouring *xthA* gene, that were both isolated from dam+ *E.coli* Rosetta^{TM} strain BL21 DE3 cells and used as the PCR starting template. 10 µl PCR products were analysed using 1% agarose gel electrophoresis (see Section 6.2.5) to test the efficiency of the reactions, and the products were then incubated with 1 µl of the restriction enzyme DpnI at 37^{o}C for 1 hour to digest the methylated template plasmid.

5 µl aliquots of the DpnI-treated DNA were used to transform chemically competent JM109 cells via heat-shock as described in Section 6.2.7. Successful colonies were selected by ampicillin-resistance, the plasmid DNAs were then obtained by miniprep (see Section 6.2.1) and subjected to specific restriction enzyme diagnostic digestions (enzymes not listed here) and finally submitted for DNA sequencing following the Cogenics' sequencing instructions (Beckman Coulter Genomics, UK) to verify that the selected clones contained the desired mutations. Manual sequence alignments

were generated using MultAlin (http://multalin.toulouse.inra.fr/multalin/). Mutant plasmids were then transformed into *E.coli* Rosetta™ strain BL21 DE3 cells to create strains capable of highly inducible over-expression of the mutant enzymes with limited proteolysis. Glycerol stocks of these strains were prepared and stored as described in Section 6.2.8.

Mutation name	Primer ID	Sequence 5'→3'	Annealing temperature (°C)
T74A	T74A_F	AGACCGTCCA**GCG**GTGCAAGC	55
	T74A_R	GCTTGCAC**CGC**TGGACGGTCT	
V75A	V75A_F	CGTCCAACG**GCG**CAAGCCAATG	56.5
	V75A_R	CATTGGCTT**GCG**CGTTGGACGGT	
Q76A	Q76A_F	CGTCCAACGGTG**GCA**GCCAATGAAG	57
	Q76A_R	CTTCATTGGC**TGC**CACCGTTGGACG	
N78A	N78A_F	CGGTGCAA**GCC**GCTGAAGTTTCAAAGC	57
	N78A_R	GCTTTGAAACTTCAGC**GGC**TTGCACCG	
E79A	E79A_F	GGTGCAAGCCAAT**GCA**GTTTCAAAGCAGG	58
	E79A_R	CCTGCTTTGAAAC**TGC**ATTGGCTTGCACC	
S81A	S81A_F	GCAAGCCAATGAAGTT**GCA**AAGCAGGAGCTAC	60
	S81A_R	GTAGCTCCTGCTT**TGC**AACTTCATTGGCTTGC	
K82A	K82A_F	CCAATGAAGTTTCA**GCG**CAGGAGCTACACG	60
	K82A_R	CGTGTAGCTCCTG**CGC**TGAAACTTCATTGG	
Q83A	Q83A_F	GAAGTTTCAAAG**GCG**GAGCTACACGCAG	58
	Q83A_R	CTGCGTGTAGCTC**CGC**CTTTGAAACTTC	
S22M	S22M_F	CCCTGGGCTT**ATG**TCCGCCGGGA	57
	S22M_R	TCCCGGCGGA**CAT**AAGCCCAGGG	
G25Y	G25Y_F	GGCTTTCATCCGCC**TAC**ACTGGTTTTCCC	58
	G25Y_R	GGGAAAACCAGT**GTA**GGCGGATGAAAGCC	
V75T	V75T_F	ACCGTCCAACG**ACG**CAAGCCAATG	57
	V75T_R	CATTGGCTTG**CGT**CGTTGGACGGT	
Q76P	Q76P_F	CGTCCAACGGTG**CCA**GCCAATGAAG	57
	Q76P_R	CTTCATTGGC**TGG**CACCGTTGGACG	
D151N	D151N_F	GATTATGGGC**AAT**ATGAATATCAGCCC	60
	D151N_R	GGGCTGATATTCAT**ATT**GCCCATAATC	

Table 10. Primers for site-directed mutagenesis of MUG mutants and ExoIII D151N mutant.

6.5 Protein structural modelling

All the protein structural models in this study were generated using PyMOL (V0.99, DeLano Scientific LLC) from the existing structures (PDB entries 1EMH, 1MWJ, 1AKO, 1QTW and 2RBA). Different key amino acid residues and DNA lesions in the protein structures were highlighted.

Chapter 7

References

1. Avery, O.T., C.M. Macleod, and M. McCarty, *Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types : Induction of Transformation by a Deoxyribonucleic Acid Fraction Isolated from Pneumococcus Type Iii.* J Exp Med, 1944. **79**(2): p. 137-158.
2. Watson, J.D. and F.H. Crick, *Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid.* Nature, 1953. **171**(4356): p. 737-738.
3. Levene, P.A.T. and L.W. Bass, *Nucleic acids*. 1931, New York: The Chemical Catalog Company, Inc.
4. Zephyris, *DNA Structure*. 2011. URL:http://upload.wikimedia.org/wikipedia/commons/4/4c/DNA_Structure%2BKey%2BLabelled.pn_NoBB.png.
5. Brenner, S., F. Jacob, and M. Meselson, *An unstable intermediate carrying information from genes to ribosomes for protein synthesis.* Nature, 1961. **190**: p. 576-581.
6. Crick, F.H., et al., *General nature of the genetic code for proteins.* Nature, 1961. **192**: p. 1227-1232.
7. Hood, L. and D. Galas, *The digital code of DNA*. Nature, 2003. **421**(6921): p. 444-448.
8. Lindahl, T., *Instability and decay of the primary structure of DNA.* Nature, 1993. **362**(6422): p. 709-715.
9. Fitzgerald, M.E. and A.C. Drohat, *Coordinating the initial steps of base excision repair. Apurinic/apyrimidinic endonuclease 1 actively stimulates thymine DNA glycosylase by disrupting the product complex.* J Biol Chem, 2008. **283**(47): p. 32680-32690.
10. Jackson, S.P. and J. Bartek, *The DNA-damage response in human biology and disease.* Nature, 2009. **461**(7267): p. 1071-1078.
11. Hanawalt, P.C., *Controlling the efficiency of excision repair.* Mutat Res, 2001. **485**(1): p. 3-13.
12. Kunkel, T.A., *DNA replication fidelity.* J Biol Chem, 2004. **279**(17): p. 16895-16898.
13. Singer, B., *Alkylation of the O6 of guanine is only one of many chemical events that may initiate carcinogenesis.* Cancer Invest, 1984. **2**(3): p. 233-238.
14. Lindahl, T., *An N-glycosidase from Escherichia coli that releases free uracil from DNA containing deaminated cytosine residues.* Proc Natl Acad Sci U S A, 1974. **71**(9): p. 3649-3653.
15. Frederico, L.A., T.A. Kunkel, and B.R. Shaw, *A sensitive genetic assay for the detection of cytosine deamination: determination of rate constants and the activation energy.* Biochemistry, 1990. **29**(10): p. 2532-2537.
16. Lindahl, T., *DNA glycosylases, endonucleases for apurinic/apyrimidinic sites, and base excision-repair.* Prog Nucleic Acid Res Mol Biol, 1979. **22**: p. 135-192.
17. Greer, S. and S. Zamenhof, *Studies on depurination of DNA by heat.* J Mol Biol, 1962. **4**: p. 123-141.
18. Nakamura, J., et al., *Highly sensitive apurinic/apyrimidinic site assay can detect*

spontaneous and chemically induced depurination under physiological conditions. Cancer Res, 1998. **58**(2): p. 222-225.
19. Lindahl, T. and O. Karlstrom, *Heat-induced depyrimidination of deoxyribonucleic acid in neutral solution.* Biochemistry, 1973. **12**(25): p. 5151-5154.
20. Ames, B.N. and L.S. Gold, *Endogenous mutagens and the causes of aging and cancer.* Mutat Res, 1991. **250**(1-2): p. 3-16.
21. Breimer, L.H., *Molecular mechanisms of oxygen radical carcinogenesis and mutagenesis: the role of DNA base damage.* Mol Carcinog, 1990. **3**(4): p. 188-197.
22. Imlay, J.A. and S. Linn, *DNA damage and oxygen radical toxicity.* Science, 1988. **240**(4857): p. 1302-1309.
23. Cooke, M.S., et al., *Oxidative DNA damage: mechanisms, mutation, and disease.* FASEB J, 2003. **17**(10): p. 1195-1214.
24. Bjelland, S. and E. Seeberg, *Mutagenicity, toxicity and repair of DNA base damage induced by oxidation.* Mutat Res, 2003. **531**(1-2): p. 37-80.
25. Breen, A.P. and J.A. Murphy, *Reactions of oxyl radicals with DNA.* Free Radic Biol Med, 1995. **18**(6): p. 1033-1077.
26. Demple, B. and L. Harrison, *Repair of oxidative damage to DNA: enzymology and biology.* Annu Rev Biochem, 1994. **63**: p. 915-948.
27. Kasai, H., et al., *8-Hydroxyguanine, a DNA adduct formed by oxygen radicals: its implication on oxygen radical-involved mutagenesis/carcinogenesis.* J Toxicol Sci, 1991. **16 Suppl 1**: p. 95-105.
28. Ward, J.F., *The yield of DNA double-strand breaks produced intracellularly by ionizing radiation: a review.* Int J Radiat Biol, 1990. **57**(6): p. 1141-1150.
29. Sanche, L., *Mechanisms of low energy electron damage to condensed biomolecules and DNA.* Radiat Prot Dosimetry, 2002. **99**(1-4): p. 57-62.
30. Ward, J.F., *DNA damage produced by ionizing radiation in mammalian cells: identities, mechanisms of formation, and reparability.* Prog Nucleic Acid Res Mol Biol, 1988. **35**: p. 95-125.
31. Teoule, R., *Radiation-induced DNA damage and its repair.* Int J Radiat Biol Relat Stud Phys Chem Med, 1987. **51**(4): p. 573-589.
32. Chetsanga, C.J., et al., *Purification and characterization of Escherichia coli formamidopyrimidine-DNA glycosylase that excises damaged 7-methylguanine from deoxyribonucleic acid.* Biochemistry, 1981. **20**(18): p. 5201-5207.
33. Lett, J.T., K.A. Stacey, and P. Alexander, *Crosslinking of dry deoxyribonucleic acids by electrons.* Radiat Res, 1961. **14**: p. 349-362.
34. Sonntag, C.v., *The chemical basis of radiation biology*. 1987, London ; New York: Taylor & Francis. 515 p.
35. Friedberg, E.C., *DNA repair and mutagenesis*. 2nd ed. 2006, Washington, D.C.: ASM Press. xxix, 1118 p.
36. Sinha, R.P. and D.P. Hader, *UV-induced DNA damage and repair: a review.* Photochem Photobiol Sci, 2002. **1**(4): p. 225-236.
37. Rastogi, R.P., et al., *Molecular mechanisms of ultraviolet radiation-induced DNA*

damage and repair. J Nucleic Acids, 2010. **2010**: p. 592980.
38. Britt, A.B., *Repair of DNA damage induced by ultraviolet radiation.* Plant Physiol, 1995. **108**(3): p. 891-896.
39. Setlow, R.B., *The photochemistry, photobiology, and repair of polynucleotides.* Prog Nucleic Acid Res Mol Biol, 1968. **8**: p. 257-295.
40. Duker, N.J. and P.E. Gallagher, *Purine photoproducts.* Photochem Photobiol, 1988. **48**(1): p. 35-39.
41. Singer, B. and J.T. Kusmierek, *Chemical mutagenesis.* Annu Rev Biochem, 1982. **51**: p. 655-693.
42. Singer, B. and B. Hang, *What structural features determine repair enzyme specificity and mechanism in chemically modified DNA?* Chem Res Toxicol, 1997. **10**(7): p. 713-732.
43. Singer, B., *O-alkyl pyrimidines in mutagenesis and carcinogenesis: occurrence and significance.* Cancer Res, 1986. **46**(10): p. 4879-4885.
44. Geiduschek, E.P., *"Reversible" DNA.* Proc Natl Acad Sci U S A, 1961. **47**: p. 950-955.
45. Borowy-Borowski, H., R. Lipman, and M. Tomasz, *Recognition between mitomycin C and specific DNA sequences for cross-link formation.* Biochemistry, 1990. **29**(12): p. 2999-3006.
46. Fox, M. and D. Scott, *The genetic toxicology of nitrogen and sulphur mustard.* Mutat Res, 1980. **75**(2): p. 131-168.
47. Roberts, J.J. and J.M. Pascoe, *Cross-linking of complementary strands of DNA in mammalian cells by antitumour platinum compounds.* Nature, 1972. **235**(5336): p. 282-284.
48. Cole, R.S., *Light-induced cross-linking of DNA in the presence of a furocoumarin (psoralen). Studies with phage lambda, Escherichia coli, and mouse leukemia cells.* Biochim Biophys Acta, 1970. **217**(1): p. 30-39.
49. Crick, F., *The double helix: a personal view.* Nature, 1974. **248**(5451): p. 766-769.
50. Lindahl, T. and D.E. Barnes, *Repair of endogenous DNA damage.* Cold Spring Harb Symp Quant Biol, 2000. **65**: p. 127-133.
51. Nakano, T., et al., *Nucleotide excision repair and homologous recombination systems commit differentially to the repair of DNA-protein crosslinks.* Mol Cell, 2007. **28**(1): p. 147-158.
52. Le May, N., J.M. Egly, and F. Coin, *True lies: the double life of the nucleotide excision repair factors in transcription and DNA repair.* J Nucleic Acids, 2010. **2010**.
53. Liu, Y., et al., *Coordination of steps in single-nucleotide base excision repair mediated by apurinic/apyrimidinic endonuclease 1 and DNA polymerase beta.* J Biol Chem, 2007. **282**(18): p. 13532-13541.
54. Lieber, M.R., *The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway.* Annu Rev Biochem, 2010. **79**: p. 181-211.
55. McVey, M. and S.E. Lee, *MMEJ repair of double-strand breaks (director's cut):*

deleted sequences and alternative endings. Trends Genet, 2008. **24**(11): p. 529-538.
56. Wyman, C., D. Ristic, and R. Kanaar, *Homologous recombination-mediated double-strand break repair.* DNA Repair (Amst), 2004. **3**(8-9): p. 827-833.
57. Ishino, Y., T. Nishino, and K. Morikawa, *Mechanisms of maintaining genetic stability by homologous recombination.* Chem Rev, 2006. **106**(2): p. 324-339.
58. Kowalczykowski, S.C., et al., *Biochemistry of homologous recombination in Escherichia coli.* Microbiol Rev, 1994. **58**(3): p. 401-465.
59. Kennedy, R.D. and A.D. D'Andrea, *The Fanconi Anemia/BRCA pathway: new faces in the crowd.* Genes Dev, 2005. **19**(24): p. 2925-2940.
60. McGlynn, P. and R.G. Lloyd, *Recombinational repair and restart of damaged replication forks.* Nat Rev Mol Cell Biol, 2002. **3**(11): p. 859-870.
61. Northam, M.R., et al., *DNA polymerases zeta and Rev1 mediate error-prone bypass of non-B DNA structures.* Nucleic Acids Res, 2013.
62. Wang, Z., *Translesion synthesis by the UmuC family of DNA polymerases.* Mutat Res, 2001. **486**(2): p. 59-70.
63. Radman, M., *SOS repair hypothesis: phenomenology of an inducible DNA repair which is accompanied by mutagenesis.* Basic Life Sci, 1975. **5A**: p. 355-367.
64. Bridges, B.A., *Error-prone DNA repair and translesion DNA synthesis. II: The inducible SOS hypothesis.* DNA Repair (Amst), 2005. **4**(6): p. 725-726, 739.
65. Erill, I., S. Campoy, and J. Barbe, *Aeons of distress: an evolutionary perspective on the bacterial SOS response.* FEMS Microbiol Rev, 2007. **31**(6): p. 637-656.
66. Michel, B., *After 30 years of study, the bacterial SOS response still surprises us.* PLoS Biol, 2005. **3**(7): p. e255.
67. Huisman, O., R. D'Ari, and S. Gottesman, *Cell-division control in Escherichia coli: specific induction of the SOS function SfiA protein is sufficient to block septation.* Proc Natl Acad Sci U S A, 1984. **81**(14): p. 4490-4494.
68. Tang, M., et al., *UmuD'(2)C is an error-prone DNA polymerase, Escherichia coli pol V.* Proc Natl Acad Sci U S A, 1999. **96**(16): p. 8919-8924.
69. Iwasaki, H., et al., *The Escherichia coli polB gene, which encodes DNA polymerase II, is regulated by the SOS system.* J Bacteriol, 1990. **172**(11): p. 6268-6273.
70. Wagner, J., et al., *The dinB gene encodes a novel E. coli DNA polymerase, DNA pol IV, involved in mutagenesis.* Mol Cell, 1999. **4**(2): p. 281-286.
71. Napolitano, R., et al., *All three SOS-inducible DNA polymerases (Pol II, Pol IV and Pol V) are involved in induced mutagenesis.* EMBO J, 2000. **19**(22): p. 6259-6265.
72. Eller, M.S., et al., *Enhancement of DNA repair in human skin cells by thymidine dinucleotides: evidence for a p53-mediated mammalian SOS response.* Proc Natl Acad Sci U S A, 1997. **94**(23): p. 12627-12632.
73. Fu, Y., et al., *Rad6-Rad18 mediates a eukaryotic SOS response by ubiquitinating the 9-1-1 checkpoint clamp.* Cell, 2008. **133**(4): p. 601-611.
74. Morita, T., et al., *A mouse homolog of the Escherichia coli recA and Saccharomyces cerevisiae RAD51 genes.* Proc Natl Acad Sci U S A, 1993. **90**(14): p. 6577-6580.
75. Lindahl, T. and R.D. Wood, *Quality control by DNA repair.* Science, 1999. **286**(5446):

p. 1897-1905.
76. Engelward, B.P., et al., *A chemical and genetic approach together define the biological consequences of 3-methyladenine lesions in the mammalian genome.* J Biol Chem, 1998. **273**(9): p. 5412-5418.
77. Friedberg, E.C. and T. Lindahl, *Inroads into base excision repair II. The discovery of DNA glycosylases. "An N-glycosidase from Escherichia coli that releases free uracil from DNA containing deaminated cytosine residues," Proc. Nat. Acad. Sci. USA, 1974.* DNA Repair (Amst), 2004. **3**(11): p. 1532-1536; discussion 1531-1532.
78. Dianov, G. and T. Lindahl, *Reconstitution of the DNA base excision-repair pathway.* Curr Biol, 1994. **4**(12): p. 1069-1076.
79. Kubota, Y., et al., *Reconstitution of DNA base excision-repair with purified human proteins: interaction between DNA polymerase beta and the XRCC1 protein.* EMBO J, 1996. **15**(23): p. 6662-6670.
80. Podlutsky, A.J., et al., *DNA synthesis and dRPase activities of polymerase beta are both essential for single-nucleotide patch base excision repair in mammalian cell extracts.* Biochemistry, 2001. **40**(3): p. 809-813.
81. Piersen, C.E., A.K. McCullough, and R.S. Lloyd, *AP lyases and dRPases: commonality of mechanism.* Mutat Res, 2000. **459**(1): p. 43-53.
82. Nagorska, K., et al., *A network of enzymes involved in repair of oxidative DNA damage in Neisseria meningitidis.* Mol Microbiol, 2012. **83**(5): p. 1064-1079.
83. Silhan, J., et al., *Specialization of an Exonuclease III family enzyme in the repair of 3' DNA lesions during base excision repair in the human pathogen Neisseria meningitidis.* Nucleic Acids Res, 2012. **40**(5): p. 2065-2075.
84. Vidal, A.E., et al., *XRCC1 coordinates the initial and late stages of DNA abasic site repair through protein-protein interactions.* EMBO J, 2001. **20**(22): p. 6530-6539.
85. Fortini, P. and E. Dogliotti, *Base damage and single-strand break repair: mechanisms and functional significance of short- and long-patch repair subpathways.* DNA Repair (Amst), 2007. **6**(4): p. 398-409.
86. Fortini, P., et al., *The type of DNA glycosylase determines the base excision repair pathway in mammalian cells.* J Biol Chem, 1999. **274**(21): p. 15230-15236.
87. Tom, S., et al., *Regulatory roles of p21 and apurinic/apyrimidinic endonuclease 1 in base excision repair.* J Biol Chem, 2001. **276**(52): p. 48781-48789.
88. Balakrishnan, L., et al., *Long patch base excision repair proceeds via coordinated stimulation of the multienzyme DNA repair complex.* J Biol Chem, 2009. **284**(22): p. 15158-15172.
89. Zhang, W., et al., *Roles of sequential ubiquitination of PCNA in DNA-damage tolerance.* FEBS Lett, 2011. **585**(18): p. 2786-2794.
90. Bhattacharyya, N. and S. Banerjee, *A variant of DNA polymerase beta acts as a dominant negative mutant.* Proc Natl Acad Sci U S A, 1997. **94**(19): p. 10324-10329.
91. Sung, J.S. and D.W. Mosbaugh, *Escherichia coli uracil- and ethenocytosine-initiated base excision DNA repair: rate-limiting step and patch size distribution.* Biochemistry, 2003. **42**(16): p. 4613-4625.

92. Baker, D., et al., *Characterization of the substrate specificity of a human 5-hydroxymethyluracil glycosylase activity.* Chem Res Toxicol, 2002. **15**(1): p. 33-39.
93. Barrett, T.E., et al., *Crystal structure of a thwarted mismatch glycosylase DNA repair complex.* EMBO J, 1999. **18**(23): p. 6599-6609.
94. Guliaev, A.B., B. Singer, and B. Hang, *Chloroethylnitrosourea-derived ethano cytosine and adenine adducts are substrates for Escherichia coli glycosylases excising analogous etheno adducts.* DNA Repair (Amst), 2004. **3**(10): p. 1311-1321.
95. Hang, B., et al., *Novel activity of Escherichia coli mismatch uracil-DNA glycosylase (Mug) excising 8-(hydroxymethyl)-3,N4-ethenocytosine, a potential product resulting from glycidaldehyde reaction.* Biochemistry, 2002. **41**(7): p. 2158-2165.
96. Saparbaev, M., et al., *1,N(2)-ethenoguanine, a mutagenic DNA adduct, is a primary substrate of Escherichia coli mismatch-specific uracil-DNA glycosylase and human alkylpurine-DNA-N-glycosylase.* J Biol Chem, 2002. **277**(30): p. 26987-26993.
97. Lee, H.W., et al., *Identification of Escherichia coli mismatch-specific uracil DNA glycosylase as a robust xanthine DNA glycosylase.* J Biol Chem, 2010. **285**(53): p. 41483-41490.
98. Patrick, M.H., *Substrate specificity of Micrococcus luteus UV endonuclease and its overlap with DNA photolyase activity.* Basic Life Sci, 1975. **5A**: p. 205-208.
99. Lindahl, T., et al., *DNA N-glycosidases: properties of uracil-DNA glycosidase from Escherichia coli.* J Biol Chem, 1977. **252**(10): p. 3286-3294.
100. Rubinson, E.H. and B.F. Eichman, *Nucleic acid recognition by tandem helical repeats.* Curr Opin Struct Biol, 2012. **22**(1): p. 101-109.
101. Klimasauskas, S., et al., *HhaI methyltransferase flips its target base out of the DNA helix.* Cell, 1994. **76**(2): p. 357-369.
102. Parker, J.B., et al., *Enzymatic capture of an extrahelical thymine in the search for uracil in DNA.* Nature, 2007. **449**(7161): p. 433-437.
103. Vassylyev, D.G., et al., *Atomic model of a pyrimidine dimer excision repair enzyme complexed with a DNA substrate: structural basis for damaged DNA recognition.* Cell, 1995. **83**(5): p. 773-782.
104. Bellamy, S.R. and G.S. Baldwin, *A kinetic analysis of substrate recognition by uracil-DNA glycosylase from herpes simplex virus type 1.* Nucleic Acids Res, 2001. **29**(18): p. 3857-3863.
105. Pearl, L.H., *Structure and function in the uracil-DNA glycosylase superfamily.* Mutat Res, 2000. **460**(3-4): p. 165-181.
106. Hitomi, K., S. Iwai, and J.A. Tainer, *The intricate structural chemistry of base excision repair machinery: implications for DNA damage recognition, removal, and repair.* DNA Repair (Amst), 2007. **6**(4): p. 410-428.
107. Brooks, S.C., et al., *Recent advances in the structural mechanisms of DNA glycosylases.* Biochim Biophys Acta, 2013. **1834**(1): p. 247-271.
108. Kavli, B., et al., *Uracil in DNA--general mutagen, but normal intermediate in acquired immunity.* DNA Repair (Amst), 2007. **6**(4): p. 505-516.
109. Savva, R., et al., *The structural basis of specific base-excision repair by uracil-DNA*

glycosylase. Nature, 1995. **373**(6514): p. 487-493.
110. Mol, C.D., et al., *Crystal structure and mutational analysis of human uracil-DNA glycosylase: structural basis for specificity and catalysis.* Cell, 1995. **80**(6): p. 869-878.
111. Slupphaug, G., et al., *Properties of a recombinant human uracil-DNA glycosylase from the UNG gene and evidence that UNG encodes the major uracil-DNA glycosylase.* Biochemistry, 1995. **34**(1): p. 128-138.
112. Huffman, J.L., O. Sundheim, and J.A. Tainer, *DNA base damage recognition and removal: new twists and grooves.* Mutat Res, 2005. **577**(1-2): p. 55-76.
113. Slupphaug, G., et al., *A nucleotide-flipping mechanism from the structure of human uracil-DNA glycosylase bound to DNA.* Nature, 1996. **384**(6604): p. 87-92.
114. Xiao, G., et al., *Crystal structure of Escherichia coli uracil DNA glycosylase and its complexes with uracil and glycerol: structure and glycosylase mechanism revisited.* Proteins, 1999. **35**(1): p. 13-24.
115. Leiros, I., et al., *The structure of uracil-DNA glycosylase from Atlantic cod (Gadus morhua) reveals cold-adaptation features.* Acta Crystallogr D Biol Crystallogr, 2003. **59**(Pt 8): p. 1357-1365.
116. Krokan, H.E., R. Standal, and G. Slupphaug, *DNA glycosylases in the base excision repair of DNA.* Biochem J, 1997. **325 (Pt 1)**: p. 1-16.
117. Parikh, S.S., et al., *Base excision repair initiation revealed by crystal structures and binding kinetics of human uracil-DNA glycosylase with DNA.* EMBO J, 1998. **17**(17): p. 5214-5226.
118. Wong, I., et al., *Presteady-state analysis of a single catalytic turnover by Escherichia coli uracil-DNA glycosylase reveals a "pinch-pull-push" mechanism.* J Biol Chem, 2002. **277**(22): p. 19424-19432.
119. Dong, J., et al., *Raman spectroscopy of uracil DNA glycosylase-DNA complexes: insights into DNA damage recognition and catalysis.* Biochemistry, 2000. **39**(43): p. 13241-13250.
120. Parikh, S.S., et al., *Uracil-DNA glycosylase-DNA substrate and product structures: conformational strain promotes catalytic efficiency by coupled stereoelectronic effects.* Proc Natl Acad Sci U S A, 2000. **97**(10): p. 5083-5088.
121. Drohat, A.C., et al., *Role of electrophilic and general base catalysis in the mechanism of Escherichia coli uracil DNA glycosylase.* Biochemistry, 1999. **38**(37): p. 11866-11875.
122. Ingraham, H.A., B.Y. Tseng, and M. Goulian, *Mechanism for exclusion of 5-fluorouracil from DNA.* Cancer Res, 1980. **40**(4): p. 998-1001.
123. Zastawny, T.H., P.W. Doetsch, and M. Dizdaroglu, *A novel activity of E. coli uracil DNA N-glycosylase excision of isodialuric acid (5,6-dihydroxyuracil), a major product of oxidative DNA damage, from DNA.* FEBS Lett, 1995. **364**(3): p. 255-258.
124. Hatahet, Z., et al., *New substrates for old enzymes. 5-Hydroxy-2'-deoxycytidine and 5-hydroxy-2'-deoxyuridine are substrates for Escherichia coli endonuclease III and formamidopyrimidine DNA N-glycosylase, while 5-hydroxy-2'-deoxyuridine is a*

substrate for uracil DNA N-glycosylase. J Biol Chem, 1994. **269**(29): p. 18814-18820.

125. Dizdaroglu, M., et al., *Novel activities of human uracil DNA N-glycosylase for cytosine-derived products of oxidative DNA damage.* Nucleic Acids Res, 1996. **24**(3): p. 418-422.
126. Dizdaroglu, M., A.E. Karakaya, and North Atlantic Treaty Organization. Scientific Affairs Division., *Advances in DNA damage and repair : oxygen radical effects, cellular protection, and biological consequences*. NATO ASI series. Series A, Life sciences. 1999, New York ; London: Kluwer Academic/Plenum. x, 511 p.
127. Gallinari, P. and J. Jiricny, *A new class of uracil-DNA glycosylases related to human thymine-DNA glycosylase.* Nature, 1996. **383**(6602): p. 735-738.
128. Barrett, T.E., et al., *Crystal structure of a G:T/U mismatch-specific DNA glycosylase: mismatch recognition by complementary-strand interactions.* Cell, 1998. **92**(1): p. 117-129.
129. O'Neill, R.J., et al., *Mismatch uracil glycosylase from Escherichia coli: a general mismatch or a specific DNA glycosylase?* J Biol Chem, 2003. **278**(23): p. 20526-20532.
130. Haushalter, K.A., et al., *Identification of a new uracil-DNA glycosylase family by expression cloning using synthetic inhibitors.* Curr Biol, 1999. **9**(4): p. 174-185.
131. Wibley, J.E., et al., *Structure and specificity of the vertebrate anti-mutator uracil-DNA glycosylase SMUG1.* Mol Cell, 2003. **11**(6): p. 1647-1659.
132. Masaoka, A., et al., *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, 2003. **42**(17): p. 5003-5012.
133. Krokan, H.E., F. Drablos, and G. Slupphaug, *Uracil in DNA--occurrence, consequences and repair.* Oncogene, 2002. **21**(58): p. 8935-8948.
134. Sandigursky, M. and W.A. Franklin, *DNA deoxyribophosphodiesterase of Escherichia coli is associated with exonuclease I.* Nucleic Acids Res, 1992. **20**(18): p. 4699-4703.
135. Sandigursky, M. and W.A. Franklin, *Uracil-DNA glycosylase in the extreme thermophile Archaeoglobus fulgidus.* J Biol Chem, 2000. **275**(25): p. 19146-19149.
136. Sartori, A.A., et al., *A novel uracil-DNA glycosylase with broad substrate specificity and an unusual active site.* EMBO J, 2002. **21**(12): p. 3182-3191.
137. Starkuviene, V. and H.J. Fritz, *A novel type of uracil-DNA glycosylase mediating repair of hydrolytic DNA damage in the extremely thermophilic eubacterium Thermus thermophilus.* Nucleic Acids Res, 2002. **30**(10): p. 2097-2102.
138. Hinks, J.A., et al., *An iron-sulfur cluster in the family 4 uracil-DNA glycosylases.* J Biol Chem, 2002. **277**(19): p. 16936-16940.
139. Cunningham, R.P., et al., *Endonuclease III is an iron-sulfur protein.* Biochemistry, 1989. **28**(10): p. 4450-4455.
140. Michaels, M.L., et al., *MutY, an adenine glycosylase active on G-A mispairs, has homology to endonuclease III.* Nucleic Acids Res, 1990. **18**(13): p. 3841-3845.

141. Yang, H., et al., *Characterization of a thermostable DNA glycosylase specific for U/G and T/G mismatches from the hyperthermophilic archaeon Pyrobaculum aerophilum.* J Bacteriol, 2000. **182**(5): p. 1272-1279.
142. Fu, W., et al., *The role of the iron-sulfur cluster in Escherichia coli endonuclease III. A resonance Raman study.* J Biol Chem, 1992. **267**(23): p. 16135-16137.
143. Brzoska, K., S. Meczynska, and M. Kruszewski, *Iron-sulfur cluster proteins: electron transfer and beyond.* Acta Biochim Pol, 2006. **53**(4): p. 685-691.
144. Neddermann, P. and J. Jiricny, *The purification of a mismatch-specific thymine-DNA glycosylase from HeLa cells.* J Biol Chem, 1993. **268**(28): p. 21218-21224.
145. Brown, T.C. and J. Jiricny, *A specific mismatch repair event protects mammalian cells from loss of 5-methylcytosine.* Cell, 1987. **50**(6): p. 945-950.
146. Waters, T.R. and P.F. Swann, *Kinetics of the action of thymine DNA glycosylase.* J Biol Chem, 1998. **273**(32): p. 20007-20014.
147. Neddermann, P. and J. Jiricny, *Efficient removal of uracil from G.U mispairs by the mismatch-specific thymine DNA glycosylase from HeLa cells.* Proc Natl Acad Sci U S A, 1994. **91**(5): p. 1642-1646.
148. Hardeland, U., et al., *Thymine DNA glycosylase.* Prog Nucleic Acid Res Mol Biol, 2001. **68**: p. 235-253.
149. Saparbaev, M. and J. Laval, *3,N4-ethenocytosine, a highly mutagenic adduct, is a primary substrate for Escherichia coli double-stranded uracil-DNA glycosylase and human mismatch-specific thymine-DNA glycosylase.* Proc Natl Acad Sci U S A, 1998. **95**(15): p. 8508-8513.
150. Sard, L., et al., *Chromosomal localizations and molecular analysis of TDG gene-related sequences.* Genomics, 1997. **44**(2): p. 222-226.
151. Borodovsky, M., et al., *Detection of new genes in a bacterial genome using Markov models for three gene classes.* Nucleic Acids Res, 1995. **23**(17): p. 3554-3562.
152. Orengo, C.A. and W.R. Taylor, *SSAP: sequential structure alignment program for protein structure comparison.* Methods Enzymol, 1996. **266**: p. 617-635.
153. Kavli, B., et al., *Excision of cytosine and thymine from DNA by mutants of human uracil-DNA glycosylase.* EMBO J, 1996. **15**(13): p. 3442-3447.
154. Grippon, S., et al., *Differential modes of DNA binding by mismatch uracil DNA glycosylase from Escherichia coli: implications for abasic lesion processing and enzyme communication in the base excision repair pathway.* Nucleic Acids Res, 2011. **39**(7): p. 2593-2603.
155. Vidal, A.E., et al., *Mechanism of stimulation of the DNA glycosylase activity of hOGG1 by the major human AP endonuclease: bypass of the AP lyase activity step.* Nucleic Acids Res, 2001. **29**(6): p. 1285-1292.
156. Sung, J.S. and D.W. Mosbaugh, *Escherichia coli double-strand uracil-DNA glycosylase: involvement in uracil-mediated DNA base excision repair and stimulation of activity by endonuclease IV.* Biochemistry, 2000. **39**(33): p. 10224-10235.
157. Porello, S.L., A.E. Leyes, and S.S. David, *Single-turnover and pre-steady-state*

kinetics of the reaction of the adenine glycosylase MutY with mismatch-containing DNA substrates. Biochemistry, 1998. **37**(42): p. 14756-14764.
158. Cortazar, D., et al., *The enigmatic thymine DNA glycosylase.* DNA Repair (Amst), 2007. **6**(4): p. 489-504.
159. Farez-Vidal, M.E., et al., *Characterization of uracil-DNA glycosylase activity from Trypanosoma cruzi and its stimulation by AP endonuclease.* Nucleic Acids Res, 2001. **29**(7): p. 1549-1555.
160. Hill, J.W., et al., *Stimulation of human 8-oxoguanine-DNA glycosylase by AP-endonuclease: potential coordination of the initial steps in base excision repair.* Nucleic Acids Res, 2001. **29**(2): p. 430-438.
161. Pope, M.A., S.L. Porello, and S.S. David, *Escherichia coli apurinic-apyrimidinic endonucleases enhance the turnover of the adenine glycosylase MutY with G:A substrates.* J Biol Chem, 2002. **277**(25): p. 22605-22615.
162. Waters, T.R., et al., *Human thymine DNA glycosylase binds to apurinic sites in DNA but is displaced by human apurinic endonuclease 1.* J Biol Chem, 1999. **274**(1): p. 67-74.
163. Atamna, H., I. Cheung, and B.N. Ames, *A method for detecting abasic sites in living cells: age-dependent changes in base excision repair.* Proc Natl Acad Sci U S A, 2000. **97**(2): p. 686-691.
164. Otterlei, M., et al., *Repair of chromosomal abasic sites in vivo involves at least three different repair pathways.* EMBO J, 2000. **19**(20): p. 5542-5551.
165. Kow, Y.W., *Mechanism of action of Escherichia coli exonuclease III.* Biochemistry, 1989. **28**(8): p. 3280-3287.
166. Daley, J.M., C. Zakaria, and D. Ramotar, *The endonuclease IV family of apurinic/apyrimidinic endonucleases.* Mutat Res, 2010. **705**(3): p. 217-227.
167. Walkup, L.K. and T. Kogoma, *Escherichia coli proteins inducible by oxidative stress mediated by the superoxide radical.* J Bacteriol, 1989. **171**(3): p. 1476-1484.
168. Chan, E. and B. Weiss, *Endonuclease IV of Escherichia coli is induced by paraquat.* Proc Natl Acad Sci U S A, 1987. **84**(10): p. 3189-3193.
169. Weiss, B., *Endonuclease II of Escherichia coli is exonuclease III.* J Biol Chem, 1976. **251**(7): p. 1896-1901.
170. Henner, W.D., S.M. Grunberg, and W.A. Haseltine, *Enzyme action at 3′ termini of ionizing radiation-induced DNA strand breaks.* J Biol Chem, 1983. **258**(24): p. 15198-15205.
171. Doetsch, P.W., et al., *A highly conserved endonuclease activity present in Escherichia coli, bovine, and human cells recognizes oxidative DNA damage at sites of pyrimidines.* Mol Cell Biol, 1987. **7**(1): p. 26-32.
172. Mol, C.D., et al., *Structure and function of the multifunctional DNA-repair enzyme exonuclease III.* Nature, 1995. **374**(6520): p. 381-386.
173. Rogers, S.G. and B. Weiss, *Exonuclease III of Escherichia coli K-12, an AP endonuclease.* Methods Enzymol, 1980. **65**(1): p. 201-211.
174. Ljungquist, S., *A new endonuclease from Escherichia coli acting at apurinic sites in*

DNA. J Biol Chem, 1977. **252**(9): p. 2808-2814.
175. Gorman, M.A., et al., *The crystal structure of the human DNA repair endonuclease HAP1 suggests the recognition of extra-helical deoxyribose at DNA abasic sites*. EMBO J, 1997. **16**(21): p. 6548-6558.
176. Demple, B., T. Herman, and D.S. Chen, *Cloning and expression of APE, the cDNA encoding the major human apurinic endonuclease: definition of a family of DNA repair enzymes*. Proc Natl Acad Sci U S A, 1991. **88**(24): p. 11450-11454.
177. Robson, C.N., et al., *Structure of the human DNA repair gene HAP1 and its localisation to chromosome 14q 11.2-12*. Nucleic Acids Res, 1992. **20**(17): p. 4417-4421.
178. Seki, S., et al., *cDNA cloning, sequencing, expression and possible domain structure of human APEX nuclease homologous to Escherichia coli exonuclease III*. Biochim Biophys Acta, 1992. **1131**(3): p. 287-299.
179. Xanthoudakis, S., et al., *Redox activation of Fos-Jun DNA binding activity is mediated by a DNA repair enzyme*. EMBO J, 1992. **11**(9): p. 3323-3335.
180. Cunningham, R.P., et al., *Endonuclease IV (nfo) mutant of Escherichia coli*. J Bacteriol, 1986. **168**(3): p. 1120-1127.
181. Kaneda, K., J. Sekiguchi, and T. Shida, *Role of the tryptophan residue in the vicinity of the catalytic center of exonuclease III family AP endonucleases: AP site recognition mechanism*. Nucleic Acids Res, 2006. **34**(5): p. 1552-1563.
182. Lu, D., et al., *Structural basis for the recognition and cleavage of abasic DNA in Neisseria meningitidis*. Proc Natl Acad Sci U S A, 2012. **109**(42): p. 16852-16857.
183. Walker, L.J., et al., *Identification of residues in the human DNA repair enzyme HAP1 (Ref-1) that are essential for redox regulation of Jun DNA binding*. Mol Cell Biol, 1993. **13**(9): p. 5370-5376.
184. Levin, J.D., R. Shapiro, and B. Demple, *Metalloenzymes in DNA repair. Escherichia coli endonuclease IV and Saccharomyces cerevisiae Apn1*. J Biol Chem, 1991. **266**(34): p. 22893-22898.
185. Ide, H., et al., *Alpha-deoxyadenosine, a major anoxic radiolysis product of adenine in DNA, is a substrate for Escherichia coli endonuclease IV*. Biochemistry, 1994. **33**(25): p. 7842-7847.
186. Levin, J.D. and B. Demple, *In vitro detection of endonuclease IV-specific DNA damage formed by bleomycin in vivo*. Nucleic Acids Res, 1996. **24**(5): p. 885-889.
187. Liochev, S.I., et al., *Induction of the soxRS regulon of Escherichia coli by superoxide*. J Biol Chem, 1999. **274**(14): p. 9479-9481.
188. Garcin, E.D., et al., *DNA apurinic-apyrimidinic site binding and excision by endonuclease IV*. Nat Struct Mol Biol, 2008. **15**(5): p. 515-522.
189. Fried, M.G., et al., *DNA binding mechanism of O6-alkylguanine-DNA alkyltransferase: stoichiometry and effects of DNA base composition and secondary structure on complex stability*. Biochemistry, 1996. **35**(48): p. 15295-15301.
190. Rasimas, J.J., et al., *Interactions of human O6-alkylguanine-DNA alkyltransferase (AGT) with short single-stranded DNAs*. J Biol Chem, 2007. **282**(5): p. 3357-3366.

191. Maiti, A., et al., *Crystal structure of human thymine DNA glycosylase bound to DNA elucidates sequence-specific mismatch recognition.* Proc Natl Acad Sci U S A, 2008. **105**(26): p. 8890-8895.
192. Morgan, M.T., et al., *Stoichiometry and affinity for thymine DNA glycosylase binding to specific and nonspecific DNA.* Nucleic Acids Res, 2011. **39**(6): p. 2319-2329.
193. Grippon, A.S. *Protein complexes in base excision repair [electronic resource] : biochemical and kinetic analysis of mismatch uracil DNA glycosylase*. 2010; URL: http://hdl.handle.net/10044/1/5666.
194. Heyduk, T. and J.C. Lee, *Application of fluorescence energy transfer and polarization to monitor Escherichia coli cAMP receptor protein and lac promoter interaction.* Proc Natl Acad Sci U S A, 1990. **87**(5): p. 1744-1748.
195. Cornish-Bowden, A., *Fundamentals of enzyme kinetics*. 4th rev. ed. ed. 2012, Weinheim: Wiley-VCH.
196. Hardeland, U., et al., *Separating substrate recognition from base hydrolysis in human thymine DNA glycosylase by mutational analysis.* J Biol Chem, 2000. **275**(43): p. 33449-33456.
197. Sidorova, N.Y., S. Muradymov, and D.C. Rau, *Trapping DNA-protein binding reactions with neutral osmolytes for the analysis by gel mobility shift and self-cleavage assays.* Nucleic Acids Res, 2005. **33**(16): p. 5145-5155.
198. Record, M.T., Jr., M.L. Lohman, and P. De Haseth, *Ion effects on ligand-nucleic acid interactions.* J Mol Biol, 1976. **107**(2): p. 145-158.
199. Otterlei, M., et al., *Post-replicative base excision repair in replication foci.* EMBO J, 1999. **18**(13): p. 3834-3844.
200. Muller-Weeks, S.J. and S. Caradonna, *Specific association of cyclin-like uracil-DNA glycosylase with the proliferating cell nuclear antigen.* Exp Cell Res, 1996. **226**(2): p. 346-355.
201. Schrofelbauer, B., et al., *Human immunodeficiency virus type 1 Vpr induces the degradation of the UNG and SMUG uracil-DNA glycosylases.* J Virol, 2005. **79**(17): p. 10978-10987.
202. Tini, M., et al., *Association of CBP/p300 acetylase and thymine DNA glycosylase links DNA repair and transcription.* Mol Cell, 2002. **9**(2): p. 265-277.
203. Shimizu, Y., et al., *Xeroderma pigmentosum group C protein interacts physically and functionally with thymine DNA glycosylase.* EMBO J, 2003. **22**(1): p. 164-173.
204. Campalans, A., et al., *XRCC1 interactions with multiple DNA glycosylases: a model for its recruitment to base excision repair.* DNA Repair (Amst), 2005. **4**(7): p. 826-835.
205. Miao, F., et al., *3-Methyladenine-DNA glycosylase (MPG protein) interacts with human RAD23 proteins.* J Biol Chem, 2000. **275**(37): p. 28433-28438.
206. Watanabe, S., et al., *Methylated DNA-binding domain 1 and methylpurine-DNA glycosylase link transcriptional repression and DNA repair in chromatin.* Proc Natl Acad Sci U S A, 2003. **100**(22): p. 12859-12864.
207. Marsin, S., et al., *Role of XRCC1 in the coordination and stimulation of oxidative*

DNA damage repair initiated by the DNA glycosylase hOGG1. J Biol Chem, 2003. **278**(45): p. 44068-44074.

208. Tuo, J., et al., *Functional crosstalk between hOgg1 and the helicase domain of Cockayne syndrome group B protein.* DNA Repair (Amst), 2002. **1**(11): p. 913-927.
209. Yang, H., et al., *Enhanced activity of adenine-DNA glycosylase (Myh) by apurinic/apyrimidinic endonuclease (Ape1) in mammalian base excision repair of an A/GO mismatch.* Nucleic Acids Res, 2001. **29**(3): p. 743-752.
210. Parker, A., et al., *Human homolog of the MutY repair protein (hMYH) physically interacts with proteins involved in long patch DNA base excision repair.* J Biol Chem, 2001. **276**(8): p. 5547-5555.
211. Bessho, T., *Nucleotide excision repair 3' endonuclease XPG stimulates the activity of base excision repairenzyme thymine glycol DNA glycosylase.* Nucleic Acids Res, 1999. **27**(4): p. 979-983.
212. Klungland, A., et al., *Base excision repair of oxidative DNA damage activated by XPG protein.* Mol Cell, 1999. **3**(1): p. 33-42.
213. Das, A., et al., *NEIL2-initiated, APE-independent repair of oxidized bases in DNA: Evidence for a repair complex in human cells.* DNA Repair (Amst), 2006. **5**(12): p. 1439-1448.
214. Fan, Z., et al., *Cleaving the oxidative repair protein Ape1 enhances cell death mediated by granzyme A.* Nat Immunol, 2003. **4**(2): p. 145-153.
215. Dianova, II, V.A. Bohr, and G.L. Dianov, *Interaction of human AP endonuclease 1 with flap endonuclease 1 and proliferating cell nuclear antigen involved in long-patch base excision repair.* Biochemistry, 2001. **40**(42): p. 12639-12644.
216. Ranalli, T.A., S. Tom, and R.A. Bambara, *AP endonuclease 1 coordinates flap endonuclease 1 and DNA ligase I activity in long patch base excision repair.* J Biol Chem, 2002. **277**(44): p. 41715-41724.
217. Bennett, R.A., et al., *Interaction of human apurinic endonuclease and DNA polymerase beta in the base excision repair pathway.* Proc Natl Acad Sci U S A, 1997. **94**(14): p. 7166-7169.
218. Masuda, Y., R.A. Bennett, and B. Demple, *Rapid dissociation of human apurinic endonuclease (Ape1) from incised DNA induced by magnesium.* J Biol Chem, 1998. **273**(46): p. 30360-30365.
219. Gaiddon, C., N.C. Moorthy, and C. Prives, *Ref-1 regulates the transactivation and pro-apoptotic functions of p53 in vivo.* EMBO J, 1999. **18**(20): p. 5609-5621.
220. Prasad, R., et al., *DNA polymerase beta -mediated long patch base excision repair. Poly(ADP-ribose)polymerase-1 stimulates strand displacement DNA synthesis.* J Biol Chem, 2001. **276**(35): p. 32411-32414.
221. Dantzer, F., et al., *Base excision repair is impaired in mammalian cells lacking Poly(ADP-ribose) polymerase-1.* Biochemistry, 2000. **39**(25): p. 7559-7569.
222. Whitehouse, C.J., et al., *XRCC1 stimulates human polynucleotide kinase activity at damaged DNA termini and accelerates DNA single-strand break repair.* Cell, 2001. **104**(1): p. 107-117.

223. Luo, H., et al., *A new XRCC1-containing complex and its role in cellular survival of methyl methanesulfonate treatment.* Mol Cell Biol, 2004. **24**(19): p. 8356-8365.
224. Harrigan, J.A., et al., *The Werner syndrome protein stimulates DNA polymerase beta strand displacement synthesis via its helicase activity.* J Biol Chem, 2003. **278**(25): p. 22686-22695.
225. Schreiber, V., et al., *Poly(ADP-ribose) polymerase-2 (PARP-2) is required for efficient base excision DNA repair in association with PARP-1 and XRCC1.* J Biol Chem, 2002. **277**(25): p. 23028-23036.
226. Hasan, S., et al., *Acetylation regulates the DNA end-trimming activity of DNA polymerase beta.* Mol Cell, 2002. **10**(5): p. 1213-1222.
227. El-Andaloussi, N., et al., *Arginine methylation regulates DNA polymerase beta.* Mol Cell, 2006. **22**(1): p. 51-62.
228. El-Andaloussi, N., et al., *Methylation of DNA polymerase beta by protein arginine methyltransferase 1 regulates its binding to proliferating cell nuclear antigen.* FASEB J, 2007. **21**(1): p. 26-34.
229. Prasad, R., et al., *Specific interaction of DNA polymerase beta and DNA ligase I in a multiprotein base excision repair complex from bovine testis.* J Biol Chem, 1996. **271**(27): p. 16000-16007.
230. Zhou, J., et al., *A role for p53 in base excision repair.* EMBO J, 2001. **20**(4): p. 914-923.
231. Muftuoglu, M., et al., *Telomere repeat binding factor 2 interacts with base excision repair proteins and stimulates DNA synthesis by DNA polymerase beta.* Cancer Res, 2006. **66**(1): p. 113-124.
232. Kedar, P.S., et al., *Direct interaction between mammalian DNA polymerase beta and proliferating cell nuclear antigen.* J Biol Chem, 2002. **277**(34): p. 31115-31123.
233. Naryzhny, S.N. and H. Lee, *The post-translational modifications of proliferating cell nuclear antigen: acetylation, not phosphorylation, plays an important role in the regulation of its function.* J Biol Chem, 2004. **279**(19): p. 20194-20199.
234. Lavrik, O.I., et al., *Photoaffinity labeling of mouse fibroblast enzymes by a base excision repair intermediate. Evidence for the role of poly(ADP-ribose) polymerase-1 in DNA repair.* J Biol Chem, 2001. **276**(27): p. 25541-25548.
235. Levin, D.S., et al., *Interaction between PCNA and DNA ligase I is critical for joining of Okazaki fragments and long-patch base-excision repair.* Curr Biol, 2000. **10**(15): p. 919-922.
236. Tom, S., et al., *DNA ligase I and proliferating cell nuclear antigen form a functional complex.* J Biol Chem, 2001. **276**(27): p. 24817-24825.
237. Caldecott, K.W., et al., *XRCC1 polypeptide interacts with DNA polymerase beta and possibly poly (ADP-ribose) polymerase, and DNA ligase III is a novel molecular 'nick-sensor' in vitro.* Nucleic Acids Res, 1996. **24**(22): p. 4387-4394.
238. Leppard, J.B., et al., *Physical and functional interaction between DNA ligase IIIalpha and poly(ADP-Ribose) polymerase 1 in DNA single-strand break repair.* Mol Cell Biol, 2003. **23**(16): p. 5919-5927.

239. Hasan, S., et al., *Regulation of human flap endonuclease-1 activity by acetylation through the transcriptional coactivator p300.* Mol Cell, 2001. **7**(6): p. 1221-1231.
240. Wu, X., et al., *Processing of branched DNA intermediates by a complex of human FEN-1 and PCNA.* Nucleic Acids Res, 1996. **24**(11): p. 2036-2043.
241. Brosh, R.M., Jr., et al., *Werner syndrome protein interacts with human flap endonuclease 1 and stimulates its cleavage activity.* EMBO J, 2001. **20**(20): p. 5791-5801.
242. Sharma, S., et al., *Stimulation of flap endonuclease-1 by the Bloom's syndrome protein.* J Biol Chem, 2004. **279**(11): p. 9847-9856.
243. Sharma, S., et al., *WRN helicase and FEN-1 form a complex upon replication arrest and together process branchmigrating DNA structures associated with the replication fork.* Mol Biol Cell, 2004. **15**(2): p. 734-750.
244. Frouin, I., et al., *Human proliferating cell nuclear antigen, poly(ADP-ribose) polymerase-1, and p21waf1/cip1. A dynamic exchange of partners.* J Biol Chem, 2003. **278**(41): p. 39265-39268.
245. Tom, S., L.A. Henricksen, and R.A. Bambara, *Mechanism whereby proliferating cell nuclear antigen stimulates flap endonuclease 1.* J Biol Chem, 2000. **275**(14): p. 10498-10505.
246. Gary, R., et al., *The DNA repair endonuclease XPG binds to proliferating cell nuclear antigen (PCNA) and shares sequence elements with the PCNA-binding regions of FEN-1 and cyclin-dependent kinase inhibitor p21.* J Biol Chem, 1997. **272**(39): p. 24522-24529.
247. Hasan, S., et al., *Transcription coactivator p300 binds PCNA and may have a role in DNA repair synthesis.* Nature, 2001. **410**(6826): p. 387-391.
248. Hong, R. and D. Chakravarti, *The human proliferating Cell nuclear antigen regulates transcriptional coactivator p300 activity and promotes transcriptional repression.* J Biol Chem, 2003. **278**(45): p. 44505-44513.
249. Vairapandi, M., et al., *Human DNA-demethylating activity: a glycosylase associated with RNA and PCNA.* J Cell Biochem, 2000. **79**(2): p. 249-260.
250. Boldogh, I., et al., *hMYH cell cycle-dependent expression, subcellular localization and association with replication foci: evidence suggesting replication-coupled repair of adenine:8-oxoguanine mispairs.* Nucleic Acids Res, 2001. **29**(13): p. 2802-2809.
251. Dotto, G.P., *p21(WAF1/Cip1): more than a break to the cell cycle?* Biochim Biophys Acta, 2000. **1471**(1): p. M43-56.
252. Plo, I., et al., *Association of XRCC1 and tyrosyl DNA phosphodiesterase (Tdp1) for the repair of topoisomerase I-mediated DNA lesions.* DNA Repair (Amst), 2003. **2**(10): p. 1087-1100.
253. Clements, P.M., et al., *The ataxia-oculomotor apraxia 1 gene product has a role distinct from ATM and interacts with the DNA strand break repair proteins XRCC1 and XRCC4.* DNA Repair (Amst), 2004. **3**(11): p. 1493-1502.
254. Date, H., et al., *The FHA domain of aprataxin interacts with the C-terminal region of XRCC1.* Biochem Biophys Res Commun, 2004. **325**(4): p. 1279-1285.

255. Gueven, N., et al., *Aprataxin, a novel protein that protects against genotoxic stress.* Hum Mol Genet, 2004. **13**(10): p. 1081-1093.
256. Sano, Y., et al., *Aprataxin, the causative protein for EAOH is a nuclear protein with a potential role as a DNA repair protein.* Ann Neurol, 2004. **55**(2): p. 241-249.
257. Masson, M., et al., *XRCC1 is specifically associated with poly(ADP-ribose) polymerase and negatively regulates its activity following DNA damage.* Mol Cell Biol, 1998. **18**(6): p. 3563-3571.
258. von Kobbe, C., et al., *Central role for the Werner syndrome protein/poly(ADP-ribose) polymerase 1 complex in the poly(ADP-ribosyl)ation pathway after DNA damage.* Mol Cell Biol, 2003. **23**(23): p. 8601-8613.
259. Hassa, P.O., et al., *Transcriptional coactivation of nuclear factor-kappaB-dependent gene expression by p300 is regulated by poly(ADP)-ribose polymerase-1.* J Biol Chem, 2003. **278**(46): p. 45145-45153.
260. Flohr, C., et al., *Poly(ADP-ribosyl)ation accelerates DNA repair in a pathway dependent on Cockayne syndrome B protein.* Nucleic Acids Res, 2003. **31**(18): p. 5332-5337.
261. Galande, S. and T. Kohwi-Shigematsu, *Poly(ADP-ribose) polymerase and Ku autoantigen form a complex and synergistically bind to matrix attachment sequences.* J Biol Chem, 1999. **274**(29): p. 20521-20528.
262. Xia, L., et al., *Human 3-methyladenine-DNA glycosylase: effect of sequence context on excision, association with PCNA, and stimulation by AP endonuclease.* J Mol Biol, 2005. **346**(5): p. 1259-1274.
263. Almeida, K.H. and R.W. Sobol, *A unified view of base excision repair: lesion-dependent protein complexes regulated by post-translational modification.* DNA Repair (Amst), 2007. **6**(6): p. 695-711.
264. Marenstein, D.R., et al., *Substrate specificity of human endonuclease III (hNTH1). Effect of human APE1 on hNTH1 activity.* J Biol Chem, 2003. **278**(11): p. 9005-9012.
265. Pope, M.A., N.H. Chmiel, and S.S. David, *Insight into the functional consequences of hMYH variants associated with colorectal cancer: distinct differences in the adenine glycosylase activity and the response to AP endonucleases of Y150C and G365D murine MYH.* DNA Repair (Amst), 2005. **4**(3): p. 315-325.
266. Sidorenko, V.S., G.A. Nevinsky, and D.O. Zharkov, *Mechanism of interaction between human 8-oxoguanine-DNA glycosylase and AP endonuclease.* DNA Repair (Amst), 2007. **6**(3): p. 317-328.
267. Li, X., P.M. Wright, and A.L. Lu, *The C-terminal domain of MutY glycosylase determines the 7,8-dihydro-8-oxo-guanine specificity and is crucial for mutation avoidance.* J Biol Chem, 2000. **275**(12): p. 8448-8455.
268. Pettersen, H.S., et al., *Uracil-DNA glycosylases SMUG1 and UNG2 coordinate the initial steps of base excision repair by distinct mechanisms.* Nucleic Acids Res, 2007. **35**(12): p. 3879-3892.
269. Kavli, B., et al., *hUNG2 is the major repair enzyme for removal of uracil from U:A matches, U:G mismatches, and U in single-stranded DNA, with hSMUG1 as a broad*

specificity backup. J Biol Chem, 2002. **277**(42): p. 39926-39936.
270. Mohan, R.D., et al., *Opposing regulatory roles of phosphorylation and acetylation in DNA mispair processing by thymine DNA glycosylase.* Nucleic Acids Res, 2010. **38**(4): p. 1135-1148.
271. Lirussi, L., et al., *Nucleolar accumulation of APE1 depends on charged lysine residues that undergo acetylation upon genotoxic stress and modulate its BER activity in cells.* Mol Biol Cell, 2012. **23**(20): p. 4079-4096.
272. Fantini, D., et al., *Critical lysine residues within the overlooked N-terminal domain of human APE1 regulate its biological functions.* Nucleic Acids Res, 2010. **38**(22): p. 8239-8256.
273. Hardeland, U., et al., *Modification of the human thymine-DNA glycosylase by ubiquitin-like proteins facilitates enzymatic turnover.* EMBO J, 2002. **21**(6): p. 1456-1464.
274. Johnson, K.A., *Transient State Kinetic Analysis of Enzyme Reaction Pathways.* The Enzymes, 1992. **20**: p. 61.
275. Privezentzev, C.V., M. Saparbaev, and J. Laval, *The HAP1 protein stimulates the turnover of human mismatch-specific thymine-DNA-glycosylase to process 3,N(4)-ethenocytosine residues.* Mutat Res, 2001. **480-481**: p. 277-284.
276. Henikoff, S., *Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing.* Gene, 1984. **28**(3): p. 351-359.
277. Halford, S.E. and J.F. Marko, *How do site-specific DNA-binding proteins find their targets?* Nucleic Acids Res, 2004. **32**(10): p. 3040-3052.
278. Halford, S.E., *An end to 40 years of mistakes in DNA-protein association kinetics?* Biochem Soc Trans, 2009. **37**(Pt 2): p. 343-348.
279. Pray, T.R., D.S. Burz, and G.K. Ackers, *Cooperative non-specific DNA binding by octamerizing lambda cI repressors: a site-specific thermodynamic analysis.* J Mol Biol, 1998. **282**(5): p. 947-958.
280. Weinberg, R.L., D.B. Veprintsev, and A.R. Fersht, *Cooperative binding of tetrameric p53 to DNA.* J Mol Biol, 2004. **341**(5): p. 1145-1159.
281. Wong, I., et al., *A dimeric mechanism for contextual target recognition by MutY glycosylase.* J Biol Chem, 2003. **278**(4): p. 2411-2418.
282. Lu, A.L., *Repair of A/G and A/8-oxoG mismatches by MutY adenine DNA glycosylase.* Methods Mol Biol, 2000. **152**: p. 3-16.
283. Bartsch, H. and J. Nair, *New DNA-based biomarkers for oxidative stress and cancer chemoprevention studies.* Eur J Cancer, 2000. **36**(10): p. 1229-1234.
284. Lutsenko, E. and A.S. Bhagwat, *The role of the Escherichia coli mug protein in the removal of uracil and 3,N(4)-ethenocytosine from DNA.* J Biol Chem, 1999. **274**(43): p. 31034-31038.
285. Lieb, M. and A.S. Bhagwat, *Very short patch repair: reducing the cost of cytosine methylation.* Mol Microbiol, 1996. **20**(3): p. 467-473.
286. Bhagwat, A.S. and M. Lieb, *Cooperation and competition in mismatch repair: very short-patch repair and methyl-directed mismatch repair in Escherichia coli.* Mol

Microbiol, 2002. **44**(6): p. 1421-1428.
287. Mokkapati, S.K., A.R. Fernandez de Henestrosa, and A.S. Bhagwat, *Escherichia coli DNA glycosylase Mug: a growth-regulated enzyme required for mutation avoidance in stationary-phase cells.* Mol Microbiol, 2001. **41**(5): p. 1101-1111.
288. Almiron, M., et al., *A novel DNA-binding protein with regulatory and protective roles in starved Escherichia coli.* Genes Dev, 1992. **6**(12B): p. 2646-2654.
289. Loewen, P.C., J. Switala, and B.L. Triggs-Raine, *Catalases HPI and HPII in Escherichia coli are induced independently.* Arch Biochem Biophys, 1985. **243**(1): p. 144-149.
290. Bridges, B.A., *The role of DNA damage in stationary phase ('adaptive') mutation.* Mutat Res, 1998. **408**(1): p. 1-9.
291. Eisenstark, A., et al., *Role of Escherichia coli rpoS and associated genes in defense against oxidative damage.* Free Radic Biol Med, 1996. **21**(7): p. 975-993.
292. Saint-Ruf, C., et al., *Causes and consequences of DNA repair activity modulation during stationary phase in Escherichia coli.* Crit Rev Biochem Mol Biol, 2007. **42**(4): p. 259-270.
293. Sak, B.D., A. Eisenstark, and D. Touati, *Exonuclease III and the catalase hydroperoxidase II in Escherichia coli are both regulated by the katF gene product.* Proc Natl Acad Sci U S A, 1989. **86**(9): p. 3271-3275.
294. Withers, H.L. and K. Nordstrom, *Quorum-sensing acts at initiation of chromosomal replication in Escherichia coli.* Proc Natl Acad Sci U S A, 1998. **95**(26): p. 15694-15699.
295. Tanaka, K., et al., *Heterogeneity of the principal sigma factor in Escherichia coli: the rpoS gene product, sigma 38, is a second principal sigma factor of RNA polymerase in stationary-phase Escherichia coli.* Proc Natl Acad Sci U S A, 1993. **90**(8): p. 3511-3515.
296. Chung, J.D. and G. Stephanopoulos, *Studies of transcriptional state heterogeneity in sporulating cultures of bacillus subtilis.* Biotechnol Bioeng, 1995. **47**(2): p. 234-242.
297. Bellamy, S.R., *A kinetics analysis of substrate recognition by uracil DNA glycosylase from herpes simplex virus type-1.* Department of Biochemistry, 2002. **Bristol**: p. 223.