5-2007

# BIOCHEMICAL STUDY OF ENDONUCLEASE V AND ITS APPLICATION IN MUTATION SCANNING

Honghai Gao
*Clemson University*, honghag@clemson.edu

# BIOCHEMICAL STUDY OF ENDONUCLEASE V AND ITS APPLICATION IN MUTATION SCANNING

---

A Dissertation
Presented to
the Graduate School of
Clemson University

---

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
Biochemistry

---

by
Honghai Gao
May 2007

---

Accepted by:
Weiguo Cao, Committee Chair
William Marcotte
Brandon Moore
Gary Powell

## ABSTRACT

The integrity of the genetic information encoded by DNA is essential to all living organisms, yet the reactive bases of DNA are constantly attacked by endogenous and exogenous agents resulting in as many as one million individual molecular lesions per cell per day. Excessive DNA damage or deficiency in DNA repair enzymes may cause cancer, premature aging, and neurodegenerative diseases.

Endonuclease V (Endo V) is a DNA repair enzyme which can recognize all four types of DNA deamination products, specifically, uracil, hypoxanthine, xanthine and oxanine. It was also shown that endo V can recognize mismatches. We screened about 60 mutants of endo V from *Thermotoga maritima* and found some mutants had altered base preferences for mismatches. *Tma* endo V Y80A was shown to become a C-specific mismatch endonuclease. G13D mutation in K-*ras* oncogene which was not recognized by wild type *Tma* endo V was successfully cleaved by *Tma* endo V Y80A. This study provides valuable information on base recognition and active site organization of *Tma* endo V. *Tma* endo V mutants can be used for cancer mutation scanning and mutation recognition.

In order to further understand the role of Y80 of endo V in base recognition, we substituted the Y80 with sixteen amino acids. Together with three Y80 mutants isolated before, we characterized all nineteen mutants of *Tma* endo V Y80 using deaminated base-containing DNA substrates and mismatch-containing DNA substrates. This comprehensive amino acid substitution at a single site (Y80) underlines the importance of aromatic ring and hydrogen bond donor capacity in base recognition by endo V, reveals

additional Y80 mutants with altered base preferences in mismatch cleavage, and offers new insight on the role of Y80 in base recognition.

Though endo V was shown to be important for repair of deaminated lesions *in vivo*, its DNA repair pathway remains unknown. In order to understand the DNA repair pathway mediated by endo V, we have developed a cell-free system from *Escherichia coli*. The preliminary results indicated that the repair patch of endo V mediated DNA repair pathways may consist of a long patch and a short patch repair pathway.

## DEDICATION

This work is dedicated to my parents.

## ACKNOWLEDGMENTS

There are lots of people I would like to thank for a huge variety of reasons.

First, I would like to thank my advisor, Dr. Weiguo Cao who have been supporting and advising me during my study and research at Clemson University. His time is shared between many projects and teaching load but still he has managed to guide my work at the right points with his many good ideas. I am deeply grateful for his guidance and encouragement.

I would also like to thank the members of my graduate committee (Dr. Gary Powell, Dr. William Marcotte, and Dr. Brandon Moore) for their advice and help during my Ph.D. training at Clemson University. I am also indebted to Dr. Richard Hilderman, Dr. Gary Powell, and Dr. Weiguo Cao for their excellent teaching. I want to say thanks to all the Cao lab current and previous members (Dr. Hong Feng, Dr. Thomas Hitchcock, Rongjuan Mi, Liang Dong, Hyun-Wook Lee, and Sung-Hyun Park) and our collaborators (Dr. Francis Barany, Dr. Jianmin Huang, Dr. Pu-Cun Ke and Jun Lin) for their help.

Finally, I would like to thank my parents and my wife for their support and encouragement when I needed them most.

## TABLE OF CONTENTS

**Table of Contents (Continued)**

**Table of Contents (Continued)**

**Table of Contents (Continued)**

## LIST OF FIGURES

**List of Figures (Continued)**

## ABBREVIATIONS

Alkyladenine DNA glycosylase gene, AAG

Activation-Induced cytidine Deaminase, AID

BER, Base Excision Repair

Class-Switch Recombination, CSR

DNA, deoxyribonucleic acid

dATP, deoxyadenosine triphosphate

dCTP, deoxycytidine triphosphate

dGTP, deoxyguanosine triphosphate

dITP,  deoinosine triphosphate

dOTP, deoxyoxanosine triphosphate

dTTP, deoxythymidine triphosphate

dUTP, deoxyuridine triphosphate

dXTP, deoxyxanthosine triphosphate

Endonuclease V, Endo V

Formamidopyrimidine-DNA glycosylase, Fpg

3-methyladenine-DNA glycosylase II, AlkA

mCpG binding domain protein 4, MBD4

Mismatch-specific Uracil-DNA glycosylase, MUG

Oxanine, Oxa

Polymerase, Pol

Reactive nitrogen species, RNS

**Abbreviations (Continued)**

Reactive oxygen species, ROS

Single strand-selective Monofunctional Uracil-DNA Glycosylase, SMUG1

Somatic hypermutation, SHM

Thymine DNA N-glycosylase, TDG

Uracil DNA Glycosylase, UDG

Xanthine, Xan

# CHAPTER ONE

## Repair of Deaminated Lesions and Applications of Endonuclease V

### I. Introduction

DNA is constantly being attacked by exogenous and endogenous agents. Over one million damage incidences occur in DNA per cell per day (1), mostly by oxidation, ionizing radiation, UV radiation, alkylation, hydrolysis, thermal disruption, and deamination. In the following reviews, I will address the DNA damages caused by deamination and the enzymes involved in deamination DNA repair; in particular, endonuclease V, a DNA repair enzyme capable of repairing all four deaminated DNA lesions will be discussed in detail. Last, I will cover DNA damage recognition mechanisms.

### II. Deaminated lesions and related enzymes

Guanine has two major deamination products: xanthine (Xan) and oxanine (Oxa), while the deamination of adenine, cytosine, and 5-methylcytosine leads to hypoxanthine, uracil, and thymine respectively. In this review, I will focus on nitrosative deamination of DNA and related repair enzymes and pathways.

## A. How does deamination occur?

### 1. Spontaneous hydrolytic deamination

There are many sources of deamination in DNA. The most common one is spontaneous hydrolytic deamination. Hydrolytic deamination of cytosine has been estimated to introduce between 100 and 500 uracil residues in the form of U:G mismatches per cell per day (2). Deamination of adenine to hypoxanthine is about 10-fold less frequent (2, 3). In the spontaneous hydrolytic deamination, protonated cytosine is thought to undergo direct attack by hydroxyl group from water at the fourth position of the pyrimidine, this may be the main reaction under physiological condition (Figure 1.1) (4). Another pathway involves formation of dihydrocytosine (5, 6).

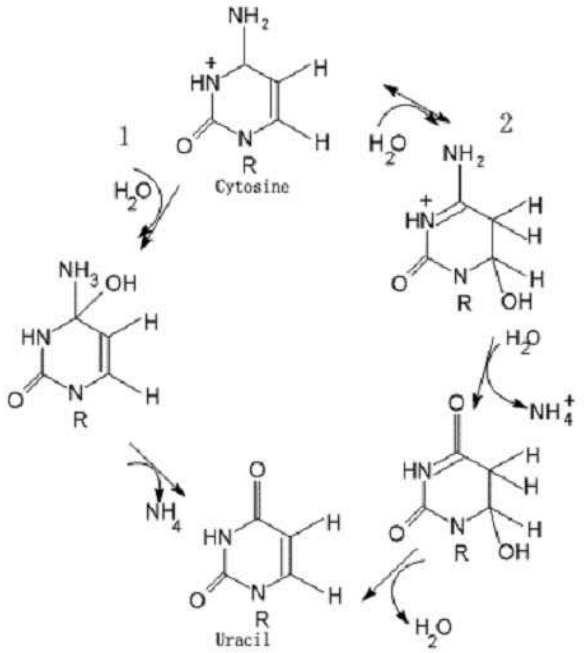



**Figure 1.1 Pathways of cytosine hydrolytic deamination.** 1. The first pathway includes direct attack of the fourth position of the pyrimidine by hydroxylion. This is the main mode of cytosine deamination under physiological conditions. 2. The second pathway occurs via dihydrocytosine formation as an intermediate. Taken from (4).

Spontaneous hydrolytic deamination is enhanced by heat, acidic pH and basic pH. The rate constant of the cytosine deamination in single strand DNA is $2 \cdot 10^{-10}$ $sec^{-1}$ at 37°C. In double-stranded DNA, the rate of cytosine deamination *in vitro* is about 3000-fold less (7). Pyrimidine bases are more susceptible to spontaneous deamination than are the purine bases (5, 7). Purine bases, however, are more easily deaminated by nitrous acid (8).

**2. Nitrosative deamination**

Reactive oxygen species (ROS) and reactive nitrogen species (RNS) are continuously produced in cells as by-products of aerobic metabolism or in response to stresses. ROS/RNS are known to play a dual role in biological systems, since they can be either harmful or beneficial to living organisms. ROS/RNS are integral parts of the host defense and are released by activated white blood cells in response to bacterial and viral infection in human (9). ROS/RNS also act as messengers in signaling pathways (10, 11). ROS includes several oxygen radicals, superoxide ($O_2^{\cdot -}$) and its protonated form, hydroperoxyl ($HO_2^{\cdot}$), hydroxyl ($OH^{\cdot}$), peroxyl ($RO_2^{\cdot}$), alkoxyl ($RO^{\cdot}$). ROS also includes some nonradicals such as hydrogen peroxide ($H_2O_2$), hypochlorous acid (HOCl), and ozone ($O_3$), which are oxidizing agents and are easily converted into radicals. DNA oxidative damage can result in base or sugar adducts, single and double strand breaks, as well as cross-links to other molecules. Most of these DNA modifications are mutagenic, and thought to be related to cancer, aging and neurodegenerative diseases (12-14). Oxidative damage to pyrimidine and purine bases has been well characterized. The most common pyrimidine lesion, thymine glycol can block replication and transcription (15).

The most common and extensively studied purine modification is 8-oxodeoxyguanine (8-oxodG). It is formed by the addition of $HO^{\cdot}$ to C-8 of guanine. It can cause G:C to A:T transversion mutation (16). RNS includes nitric oxide ($NO^{\cdot}$), nitrate ($NO_3^-$), nitrite ($NO_2^-$), peroxynitrite (ONOO–), nitrogen dioxide, and peroxynitrous acid (ONOOH). They can be converted to more active $N_2O_3$ which can induce nitrosative deamination of DNA (Figure 1.2) (10, 17).

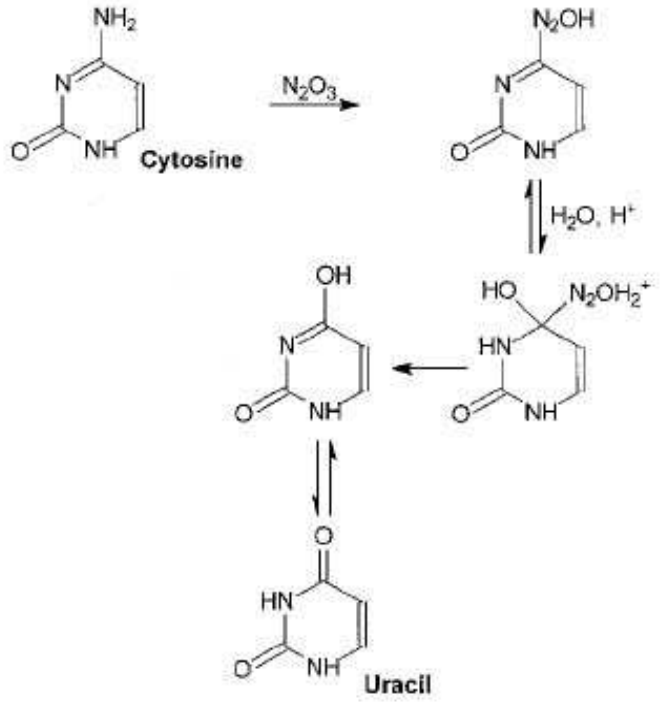



**Figure 1.2 Nitrosative deamination of cytosine.** Taken from (17).

Increasing evidence shows that chronic inflammation is a risk factor for the development of a variety of human cancers (18-20). Although the mechanisms of carcinogenesis associated with inflammation are not fully understood, it is suggested that DNA damage by reactive oxygen species and reactive nitrogen species secreted by

activated macrophages and neutrophils is one of the mechanisms. Nitric oxide (NO) can be generated endogenously through the oxidation of L-arginine by nitric oxide synthases (21). NO exists in various cells after infection, chemical stimulation, and inflammation. NO is produced at the surface of the endothelial cells following the addition of bradykinin and in activated macrophages (22). Activated macrophages generate NO at a rate of ~6 pmol $s^{-1}$ per $10^6$ cells (23). NO also can be formed by the combustion of fossil fuels and smoking of cigarettes (24). NO concentrations in the atmosphere of large cities range from 0.040 ppm to 0.077 ppm due to emissions from motor vehicle (24, 25). Cigarettes smoke contains 80–110 ppm NO (26). Approximately 80% of NO was absorbed during human breathing (27). The NO reacts quickly with $O_2$ to generate nitrous anhydride ($N_2O_3$), a potent nitrosating agent that causes deamination of aromatic amines by formation of an aryl diazonium ion (28). G:C to A:T mutations at CpG sites containing 5'-methylcytosine may result from the deamination of 5-methylcytosine to thymine (29). The deamination of cytosine to uracil causes G:C to A:T mutations by pairing of uracil with adenine (30). Deamination of guanine to xanthine causes G:C to A:T mutations by pairing of xanthine with thymine (17). The deamination of adenine to hypoxanthine causes A:T to G:C mutations by pairing of hypoxanthine with cytosine (17).

**3. Deaminase caused deamination**

In 1999, Muramatsu and colleagues found a novel RNA-editing deaminase named activation-induced cytidine deaminase (AID) (31). Later, it was found that AID can catalyze deamination of dC residues on single-stranded DNA *in vitro* (32). The dC to dU

deamination activity was most avid on double-stranded DNA substrates containing a small "transcription-like" single-stranded DNA bubble (32).

When human AID is expressed in yeast, which does not have a DNA deaminase, it causes increased C:G to U:G mutation (33). Recently, researchers showed that AID is necessary in mouse and human for the somatic hypermutation (SHM) and class-switch recombination (CSR) of immunoglobulins (Igs) (34, 35). Patients with AID deficiencies suffer recurrent infections due to autoimmune disorders (34, 36). AID is not only required for a healthy immune system to fight off pathological invaders but also necessary for bioenergetics (37, 38, 39). Though we need deaminases to help immune system to fight pathogens, too many of them may increase the chance of cancer.

I will discuss the deaminated lesions and related enzymes in the following sections.

**B. Occurrence of uracil in DNA and related repair enzymes**

**1. dUTP formation and regulation in DNA**

dUTP (2´-deoxyuridine 5´-triphosphate) is one of the nucleotides in living organisms. It takes part in several chemical reactions in cells. dUTP can be formed from the reaction of a DNP kinase on dUDP (40). It can be generated from UTP by a ribonucleotide reductase in *E. coli* (41). dUTP also can be produced by dCTP deaminase through deamination of dCTP (42). dUTP is consumed in the dUTPase reaction where the dUMP is utilized as precursor for the synthesis of thymine nucleotides by the enzyme thymidylate synthase (43). Anticancer agents, such as 5-fluorouracil (5-FU), fluorodeoxyuridine (FUdR), and xeloda (capecitabine) can inhibit the thymidylate

synthase activity, leading to dUTP accumulation, uracil misincorporation into DNA, and uracil-DNA glycosylase-induced strand breaks, ultimately result in cell death (44). Uracil in DNA results from either deamination of cytosine or misincorporation of dUTP, giving rise to mutagenic U:G mispairs and less harmful U:A respectively. U:A pairs, though not miscoding, may yield cytotoxic and potentially mutagenic abasic sites (45). The concentration of dUTP *in vivo* is regulated by DNP kinase, ribonucleotide reductase, dCTP deaminase, and dUTPase (Figure 1.3) (46).

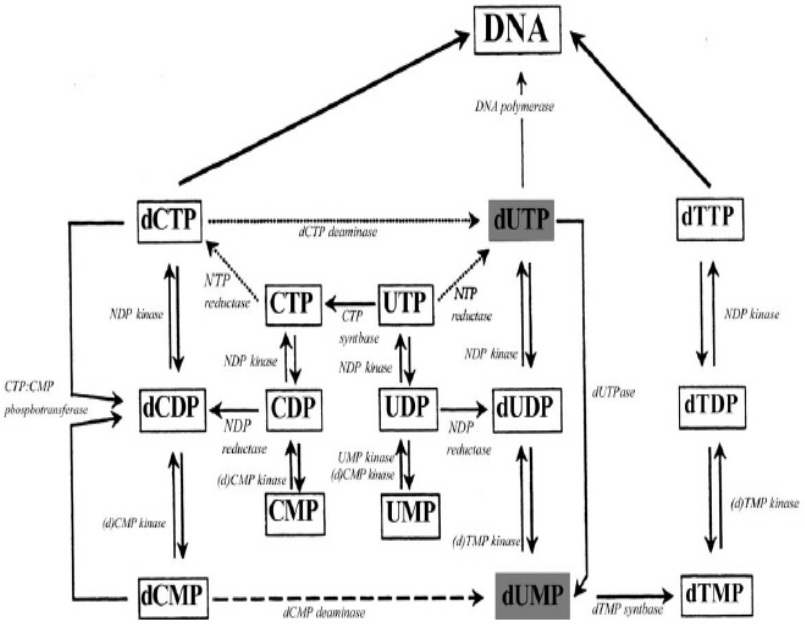



**Figure 1.3 The *de novo* pathway of pyrimidine deoxynucleotide synthesis.** Individual enzymes are identified. Dotted arrows indicate enzymic activities, so far only found in *E. coli* and other enterobacteria. Dashed arrow indicates enzymic activities not found in *E. coli* but observed in a majority of other organisms. Taken from (46).

Deoxyuridine triphosphate nucleotidohydrolase (dUTPase) is the key regulator of cellular dUTP pool. dUTPase acts in hydrolyzing dUTP to dUMP and inorganic pyrophosphate. This reaction not only provides the precursor for de novo dTMP synthesis

but also decreases intracellular dUTP level and reduces the probability that dUTP will be incorporated into DNA by DNA polymerases during replication and repair processes (47). dUTPase can be classified into two groups, one is extremely specific to dUTP while the other has broader enzyme specificity. The dUTPases that work exclusively on dUTP are almost ubiquitous in nature. dUTPases from animals, higher plants, fungi, bacteria and viruses (DNA and RNA viruses) have been characterized, which consist of three identical subunit polypeptides, each of approximately 150 amino acid residues (47-50). The other group of dUTPases which is less well known consists of the enzymes from protozoa like *Leishmania* and *Trypanosoma* (51, 52). The members of this group differ in sequence from the dUTPase in the first group. The active form of dUTPase of this group is homodimer and the enzymes have broader specificity. dUTPase from *Leishmani* and *Trypanosoma* can hydrolyze both dUTP and dUDP (53, 54). The activity of dUTPase is not sufficient to protect against the presence of uracil in the genome. A small, steady-state concentration of dUTP is present in cells at all times that can be incorporated into DNA by polymerases. In addition to this, cytosine deamination also cause considerable amount of the uracil in the genome. Cells need a mechanism to remove the uracil from the DNA.

**2. Enzymes that recognize uracil in DNA**

Uracil lesion in DNA is mainly repaired by uracil glycosylases through Base Excision Repair pathway (55).

**a. Uracil-DNA glycosylase (UDG)**

In *E. coli,* Uracil-DNA glycosylase (UDG) and mismatch-specific uracil-DNA glycosylase (MUG) can recognize uracil residues in DNA and cleave *N*-glycosidic bond between uracil and deoxyribose sugar. In 1974, Lindahl first discovered the *E. coli* uracil DNA glycosylase which is the first glycosylase discovered (56). The *E. coli* UDG gene (*ung*) was also the first uracil DNA glycosylase cloned (57) and sequenced (58). UDG is inhibited by free uracil and some of its derivatives (59, 60). Interestingly, a family of bacteriophages including PBS1 and PBS2 which infect the bacteria *Bacillus subtilis* incorporate uracils instead of thymines into their genomes (61, 62). PBS1 and PBS2 have a UDG inhibitor protein (Ugi) to avoid these futile repair cycles by the host UDG (63, 64). *E. coli* UDG removes uracil both from single-stranded and duplex DNA (65). Researchers have shown uracil-DNA glycosylase binds, kinks, and compresses the duplex DNA backbone while scanning the minor groove for a uracil residue (66). Once bound with the uracil, the enzyme uses a "pinch-push-pull" mechanism to extract the uracil nucleotide from the DNA base stack and position it into the active site (66-68).

**b. Mismatch-specific uracil-DNA glycosylase (MUG)**

*E. coli* mismatch-specific uracil-DNA glycosylase (MUG) was first discovered by Gallinari and Jiricny (69). It shares homology with the mammalian thymine-DNA glycosylase and can excise uracil from U:G pairs. Thus, it was thought to be a backup enzyme for uracil DNA glycosylase. MUG was found to excise 3, N4-ethenocytosine (εC) from εC:G pairs and had much better activity for εC from εC:G pairs than uracil from U:G pairs (70). Some evidence showed that neither εC nor U may be the

biologically relevant substrate for MUG. Ethenocytosine is not detected in *E. coli* and the reversion assay showed no increase in mutations when mug was inactive (71, 72). Mug was later shown to have wide substrate specificity: it can cleave T (73), 5-hydroxymethyluracil (74), 1,$N^2$-ethenoguanine (75), and 8-(hydroxymethyl)-3,$N^4$-ethenocytosine (76).

In human, it was reported that human uracil DNA glycosylase 1 (UNG1), human uracil DNA glycosylase 2 (UNG2), single strand-selective monofunctional uracil-DNA glycosylase (SMUG1), thymine DNA N-glycosylase (TDG), and mCpG binding domain protein 4 (MBD4) can recognize the uracil in DNA.

**c. Uracil DNA glycosylase 1 and 2 (UNG1, UNG2)**

In 1976, a human uracil-DNA glycosylase activity was first reported by Sekiguchi and colleagues (77). Olsen and colleagues cloned human UNG cDNA based on the amino acid sequence of N-terminal end of human placental UNG (78, 79). The human UNG gene is located at 12q24.1 and spans approximately 13.8 kb (80). The human UNG-gene encodes the mitochondrial form UNG1 and nuclear form UNG2 using differentially regulated promoters (PA and PB) and alternative splicing. UNG1 and UNG2 have a common catalytic domain, but different N-terminal sequences (81). The N-terminal end of UNG1 has a classical and very strong mitochondrial targeting signal (MTS) which is not found in UNG2. Most of the MTS will be cleaved by mitochondrial processing peptidase once entering the mitochondria (82). About 100 amino acids of the N-terminal of UNG2 are required for complete nuclear targeting, though the unique N-terminal 44 amino acids have the most important information in the nuclear localization signal

(NLS). The N-terminal part of UNG2 contains a motif for binding with proliferating cell nuclear antigen (PCNA) (83). It also contains a motif which has been demonstrated to bind replication protein A (RPA) (84). The interactions of UNG2 with PCNA and RPA take place in replication foci (83).

UNG2 plays an important role in removing misincorpated uracil (U:A pairs) in DNA. It is supported by the fact that $Ung^{-/-}$ mice have very slow rate of removal of incorporated uracil in nuclei (85) and anti-UNG2 antibodies can inhibit immediate post-replicative removal of incorporated uracil in isolated nuclei (86). UNG2 has a turn over number of 600-1000 per min which is important for removal of misincorporated uracil close to the fast moving replication fork. UNG2 is also found in neucleoplasm and may repair U:G mispairs generated from cytosine deamination (87). UNG2 has greater affinity for the abasic site than for the uracil in DNA and remains bound to the abasic site after the removal of uracil (88). The abasic site is next processed by apurinic/apyrimidinic endonuclease (APE1) which nicks at 5′ side of the abasic site, resulting in a 5′-deoxyribose phosphate group that is a substrate for subsequent repair by repair enzymes.

**d. UNG2 in immunoglobulin diversity**

B lymphocytes are specialized in antibody production. Functional antibody-encoding immunoglobulin variable (V) genes are assembled from non-functional V, diversity (D) and joining (J) gene segments only in B cells (89). After antigen encounter, these cells can then further modify their V genes by somatic hypermutation (SHM), generating antibodies with improved affinity. In addition, the constant region, Cμ in virgin B cells, can be changed to Cγ, Cα, or Cε by class switch recombination (CSR)

after activation of B cells. Both AID and UNG are essential for somatic hypermutation (Figure 1.4) and class-switch recombination (Figure 1.5) (35). Compared to nuclear form UNG2, mitochondrial form UNG1 is less studied. The role of UNG1 in mitochondria is still unclear. Recent research shows that inhibition of UNG1 in mitochondria by uracil glycosylase inhibitor (UGI) did not lead to either spontaneous or induced mutations in mtDNA suggesting that backup mechanism(s) may exist in the mitochondria (91).

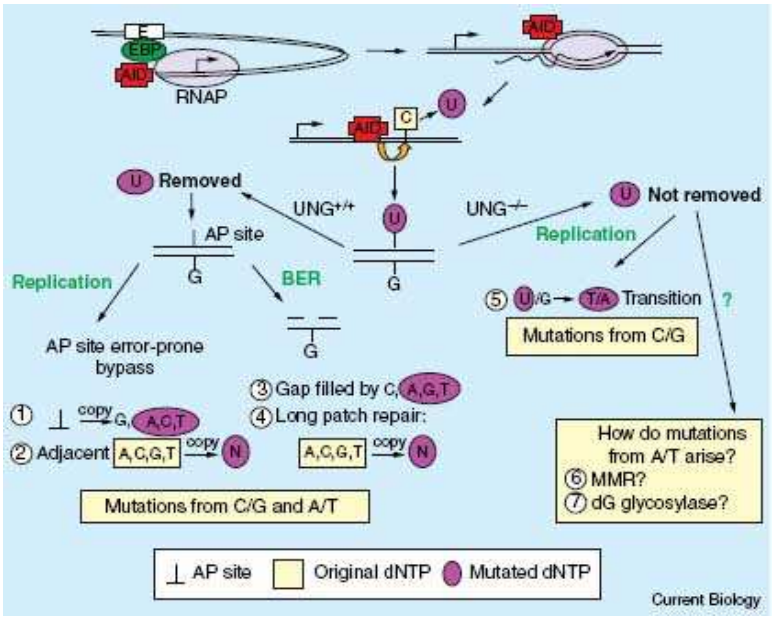



**Figure 1.4 Model for somatic hypermutation.** The dU bases resulting from AID-mediated dC deamination can be removed by uracil DNA glycosylases. If removed, mutations from G/C or A/T can arise via replication or base excision repair. If U is not removed the U template can be copied into A by replication, leading to a C/G to T/A transition in one daughter cell. The other daughter cell will have the unmutated C/G. AID, activation-induced deaminase; E, Ig enhancer; EBP, enhancer-binding proteins; RNAP, RNA polymerase II; bent arrow, transcription start site; Taken from (90)

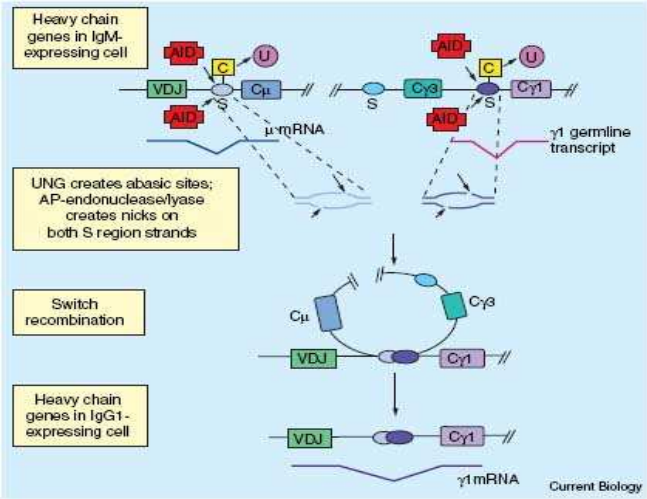



**Figure 1.5 Model for class switch recombination.** The top line shows the mouse Ig heavy chain genes in B cells expressing IgM. Small arrows indicate that AID deaminates dC within Sμ and a downstream S region (Sγ1 here) to initiate CSR. UNG excises U, AP-endonuclease/lyase activities create single-strand nicks on top and bottom strands of Sμ and Sγ1. Also shown are germline transcripts which must be transcribed in *cis* from the unrearranged S–CH segment in order to obtain switching. Recombination occurs by an intrachromosomal deletional end-joining process. The bottom line shows the heavy chain chromosome after CSR to IgG1. Taken from (90).

**e. Single strand-selective Monofunctional Uracil-DNA Glycosylase (SMUG1)**

hSMUG1 was first discovered in 1999 (92). It has wide spectrum of substrates. It removes uracil, 5-hydroxymethyluracil (93) and 5-formyluracil from single- and double-stranded DNA (94). hSMUG1 is not indicated to have a role in removal of incorporated uracil and it does not accumulate in replication foci. It is proposed that SMUG1 evolved as a necessary and separate mechanism on premutagenic U:G lesions resulting from genome-wide hydrolytic deamination of cytosine (95). Overexpression of SMUG1 does not restore detectable *in vitro* class switching in UNG-deficient mouse B lymphocytes suggesting that SMUG1 is not involved in the diversification of antibody genes (96). In mammalian cells, SMUG1 is only expressed at low levels and is not detected in cell-free

extracts by western blotting with SMUG1 antibodies (97). UNG1, UNG2, and SMUG1 are the only known DNA glycosylases with preference for single-stranded DNA.

**f. mCpG binding domain protein 4 (MBD4)**

Human MBD4 was first discovered in 1998 as a protein which binds specifically to methylated DNA *in vitro* (98). It has 580 amino acids which contains an N-terminal methyl-binding domain, MBD (residues 82-147) and a C-terminal glycosylase domain (residues 401-580) with homology to *E. coli* thymine glycol glycosylase (Endo III) and 8-oxoG:A specific adenine glycosylase (MutY). MBD4 recognizes and removes thymine and uracil from a G:T and G:U mispair respectively at unmethylated CpG sequences (99, 100). MBD4 also removes the uracil analog, 5-fluorouracil from G:5-fluorouracil mispair (101). It has been shown to interact with MLH1, a protein involved in mismatch repair, by using a yeast two-hybrid system (102). Further experiments demonstrated that loss of MBD4 function causes several MMR proteins to be down-regulated (103). It mediates the apoptotic response to DNA damage in cells. MBD4 deficiency reduced the normal apoptotic response to a range of cytotoxic agents including gamma-irradiation, cisplatin, temozolomide and 5-fluorouracil (5-FU) (104). MBD4 is frequently mutated in a large number of MMR-deficient tumors that exhibit microsatellite instability (MSI). However, MBD4-null mice are viable and fertile, and have no apparent increase in tumor susceptibility (105).

**g. Thymine DNA N-glycosylase (TDG)**

Human thymine DNA glycosylase (hTDG) was first purified from Hela cells by Neddermann and colleagues (106). It was cloned in 1996 and showed a molecular weight of 46 kDa (107). hTDG was found to have activity on G:U and G:T mispairs while it removes U from U:G mispairs with 10-fold higher $k_{cat}$ than T from G:T mispairs. hTDG can cleave 5-fluorouracil (5-FU) in both double-stranded and single-stranded DNA. It also has activity on εC in DNA which may result from exposure to vinylchloride or lipid peroxidation (108). After hTDG cleaves the U from U:G mispair, it remains bound to the products suggesting it has a very low catalytic turnover number. The rate-limiting step in this reaction is the release of the product due to strong interactions between the enzyme and the Watson-Crick face of the guanine opposite the AP-site (109). APE1 strongly stimulates the activity of TDG by displacing TDG from the AP site due to its higher affinity for the AP site (110). It has been demonstrated that a large portion of cellular TDG is covalently modified by the ubiquitin-like proteins SUMO-1 and SUMO-2/3. *In vitro* sumoylation of TDG leads to the reduction of TDG's affinity for both substrates and products. It was suggested that TDG binds its substrate in the unmodified state, and the cleavage, sumoylation allows TDG detach from the product AP-site due to reduced affinity (111).

**C. Occurrence of hypoxanthine in DNA and related repair enzymes**

**1. dITP formation and regulation**

dITP exists in all cells in a low concentration. dITP may be generated from dATP by spontaneous and nitrosative deamination (112). dITP is potentially mutagenic, and the

concentration of the nucleotide in cell is controlled by inosine triphosphate pyrophosphatase (ITPase). The gene coding for ITPase, *ITPA*, is located on the short arm of chromosome 20 (113). ITPase hydrolyzes ITP/dITP to IMP/dIMP and PPi (114). ITPase deficiency is a common inherited condition characterized by the abnormal accumulation of inosine triphosphate (ITP) in erythrocytes that can lead to thiopurine drug toxicity (115, 116). An ITPase which hydrolyze ITP, XTP, dITP was first discovered in *Methanococcus jannaschii* by a structure-based approach (117). In 2001, Lin and colleagues cloned a 21.5-kDa human inosine triphosphate pyrophosphatase (hITPase) which can hydrolyze ITP, dITP, and xanthosine 5'-triphosphate to their respective monophosphates, at optimal pH of 10 (118). A dITP- and XTP-hydrolyzing protein was found in *E. coli* by Chung and colleagues (119). Later, a protein of 184 amino acid encoded by y*jjX* gene in *E. coli* was found to hydrolyze ITP and XTP (120). ITPase may play a role in protecting the genome from incorporation of rogue nucleotides, such as ITP, dITP and XTP, into DNA and RNA (121).

**2. Enzymes that recognize hypoxanthine in DNA**

**a. 3-methyladenine-DNA glycosylase II (AlkA)**

3-methyladenine-DNA glycosylase II (AlkA) was first purified from *E. coli* by Riazuddin and Lindahl in 1978 (122). The enzyme was later cloned by two groups at about the same time (123, 124). The *E. coli* AlkA protein is a member of the glycosylase family which includes the endonuclease III, MutY, and 8-oxoguanine glycosylase (OGG) proteins (125, 126). These proteins have little similarity in amino acid sequence except in the region of the Helix-hairpin-Helix (HhH) motif and a conserved catalytic aspartate, but

share a very similar three-dimensional structure. *E. coli* AlkA can remove alkylated bases such as 3-methyladenine (3-MeA), 3-methylguanine (3-MeG), 7-methylguanine (7-MeG), 7-methyladenine (7-MeA) among which 3-MeA is a preferred substrate (127). AlkA also recognizes formyluracil (128), ethanobases (129, 130), hypoxanthine (131) and xanthine (132). The $K_{cat}$ value of AlkA for xanthine was 5-fold lower than that for 7-methylguanine (132). The crystal structure of AlkA suggests the widening of the active site located in the cleft between two domains of the AlkA protein could explain the broad spectrum of substrates of this enzyme (133). Though AlkA excise hypoxanthine in DNA *in vitro*, it is not likely the primary enzyme that involved in repair of hypoxanthine *in vivo*. *E. coli* mutant cells which lack AlkA protein did not exhibit significant increase in mutation frequency when exposed to nitrous acid (134).

**b. Endonuclease V (Endo V)**

Endonuclease V (Endo V) was first discovered as an enzyme that can recognize DNA treated with osmium tetroxide, acid, and base (135). Endo V was shown to be a primary enzyme that deals with hypoxanthine (136). We will review this enzyme in detail later (Section II and Section III).

**c. Alkyladenine DNA glycosylase gene (AAG)**

In 1991, three groups published data on cloning and expression of human alkyladenine DNA glycosylase gene (hAAG) by phenotypic screening of *E. coli* (tag-alkA-) cells exposed to methylmethane sulfonate (137-139). AAG is also named MPG (138) and ANPG (139). Besides alkylation damages, hAAG recognizes a wide variety of

other damaged bases in DNA, including hypoxanthine (131), xanthine (140), 8-oxoguanine (141), oxanine (143), cyclic etheno adducts (144, 145), and various adducts of nitrogen mustards used in cancer chemotherapy (146). The structure of AAG complexed with DNA containing a modified apurinic/apyrimidinic (AP) site showed that the AP site analogue was flipped out of the DNA helix and into the active site of the enzyme (147). Using cell extracts from wild-type and AAG knockout mice, researchers found that AAG is the principal enzyme repair 3-methyladenine, 1, $N^6$-ethenoadenine, and hypoxanthine (148, 149). hAAG activity can be stimulated by AP endonuclease, proliferating cell nuclear antigen (PCNA) (150), hRAD23 (151), estrogen receptor α (ERα) (152). ERα can increase acetylation of hAAG, stabilize the binding of hAAG with hypoxanthine-containing oligos, and enhance removal of hypoxanthine from DNA by hAAG. In addition, hAAG can decrease p300-mediated acetylation of estrogen receptor, stabilize the interaction of ERα with estrogen response element-containing oligos and reduce transcription of simple and complex ERE (estrogen response element)-containing reporter plasmids (152). Interestingly, the interaction of hAAG with MBD1 studied *in vivo* showed that the MBD1–hAAG complex normally exists on the methylated gene promoter suggesting that hAAG may cooperate with MBD1 for transcriptional regulation (153). Recently AAG from *Bacillus subtilis* was cloned and characterized (154). It was shown that the AAG from *Bacillus subtilis* removes hypoxanthine much faster than human AAG with a 10-fold higher value for $k_{cat}$ than that of hAAG. We have cloned the AAG from *Mycobacterium tuberculosis* which has stronger activity toward hypoxanthine than that of hAAG (H. Gao and W. Cao, unpublished data).

## D. Occurrence of xanthine in DNA and related repair enzymes

### 1. dXTP formation and regulation

Xanthine, is a product on the pathway of purine metabolism and is converted to uric acid by the action of the xanthine oxidase enzyme. 2'-deoxyxanthosine (dX) results from the deamination of deoxyguanine by hydrolytic or nitrosative deamination. dX is a relatively stable lesion at pH 7 and could play a role in deamination-induced mutagenesis (154). Xanthine in DNA under acidic conditions is unstable and tends to depurinate (155-157). Oligos containing xanthine can be prepared post-synthetically (156). It also can be prepared by phosphoramidite chemical synthesis (157). The miscoding properties of dX have been studied by several groups using various polymerases (157-160). Among the polymerase investigated, human immunodeficiency virus I reverse transcriptase, exo$^-$ Klenow fragment of DNA pol I, and Drosophila DNA pol α tends to incorporate dCMP and dTMP opposite the dX lesion. Taq pol, rat pol β and human pol β insert only dCMP, the correct base, opposite the lesion, human pol α, η and κ preferentially incorporate dTMP opposite the lesion.

### 2. Enzymes that recognize xanthine in DNA

In 2000, Kow and colleague first reported that deoxyxanthosine in DNA can be recognized by *E. coli* endo V (161). *E. coli* endo V cleaves both deoxyinosine and deoxyuridine in single or double-stranded DNA, while it only recognizes deoxyxanthosine in double-stranded oligos. It was shown endo V cleaved DNA containing deoxyxanthosine at the second phosphodiester bond 3′ to the lesion just like the cleavage of DNA containing deoxyinosine or single base mismatches (161-163).

Terato and colleagues showed that AlkA and endo VIII also can recognize deoxyxanthosine. The $K_m$ values of AlkA and endo VIII for xanthine were 53 nM and 124 nM respectively (164). AAG, endonuclease III, and formamidopyrimidine-DNA glycosylase (Fpg) also cleave deoxyxanthosine from DNA but limited to X:C mispair (140).

**a. Endonuclease V**

See Section II-B, page 27

**b. 3-methyladenine-DNA glycosylase II (AlkA)**

See section I-C-2-a, page 17

**c. Alkyladenine DNA glycosylase gene (AAG)**

See section I-C-2-c, page 18

**d. Endonuclease VIII**

*nei*, the gene for *E. coli* endonuclease VIII, was first cloned in 1997 (164). It encodes a 263 amino acid protein that has significant similarity in both the N-terminal and C-terminal regions with bacterial Fpg proteins (164). *E. coli* single *nei* mutant has no obvious phenotype while *E. coli* double mutants of *nth* (endonuclease III), *nei*, are hypersensitive to the lethal effects of ionizing radiation and hydrogen peroxide (164-166). *nei nth* double mutant has 20-fold spontaneous C to T transitions above background (167). *E. coli* endonuclease VIII (*nei*) is a bifunctional glycosylase with both *N*-

glycosylase and AP lyase activities. It excises modified pyrimidines, including thymine glycol, dihydrothymine, β-ureidoisobutyric acid, urea residues, hydroxycytosine, 5-hydroxyuracil, and uracil glycol from DNA (168-170). Endonuclease VIII also can recognize 8-oxoguanine (8-oxoG) (171), 2,6-diamino-4-hydroxy-5*N*-methyl-formamidopyrimidine (171) and xanthine (132). The $K_{cat}$ value of endonuclease for xanthine was 50-fold lower than that for thymine glycol (132). Human *nei* homologs *hNEI1* and *hNEI2* was cloned recently (172, 173). hNEI1 cleaves substrates containing thymine glycol, dihydrothymine, dihydrouracil, 5-hydroxycytosine and 5-hydroxyuracil (174). It was demonstrated that hNEI1 can recognize FAPY-A and FAPY-G from X-irradiated DNA and have weak activity against 8-oxoG (175).

**e. Endonuclease III**

Endonuclease III (*nth*) initiates base excision repair of oxidatively damaged pyrimidine bases in DNA (176). The primary substrates for *E. coli* Endonuclease III are 5,6-saturated pyrimidines, such as 5,6-dihydrothymine, 5,6-dihydro-5-hydroxythymine and thymine glycol (5,6-dihydro-5,6-dihydroxythymine), and 5-hydroxy-5-methylhydantoin formed by γ-irradiation in DNA (177). *E. coli* Endonuclease III also excise a pyrimidine ring-opened derivative of 1, $N^6$-ethenoadenine (178), 8-oxoguanine in 8-oxoG:G mispair (179), hydantoins derived by further oxidation of 8-oxoguanine (180), apurinic/apyrimidinic (AP) site, thymine glycols, urea residues (181), and 5-hydroxycytosine and 5-hydroxyuracil (182). *E. coli* endonuclease III first excises the damaged base by hydrolyzing the N-glycosylic bond between the damaged base and the DNA backbone, and then cleaves the phosphodiester linkage on the 3' side of the

resultant AP site by a ß-elimination reaction to generate a strand scission product containing a 5' phosphate and a 3' unsaturated aldehyde (183-186). Endonuclease III contains a [4Fe-4S] cluster which is not directly involved in the substrate binding (187-190). It is reported that endonuclease III may use the [4Fe-4S] cluster to mediate redox chemistry as part of a signaling mechanism to detect base lesions (191).

**f. Formamidopyrimidine-DNA glycosylase (Fpg)**

*E. coli* formamidopyrimidine-DNA glycosylase (Fpg, also known as MutM) is a bifunctional glycosylase which has 269 amino acids with a molecular weight of 30.2 kDa. The monomeric zinc finger repair enzyme excise damaged bases such as 8-oxoguanine (8-oxoG) (192, 193), 2, 6-diamino-4-hydroxy-5-formamido pyrimidine (Fapy-G) (194, 195), which were indicated as the major substrate *in vivo*. Fpg also remove 7-hydro-8-oxoadenine (8-oxoA), Fapy-A (196), apurinic/apyrimidinic (AP) sites (197), 5-hydroxycytosine, 5-hydroxyuracil (5-OHU) (181), dihydrouracil (DHU) (198), uracil glycol (170), thymine glycol (199), and hydantion (200). Fpg has three enzymatic activities, DNA glycosylase activity that remove the damaged base from the DNA to generate an abasic (AP) site (192), AP lyase activity that cleaves the 3′- and 5′-phosphodiester bond at AP sites via a β,δ-elimination mechanism (201, 202), deoxyribophosphodiesterase (dRpase) activity that removes the 5′-deoxyribose phosphate moiety (Figure 1.6) (202).

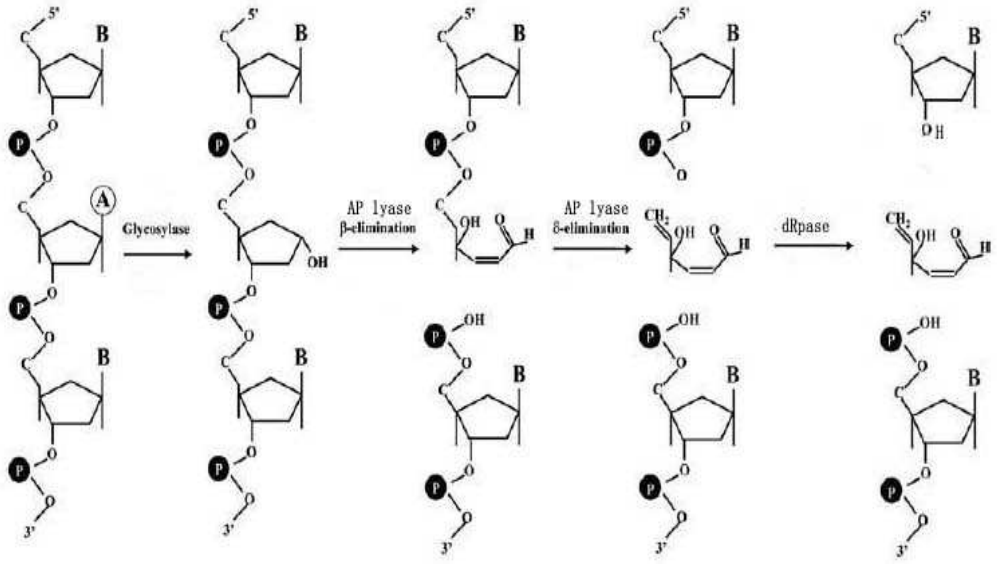



**Figure 1.6 Fpg activities.** 1. Glycosylase activity. 2. AP lyase activity. 3. dRpase activity. Adapted from (202).

**E. Discovery of oxanine in DNA and related repair enzymes**

**1. dOTP discovery and enzymes that recognize oxanosine in DNA**

Oxanosine was discovered as a new antibiotic in 1981 from *Streptomyces capreolus* MG265-CF3 (204). Oxanosine suppresses the growth of L-1210 leukemia in mice, HeLa cells *in vitro*, and exhibits antibacterial activity against *E. coli* K-12 (204-206). The antitumor studies of oxanine using rat kidney cells infected with mutant *Rous sarcoma* virus showed that oxanosine is more cytotoxic to tumor cells than to normal cells (207). Oxanosine also can reverse K-*ras*-transformed rat kidney cells to normal phenotype (208). Majumdar and colleagues showed oxanosine and 2'-deoxyoxanosine are substrates of adenosine deaminase (ADA) (209). In 1996, Suzuki and colleagues first reported deoxyoxanosine (dO) which is generated through nitric oxide $(NO)^-$ or nitrous acid $(HNO_2)^-$ induced nitrosative oxidation of deoxyguanosine (210). Compare to deoxyuridine, deoxythymidine, deoxyinosine and deoxyxanthosine, deoxyoxanosine is a

unique deaminated form in which an endocyclic nitrogen atom of guanine is substituted by an oxygen atom (210-212). Several groups demonstrated that the ring-opened intermediate of guanine, after loss of dinitrogen in guanine diazonium ions, results the formation of oxanine as well as xanthine (212-214). Whether dO exists in the cellular system is still under debate. Dong and colleagues showed that there are no 2'-deoxyoxanosine in DNA exposed to nitric oxide at controlled physiological concentrations (215) while Shuker and Glaser showed dO exists at physiological condition (211, 214). Several groups have used T4 polymerase (exo$^-$) or Pol I Klenow fragment to prepare O-containing DNA (142, 216-218). Recently, a chemical method for preparing O-containing oligos also has been developed (219). To assess genotoxic potential of dO, *in vitro* DNA polymerase studies have been conducted by different groups. It was shown that dOTP can pair with C or T in DNA by the Klenow fragment of DNA polymerase I using O-containing template, resulting in G/C to A/T transition mutations (216). Recent studies showed that cytosine, adenine and thymine all can be incorporated opposite oxanine. Deoxyoxanosine triphosphate only pairs with cytosine using human DNA polymerase β (142). Several enzymes, including endo V, hAAG, AlkA, Fpg, Endo VIII, which recognize oxanine in DNA have been reported (142, 132, 217, 218). hAAG can cleave oxanine from O-containing oligos, but only 30% substrate can be cleaved using excess of enzyme (E:S=10:1). Oxanine DNA glycosylase (ODG) activity was detected in spleen cell extracts of wild type age-matched mice but not in AAG-knockout mice, suggesting that AAG is the primary enzyme for oxanine (142). Other studies showed that hAAG can cleave only 9% of the O-containing substrates using hypoxanthine-containing substrates as control (218). It was reported that DNA-

binding proteins such as histone, high mobility group (HMG) protein, and DNA glycosylases can crosslink with O-containing DNA duplex to form DNA-protein cross-links (DPCs) which are potentially carcinogenic (220). UvrABC nuclease was demonstrated to have activity towards Oxanine–spermine cross-link lesions (218). Endo V can cleave O-containing DNA at the second phosphodiester bond 3' to the lesion in the presence of $Mg^{2+}$ or $Mn^{2+}$. The cleavage of T/O substrate can reach 70% when enzyme is in excess. Both Endo V and hAAG have similar cleavage efficiency towards all four oxanosine-containing base pairs (A/O, T/O, C/O and G/O) (142, 217).

## III. Endonuclease V initiated DNA repair

### A. History of research on endonuclease V

*E. coli* Endonuclease V was first discovered and characterized by Linn and colleagues (135, 221). During the course of purifying the RecBC deoxyribonuclease of *E. coli K-12*, the researchers found that endodeoxyribonuclease activity eluted from DEAE-cellulose prior to the RecBC enzyme. Unlike the RecBC enzyme the activity was not stimulated by adding ATP. It differed from *E. coli* endonuclease II by requiring $Mg^{2+}$. It is also distinguishable from the restriction enzymes of *E. coli* which require ATP and require, or are stimulated by S-adenosylmethionine. Further characterization of this enzyme by Linn's group found it has endonuclease activity toward DNA, treated with osmium tetroxide, or 7-bromomethyl-benz[a]anthracene, irradiated with ultraviolet light, or exposed to pH 5.0. The uracil-containing duplex DNA was a good substrate. The enzyme was then designated as endonuclease V of *E. coli*. It was thought the enzyme

may provide an alternative mechanism for remove the uracil residue from DNA due to its high activity towards uracil-containing duplex DNA (221).

In 1988, Sperling's group published a paper which suggested there was a hypoxanthine-DNA glycosylase in *E. coli*. The enzyme had a molecular mass of 56 kDa. This hypoxanthine-DNA glycosylase from *E. coli* requires $Mg^{2+}$ and is totally inhibited in the presence of EDTA (222). In an effort to clone the gene encoding the hypoxanthine-DNA glycosylase described by Sperling and colleagues, Kow's group started to purify the activity. During the purification they detected a deoxyinosine-specific endonuclease. SDS-PAGE showed that the enzyme had an apparent molecular mass of 25 kDa. Since the enzyme made an incision at the second phosphodiester bond 3' to a deoxyinosine this enzyme was named as deoxyinosine 3' endonuclease (136). In late 1996 and early 1997, two papers were published by Kow's group and Weiss's group, both groups cloned endonuclease V gene (*nfi*), which was the same gene encoding deoxyinosine 3' endonuclease (161, 223). The *nfi* gene was located in a cluster of four codirectional genes, *yja*D-*hem*E-*nfi*-*yia*G and expressed independently (223). Since then Weiss and colleagues set out to study the function of endonuclease V in *E. coli*. In order to understand its role in DNA repair pathways, they constructed *nfi* insertion mutants and over-expression strains. A *nfi* mutant displayed a 12- to 1,000- fold in the frequency of nitrite-induced A:T to G:C transition mutation compare to the wild-type strain. A *nfi xth dut* (dutpase) mutant also had increased lethality compare to an *xth* and *dut* mutant. These results suggested that endo V played an important role in the repair of deoxyinosine and abasic sites in DNA (224). Further study also show endo V was important in remove deaminated guanine (xanthine), while AlkA did not contribute to repair deaminated base

*in vivo*, though it has hypoxanthine-DNA glycosylase activity (225). Endo V from *Thermotoga maritime* was first cloned and studied by Cao and colleagues (217, 226, 227).

**B. Properties of endonuclease V from different species**

**1. *Escherichia coli* endonuclease V (*E. coli* endo V)**

Endonuclease V (Endo V) is a ubiquitous enzyme. Its homologues have been found in different species from bacteria to eukaryotes including *Escherichia coli, Streptomyces coelicolor, Bacillus subtilis, Archaeoglobus fulgidus, Thermoplasma acidophilum, Ferroplasma acidarmanus, Sulfolobus solfataricus, Pyrococcus furiosus, Salmonella typhimurium, Yersinia pestis Schizosaccharomyces pombe, Caenorhabditis elegans, Arabidopsis thaliana*, rats, mice, and humans. *E. coli* endo V is the first endonuclease V discovered (135, 221). Characterization of *E. coli* endo V has demonstrated that the enzyme recognizes a wide variety of DNA lesions including deoxyinosine, deoxyuridine, AP sites, tetrahydrofuran (an AP site homolog), base mismatches (136, 162, 228, 229), deoxyxanthosine (161), deoxyoxanine (217), *N*-6-hydroxylaminopurine (HAP) (230), loops, hairpins, pseudo-Y and flap structures (162). Endo V cleaves the lesion containg DNA at the second phosphodiester bond 3′ to the DNA damage, results in a nick with 3′ hydroxyl and 5′-phosphoryl end groups (163). The enzyme forms stable complexes with deoxyinosine containing DNA before and after cleavage, exhibiting similar affinity to the substrate and the product (228). This property is also observed in endo V from other species (217). Based on the tight binding of endo V with the substrate, it was suggested that the enzyme may function to target other repair

protein(s) to the lesion to continue the subsequent repair (163). It was reported that *E. coli* endo V cleaves mismatch containing oligos in a strand-specific manner; it cleaves the DNA strand containing mismatch closer to the 5' terminus (162). Using *Tma* endo V, the strand-specific manner of cleavage was not obvious (226). Mismatch activity of *E. coli* endo V is much higher in the presence of $Mn^{2+}$ than $Mg^{2+}$ (162). Deoxyinosine-specific activity of *E. coli* endo V is not affected by the sequence context of the deoxyinosine containing oligo while the mismatch-specific activity of the enzyme is reduced when the flanking sequence of the mismatch is G:C pairs. Based on these findings, it was speculated *E. coli* endo V adopts different modes of interaction between DNA containing deoxyinosine and mismatches (162). *E. coli* endo V cleaves deoxyuridine in the duplex efficiently in the presence of $Mg^{2+}$ or $Mn^{2+}$, but it has a very weak activity on deoxyuridine in single-stranded DNA (162).

**2. *Salmonella typhimurium* endonuclease V (Sty endo V)**

*Salmonella typhimurium* endonuclease V (*Sty* endo V) was recently cloned and characterized. *Sty* endo V shares 93% amino acid sequence identities with *E. coli* endo V (231). The study of *Sty* endo V's activity towards deaminated lesions revealed that *Sty* endo V possesses single-stranded deoxyxanthosine endonuclease activity while *E. coli* endo V doesn't. *Sty* endo V turns over double-stranded deoxyuridine-containing substrates very slowly (231) while *Tma* endo V shows rapid turn-over of deoxyuridine-containing duplex DNA when the substrate is in excess (226). The results demonstrated that *Sty* endo V binds deoxyuridine-containing products much better than does *Tma* endo V.

### 3. *Thermotoga maritime* endonuclease V (*Tma* endo V)

Endonuclease V from *Thermotoga maritime* (*Tma* endo V) is a thermal stable enzyme, prolonged incubation at 65°C which is an optimum reaction temperature for 8 hours doesn't cause significant loss of the enzyme activity (226). The enzyme recognizes inosine, abasic site (AP site), uracil, mismatches (226), oxanosine (217), and xanthosine (W Cao, unpublished data). Excess of *Tma* endo V in the reaction may lead to a second nicking event on the complementary strand and generate a double-stranded break (227). Study of mismatch cleavage activity of *Tma* endonuclease V showed that purine bases (A, G) in the mismatch are cleaved preferentially; cytosine is the most difficult one to cleave for *Tma* endo V (226). Mutational study of *Tma* endo V revealed residues important for the functions of the enzyme. Seven motifs universal to all endo V family proteins have been identified (227, 232). D43 in motif II, E89 in motif III, and D110 in motif IV are residues essential for coordination of catalytic metal ions. Y80, G83, and L85 in motif III, G113, H116, and G121 in motif IV, G136 and A138 in motif V, and S182 in motif VI are involved in both substrate and product binding. Interestingly, some *Tma* endo V mutants defective in DNA binding showed 3'-exonuclease activity, based on the results, an alternative model of endonuclease V initiated DNA repair was proposed by Cao and colleagues (232).

### 4. *Archaeoglobus fulgidus* endonuclease V (*Afu* endonuclease V)

Endonuclease V from *Archaeoglobus fulgidus* (*Afu* endo V) which is a hyperthermophilic archaea shows 39% identity to the *E. coli nfi* gene (231). In contrast to *E. coli* endo V, *A. fulgidus* endo V only recognizes deoxyinosine. Abasic site,

deoxyxanthosine, deoxyuridine and base mismatches are not cleaved by the enzyme. The fact that *A. fulgidus* endo V recognizes exclusively deoxyinosine indicates that deoxyinosine activity may be a primordial activity for endo V (233).

**5. Mouse endonuclease V**

Moe and colleagues recently cloned and expressed mouse endonuclease V, a 37-kDa protein which has weak activity toward hypoxanthine and uracil (234). Expression of the mouse protein in an *E. coli* mutant strain (*nfi alkA*) suppresses its spontaneous mutator phenotype. The study showed mouse endo V does not have broad substrate specificity as has been described for *E. coli* endo V. It has the following substrate preference: ss DNA containing hypoxanthine > DNA ds containing hypoxanthine ≥ ss DNA containing uracil (234). No obvious activity was monitored on uracil residues in double-stranded DNA, nor against 8-oxoguanine, AP-sites or 5' flap structures. Like other endo V, the mouse endo V is a metal dependent enzyme (234).

**6. *Ferroplasma acidarmanus* endonuclease V (AGTendo V)**

AGTendonuclease V is a DNA repair protein of 303 amino acids from the archaeal organism *Ferroplasma acidarmanus* (235). It consists of two domains, the C-terminal active site domain of $O^6$-alkylguanine-DNA alkyltransferase (AGT) and an endonuclease V domain. AGT is a DNA repair enzyme which excises alkyl damages from the $O^6$-position of guanine and $O^4$-position of thymine in DNA (236, 237). AGTs directly repair alkyl adducts by transferring the alkyl group to a cysteine amino acid residue at the active site of the protein and restore the damaged DNA to its unmodified

form. The enzyme thus loses the activity permanently. AGTendoV repairs $O^6$-methylguanine lesions but not $O^4$-methylthymine adducts in DNA. The enzyme shows optimal AGT activity at pH 6.6 and 46°C. AGTendoV also recognize uracil, hypoxanthine, or xanthine in DNA. It exhibited good endo V activity at pH 6.6-pH 7.6 and 46°C in the presence of $Mg^{2+}$. Inactivation of the AGT domain of AGTendo V by formation of *S*-methlycysteine at Cys-58 did not alter the endo V activity of AGTendoV (236, 237).

Endonuclease V was studied extensively by different researchers. Unfortunately, the repair pathway initiated by the enzyme remains unknown after one decade of research, mostly due to no other component was found downstream of endo V. Several repair models have been proposed, I will discuss these models in detail in the following sections.

**C. Proposed models for endonuclease V mediated pathways**

Genetic studies done by Weiss's group clearly suggested that *E. coli* endonuclease V prevents mutations from nitrosative deamination *in vivo* (238, 239). *In vitro* study also showed that deoxyinosine and deoxyxanthosine are two major substrates for endo V (136, 161, 226, 227). Based on genetic and biochemical data, Kow proposed a model for the repair pathway of *E. coli* endo V (161, 240). Endo V first recognizes hypoxanthine or xanthine and makes a nick at the second phosphodiester bond 3' to the lesion. It does not remove the lesion, but only initiates repair. The initiation of the repair resembles that of the VSR protein in the repair of T:G mismatch in DNA which showed that the subsequent step of repair involves a 3'-5' exonuclease (Figure 1.7) (241, 242). So it is

likely that the repair initiated by endo V may also require a 3'-5' exonuclease to remove the lesion followed by polymerase and ligase to complete the repair. It is also likely that an unknown 5' endonuclease that cleaves the phosphodiester bond 5' to the deaminated purine. However, despite great efforts, no such 5' endonuclease was found (240).

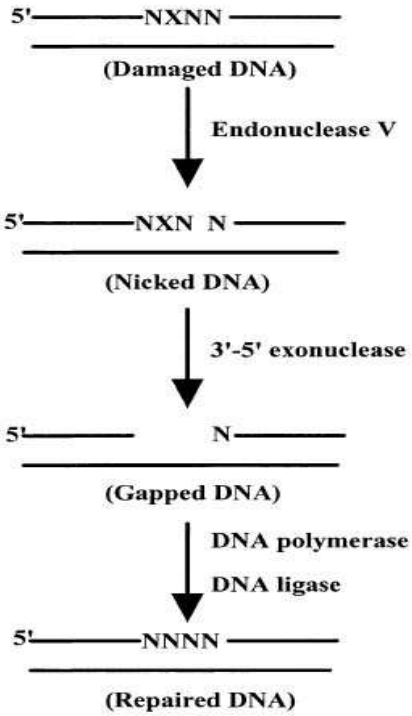



**Figure 1.7 Alternative repair pathway of endonuclease V proposed by Kow.** Taken from (241).

Klungland and colleagues proposed a similar model for mouse endo V initiated DNA repair scheme, incision at hypoxanthine residues in DNA by a mammalian homologue of the *E. coli* endo V (234). In this model, mouse endo V first generate a nick at the second phosphodiester bond 3' to the deaminated purine, then APE1 which posses 3' mismatch-specific exonucleolytic activity or Mus81 which posses a 3' flap specific endonuclease activity remove a short 3' region containing the lesion. Finally polymerase

fills in and ligase seals the nick. Cunningham and colleagues proposed a model for endo V to remove purine base *N*-6-hydroxylaminopurine (HAP) based on genetic studies (230). They suspected that endo V initiated repair may be a slow process in which the repair intermediates is subject to recombinational repair and induce SOS response.

Recently Cao and colleagues found an interesting phenomenon in *Tma* endo V (232). In the presence of $Mn^{2+}$, this enzyme displays a 3' exonuclease activity. A model was proposed that endo V may play a dual role in the repair pathway, after endo V cleaves the lesion-containing DNA and forms a complex with the product, downstream proteins may be recruited to the repair complex, and change the conformation of endo V. This may trigger the endo V's exonuclease activity and remove the lesion and generate a gap for polymerase and ligase to finish the repair (Figure 1.8).

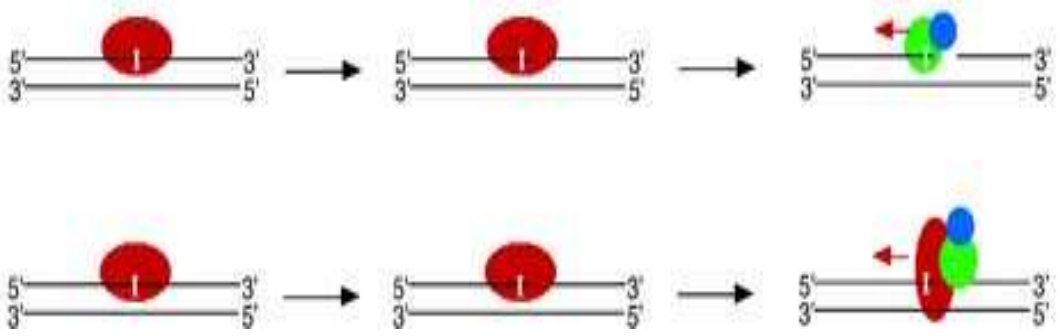



**Figure 1.8 Proposed models for endonuclease V-mediated repair.** I represents deoxyinosine. The ovals represent endo V and the circles putative endo V interacting proteins. After the inosine strand is nicked, the model shown at the top suggests that endo V is displaced by downstream protein(s) from the lesion and a 3'-exonuclease other than endo V initiates the removal of the deaminated lesion. An alternative model shown at the bottom based on data obtained from this study illustrates that endo V recruits other protein(s) to the repair site, which switch endo V from endonuclease mode to 3'-exonuclease mode for removal of the deaminated lesion from DNA. Taken from (232)

## IV. Endonuclease V related techniques

### A. Alternative method for DNA fragmentation with endonuclease V

#### 1. DNA shuffling and its applications

DNA shuffling, also called sexual polymerase chain reaction (PCR) was first developed by Stemmer and colleagues (243, 244). This method generates libraries by random fragmentation of one gene or a pool of related genes, followed by the reassembly of the fragments into full length chimeric sequence by PCR reaction (245). The libraries are then screened to identify desired recombinants. The methods of DNA shuffling and screening which develop novel chimeric DNA and proteins with desirable characteristics are also known as Molecular Breeding$^{TM}$(maxygen, Inc., Redwood City, CA), directed molecular evolution (246). Directed evolution is now widely used in academic and industrial labolatories to enhance protein stability and improve the activity or overall characteristics of enzymes and organisms or to alter enzyme substrate specificity and to obtain new activity. An emerging field in biotechnology is to modify DNA-modifying proteins which can lead to novel application in genetic engineering, functional genomics, and gene therapy. One of the enzymes which have been successfully modified by directed evolution is DNA polymerase. The modified enzymes have enhanced activity (247, 248) or have been converted to RNA polymerases (249). Restriction endonucleases are good targets for directed evolution because they are widely used in modern molecular biology. Many restriction enzymes with altered substrate specificity have been developed with the method (250-252). Other enzymes, such as transposase (253), integrase/recombinase (254-258) have been modified using directed evolution to obtain desired properties. Directed evolution technique is not limited to the research, the method also has great

applications in industry. Many novel industrial enzymes variants have been developed by the directed evolution technique to yields desired properties, including proteolytic enzymes (259-268), cellulolytic enzymes (269-271), lipases and esterases (272-278) and so on.

Directed evolution has played an important role in pharmaceuticals. Directed evolution can be used to generate improved antigens or other immunomodulatory molecules, and DNA vaccines (279-288). Therapeutic antibodies represent the fastest growing area in pharmaceutical industry. High-affinity antibodies have been developed *in vitro* by direct evolution (289-295). Directed evolution has great impact on the development of therapeutic proteins with better properties, such as P53 (296-297), insulin (298), hormones and hormone receptors (299-301). Stem's method to use DNase I to generate a pool of random DNA fragments has proven to be very useful in directed evolution (244), but the experimental procedure is extremely labor intensive and time consuming. DNA fragmentation by DNase I is problematic since the reaction has to be carefully controlled in order to obtain fragments of appropriate length. In addition, the length of the fragment generated by DNase I digestion varies greatly with minor changes in conditions, including the amount of nuclease, the source or lot of nuclease, the reaction temperature and the purity of DNA substrates. Therefore, using fragments generated by DNase I digestion may induce a sequence bias into the recombination (302, 303)

**2. Endonuclease V as an enzyme for DNA fragmentation**

Endonuclease V can cleave uracil containing DNA. Miyazaki prepared uracil-containing recombinant templates by PCR in the presence of dUTP (304). The *E. coli*

endo V was incubated with the PCR products thus generating random fragments. The length of random fragments can be controlled by adjusting the number of the uracil residues in the DNA simply by changing the concentration of dUTP in the PCR. Using the random fragments generated by endo V digestion, successful DNA shuffling was achieved with shuffling efficiency equivalent to DNase I (304).

## B. Balanced Linear DNA amplification by Endonuclease V and Polymerase

### 1. *In vitro* nucleic acid amplification methods

*In vitro* nucleic acid amplification techniques can be classified into two categories, isothermal systems which include Nucleic Acid Sequence-Based Amplification (NASBA) (305), Strand Displacement Amplification (SDA) (306), and those requiring temperature cycling, which include Polymerase Chain Reaction (PCR) (307) and Ligase Chain Reaction (Ligase Chain Reaction) (308, 309). Here I will focus on a strand displacement amplification method that employs endo V combined with polymerase.

### 2. Strand displacement amplification (SDA)

Strand displacement amplification was first developed in 1992 by Walker and colleagues (310). In 1994, walker and colleagues developed a multiplex form of SDA which allowed two target sequences and an internal amplification control to be co-amplified by a single pair of primers after common priming sequences are spontaneously appended to the end of target sequences. Although multiplex amplification has its

advantage, it often leads to large decreases in amplification efficiency (311), nonspecific amplification, especially when low concentrations of template DNA are used (312).

### 3. Modified SDA method using *Thermotoga maritima* endonuclease V

Recently, Barany and colleagues developed a method of SDA by nicking and extending that relies on the cooperation of *Thermotoga maritima* (*Tma*) endonuclease V (endo V), a DNA mismatch-cleavage enzyme, and *Bacillus stearothermophilus* (*Bst*) polymerase (313).

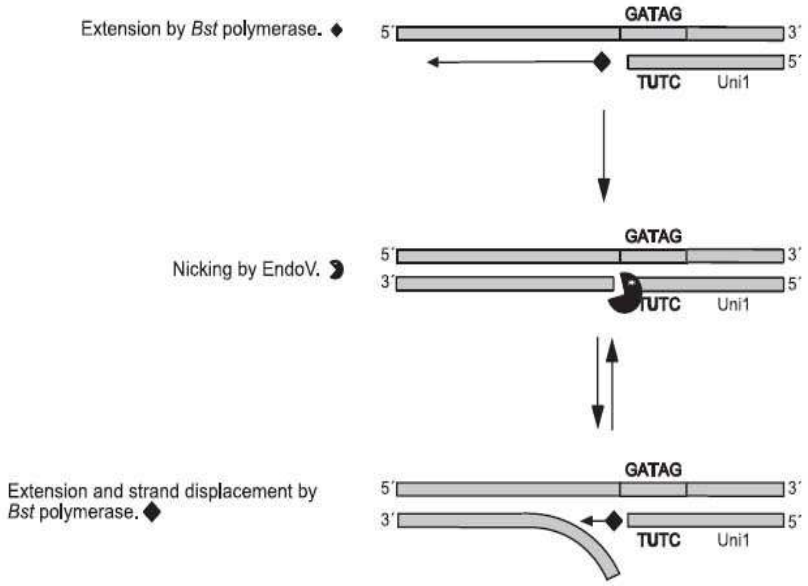



**Figure 1.9 Endo V facilitated DNA amplification.** Primer Uni1 anneals to a template containing the pentanucleotide GATAG, creating a U:T mismatch. The primer is extended by *Bst* polymerase and is nicked by *Thermotoga maritima* (*Tma*) endonuclease V (Endo V), 2 phosphodiester bonds to the 3′ side of the mismatch, leaving the mismatch intact. The top strand is unaffected. Because the primer contains uracil, after cleaving the strand, Endo V dissociates. The newly generated 3′ end can again be extended. *Bacillus stearothermophilus* (*Bst*) polymerase is able to displace the strand in front as it extends the primer. Cycles of nicking and extending amplify the template molecule. Product molecules can be detected and quantified by the ligase-detection reaction (LDR) followed by capillary electrophoresis. Taken from (313).

Compared to multiplex strand-displacement amplification, this method minimizes the chance for amplification of primer artifacts and other nonspecific products by using single primer, coupled with a high reaction temperature (65°C). In addition, unlike the restriction enzymes used in strand-displacement amplification, endo V can nick one strand of duplex DNA without requiring the presence of a phosphorothioate nucleotide in the template strand. In this method, a uracil-containing primer is annealed to a DNA template generating a U:N mismatch. *Tma* endo V will only nick the uracil-containing strand preferentially if the mismatch is in an appropriate position and sequence context. The enzyme dissociates after nicking and produces a 3′ end that can be extended by *Bst* polymerase, resulting in repeated cycles of extension followed by nicking, which amplify the template in a linear fashion (Figure 1.9). SDA performed by this way has more products than conventional SDA which use restriction endonuclease *Taq* I to nick the primers that contain a phosphorothioate linkage. Although the phosphorothioate bond can prevent cleavage of the template strand, it also hampers the amplification. The method currently allows 100-fold multiplexed amplification of target molecules to be performed isothermally, with an average change of <1.3-fold in their original representation.

**C. Mutation scanning by endonuclease V**

**1. Mutation detection methods**

The detection of sequence in DNA is important in the diagnosis of both genetic and somatically acquired diseases. Many techniques have been developed during the past two decades in an effort to detect minor changes in DNA. Such changes may take the

form of insertions or deletions, or single base changes. The following review will focus on single base mutation detection.

**a. Sequencing method**

The most direct way to detect a mutation is to sequence the DNA using the Sanger & Coulson dideoxy chain termination method. DNA sequencing has been a very accurate and reliable tool for pinpointing mutated sequences. However, DNA sequencing can involve time-consuming cloning and purification steps. These additional steps make rapid and routine high-volume screening very difficult. Thus, methods based upon non-sequencing mutation detection principles are needed as an alternative to DNA sequencing which is expensive and laborious.

**b. Nonsequencing mutation detection techniques**

Several methods that based on melting temperature change of the DNA have been developed, such as denaturing high-performance liquid chromatography (DHPLC) (314, 315), single-stranded conformational polymorphism (SSCP) (316), heteroduplex analysis (HA) (317), and denaturing gradient gel electrophoresis (DGGE) (318). Several chemicals like osmium tetroxide, $KMnO_4$ and hydroxylamine can react with mismatched thymine or cytosine. Based on the findings, Cotton et al. (319), Gogos et al. (320) Lambrinakos et al. (321) developed the chemical cleavage of mismatch (CCM) method which was first a solution-based reaction and recently modified as support-based assay of more convenience (322-325). This methods offer greater sensitivity than the melting temperature-based systems.

The enzyme mismatch cleavage (EMC) of mutation detection is similar to chemical cleavage of mismatch method. Both methods require amplification of the mutant and wild type using fluorescence labeled primers or radiolabeled primers. Heteroduplexes are formed by annealing mutant DNA with wild type DNA. Mismatch signals can be subsequently detected after chemical or enzyme cleavage. However the enzymatic technique offers greater simplicity and is non-toxic compared to the chemical method. Bacteriophage T7 endonuclease I is a 149 amino acid protein that is a Holliday junction-resolving enzyme (resolvase) (326, 327). Endonuclease VII, the product of gene 49 of phage T4, was shown to have activity on Holiday junctions (328), Y junctions (329), and heteroduplex loops (330). Later, both enzymes were found to posses activities towards mismatches, and were useful for mutation screening (331-333). The phage resolvase-based mutation detection systems tend to have high noise background because both enzymes naturally function to recognize Holiday junctions instead of mismatches. Recently several DNA N-glycosylases have been used to detect DNA mismatches, including MutY (334) and Thymine DNA glycosylase (335). DNA glycosylases are highly specific DNA repair enzymes that cleave the N-glycosidic bond between the base and the sugar deoxyribose in a DNA molecule and generate an apyrimidinic (AP) site. The AP site can subsequently be cleaved chemically by alkali or an AP endocuclease (336, 337). The MutY protein and Thymine DNA glycosylase are involved in a DNA repair pathways which convert A:G or T:G mismatches respectively to C:G base pair (338-340). CEL 1 nuclease from celery is a member of the S1 nuclease family. It was found that the enzyme can cleave mismatches and insertions/deletions (341).

**2. Endonuclease V-based mutation scanning method**

The enzymatic detection methods described above can identify the mutations and polymorphisms of DNA fragments. However these enzymes also nick matched DNA resulting in high background in the assays. Barany and colleagues in Cornell University developed an endonuclease V (endo V)/ligase mutation scanning method which has very high sensitivity and is suitable for low-frequency known or unknown mutation detection (342). This method has been shown to detect mismatches, insertions/deletions in DNA fragments up to 1.7 kb (342). Endonuclease V recognizes a wide variety of DNA lesions including base mismatches, insertions/deletions. Like other mismatch cleavage enzymes, endo V also nicks matched base pairs slightly, but in contrast to other enzymes, it leaves ligatable ends. TAK16D ligase is a high fidelity thermostable ligase from *Thermus* species (342). By using the two enzymes sequentially, they can religate the nonspecific nicks, while maintain the desired nicks (Figure 1.10).

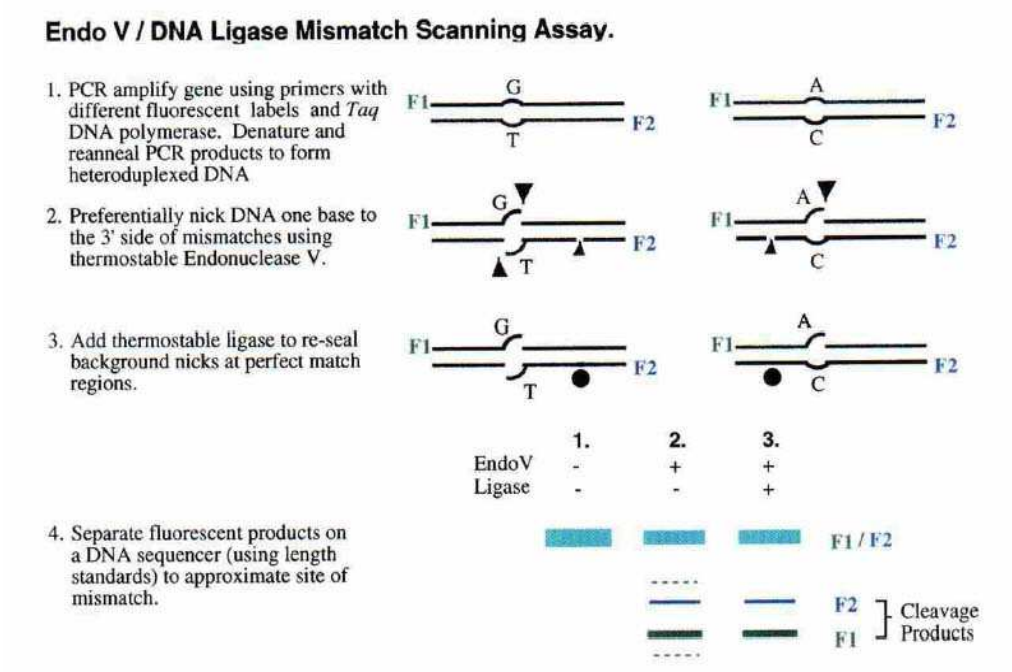



**Figure 1.10 Endonuclease V/DNA ligase mismatch scanning assay for scoring unknown mutations.** Heteroduplexes are formed from PCR amplicons both normal and variant sequence. Mixed amplicons of variant sequence can be obtained by amplifying heterozygous germline samples, or from tumor samples where stromal cell contamination provides sufficient amount of wild-type DNA, or by mixing the PCR products from unknown and wild-type samples in a 1:1 ratio,. *Tma* Endo V nicks DNA one base 3' to the mismatch (big triangle), and it can also generate non-specific nicks in homoduplex DNA with minor activity (small triangle). Addition of *Thermus. Sp.* AK 16D DNA ligase (solid circle) seals these non-specific nicks, providing a proofreading mechanism to improve signal-to-noise. Both top and bottom strand PCR primers are 5' end-labeled with different fluorescent dyes (6-FAM and TET, respectively) allowing for cleavage products to be distinguished on a denaturing polyacrylamide gel. The approximate position of the mutation can be determined from the resultant fragment lengths. Taken from (342).

## V. DNA lesion recognition mechanisms

A fundamental question in DNA repair is how mismatched or modified bases are located within the vast excess of normal base pairs. On the basis of the frequency at which spontaneous DNA lesion forms in mammalian cells, it is estimated that one DNA repair enzyme need to survey 10,000 to 100,000 in order to locate a single lesion (344). The lesion recognition mechanism may be considered in two parts: how a DNA repair

enzyme finds a lesion embedded in genome, and how a lesion, once found, is accommodated in the enzyme's active site.

**A. Models for enzymes to locate 'targets'**

It is known that proteins locate their specific targets on DNA up to two orders-of-magnitude faster than the three-dimensional diffusion rate. A general explanation of this fact is that proteins are randomly bound with DNA, and sliding along DNA (facilitated diffusion) provides for the faster one-dimensional scanning (344-346). Several models including 'sliding', 'hopping' and 'intersegmental transfer' have been put forward to explain how facilitated diffusion of a protein to a target actually occurs (Figure 1.11) (347).

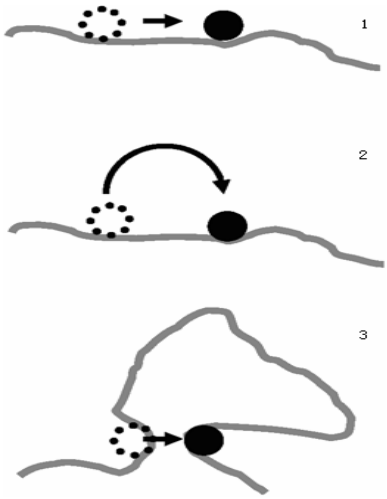



**Figure 1.11 Models for target site location.** (1) 'sliding', a protein might 'slide' along the double helix, transferring from one base pair position to the adjacent one without dissociating from the DNA (2) 'hopping', if dissociation occurs, the protein might re-encounter the same DNA, but at a new contour position. (3) 'intersegmental transfer', on scales beyond the persistence length of the DNA double helix, 150 bp (50 nm), the DNA can run into itself as a result of its random thermally excited bending. Such encounters permit the protein to move from one DNA site to another via an intermediate in which the protein is bound transiently to both sites. Taken from (347)

For proteins which do not use biochemical energy to facilitate their diffusion, target search along DNA is a thermally activated and directionally unbiased process. Sliding is a process which the protein undergoes diffusion while remaining bound to the DNA. Hopping refers to the process that the protein repeatedly dissociates from the DNA and rebinds at a new location on the DNA. Intersegmental transfer is a process that the protein molecule bound at one end jumps to another end of the same DNA lattice. Whether proteins use 'sliding' or 'hopping' to move along DNA remains controversial. Studies with restriction enzymes have demonstrated that hopping is the primary mode of translocation and that sliding, if it occurs, only contributes to movement on length scales of <173 bp (349, 350). Recently, Blainey and colleagues proved that at near-physiologic pH and salt concentration, human oxoguanine DNA glycosylase 1 (hOgg1) has a subsecond DNA-binding time and slides with a diffusion constant as high as $5 \times 10^6$ $bp^2/s$ (351). However, the searching for the targets by enzymes may be not just a naive one-dimensional sliding or hopping but rather a delicately weighted mixture of 'sliding' 'hopping', 'intersegmental transfer' and three-dimensional diffusion. In the following review, I will focus on how DNA mismatches or damages are recognized by the enzymes through different mechanisms.

## B. Models for DNA repair enzymes to recognize DNA damages

### 1. Structure-based recognition

Early hypotheses which tried to explain the recognition of damaged or mismatched DNA focused on the structural detail of the lesion and the difference between lesion and normal DNA. The X-ray crystallography and NMR study of damaged

and mismatched DNA showed that the incorporation of a base mismatch does not result in large changes in the overall conformation of a B-form DNA helix (352-362). The structure of an oligonucleotide containing the 8-oxo-guanine lesion showed almost no structural differences from B-DNA (363). Bulky lesion-containing DNA showed more structural distortion as revealed by NMR studies (364-369). The small differences between DNA lesion and normal DNA may be used by DNA repair enzymes to identify the lesion through direct read out (direct base–amino acid interactions). DNA repair enzyme may also detect bulky lesions which distort DNA by indirect read out (conformational properties of the DNA). 8-oxo-guanine and guanine differ in chemical composition at only two positions: C8 (O versus H) and N7 (H versus lone electron pair). The crystal structural studies showed that the hydrogen bond made by the N7 H of oxoG to the main-chain carbonyl of Gly 42 of human 8-oxogunanine DNA glycosylase I (hOGG), which would be missing with G, contribute to the discrimination (370). Hanawalt and Haynes proposed a model of DNA damage recognition mechanism for nucleotide excision repair, termed ‘close fitting sleeve’ model, which postulated that DNA repair is dependent on the difference between the structure of the DNA surrounding the lesion and that of the normal Watson–Crick double helix (371). Gunz and colleagues demonstrated that the efficiency of bulky lesion recognition by the human NER pathway can vary over several orders of magnitude and the efficiency of repair of a particular lesion is related to the amount of helical destabilization it can cause (372). Crystal structure of T4 endonuclease V (a DNA repair enzyme not related to endonuclease V family) complexed with thymine dimer revealed a sharp kink at the central thymine dimer in the bound DNA that split the DNA duplex into two halves of B-DNA, which

may facilitate the recognition (373). Further studies on different DNA repair enzymes indicate that structural discrimination of damage and mismatch in DNA do not explain the full range of mechanisms applied in DNA damage recognition. In addition, some DNA repair enzymes, such as AAG, AlkA, MutY etc, have very wide substrate specificity (374, 375). It is noteworthy that AlkA cleaves not only damaged nucleobases, but also releases undamaged native bases at a low background level (376). Models based upon structural recognition cannot explain wide substrate specificity of these DNA repair enzymes which identify and correct lesions of different sizes and structures.

**2. Thermodynamic and kinetic aspect**

Thermodynamic studies of base lesions and mismatches indicate that the damages tend to induce destabilizing influences on the DNA helix (376-378). This destabilization can be measured through a reduction in the denature temperature ($T_m$) compared to their undamaged counterpart. In addition, the thermodynamic consequences of the damage are sequence dependent (379-380). Also, damage can induce a kinetic destabilization of the helix lead to the increase of the rate of base pair opening events at the aberrant site (381). NMR studies showed that for some mismatches, such as G:G, interconversion between different conformations of the mispair is fast, approx. $10^4$ $s^{-1}$ at 303 K (30°C), while the G:A mismatch had a conformational exchange rate of only 100 $s^{-1}$ at 303 K comparable to the rates of opening for Watson–Crick base pairs (382). Modified DNA base pairs tend to adopt extra-helical conformations. The non-planar thymine glycol lesion has been shown to undergo significant motion by NMR studies and may be completely extra-helical (383). It is generally believed that both a distortion of the DNA helix and a

modification to the DNA chemical structure are needed for a NER substrate (384, 385). DNA helix distortions, including disruption of Watson-Crick base pairing, DNA bending, and unwinding of the DNA strands, are thought to play an important role in the NER recognition process (386-388). Simple mismatches or bubbles are not processed by the NER pathway, indicating the local thermodynamic destabilization of duplex DNA is not sufficient to be considered as a NER substrate. The human NER complex remains inactive on DNA substrates in which only the backbone of the duplex has been modified (389). However, the presence of covalent DNA backbone modifications in conjunction with mismatches results in a robust response by the NER machinery (389).

**3. Base unstacking initiated recognition**

Yang recently proposed that DNA repair proteins may first recognize poor base stacking in the search for DNA damages among a large excess of normal bases (390). The hypothesis suggests that all lesions have weakened base stacking. It results in local flexibility, reduced melting temperature, changes in helical parameters, steric clashes, duplex opening, bending of helical axis, and base flipping in the lesion containing DNA (391-394). The structure of the DNA helix is stabilized mainly by hydrogen bonding of coplanar base-pairs and vertical stacking potential between base-pairs. Wobble base pairing between mismatched bases requires base displacement (a shift in the plane of base pair), which may cause the base to unstack. In the crystal structures of oligonucleotides containing a G:T wobble base pair, base stacking in the neighborhood of the mismatch is perturbed and helical parameters are altered (395). Yakovchuk and colleagues recently demonstrated that in DNA duplexes with solitary nicks or gaps, base

stacking is the dominant stabilizing factor while A:T pairing is always destabilizing and G:C pairing contributes almost no stabilization (396). It may explain why DNA lesions are favorably repaired in some sequences, since base stacking stabilization is sequence dependent. DNA lesions in less stabilized sequences may be better recognized by DNA repair enzymes when they search for the 'unstacking signal'. Fluctuations in local helical conformation of DNA, known as 'DNA breathing', is also sequence dependent (397-398). Lesions in some sequences may open more frequently and result in favorable repair. Crystal structures of *E. coli* MutS complexed with a variety of mismatches reveal a common recognition mechanism. Mismatch binding by *E. coli* MutS involves the stacking of a phenylalanine residue (Phe 36) of one of the monomers, onto one of the mismatched bases and causes the mismatched base unstack with neighboring bases. The same base is reoriented and brought into proximity to the glutamate (Glu 38) and forms a hydrogen bond with this amino acid (399). VSR endonuclease recognizes T:G mismatches in a particular sequence context [5′-CT(A or T) GG-3′ paired with 5′-CC(T or A)GG-3′] and cleaves the DNA 5′ of the mispaired T. Tsutakawa and colleagues (400) showed that in the crystal structure of VSR-DNA complexes, three aromatic residues intercalate into the DNA next to the T:G mismatch and the T:G mismatch forms a wobble base pair, which is completely unstacked from the adjacent base pair at the 5′ side of the mismatched guanine. Mismatch Uracil DNA glycosylase (MUG) excises uracil or thymine base from U:G or T:G mismatches. The crystal structure of *E. coli* Mug-DNA complex shows that the U:G mismatch is unstacked from the neighboring base pair. Phe 30 of Mug makes π-π stacking interaction with uracil. The guanine opposite this uracil is separated from its 5′ neighbor by Arg-146 side chain (401).

**4. Nucleotide flipping mechanism**

Nucleotide flipping was first observed in cytosine-5-methyltransferases bound to their target DNA sequences (402). Nucleotide flipping, a process in which a target DNA nucleotide is flipped out of the DNA base stack, has been demonstrated to occur in some damage-specific DNA repair enzymes. Accumulating evidence shows that many repair enzymes use nucleotide flipping mechanism to recognize and remove damaged DNA (403-407). Crystal structures of repair enzymes complexed with damaged DNA exhibit a flipped-out base accompanied by substantial bending and distortion of the DNA. DNA bending and distortion by the enzymes may facilitate base-pair opening (408-410). In the processive mechanism, the enzyme scans along the duplex and samples every base it encounters. Normal bases are released and reanneal within the duplex. When the enzyme reaches a lesion that displays enhanced affinity for the enzymatic binding pocket, it stops and removes the base. This scenario applies to the lesions, such as 8-oxoG, which induce little destabilization of the duplex and show little extrahelical tendency. Lesions causing destabilization of the DNA duplex adopt extrahelical conformation frequently. These lesions are subject to the repair of nonprocessive mechanism in which the random extrusion of lesions from the duplex substrate are accommodated and processed by the repair enzymes. Nucleotide flipping is thought to be the primary mechanism for base excision repair enzymes to detect and remove a variety of base lesions within a large pool of undamaged DNA (411). Crystal structures of uracil-DNA glycosylase (407), alkyladenine glycosylase (412) and 8-oxoguanine-DNA glycosylase (413) complexed with lesions have shown that the damaged base exists in an extrahelical mode. Uracil DNA glycosylase (UDG) is a highly specific DNA repair enzyme which only removes

uracil lesions from DNA. Crystal structure of human UDG bound to uracil-containing DNA shows that UDG uses pinch-push-pull mechanism to flip the uracil out (407). UDG rapid scanning by DNA backbone compression (pinch) slightly bends the DNA. The Leu 272 side chain (L of the HPSPLS motif) penetrates into the base stack through the minor groove, causing extrahelical localization ("push") of the target uridine, which was then accommodated ("pull") into the active site pocket by specific interaction between them (407). However nucleotide flipping is not confined to base excision repair enzymes. Photolyase in direct reveral (414) and uvrB in NER (415) also use nucleotide flipping mechanism to facilitate the DNA repair. DNA photolyases use the energy of a blue-light photon to transfer an electron onto UV-damaged DNA, such as the *cis-syn* cyclobutane pyrimidine dimers (CPDs) and the resulting radical anion then splits into two pyrimidines and transfers back the excess electron to the enzyme (414). Crystal structure of a photolyase bound to a CPD-Like DNA lesion shows the thymine dimer is specifically recognized in the active site by being completely flipped out of the duplex DNA (416). The DNA bends about 50°C at the CPD bend. The complementary adenines stack with neighboring bases but do not stack on top of each other because of the large intrahelical bend (416). UvrB, a DNA helicase, is a central component of the bacterial NER system participating in damage recognition, strand excision and repair synthesis (417). Crystal structure of uvrB complexed with fluorescent adenine analogue 2-aminopurine indicates that binding of UvrB to the damage containing DNA moves the base adjacent to the lesion at the 3' side into an extrahelical position and the base opposite this flipped base is also extruded from the DNA helix. Tyr 95 in the haipin region is involved in base

flipping in the non-damaged strand. This conformational change in the non-damaged strand may be critical for 3' incision by UvrC (417).

## VI. References


1. Lodish H, Berk A, Matsudaira P, Kaiser CA, Krieger M, Scott MP, Zipursky SL, Darnell J. (2004). Molecular Biology of the Cell, p963. WH Freeman: New York, NY. 5th ed

2. Lindahl,T. (1993). Instability and decay of the primary structure of DNA. *Nature*, 362, 709–7

3. Shen,J.C., Rideout,W.M. 3rd and Jones,P.A. (1994) The rate of hydrolytic deamination of 5-methylcytosine in double-stranded DNA. *Nucleic Acids Res.*, 22, 972–976

4. Vasilenko NL, Nevinsky GA. Pathways of accumulation and repair of deoxyuridine residues in DNA of higher and lower organisms Biochemistry (Mosc). 2003 Feb;68(2):135-51

5. Shapiro, R., and Klein, R. S. (1966) The deamination of cytidine and cytosine by acidic buffer solutions. Mutagenic implications *Biochemistry*, 5, 2358_ 2363.

6. Shapiro, R. (1981) Chromosome Damage and Repair (E. Seeberg and K. Kleppe, eds.) Plenum, N. Y., pp. 3_10.

7. Lindahl, T., and Nyberg, B. Heat-induced deamination of cytosine residues in deoxyribonucleic acid. (1974) *Biochemistry* 13, 3405-3410

8. Shapiro, R., and Yamaguchi, H. Nucleic acid reactivity and conformation. I. Deamination of cytosine by nitrous acid (1972) *Biochim. Biophys. Acta* 281, 501-506.

9. Lindgren H, Stenmark S, Chen W, Tarnvik A, Sjostedt A. Distinct roles of reactive nitrogen and oxygen species to control infection with the facultative intracellular bacterium Francisella tularensis Infect Immun. 2004 Dec;72(12):7172-82.

10. Mikkelsen RB, Wardman P. Biological chemistry of reactive oxygen and nitrogen and radiation-induced signal transduction mechanisms. Oncogene. 2003 Sep 1;22(37):5734-54

11. Bauer G. Reactive oxygen and nitrogen species: efficient, selective, and interactive signals during intercellular induction of apoptosis Anticancer Res. 2000 Nov-Dec;20(6B):4115-39

12. V.A. Bohr, Repair of oxidative DNA damage in nuclear and mitochondrial DNA, and some changes with aging in mammalian cells, *Free Radic. Biol. Med.* 32 (2002), pp. 804–812

13. B.S. Berlett and E.R. Stadtman, Protein oxidation in aging, disease and oxidative stress, *J. Biol. Chem.* 272 (1997), pp. 20313–20316

14. M.S. Cooke, M.D. Evans, M. Dizdaroglu and J. Lunec, Oxidative DNA damage: mechanisms, mutation and disease, *FASEB J.* 17 (2003), pp. 1195–1214

15. H. Ide, Y.W. Kow and S.S. Wallace, Thymine glycols and urea residues in M13 DNA constitute replicative blocks *in vitro*, *Nucleic Acids Res.* 13 (1985), pp. 8035–8052

16. A.P. Grollman and M. Moriya, Mutagenesis by 8-oxoguanine: an enemy within, *Trends Genet.* 9 (1993), pp. 246–249

17. Caulfield, J. L., Wishnok, J. S., and Tannenbaum, S. R. (1998) Nitric oxide-induced deamination of cytosine and guanine in deoxynucleosides and oligonucleotides. *J. Biol. Chem. 273*, 12689-12695

18. Ohshima,H. and Bartsch,H. (1994) Chronic infections and inflammatory processes as cancer risk factors: possible role of nitric oxide in carcinogenesis. *Mutat. Res.*, 305, 253–264

19. Shacter,E. and Weitzman,S.A. (2002) Chronic inflammation and cancer. *Oncology*, 16, 217–226

20. Hussain SP, Hofseth LJ, Harris CC. 2003. Radical causes of cancer. Nat Rev Cancer 3:276 –285.

21. D.A. Geller and T.R. Billiar Molecular biology of nitric oxide synthases, *Cancer Metastasis Rev.* 17 (1998), pp. 7–23.

22. Malinski and Z. Taha, Nitric oxide release from a single cell measured in situ by a porphyrinic-based microsensor, *Nature* 358 (1992), pp. 676–715.

23. Lewis,R.S., Tamir,S., Tannenbaum,S.R. and Deen,W.M. (1995) Kinetic analysis of the fate of nitric oxide synthesized by macrophages *in vitro*. *J. Biol. Chem.*, 270, 29350–29355.

24. Oxides of nitrogen, *WHO Environ. Health Criteria* 4 (1977).

25. Nitrogen oxides, *WHO Environ. Health Criteria* 188 (1997).

26. D. Bokhoven and H.J. Niessen, Amounts of oxides of nitrogen and carbon monoxide in cigarette smoke, with and without inhalation, *Nature* 192 (1961), pp. 458–459.

27. K. Yoshida and K. Kasama, Biotransformation of nitric oxide, *Environ. Health Perspect.* 73 (1987), pp. 201–206.

28. Singer,B. and Grunberger,D. (1983) *Molecular Biology of Mutagens and Carcinogens*. Plenum, New York.

29. Coulondre,C., Miller,J.H., Farabaugh,P.J. and Gilbert,W. (1978) Molecular basis of base substitution hotspots in *Escherichia coli*. *Nature*, 274, 775–780.

30. Duncan,B.K. and Miller,J.H. (1980) Mutagenic deamination of cytosine residues in DNA. *Nature*, 287, 560–561.

31. Muramatsu M, Sankaranand VS, Anant S, Sugai M, Kinoshita K, Davidson NO, Honjo T. Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells J Biol Chem. 1999 Jun 25;274(26):18470-6

32. R. Bransteitter, P. Pham, M. D. Scharff, and M. F. Goodman Activation-induced cytidine deaminase deaminates deoxycytidine on single-stranded DNA but requires the action of RNase. PNAS, April 1, 2003; 100(7): 4102 – 4107

33. Poltoratsky VP, Wilson SH, Kunkel TA, Pavlov YI. Recombinogenic phenotype of human activation-induced cytosine deaminase. J Immunol 2004;172: 4308–4313.

34. Revy P, et al. Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2). Cell 2000;102: 565–575.

35. Muramatsu M, Kinoshita K, Fagarasan S, Yamada S, Shinkai Y, Honjo T. Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. Cell 2000;102:553–563.

36. Quartier P, et al. Clinical, immunologic and genetic analysis of 29 patients with autosomal recessive hyper-IgM syndrome due to Activation-Induced Cytidine Deaminase deficiency. Clin Immunol 2004;110:22–29.

37. Shikina T, et al. IgA class switch occurs in the organized nasopharynx- and gutassociated lymphoid tissue, but not in the diffuse lamina propria of airways and gut. J Immunol 2004;172:6259–6264.

38. Fagarasan S, Kinoshita K, Muramatsu M, Ikuta K, Honjo T. In situ class switching and differentiation to IgA-producing cells in the gut lamina propria. Nature 2001;413: 639–643.

39. Fagarasan S, Muramatsu M, Suzuki K, Nagaoka H, Hiai H, Honjo T. Critical roles of activation-induced cytidine deaminase in the homeostasis of gut flora. Science 2002;298: 1424–1427.

40. Kozlov AV, Svetlikova SB. Properties and composition of DNP of *Escherichia coli* B Mol Biol. 1974 Mar;7(5):519-26.

41. Williams MV, Chang CH, Cheng YC. An enzymatic method for distinguishing deoxyuridine and deoxythymidine nucleotide pools and its application for determining ribonucleotide reductase activity. J Biochem Biophys Methods. 1979 July;1(3):153-62.

42. Wang L, Weiss B. dcd (dCTP deaminase) gene of *Escherichia coli*: mapping, cloning, sequencing, and identification as a locus of suppressors of lethal dut (dUTPase) mutations. J Bacteriol. 1992 Sep;174(17):5647-53.

43. Taylor AF, Siliciano PG, Weiss B. Cloning of the dut (deoxyuridine triphosphatase) gene of *Escherichia coli*. Gene. 1980 May;9(3-4):321-36

44. Laurent PL, Tevaearai HT, Eliason JF, Givel JC, Odartchenko N. Interferon alpha and 5'-deoxy-5-fluorouridine in colon cancer: effects as single agents and in combination on growth of xenograft tumours. Eur J Cancer. 1994;30A(12):1859-65.

45. Otterlei,M., Kavli,B., Standal,R., Skjelbred,C., Bharati,S. and Krokan,H.E. (2000) Repair of chromosomal abasic sites *in vivo* involves at least three different repair pathways. *EMBO J.*19**,** 5542–5551

46. P.O.nyman. Introduction Current Protein and Peptide Science, 2001, *2*, 277-285

47. McIntosh EM, Ager DD, Gadsden MH, Haynes RH. Human dUTP pyrophosphatase: cDNA sequence and potential biological importance of the enzyme Proc Natl Acad Sci U S A. 1992 Sep 1;89(17):8020-4

48. Gadsden MH, McIntosh EM, Game JC, Wilson PJ, Haynes RH. dUTP pyrophosphatase is an essential enzyme in *Saccharomyces cerevisiae* EMBO J. 1993 Nov;12(11):4425-31

49. Taylor AF, Siliciano PG, Weiss B. Cloning of the dut (deoxyuridine triphosphatase) gene of *Escherichia coli*. Gene. 1980 May;9(3-4):321-36

50. Preston VG, Fisher FB. Identification of the herpes simplex virus type 1 gene encoding the dUTPase Virology. 1984 Oct 15;138(1):58-68

51. Bernier-Villamor V, Camacho A, Hidalgo-Zarco F, Perez J, Ruiz-Perez LM, Gonzalez-Pacanowska D. Characterization of deoxyuridine 5'-triphosphate nucleotidohydrolase from Trypanosoma cruzi. FEBS Lett. 2002 Aug 28;526(1-3):147-50

52. Camacho A, Arrebola R, Pena-Diaz J, Ruiz-Perez LM, Gonzalez-Pacanowska D. Description of a novel eukaryotic deoxyuridine 5'-triphosphate nucleotidohydrolase in Leishmania major. Biochem J. 1997 Jul 15;325 ( Pt 2):441-7.

53. Hidalgo-Zarco, F.; Camacho, A.G.; Bernier-Villamor, V.; Nord, J.; Ruiz-Perez, L.M. and González-Pacanowska, D. Kinetic properties and inhibition of the dimeric dUTPase-dUDPase from Leishmania major (2001) *Prot. Sci., 10*, 1426-1433.

54. Bernier-Villamor V, Camacho A, Hidalgo-Zarco F, Perez J, Ruiz-Perez LM, Gonzalez-Pacanowska D. Characterization of deoxyuridine 5'-triphosphate nucleotidohydrolase from *Trypanosoma cruzi*. FEBS Lett. 2002 Aug 28;526(1-3):147-50.

55. Krokan HE, Kavli B, Slupphaug G. Novel aspects of macromolecular repair and relationship to human disease. J Mol Med. 2004 May;82(5):280-97. Epub 2004 Feb 24

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

57. Duncan,B.K. and Chambers,J.A. The cloning and overproduction of *Escherichia coli* uracil-DNA glycosylase (1984) *Gene*, 28, 211-219.

58. Varshney,U., Hutcheon,T. and van de Sande,J.H. Sequence analysis, expression, and conservation of *Escherichia coli* uracil DNA glycosylase and its gene (ung). (1988) *J. Biol. Chem.*, 263, 7776-7784

59. Lindahl T. DNA repair enzymes. Annu Rev Biochem. 1982;51:61-87.

60. Krokan HE, Standal R, Slupphaug G. DNA glycosylases in the base excision repair of DNA Biochem J. 1997 Jul 1;325 ( Pt 1):1-16

61. I Takahashi and J. Marmur, Replacement of thymidylic acid by deoxyuridylic acid in the deoxyribonucleic acid of a transducing phage for *Bacillus subtilis*. *Nature* 197 (1963), pp. 794–795

62. G. Vieria, H. de Lencastre and L. Archer Restriction analysis of PBS 1-related phages. *Arch. Virol.* 106 (1989), pp. 121–126

63. E.C. Friedberg, A.K. Ganeson and K. Minton, *N*-Glycosidase activity in extracts of *Bacillus subtilis* and its inhibition after infection with bacteriophage PBS2. *J. Virol.* 16 (1975), pp. 315–321.

64. R. Cone, T. Bonura and E.C. Friedberg, Inhibitor of uracil-DNA glycosylase induced by bacteriophage PBS2. Purification and preliminary characterization. *J. Biol. Chem.* 255 (1980), pp. 10354–10358.

65. Mosbaugh, D. W., and Bennett, S. E. Uracil-excision DNA repair (1994) Prog. Nucleic Acid Res. Mol. Biol. 48, 315-370.

66. Parikh, S. P., Mol, C. D., Slupphaug, G., Bharati, S., Krokan, H. E., and Tainer, J. A. Base excision repair initiation revealed by crystal structures and binding kinetics of human uracil-DNA glycosylase with DNA. (1998) *EMBO J. 17*, 5214-5226

67. J.T. Stivers, K.W. Pankiewicz and K.A. Watanabe, Kinetic mechanism of damage site recognition and uracil flipping by *Escherichia coli* uracil-DNA glycosylase. *Biochemistry* 38 (1999), pp. 952–963.

68. C.D. Mol, S.S. Parikh, C.D. Putnam, T.P. Lo and J.A. Tainer, DNA repair mechanisms for the recognition and removal of damaged DNA bases. *Annu. Rev. Biophys. Biomol. Struct.* 28 (1999), pp. 101–128.

69. Gallinari, P., and Jiricny, J. (1996) A new class of uracil-DNA glycosylases related to human thymine-DNA glycosylase. Nature 383: 735–738.

70. Saparbaev, M., and Laval, J. (1998) 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 USA 95: 8508–8513.

71. Bartsch, H., and Nair, J. (2000) New DNA-based biomarkersfor oxidative stress and cancer chemoprevention studies. Eur J Cancer 36: 1229–1234.

72. Lutsenko, E., and Bhagwat, A.S. (1999) The role of the *Escherichia coli* mug protein in the removal of uracil and 3,N-ethenocytosine from DNA. J Biol Chem 274: 31034–31038.

73. Barrett, T. E., Savva, R., Panayotou, G., Barlow, T., Brown, T., Jiricny, J., and Pearl, L. H. Crystal structure of a G:T/U mismatch-specific DNA glycosylase: mismatch recognition by complementary-strand interactions (1998) *Cell* 92, 117–129

74. Baker, D., Liu, P., Burdzy, A., and Sowers, L. C. Characterization of the substrate specificity of a human 5-hydroxymethyluracil glycosylase activity (2002) *Chem. Res. Toxicol.* 15, 33–39

75. Saparbaev, M., Langouet, S., Privezentzev, C. V., Guengerich, F. P., Cai, H., Elder, R. H., and Laval, J. 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 (2002) *J. Biol. Chem.* **277,** 26987–26993

76. Hang, B., Downing, G., Guliaev, A. B., and Singer, B. Novel activity of Escherichia coli mismatch uracil-DNA glycosylase (Mug) excising 8-(hydroxymethyl)-3,N4-ethenocytosine, a potential product resulting from glycidaldehyde reaction. (2002) *Biochemistry*

77. Sekiguchi M, Hayakawa M, Makino F, Tanaka K and Okada Y. (1976). Biochem. Biophys. Res. Commun., 73, 293 -299.

78. Olsen LC, Aasland R,Wittwer CU, Krokan HE and Helland DE. (1989). Molecular cloning of human uracil-DNA glycosylase, a highly conserved DNA repair enzyme EMBO J., 8, 3121-3125

79. Wittwer CU, Bauw G and Krokan HE. Purification and determination of the NH2-terminal amino acid sequence of uracil-DNA glycosylase from human placenta. (1989). Biochemistry, 28, 780-784.

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

81. Nilsen H, Otterlei M, Haug T, Solum K, Nagelhus TA, Skorpen F and Krokan HE. (1997). Nucleic Acids Res., 25, 750-755.

82. Bharati S, Krokan HE, Kristiansen L, Otterlei M and Slupphaug G. Human mitochondrial uracil-DNA glycosylase preform (UNG1) is processed to two forms one of which is resistant to inhibition by AP sites (1998). Nucleic Acids Res., 26, 4953-4959

83. Nagelhus,T., Haug,T., Singh,K.K., Keshav,K.F., Skorpen,F., Otterlei,M., Bharati,S., Lindmo,T., Benichou,S., Benarous,R. and Krokan,H.E. A sequence in the N-terminal region of human uracil-DNA glycosylase with homology to XPA interacts with the C-terminal part of the 34-kDa subunit of replication protein A (1997) *J. Biol. Chem.*, 272, 6561-6566

84. Otterlei M, Warbrick E, Nagelhus TA, Haug T, Slupphaug G, AkbariM, Aas PA, Steinsbekk K, Bakke O and Krokan HE. Post-replicative base excision repair in replication foci (1999). EMBO J., 18, 3834-3844.

85. Nilsen H, Rosewell I, Robins P, Skjelbred CF, Andersen S, Slupphaug G, Daly G, Krokan HE, Lindahl T and Barnes DE. Uracil-DNA glycosylase (UNG)-deficient mice reveal a primary role of the enzyme during DNA replication. (2000). Mol. Cell, 5, 1059-1065

86. Otterlei M, Warbrick E, Nagelhus TA, Haug T, Slupphaug G, AkbariM, Aas PA, Steinsbekk K, Bakke O and Krokan HE. Post-replicative base excision repair in replication foci (1999). EMBO J., 18, 3834-3844

87. Kavli B, Sundheim O, Akbari M, Otterlei M, Nilsen H, Skorpen F, Aas PA, Hagen L, Krokan HE, Slupphaug G. hUNG2 is the major repair enzyme for removal of uracil from U:A matches, U:G mismatches, and U in single-stranded DNA, with hSMUG1 as a broad specificity backup J Biol Chem. 2002 Oct 18;277(42):39926-36.

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

89. D.G. Schatz, M.A. Oettinger and D. Baltimore, The V(D)J recombination activating gene, RAG-1. *Cell* 59 (1989), pp. 1035–1048

90. Storb U, Stavnezer J. Immunoglobulin genes: generating diversity with AID and UNG Curr Biol. 2002 Oct 29;12(21):R725-7

91. Kachhap S, Singh KK. Mitochondrial inhibition of uracil-DNA glycosylase is not mutagenic Mol Cancer. 2004 Dec 1;3:32

92. Haushalter KA, Todd Stukenberg MW, Kirschner MW, Verdine GL. Identification of a new uracil-DNA glycosylase family by expression cloning using synthetic inhibitors. : Curr Biol. 1999 Feb 25;9(4):174-85.

93. Boorstein RJ, Cummings Jr A, Marenstein DR, Chan MK, Ma Y, Neubert TA, Brown SM, Teebor GW (2001) Definitive identification of mammalian 5-hydroxymethyluracil DNA N-glycosylase activity as SMUG1. *J Biol Chem* 276: 41991–41997

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

95. An Q, Robins P, Lindahl T, Barnes DE. C --> T mutagenesis and gamma-radiation sensitivity due to deficiency in the Smug1 and Ung DNA glycosylases. EMBO J. 2005 Jun 15;24(12):2205-13. Epub 2005 May 19

96. Rada C , Di Noia JM , Neuberger MS (2004) Mismatch recognition and uracil excision provide complementary paths to both Ig switching and the a/t-focused phase of somatic mutation. *Mol Cell* 16: 163–171

97. Nilsen H, Haushalter KA, Robins P, Barnes DE, Verdine GL, Lindahl T (2001) Excision of deaminated cytosine from the vertebrate genome: role of the SMUG1 uracil-DNA glycosylase. *EMBO J* 20: 4278–4286

98. Hendrich B and Bird A. Identification and characterization of a family of mammalian methyl-CpG binding proteins (1998). Mol. Cell. Biol., 18, 6538-6547

99. Hendrich,B., Hardeland,U., Ng,H.-H., Jiricny,J. and Bird,A. (1999) The thymine glycosylase MBD4 can bind to the product of deamination at methylated CpG sites. *Nature*, 401, 301–304

100. Petronzelli,F., Riccio,A., Markham,G.D., Seeholzer,S.H., Stoerker,J., Genuardi,M., Yeung,A.T., Matsumoto,Y. and Bellacosa,A. (2000) Biphasic kinetics of the human DNA repair protein MED1 (MBD4), a mismatch-specific DNA N-glycosylase. *J. Biol. Chem.*, 275, 32422–32429

101. Petronzelli,F., Riccio,A., Markham,G.D., Seeholzer,S.H., Genuardi,M., Karbowski,M., Yeung,A.T., Matsumoto,Y. and Bellacosa,A. (2000) Investigation of the substrate spectrum of the human mismatch-specific DNA N-glycosylase MED1 (MBD4): fundamental role of the catalytic domain. *J. Cell. Physiol.*, 185, 473–480.

102. A., Cicchillitti, L., Schepis, F., Riccio, A., Yeung, A. T., Matsumoto, Y., Golemis, E. A., Genuardi, M. & Neri, G. MED1, a novel human methyl-CpG-binding endonuclease, interacts with DNA mismatch repair protein MLH1. (1999) *Proc. Natl. Acad. Sci. USA* 96, 3969–3974.

103. Cortellino, S., Turner, D., Masciullo, V., Schepis, F., Albino, D., Daniel, R., Skalka, A. M., Meropol, N. J., Alberti, C., Larue, L. & Bellacosa, A. The base excision repair enzyme MED1 mediates DNA damage response to antitumor drugs and is associated with mismatch repair system integrity (2003) *Proc. Natl. Acad. Sci. USA* 100, 15071–15076

104. Sansom OJ, Zabkiewicz J, Bishop SM, Guy J, Bird A and Clarke AR. MBD4 deficiency reduces the apoptotic response to DNA-damaging agents in the murine small intestine. (2003b). Oncogene, 22, 7130–7136.

105. Wong, E., Yang, K., Kuraguchi, M., Werling, U., Avdievich, E., Fan, K., Fazzari, M., Jin, B., Brown, A. M., Lipkin, M. & Edelmann, W. Mbd4 inactivation increases C→T transition mutations and promotes gastrointestinal tumor formation (2002) *Proc. Natl. Acad. Sci. USA* 99, 14937–14942.

106. Neddermann P, Jiricny J. The purification of a mismatch-specific thymine-DNA glycosylase from HeLa cells J Biol Chem. 1993 Oct 5;268(28):21218-24

107. Neddermann P, Gallinari P, Lettieri T, Schmid D, Truong O, Hsuan JJ, Wiebauer K, Jiricny J. Cloning and expression of human G/T mismatch-specific thymine-DNA glycosylase J Biol Chem. 1996 May 31;271(22):12767-74.

108. Hardeland U, Bentele M, Lettieri T, Steinacher R, Jiricny J, Schar P. Thymine DNA glycosylase. Prog Nucleic Acid Res Mol Biol. 2001;68:235-53.

109. Hardeland U, Bentele M, Jiricny J and Schar P. Separating substrate recognition from base hydrolysis in human thymine DNA glycosylase by mutational analysis. (2000). J. Biol. Chem., 275, 33449-33456.

110. Waters TR, Gallinari P, Jiricny J and Swann PF. Human thymine DNA glycosylase binds to apurinic sites in DNA but is displaced by human apurinic endonuclease 1. (1999). J. Biol. Chem., 274, 67-74.

111. Hardeland U, Steinacher R, Jiricny J, Schar P. Modification of the human thymine-DNA glycosylase by ubiquitin-like proteins facilitates enzymatic turnover. EMBO J. 2002 Mar 15;21(6):1456-64.

112. Vanderheiden, B. S. Inosine di- and triphosphate synthesis in erythrocytes and cell extracts (1979) *J. Cell. Physiol.* 99, 287–301

113. Mohandas, T., Sparkes, R. S., Passage, M. B., Sparkes, M. C., Miles, J. H., and Kaback, M. M. Regional mapping of ADA and ITP on human chromosome 20: cytogenetic and somatic cell studies in an X/20 translocation. (1980) *Cytogenet. Cell Genet.* 26, 28–35

114. Holmes SL, Turner BM, Hirschhorn K. Human inosine triphosphatase: catalytic properties and population studies. Clin Chim Acta. 1979 Oct 1;97(2-3):143-53.

115. Van Waeg G, Niklasson F, Ericson A, de Verdier CH. Purine metabolism in normal and ITP-pyrophosphohydrolase-deficient human erythrocytes. Clin Chim Acta. 1988 Feb 15;171(2-3):279-92.

116. Fraser JK Meyers H, Henderson JF, Brox LW, McCoy EE (1975) Individual variation in inosine triphosphate accumulation in human erythrocytes. Clin Biochem 8:353–364

117. K.Y. Hwang, J.H. Chung, S.H. Kim, Y.S. Han and Y. Cho, Structure-based identification of a novel NTPase from *Methanococcus jannaschii*, *Nat. Struct. Biol.* 6 (1999), pp. 691–696.

118. Lin, A.G. McLennan, K. Ying, Z. Wang, S. Gu, H. Jin, C. Wu, W. Liu, Y. Yuan and R. Tang *et al.*, Cloning, expression, and characterization of a human inosine triphosphate pyrophosphatase encoded by the itpa gene, *J. Biol. Chem.* 276 (2001), pp. 18695–18701.

119. J.H. Chung, H.Y. Park, J.H. Lee and Y. Jang, Identification of the dITP- and XTP-hydrolyzing protein from *Escherichia coli*, *J. Biochem. Mol. Biol.* 35 (2002), pp. 403–408.

120. Zheng J, Singh VK, Jia Z. Identification of an ITPase/XTPase in *Escherichia coli* by structural and biochemical analysis: Structure. 2005 Oct;13 (10):1511-20.

121. Wang,J.K. and Morris,A.J. (1974) The distribution of nucleoside triphosphate pyrophosphohydrolase in the tissues of the rabbit. *Arch. Biochem. Biophys*., 161, 118–124

122. Riazuddin S, Lindahl T. Properties of 3-methyladenine-DNA glycosylase from *Escherichia coli.* Biochemistry. 1978 May 30;17(11):2110-8.

123. Clarke ND, Kvaal M, Seeberg E. Cloning of *Escherichia coli* genes encoding 3-methyladenine DNA glycosylases I and II. Mol Gen Genet. 1984;197(3):368-72.

124. Nakabeppu Y, Kondo H, Sekiguchi M. Cloning and characterization of the alkA gene of *Escherichia coli* that encodes 3-methyladenine DNA glycosylase II. J Biol Chem. 1984 Nov 25;259(22):13723-9.

125. J. Labahn, O.D. Scharer, A. Long, K. Ezaz-Nikpay, G.L. Verdine and T.E. Ellenberger, Structural basis for the excision repair of alkylation-damaged DNA. *Cell* 86 (1996), pp. 321–329.

126. Y. Guan, R.C. Manuel, A.S. Arvai, S.S. Parikh, C.D. Mol, J.H. Miller, S. Lloyd and J.A. Tainer, MutY catalytic core, mutant and bound adenine structures define specificity for DNA repair enzyme superfamily. *Nat. Struct. Biol.* 5 (1998), pp. 1058–1064.

127. Thomas L, Yang CH, Goldthwait DA. Two DNA glycosylases in *Escherichia coli* which release primarily 3-methyladenine. Biochemistry. 1982 Mar 16;21(6):1162-9.

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

129. Habraken Y, Carter CA, Sekiguchi M, Ludlum DB. Release of N2,3-ethanoguanine from haloethylnitrosourea-treated DNA by *Escherichia coli* 3-methyladenine DNA glycosylase II. Carcinogenesis. 1991 Oct;12(10):1971-3.

130. Saparbaev M, Kleibl K, Laval J. *Escherichia coli, Saccharomyces cerevisiae*, rat and human 3-methyladenine DNA glycosylases repair 1,N6-ethenoadenine when present in DNA. Nucleic Acids Res. 1995 Sep 25;23(18):3750-5.

131. Saparbaev M, Laval J. Excision of hypoxanthine from DNA containing dIMP residues by the *Escherichia coli*, yeast, rat, and human alkylpurine DNA glycosylases. Proc Natl Acad Sci U S A. 1994 Jun 21;91(13):5873-7.

132. Terato H, Masaoka A, Asagoshi K, Honsho A, Ohyama Y, Suzuki T, Yamada M, Makino K, Yamamoto K, Ide H. Novel repair activities of AlkA (3-methyladenine DNA glycosylase II) and endonuclease VIII for xanthine and oxanine, guanine lesions induced by nitric oxide and nitrous acid. Nucleic Acids Res. 2002 Nov 15;30(22):4975-84.

133. Yamagata Y, Kato M, Odawara K, Tokuno Y, Nakashima Y, Matsushima N, Yasumura K, Tomita K, Ihara K, Fujii Y, Nakabeppu Y, Sekiguchi M, Fujii S. Three-dimensional structure of a DNA repair enzyme, 3-methyladenine DNA glycosylase II, from *Escherichia coli*. Cell. 1996 Jul 26;86(2):311-9.

134. O. Sidorkina, M. Saparbaev and J. Laval , Effects of nitrous acid treatment on the survival and mutagenesis of *Escherichia coli* cells lacking base excision repair (hypoxanthine-DNA glycosylase-ALK A protein) and/or nucleotide excision repair. *Mutagenesis* 12 (1997), pp. 23–28.

135. B. Demple and S. Linn , On the recognition and cleavage mechanism of Escherichia endodeoxyribonuclease V, a possible DNA repair enzyme. *J. Biol. Chem.* 257 (1982), pp. 2848–2855

136. M. Yao, Z. Hatahet, R.J. Melamede and Y.W. Kow , Purification and characterization of a novel deoxyinosine-specific enzyme, deoxyinosine 3′ endonuclease, from *Escherichia coli*. *J. Biol. Chem.* 269 (1994), pp. 16260–16268.

137. L. Samson, B. Derfler, M. Boosalis and K. Call , Cloning and characterization of a 3-methyladenine DNA glycosylase cDNA from human cells whose gene maps to chromosome 16. *Proc. Natl. Acad. Sci. U.S.A.* 83 (1991), pp. 5607–5610.

138. D. Chakravarti, G.C. Ibeanu, K. Tano and S. Mitra , Cloning expression in *Escherichia coli* of a human cDNA encoding the DNA repair protein *N*-methylpurine-DNA glycosylase. *J. Biol. Chem.* 266 (1991), pp. 15710–15715.

139. T.R. O'Connor and J. Laval , Human cDNA expressing a functional DNA glycosylase excising 3-methyladenine and 7-methylguanine. *Biochem. Biophys. Res. Commun.* 176 (1991), pp. 1170–1177.

140. G.E. Wuenschell, T.R. O'Connor and J. Termini, Stability, miscoding potential, and repair of 2′-deoxyxanthosine in DNA: implications for nitric oxide-induced mutagenesis, *Biochemistry* 42 (2003), pp. 3608–3616.

141. Bessho T, Roy R, Yamamoto K, Kasai H, Nishimura S, Tano K, Mitra S Repair of 8-hydroxyguanine in DNA by mammalian N-methylpurine-DNA glycosylase. Proc Natl Acad Sci USA 1993;90:8901–8904.

142. Hitchcock TM, Dong L, Connor EE, Meira LB, Samson LD, Wyatt MD, Cao W. Oxanine DNA glycosylase activity from Mammalian alkyladenine glycosylase. J Biol Chem. 2004 Sep 10;279(37):38177-83. Epub 2004 Jul 7.

143. Nakano T, Katafuchi A, Shimizu R, Terato H, Suzuki T, Tauchi H, Makino K, Skorvaga M, Van Houten B, Ide H. Repair activity of base and nucleotide excision repair enzymes for guanine lesions induced by nitrosative stress.Nucleic Acids Res. 2005 Apr 14;33(7):2181-91. Print 2005

144. M.K. Dosanjh, R. Roy, S. Mitra and B. Singer, 1,$N^6$-ethenoadenine is preferred over 3-methyladenine as substrate by a cloned human N-methylpurine-DNA glycosylase (3-methyladenine-DNA glycosylase). *Biochemistry* 33 (1994), pp. 1624–1628.

145. B. Hang, A. Chenna, S. Rao and B. Singer, 1,$N^6$-ethenoadenine and 3,$N^4$-ethenocytosine are excised by separate human DNA glycosylases. *Carcinogenesis* 17 (1996), pp. 155–157.

146. W.B. Mattes, C.-S. Lee, J. Laval and T.R. O'Connor, Excision of DNA adducts of nitrogen mustards by bacterial and mammalian 3-methyladenine-DNA glycosylases. *Carcinogenesis* **17** (1996), pp. 643–648.

147. Lau, A. Y., Scharer, O. D., Samson, L., Verdine, G. L., and Ellenberger, T. Crystal Structure of a Human Alkylbase-DNA Repair Enzyme Complexed to DNA (1998) *Cell 95*, 249-258.

148. B.P. Engelward, G. Weeda, M.D. Wyatt, J.L. Broekhof, J. de Wit, I. Donker, J.M. Allan, B. Gold, J.H. Hoeijmakers and L.D. Samson, Base excision repair deficient mice lacking the Aag alkyladenine DNA glycosylase. *Proc. Natl. Acad. Sci. USA* 94 (1997), pp. 13087–13092.

149. Hang, B. Singer, G.P. Margison and R.H. Elder, Targeted deletion of alkylpurine-DNA-N-glycosylase in mice eliminates repair of 1,$N^6$-ethenoadenine and hypoxanthine but not of 3,$N^4$-ethenocytosine or 8-oxoguanine. *Proc. Natl. Acad. Sci. USA* 94 (1997), pp. 12869–12874.

150. Xia L, Zheng L, Lee HW, Bates SE, Federico L, Shen B, O'Connor TR. Human 3-methyladenine-DNA glycosylase: effect of sequence context on excision, association with PCNA, and stimulation by AP endonuclease. J Mol Biol. 2005 Mar 11;346(5):1259-74. Epub 2005 Jan 20.

151. F. Miao, M. Bouziane, R. Dammann, C. Masutani, F. Hanaoka, G. Pfeifer and T.R. O'Connor, 3-Methyladenine-DNA glycosylase (MPG protein) interacts with human RAD23 proteins, *J. Biol. Chem.* 275 (2000), pp. 28433–28438.

152. Likhite VS, Cass EI, Anderson SD, Yates JR, Nardulli AM. Interaction of estrogen receptor alpha with 3-methyladenine DNA glycosylase modulates transcription and DNA repair. J Biol Chem. 2004 Apr 16;279 (16):16875-82. Epub 2004 Feb 3.

153. Watanabe, T. Ichimura, N. Fujita, S. Tsuruzoe, I. Ohki and M. Shirakawa *et al.*, Methylated DNA-binding domain 1 and methylpurine-DNA glycosylase link transcriptional repression and DNA repair in chromatin, *Proc. Natl Acad. Sci. USA* 100 (2003), pp. 12859–12864.

154. Aamodt RM, Falnes PO, Johansen RF, Seeberg E, Bjoras M. The Bacillus subtilis counterpart of the mammalian 3-methyladenine DNA glycosylase has hypoxanthine and 1,N6-ethenoadenine as preferred substrates. J Biol Chem. 2004 Apr 2;279(14):13601-6.

155. Vongchampa V, Dong M, Gingipalli L, Dedon P. Stability of 2'-deoxyxanthosine in DNA Nucleic Acids Research, 2003, Vol. 31, No. 3 1045-1051

156. Friedkin, M. Enzymic synthesis of deoxyxanthosine by the action of xanthosine phosphorylase in mammalian tissue (1952) *J. Am. Chem. Soc. 74*, 112-115

157. Suzuki, T., Matsumura, Y., Ide, H., Kanaori, K., Tajima, K., and Makino, K. Deglycosylation susceptibility and base-pairing stability of 2'-deoxyoxanosine in oligodeoxynucleotide. (1997) *Biochemistry 36*, 8013-8019

158. Eritja, R., Horowitz, D. M., Walker, P. A., Ziehler-Martin, J. P.,Boosalis, M. S., Goodman, M. F., Itakura, K., and Kaplan, B. E. Synthesis and properties of oligonucleotides containing 2'-deoxynebularine and 2'-deoxyxanthosine (1986) *Nucleic Acids Res. 14*, 8135-8153.

159. Yasui M, Suzuki N, Miller H, Matsuda T, Matsui S, Shibutani S. Translesion synthesis past 2'-deoxyxanthosine, a nitric oxide-derived DNA adduct, by mammalian DNA polymerases. J Mol Biol. 2004 Nov 26;344(3):665-74.

160. H. Kamiya, T. Sakaguchi, N. Murata, M. Fujimuro, H. Miura and H. Ishikawa *et al.*, *In vitro* replication study of modified bases in ras sequences, *Chem. Pharm. Bull.* 40 (1992), pp. 2792–2795.

161. He B, Qing H, Kow YW. Deoxyxanthosine in DNA is repaired by *Escherichia coli* endonuclease V. Mutat Res. 2000 Mar 20;459 (2):109-14.

162. M. Yao and Y.W. Kow, Cleavage of insertion/deletion mismatches, flap and pseudo-Y DNA structures by deoxyinosine 3′-endonuclease from *Escherichia coli*. *J. Biol. Chem.* 271 (1996), pp. 30672–30676.

163. M. Yao and Y.W. Kow, Further characterization of *Escherichia coli* endonuclease V. Mechanism of recognition for deoxyinosine, deoxyuridine, and base mismatches in DNA. *J. Biol. Chem.* 272 (1997), pp. 30774–30779.

164. D. Jiang, Z. Hatahet, J.O. Blaisdell, R.J. Melamede and S.S. Wallace, *Escherichia coli* endonuclease VIII: cloning, sequencing, and overexpression of the *nei* structural gene and characterization of *nei* and *nei nth* mutants. *J. Bacteriol.* 179 (1997), pp. 3773–3782

165. Y. Saito, F. Uraki, S. Nakajima, A. Asaeda, K. Ono, K. Kubo and K. Yamamoto, Characterization of endonuclease III (*nth*) and endonuclease VIII (*nei*) mutants of *Escherichia coli* K-12. *J. Bacteriol.* 179 (1997), pp. 3783–3785.

166. S.S. Wallace, L. Harrison, D. Jiang, J.O. Blaisdell, A.A. Purmal, Z. Hatahet, Processing and consequences of oxidative DNA base lesions, in: M. Dizaroglu, A.E. Karakaya (Eds.), Advances in DNA Damage and Repair: Oxygen Radical Effects, Cellular Protection, and Biological Consequences, Plenum Publishers, New York, 1999, pp. 419–430

167. J.O. Blaisdell, Z. Hatahet and S.S. Wallace, A novel role for *Escherichia coli* endonuclease VIII in the prevention of spontaneous G → T transversions. *J. Bacteriol.* 181 (1999), pp. 6396–6402.

168. Melamede R.J., Hatahet,Z., Kow,Y.W., Ide,H. and Wallace,S.S. (1994) Isolation and characterization of endonuclease VIII from *Escherichia coli*. *Biochemistry*, 33, 1255–1264.

169. D. Jiang, Z. Hatahet, R.J. Melamede, Y.W. Kow and S.S. Wallace, Characterization of *Escherichia coli* endonuclease VIII. *J. Biol. Chem.* 272 (1997), pp. 32230–32239.

170. A.A. Purmal, G.W. Lampman, J.P. Bond, Z. Hatahet and S.S. Wallace, Enzymatic processing of uracil glycol, a major oxidative product of DNA cytosine. *J. Biol. Chem.* 273 (1998), pp. 10026–10035.

171. Asagoshi, K., Yamada, T., Okada, Y., Terato, H., Ohyama, Y., Seki, S., Ide, H. (2000) Recognition of formamidopyrimidine by *Escherichia coli* and mammalian thymine glycol glycosylases. Distinctive paired base effects and biological and mechanistic implications *J. Biol. Chem.*, 275, 24781–24786

172. T.K. Hazra, T. Izumi, I. Boldogh, B. Imhoff, Y.W. Kow, P. Jaruga, M. Dizdaroglu and S. Mitra, Identification and characterization of a human DNA glycosylase for repair of modified bases in oxidatively damaged DNA. *Proc. Natl. Acad. Sci. U.S.A.* 99 (2002), pp. 3523–3528.

173. T.K. Hazra, Y.W. Kow, Z. Hatahet, B. Imhoff, I. Boldogh, S.K. Mokkapati, S. Mitra and T. Izumi, Identification and characterization of a novel human DNA glycosylase for repair of cytosine derived lesions. *J. Biol. Chem.* 277 (2002), pp. 30417–30420.

174. V. Bandaru, S. Sunkara, S.S. Wallace and J.P. Bond, A novel human DNA glycosylase that removes oxidative DNA damage and is homologous to *Escherichia coli* endonuclease VIII. *DNA Repair* 1 (2002), pp. 517–529.

175. T.K. Hazra, T. Izumi, I. Boldogh, B. Imhoff, Y.W. Kow, P. Jaruga, M. Dizdaroglu and S. Mitra, Identification and characterization of a human DNA glycosylase for repair of modified bases in oxidatively damaged DNA. *Proc. Natl. Acad. Sci. U.S.A.* 99 (2002), pp. 3523–3528.

176. Demple, B. and Linn, S. (1980) DNA *N*-glycosylases and UV repair *Nature*, 287, 203–208

177. Dizdaroglu, M., Laval, J., Boiteux, S. (1993) Substrate specificity of the *Escherichia coli* endonuclease III: excision of thymine- and cytosine-derived lesions in DNA produced by radiation-generated free radicals *Biochemistry*, 32, 12105–12111

178. Speina, E., Ciesla, J.M., Wojcik, J., Bajek, M., Kusmierek, J.T., Tudek, B. (2001) The pyrimidine ring-opened derivative of 1,$N^6$-ethenoadenine is excised from DNA by the *Escherichia coli* Fpg and Nth proteins *J. Biol. Chem*, . 276, 21821–21827

179. Matsumoto, Y., Zhang, Q.M., Takao, M., Yasui, A., Yonei, S. (2001) *Escherichia coli* Nth and human hNTH1 DNA glycosylases are involved in removal of 8-oxoguanine from 8-oxoguanine/guanine mispairs in DNA *Nucleic Acids Res*, 29, 1975–1981

180. Hazra, T.K., Muller, J.G., Manuel, R.C., Burrows, C.J., Lloyd, R.S., Mitra, S. (2001) Repair of hydantoins, one electron oxidation product of 8-oxoguanine, by DNA glycosylases of *Escherichia coli Nucleic Acids Res*, . 29**,** 1967–1974

181. Laspia, M.F. and Wallace, S.S. (1988) Excision repair of thymine glycols, urea residues, and apurinic sites in *Escherichia coli J. Bacteriol*, . 170**,** 3359–3366

182. Hatahet, Z., Kow, Y.W., Purmal, A.A., Cunningham, R.P., Wallace, S.S. (1994) 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*, . 269, 18814–18820

183. H.L. Katcher and S.S. Wallace, Characterization of the *Escherichia coli* X-ray endonuclease, endonuclease III. *Biochemistry* 22 (1983), pp. 4071–4081.

184. L.H. Breimer and T. Lindahl, DNA glycosylase activities for thymine residues damaged by ring saturation, fragmentation, or ring contraction are functions of endonuclease III in *Escherichia coli*. *J. Biol. Chem.* 259 (1984), pp. 5543–5548.

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

186. Y.W. Kow and S.S. Wallace, Mechanism of action of *Escherichia coli* endonuclease III. *Biochemistry* 26 (1987), pp. 8200–8206

187. R.P. Cunningham, H. Asahara, J.F. Bank, C.P. Scholes, J.C. Salerno, K. Surerus, E. Munck, J. McCracken, J. Peisach and M.H. Emptage, Endonuclease III is an iron-sulfur protein. *Biochemistry* 28 (1989), pp. 4450–4455.

188. W. Fu, S. O'Handley, R.P. Cunningham and M.K. Johnson, The role of the iron-sulfur cluster in *Escherichia coli* endonuclease III. A resonance Raman study. *J. Biol. Chem.* 267 (1992), pp. 16135–16137

189. C.F. Kuo, D.E. McRee, C.L. Fisher, S.F. O'Handley, R.P. Cunningham and J.A. Tainer, Atomic structure of the DNA repair [4Fe-4S] enzyme endonuclease III. *Science* 258 (1992), pp. 434–440.

190. M.M. Thayer, H. Ahern, D. Xing, R.P. Cunningham and J.A. Tainer, Novel DNA binding motifs in the DNA repair enzyme endonuclease III crystal structure. *EMBO. J.* 14 (1995), pp. 4108–4120.

191. Yavin E, Stemp ED, O'shea VL, David SS, Barton JK. Electron trap for DNA-bound repair enzymes: a strategy for DNA-mediated signaling. Proc Natl Acad Sci U S A. 2006 Mar 7;103(10):3610-4. Epub 2006 Feb 27.

192. Tchou, J., Kasai, H., Shibutani, S., Chung, M.H., Laval, J., Grollman, A.P., Nishimura, S. (1991) 8-oxoguanine (8-hydroxyguanine) DNA glycosylase and its substrate specificity *Proc. Natl Acad. Sci. USA*, 88, 4690–4696

193. Boiteux, S., O'Connor, T.R., Laval, J. (1987) Formamidopyrimidine–DNA glycosylase of *Escherichia coli*: cloning and sequencing of the fpg structural gene and overproduction of the protein *EMBO J.*, 6, 3177–3183

194. C.J. Chetsanga and T. Lindahl, Release of 7-methylguanine residues whose imidazole rings have been opened from damaged DNA by a DNA glycosylase from *Escherichia coli*, *Nucleic Acids Res.* 6 (1979), pp. 3673–3684.

195. C.J. Wiederholt, M.O. Delaney, M.A. Pope, S.S. David and M.M. Greenberg, Repair of DNA containing FapydG and its beta –C– nucleoside analogue by formamidopyrimidine DNA glycosylase and MutY, *Biochemistry* 42 (2003), pp. 9755–9760.

196. S. Boiteux, E. Gajewski, J. Laval and M. Dizdaroglu, Substrate specificity of the *Escherichia coli* Fpg protein (formamidopyrimidine-DNA glycosylase): excision of purine lesions in DNA produced by ionizing radiation or photosensitization, *Biochemistry* 31 (1992), pp. 106–110.

197. B. Castaing, S. Boiteux and C. Zelwer, AP site structural determinants for Fpg specific recognition, *Nucleic Acids Res.* 20 (1992), pp. 389–394.

198. J.C. Fromme and G.L. Verdine, DNA lesion recognition by the bacterial repair enzyme MutM, *J. Biol. Chem.* 278 (2003), pp. 51543–51548.

199. C. D'Ham, A. Romieu, M. Jaquinod, D. Gasparutto and J. Cadet, Excision of 5,6-dihydroxy-5,6-dihydrothymine, 5,6-dihydrothymine, and 5-hydroxycytosine from defined sequence oligonucleotides by *Escherichia coli* endonuclease III and Fpg proteins: kinetic and mechanistic aspects, *Biochemistry* 38 (1999), pp. 3335–3344.

200. D. Gasparutto, M. Ait-Abbas, M. Jaquinod, S. Boiteux and J. Cadet, Repair and coding properties of 5-hydroxy-5-methylhydantoin nucleosides inserted into DNA oligomers, *Chem. Res. Toxicol.* 13 (2000), pp. 575–584.

201. T.R. O'Connor and J. Laval, Physical association of the 2,6-diamino-4-hydroxy-5*N*-formamidopyrimidine-DNA glycosylase of *Escherichia coli* and an activity nicking DNA at apurinic/apyrimidinic sites, *Proc. Natl. Acad. Sci. U.S.A.* 86 (1989), pp. 5222–5226.

202. R.J. Graves, I. Felzenwalb, J. Laval and T.R. O'Connor, Excision of 5′-terminal deoxyribose phosphate from damaged DNA is catalyzed by the Fpg protein of *Escherichia coli*, *J. Biol. Chem.* 267 (1992), pp. 14429–14435.

203. Lu AL, Lee CY, Li L, Li X. Physical and functional interactions between *Escherichia coli* MutY and endonuclease VIII Biochem J. 2006 Jan 1;393(Pt 1):381-7.

204. Shimada, N., Yagisawa, N., Naganawa, H., Takita, T., Hamada, M., Takeuchi, T., and Umezawa, H. J. (1981) Oxanosine, a novel nucleoside from actinomycetes. *J. Antibiot. 34*, 1216-1218.

205. Nakamura, H., Yagisawa, N., Shimada, N., Takita, T., Umezawa, H., and Iitaka, Y. J. (1981) The X-ray structure determination of oxanosine. *J. Antibiot. 34*, 1219-1221.

206. Yagisawa, N., Shimada, N., Takita, T., Ishizuka, M., Takeuchi, T., and Umezawa, H. J. (1982) Mode of action of oxanosine, a novel nucleoside antibiotic. *J. Antibiot. 35*, 755-759.

207. Uehara, Y., Hasegawa, M., Hori, M., and Umezawa, H. (1985) Increased sensitivity to oxanosine, a novel nucleoside antibiotic, of rat kidney cells upon expression of the integrated viral src gene. *Cancer Res. 45*, 5230-5234

208. Itoh, O., Kuroiwa, K., Atsumi, S., Umezawa, K., Takeuchi, T., and Hori, M. (1989) Induction by the guanosine analogue oxanosine of reversion toward the normal phenotype of K-ras-transformed rat kidney cells. *Cancer Res. 49*, 996-1000.

209. Majumdar P, Wu H, Tipton P, Glaser R. Oxanosine is a substrate of adenosine deaminase. Implications for the quest for a toxicological marker for nitrosation activity. Chem Res Toxicol. 2005 Dec;18(12):1830-41.

210. Suzuki, T., Yamaoka, R., Nishi, M., Ide, H., Makino, K. (1996) Isolation and characterization of a novel product, 2'-deoxyoxanosine, from 2'-deoxyguanosine, oligodeoxyribonucleotide, and calf thymus DNA treated by nitrous acid and nitric oxide *J. Am. Chem. Soc.*, 118, 2515–2516

211. Lucas, L.T., Gatehouse, D., Shuker, D.G. (1999) Efficient nitroso group transfer from *N*-nitrosoindoles to nucleotides and 2'-deoxyguanosine at physiological pH. A new pathway for *N*-nitrosocompounds to exert genotoxicity *J. Biol. Chem.*, 274, 18319–18326

212. Glaser, R. and Son, M.S. (1996) Pyrimidine ring opening in the unimolecular dediazoniation of guanine diazonium ion. An *ab initio* theoretical study of the mechanism of nitrosative guanosine deamination *J. Am. Chem. Soc* ., 118, 10942–10943

213. Rayat, S. and Glaser, R. (2003) 5-Cyanoimino-4-oxomethylene-dihydroimidazole and nitrosative guanine deamination. A theoretical study of geometries, electronic structures and N-Protonation *J. Org. Chem.*, 68, 9882–9892

214. Qian, M. and Glaser, R. (2004) 5-Cyanoamino-4-imidazolecarboxamide and nitrosative guanine deamination: Experimental evidence for pyrimidine ring-opening during deamination *J. Am. Chem. Soc.*, 126, 2274–2275

215. Dong M, Wang C, Deen WM, Dedon PC. Absence of 2'-deoxyoxanosine and presence of abasic sites in DNA exposed to nitric oxide at controlled physiological concentrations. Chem Res Toxicol. 2003 Sep;16(9):1044-55.

216. Suzuki, T., Yoshida, M., Yamada, M., Ide, H., Kobayashi, M., Kanaori, K., Tajima, K., Makino, K. (1998) Misincorporation of 2'-deoxyoxanosine 5'-triphosphate by DNA polymerases and its implication for mutagenesis *Biochemistry*, 37, 11592–11598

217. Hitchcock, T.M., Gao, H., Cao, W. (2004) Cleavage of deoxyoxanosine-containing oligodeoxyribonucleotides by bacterial endonuclease V *Nucleic Acids Res.*, 32, 4071–4080

218. Nakano, T., Katafuchi, A., Shimizu, R., Terato, H., Suzuki, T., Tauchi, H., Makino, K., Skorvaga, M., Van Houten, B., Ide, H. (2005) Repair activity of base and nucleotide excision repair enzymes for guanine lesions induced by nitrosative stress *Nucleic Acids Res.*, 33, 2181–2191

219. Pack SP, Nonogawa M, Kodaki T, Makino K. Chemical synthesis and thermodynamic characterization of oxanine-containing oligodeoxynucleotides. Nucleic Acids Res. 2005 Oct 11;33(18):5771-80.

220. Nakano T, Terato H, Asagoshi K, Masaoka A, Mukuta M, Ohyama Y, Suzuki T, Makino K, Ide H. DNA-protein cross-link formation mediated by oxanine. A novel genotoxic mechanism of nitric oxide-induced DNA damage. J Biol Chem. 2003 Jul 4;278(27):25264-72.

221. Gates, F.T. and S. Linn, *Endonuclease V of Escherichia coli.* Journal of Biological Chemistry, 1977. 252(5): p. 1647-1653.

222. Harosh, I. and J. Sperling, Hypoxanthine-DNA Glycosylase from *Escherichia coli.* Journal of Biological Chemistry, 1988. 263(7): p. 3328-3334.

223. Guo, G., Y. Ding, and B. Weiss, *nfi, the Gene for Endonuclease V in Escherichea coli K-12.* Journal of Bacteriology, 1997. 179(2): p. 310-316.

224. Guo, G. and B. Weiss, *Endonuclease V (nfi) Mutant of Escherichia coli K-12.* Journal of Bacteriology, 1998. 180(1): p. 46-51.

225. Karen A. Schouten and B. Weiss, Endonuclease V protects *Escherichia coli* against specific mutations caused by nitrous acid. Mutation Research 435 (1999) 245-254

226. Huang J, Lu J, Barany F, Cao W. Multiple cleavage activities of endonuclease V from *Thermotoga maritima*: recognition and strand nicking mechanism Biochemistry. 2001 Jul 31;40 (30):8738-48.

227. Huang J, Lu J, Barany F, Cao W. Mutational analysis of endonuclease V from *Thermotoga maritima*. Biochemistry. 2002 Jul 2;41(26):8342-50.

228. M. Yao and Y.W. Kow, Interaction of deoxyinosine 3′-endonuclease from *Escherichia coli* with DNA containing deoxyinosine. *J. Biol. Chem.* 270 (1995), pp. 28609–28616

229. M. Yao and Y.W. Kow, Strand-specific cleavage of mismatch-containing DNA by deoxyinosine 3′-endonuclease from *Escherichia coli*. J. Biol. Chem. 269 (1994), pp. 31390–31396.

230. Burgis NE, Brucker JJ, Cunningham RP. Repair system for noncanonical purines in *Escherichia coli*. J Bacteriol. 2003 May;185(10):3101-10.

231. Feng H, Klutz AM, Cao W. Active site plasticity of endonuclease V from *Salmonella typhimurium*. Biochemistry. 2005 Jan 18;44(2):675-83.

232. Feng H, Dong L, Klutz AM, Aghaebrahim N, Cao W. Defining amino acid residues involved in DNA-protein interactions and revelation of 3'-exonuclease activity in endonuclease V. Biochemistry. 2005 Aug 30;44(34):11486-95.

233. Liu J, He B, Qing H, Kow YW. A deoxyinosine specific endonuclease from hyperthermophile, Archaeoglobus fulgidus: a homolog of *Escherichia coli* endonuclease V. Mutat Res. 2000 Nov 9;461(3):169-77.

234. Moe A, Ringvoll J, Nordstrand LM, Eide L, Bjoras M, Seeberg E, Rognes T, Klungland A. Incision at hypoxanthine residues in DNA by a mammalian homologue of the *Escherichia coli* antimutator enzyme endonuclease V. Nucleic Acids Res. 2003 Jul 15;31(14):3893-900.

235. Kanugula S, Pauly GT, Moschel RC, Pegg AE. A bifunctional DNA repair protein from Ferroplasma acidarmanus exhibits O6-alkylguanine-DNA alkyltransferase and endonuclease V activities. Proc Natl Acad Sci U S A. 2005 Mar 8;102(10):3617-22.

236. Margison GP, Santibanez-Koref MF. O6-alkylguanine-DNA alkyltransferase: role in carcinogenesis and chemotherapy. Bioessays. 2002 Mar;24(3):255-66.

237. Daniels DS, Woo TT, Luu KX, Noll DM, Clarke ND, Pegg AE, Tainer JA. DNA binding and nucleotide flipping by the human DNA repair protein AGT. Nat Struct Mol Biol. 2004 Aug;11 (8):714-20.

238. Weiss, B., Endonuclease V *of Escherichia coli* prevents mutations from nitrosative deamination during nitrate/nitrite respiration. Mutat Res, 2001. 461(4): p. 301-9.

239. Weiss, B., Evidence for mutagenesis by nitric oxide during nitrite metabolism in *E. coli.* Journal of Bacterology, 2006. 188(3): p. 829-833.

240. Kow YW. Repair of deaminated bases in DNA. Free Radic Biol Med. 2002 Oct 1;33(7):886-93

241. M. Lieb and A.S. Bhagwat, Very short patch repair: reducing the cost of cytosine methylation. *Mol. Microbiol.* 20 (1996), pp. 467–473

242. Monastiriakos SK, Doiron KM, Siponen MI, Cupples CG. Functional interactions between the MutL and Vsr proteins of *Escherichia coli* are dependent on the N-terminus of Vsr DNA Repair (Amst). 2004 Jun 3;3(6):639-47

243. Stemmer WPC (1994a) DNA shuffling by random fragmentation and reassembly: *in vitro* recombination for molecular evolution. Proc Natl Acad Sci USA 91:10747-10751

244. Stemmer WPC (1994b) Rapid evolution of a protein *in vitro* by DNA shuffling. Nature370:389-391

245. Locher CP et al. DNA shuffling and screening strategies for improving vaccine efficacy. DNA Cell Biol. 2005 Apr;24(4):256-63.

246. Minshull,J and Stemmer,WP.(1999). Protein evolution by molecular-breeding. Curr.Opin.Biol. 3,:284-290

247. Patel, P.H., and L.A. Loeb. 2000. DNA polymerase active site is highly mutable: evolutionary consequences. Proc. Natl. Acad. Sci. USA 97**:**5095–5100

248. Fa, M., A. Radeghieri, A. A. Henry, and F. E. Romesberg. 2004. Expanding the substrate repertoire of a DNA polymerase by directed evolution. J. Am. Chem. Soc. 126:1748-1754

249. Xia G, Chen L, Sera T, Fa M, Schultz PG, Romesberg FE. Directed evolution of novel polymerase activities: mutation of a DNA polymerase into an efficient RNA polymerase.Proc Natl Acad Sci U S A. 2002 May 14;99(10):6597-602

250. Lanio, T., A. Jeltsch, and A. Pingoud. 1998. Towards the design of rare cutting restriction endonucleases: using directed evolution to generate variants of EcoRV differing in their substrate specificity by two orders of magnitude. J. Mol. Biol. 283**:**59–69.

251. Samuelson, J. C., and S. Y. Xu**.** 2002. Directed evolution of restriction endonuclease BstYI to achieve increased substrate specificity. J. Mol. Biol. 319**:**673–683

252. Zhu, Z., J. C. Samuelson, J. Zhou, A. Dore, and S. Y. Xu. 2004. Engineering strand-specific DNA nicking enzymes from the type IIS restriction endonucleases BsaI, BsmBI, and BsmAI. J. Mol. Biol. 337**:**573–583.

253. Naumann, T. A., and W. S. Reznikoff. 2002. Tn*5* transposase with an altered specificity for transposon ends. J. Bacteriol. 184**:**233–240

254. Sclimenti, C. R., B. Thyagarajan, and M. P. Calos. 2001. Directed evolutionof a recombinase for improved genomic integration at a native human sequence. Nucleic Acids Res. 29**:**5044–5051

255. Buchholz, F., P. O. Angrand, and A. F. Stewart. 1998. Improved properties of FLP recombinase evolved by cycling mutagenesis. Nat. Biotechnol. 16**:** 657–662

256. Buchholz, F., and A. F. Stewart. 2001. Alteration of Cre recombinase site specificity by substrate-linked protein evolution. Nat. Biotechnol. 19**:**1047–1052.

257. Rufer, A. W., and B. Sauer. 2002. Non-contact positions impose site selectivity on Cre recombinase. Nucleic Acids Res. 30:2764–2771.

258. Santoro, S. W., and P. G. Schultz. 2002. Directed evolution of the site specificity of Cre recombinase. Proc. Natl. Acad. Sci. USA 99**:**4185–4190.

259. Ness, J. E., M. Welch, L. Giver, M. Bueno, J. R. Cherry, T. V. Borchert, W. P. Stemmer, and J. Minshull**.** 1999. DNA shuffling of subgenomic sequences of subtilisin. Nat. Biotechnol. 17**:**893–896.

260. Bryan, P. N. 2000. Protein engineering of subtilisin. Biochim. Biophys. Acta 1543**:**203–222.

261. Gupta, R., Q. K. Beg, and P. Lorenz**.** 2002. Bacterial alkaline proteases: molecular approaches and industrial applications. Appl. Microbiol. Biotechnol.59**:**15–32.

262. Kong, X. D., Y. M. Liu, X. J. Gou, S. H. Zhu, H. Y. Zhang, X. P. Wang, and J. Zhang. 2001. Directed evolution of alpha-aspartyl dipeptidase from Salmonella typhimurium. Biochem. Biophys. Res. Commun. 289**:**137–142.

263. Miyazaki, K., P. L. Wintrode, R. A. Grayling, D. N. Rubingh, and F. H. Arnold**.** 2000. Directed evolution study of temperature adaptation in a psychrophilic enzyme. J. Mol. Biol. 297**:**1015–1026.

264. Sroga, G. E., and J. S. Dordick. 2001. Generation of a broad esterolytic subtilisn using combined molecular evolution and periplasmic expression. Protein Eng. 14**:**929–937.

265. Taguchi, S., A. Ozaki, T. Nonaka, Y. Mitsui, and H. Momose**.** 1999. A cold-adapted protease engineered by experimental evolution system. J. Biochem. (Tokyo) 126**:**689–693.

266. Uchiyama, H., T. Inaoka, T. Ohkuma-Soyejima, H. Togame, Y. Shibanaka, T. Yoshimoto, and T. Kokubo**.** 2000. Directed evolution to improve the thermostability of prolyl endopeptidase. J. Biochem. (Tokyo) 128**:**441–447.

267. Wintrode, P. L., K. Miyazaki, and F. H. Arnold**.** 2000. Cold adaptation of a mesophilic subtilisin-like protease by laboratory evolution. J. Biol. Chem.275**:**31635–31640.

268. Zhao, H., and F. H. Arnold**.** 1999. Directed evolution converts subtilisin E into a functional equivalent of thermitase. Protein Eng. 12**:**47–53.

269. Murashima, K., and R. H. Doi**.** 2003. Selection of heat-stable Clostridium cellulovorans cellulases after *in vitro* recombination. Methods Mol. Biol. 230**:**231–237.

270. Murashima, K., A. Kosugi, and R. H. Doi**.** 2002. Thermostabilization of cellulosomal endoglucanase EngB from Clostridium cellulovorans by *in vitro* DNA recombination with non-cellulosomal endoglucanase EngD. Mol. Microbiol. 45**:**617–626.

271. Kim, Y. S., H. C. Jung, and J. G. Pan**.** 2000. Bacterial cell surface display of an enzyme library for selective screening of improved cellulase variants.Appl. Environ. Microbiol. 66**:**788–793.

272. Liebeton, K., A. Zonta, K. Schimossek, M. Nardini, D. Lang, B. W. Dijkstra, M. T. Reetz, and K. E. Jaeger**.** 2000. Directed evolution of an enantioselective lipase. Chem. Biol. **7:**709–718.

273. Kauffmann, I., and C. Schmidt-Dannert**.** 2001. Conversion of Bacillus thermocatenulatus lipase into an efficient phospholipase with increased activity towards long-chain fatty acyl substrates by directed evolution and rational design. Protein Eng. 14**:**919–928.

274. Song, J. K., and J. S. Rhee**.** 2001. Enhancement of stability and activity of phospholipase A(1) in organic solvents by directed evolution. Biochim.Biophys. Acta 1547**:**370–378.

275. Song, J. K., B. Chung, Y. H. Oh, and J. S. Rhee**.** 2002. Construction of DNA-shuffled and incrementally truncated libraries by a mutagenic and unidirectional reassembly method: changing from a substrate specificity of phospholipase to that of lipase. Appl. Environ. Microbiol. 68**:**6146–6151.

276. Moore, J. C., and F. H. Arnold. 1996. Directed evolution of a para-nitrobenzyl esterase for aqueous-organic solvents. Nat. Biotechnol. 14**:**458–467.

277. Giver, L., A. Gershenson, P. O. Freskgard, and F. H. Arnold. 1998. Directed evolution of a thermostable esterase. Proc. Natl. Acad. Sci. USA 95**:**12809–12813.

278. Griffiths, A. D., and D. S. Tawfik. 2003. Directed evolution of an extremely fast phosphotriesterase by *in vitro* compartmentalization. EMBO J. 22**:**24–35.

279. Cohen, J**.** 2001. Breeding' antigens for new vaccines. Science 293**:**236–238.

280. Locher, C. P., V. Heinrichs, D. Apt, and R. G. Whalen**.** 2004. Overcoming antigenic diversity and improving vaccines using DNA shuffling and technologies. Expert Opin. Biol. Ther. 4**:**589–597.

281. Locher, C. P., N. W. Soong, R. G. Whalen, and J. Punnonen. 2004. Developmentof novel vaccines using DNA shuffling and screening strategies.Curr. Opin. Mol. Ther. 6**:**34–39.

282. Marshall, S. H**.** 2002. DNA shuffling: induced molecular breeding to produce new generation long-lasting vaccines. Biotechnol. Adv. 20**:**229–238.

283. Patten, P. A., R. J. Howard, and W. P. Stemmer. 1997. Applications of DNA shuffling to pharmaceuticals and vaccines. Curr. Opin. Biotechnol. 8**:**724–733.

284. Punnonen, J. 2000. Molecular breeding of allergy vaccines and antiallergic cytokines. Int. Arch. Allergy Immunol. 121**:**173–182.

285. Whalen, R. G., R. Kaiwar, N. W. Soong, and J. Punnonen. 2001. DNA shuffling and vaccines. Curr. Opin. Mol. Ther. 3**:**31–36.

286. Osen, W., T. Peiler, P. Ohlschlager, S. Caldeira, S. Faath, N. Michel, M. Muller, M. Tommasino, I. Jochmus, and L. Gissmann**.** 2001. A DNA vaccine based on a shuffled E7 oncogene of the human papillomavirus type 16 (HPV 16) induces E7-specific cytotoxic T cells but lacks transforming activity. Vaccine 19**:**4276–4286.

287. Leong, S. R., J. C. Chang, R. Ong, G. Dawes, W. P. Stemmer, and J. Punnonen**.** 2003. Optimized expression and specific activity of IL-12 by directed molecular evolution. Proc. Natl. Acad. Sci. USA 100**:**1163–1168.

288. Holler, P. D., P. O. Holman, E. V. Shusta, S. O'Herrin, K. D. Wittrup, and D. M. Kranz**.** 2000. *In vitro* evolution of a T cell receptor with high affinity for peptide/MHC. Proc. Natl. Acad. Sci. USA 97**:**5387–5392.

289. Brophy, S. E., P. D. Holler, and D. M. Kranz**.** 2003. A yeast display system for engineering functional peptide-MHC complexes. J. Immunol. Methods 272**:**235–246.

290. Crameri, A., S. Cwirla, and W. P. Stemmer. 1996. Construction and evolution of antibody-phage libraries by DNA shuffling. Nat. Med. 2**:**100–102.

291. Fermer, C., I. Andersson, K. Nilsson, and O. Nilsson**.** 2004. Specificity rescue and affinity maturation of a low-affinity IgM antibody against progastrin-releasing peptide using phage display and DNA shuffling. Tumour Biol. 25**:**7–13.

292. Huls, G., D. Gestel, J. van der Linden, E. Moret, and T. Logtenberg. 2001. Tumor cell killing by *in vitro* affinity-matured recombinant human monoclonal antibodies. Cancer Immunol. Immunother. 50**:**163–171.

293. Jung, S., A. Honegger, and A. Pluckthun. 1999. Selection for improved protein stability by phage display. J. Mol. Biol. 294**:**163–180.

294. Swers, J. S., B. A. Kellogg, and K. D. Wittrup**.** 2004. Shuffled antibody libraries created by *in vivo* homologous recombination and yeast surface display. Nucleic Acids Res. 32**:**e36.

295. van den Beucken, T., N. van Neer, E. Sablon, J. Desmet, L. Celis, H. R. Hoogenboom, and S. E. Hufton. 2001. Building novel binding ligands to B7.1 and B7.2 based on human antibody single variable light chain domains. J. Mol. Biol. 310**:**591–601.

296. Matsumura, I., and A. D. Ellington**.** 1999. *In vitro* evolution of thermostable p53 variants. Protein Sci. 8**:**731–740.

297. Xirodimas, D. P., and D. P. Lane**.** 1999. Molecular evolution of the thermosensitive PAb1620 epitope of human p53 by DNA shuffling. J. Biol. Chem. 274**:**28042–28049.

298. Guo, Z. Y., L. Shen, W. Gu, A. Z. Wu, J. G. Ma, and Y. M. Feng**.** 2002. *In vitro* evolution of amphioxus insulin-like peptide to mammalian insulin. Biochemistry 41**:**10603–10607.

299. Coco, W. M., L. P. Encell, W. E. Levinson, M. J. Crist, A. K. Loomis, L. L. Licato, J. J. Arensdorf, N. Sica, P. T. Pienkos, and D. J. Monticello. 2002. Growth factor engineering by degenerate homoduplex gene family recombination. Nat. Biotechnol. 20**:**1246–1250.

300. Dell'Era, P., R. Ronca, L. Coco, S. Nicoli, M. Metra, and M. Presta**.** 2003. Fibroblast growth factor receptor-1 is essential for *in vitro* cardiomyocyte development. Circ. Res. 93**:**414–420.

301. Sun, J., J. A. Katzenellenbogen, H. Zhao, and B. S. Katzenellenbogen**.** 2003. DNA shuffling method for generating estrogen receptor alpha and beta chimeras in yeast. BioTechniques 34**:**278–280.

302. Joern,J.M., Meinhold,P. and Arnold,F.H. (2002) Analysis of shuffled gene libraries. J. Mol. Biol., 316, 643-656.

303. Miyazaki developed an alternative random digestion method which use endonuclease V to generated DNA fragments.

304. Miyazaki, K. (2002) Random DNA fragmentation with endonuclease V: application to DNA shuffling *Nucleic Acids Res.*, 30, e139

305. Guatelli J, Whitfield K, Kwoh D, Barringer K, Richman D, Gingeras T (1990) Isothermal *in vitro* amplification of nucleic acids by a multienzyme reaction modeled after retroviral replication. Proc Natl Acad Sci USA 87:1874–1878

306. Walker, G.T., Fraiser, M.S., Schram, J.L., Little, M.C., Nadeau, J.G., Malinowski, D.P., 1992. Strand displacement amplification-an isothermal, *in vitro* DNA amplification technique. Nucleic Acids Res. 20, 1691–1696.

307. Saiki RK, Scharf S, Faloona F, Mullis KB, Horn GT, Erlich HA, Arnheim N. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science.(1985) Dec 20;230 (4732):1350-4.

308. Wu and Wallace, 1989 D.Y. Wu and R.B. Wallace, The ligation amplification reaction (LAR)-amplification of specific DNA sequences using sequential rounds of template-dependent ligation, *Genomics* 4 (1989), pp. 560–569

309. Barany F. The ligase chain reaction in a PCR world. PCR Methods Appl. 1991 Aug;1 (1):5-16

310. Walker, G.T., Fraiser, M.S., Schram, J.L., Little, M.C., Nadeau, J.G., Malinowski, D.P., 1992. Strand displacement amplification-an isothermal, *in vitro* DNA amplification technique. Nucleic Acids Res. 20, 1691–1696.

311. Westin, L., Xu, X., Miller, C., Wang, L., Edman, C.F., and Nerenberg, M. 2000. Anchored multiplex amplification on a microelectronic chip array. Nat. Biotechnol. 18: 199–204.

312. Ehses, S., Ackermann, J., and McCaskill, J.S. 2005. Optimization and design of oligonucleotide setup for strand displacement amplification. J. Biochem. Biophys. Methods, 63: 170–186.

313. Turner DJ, Pingle MR, Barany F. Harnessing asymmetrical substrate recognition by thermostable EndoV to achieve balanced linear amplification in multiplexed SNP typing Biochem Cell Biol. 2006 Apr;84(2):232-42.

314. Hecker, K.H., P.D. Taylor, and D.T. Gjerde. 1999. Mutation detection by denaturing DNA chromatography using fluorescently labeled polymerase chain reaction products. Anal. Bio chem. 272:156-164.

315. Bahrami, A.R., M.J. Dickman, M.M. Matin, J.R. Ashby, P.E. Brown, M.J. Conroy, G.J.S. Fowler, J.P. Rose, et al**.** 2002. Use of fluorescent DNA-intercalating dyes in the analysis of DNA via ion-pair reversed-phase denaturing high-performance liquid chromatography. Anal. Biochem. 309:248-252.

316. Nataraj, A.J., I. Olivos-Glander N. Kusu-kawa, and W.E. Highsmith, Jr. 1999. Single-strand conformation polymorphism and heteroduplex analysis for gel-based mutation detection. Electrophoresis *20*:1177-1185.

317. Vahedi, G., K. Kaler, and C.J. Backhouse. 2004. An integrated method for mutation detection using on-chip sample preparation, single-stranded conformation polymorphism and heteroduplex analysis. Electrophoresis 25:2346-2356.

318. Fischer SG, Lerman LS DNA fragments differing by single base-pair substitutions are separated in denaturing gradient gels: correspondence with melting theory (1983) Proc Natl Acad Sci USA 80:1579–1583

319. Cotton, R.G.H., H.R. Rodrigues, and R.D. Campbell. 1988. Reactivity of cytosine and thymine in single base-pair mismatches with hydroxylamine and osmium tetroxide and its application to the study of mutations. Proc Natl Acad Sci U S A. 1988 Jun;85(12):4397-401.

320. Gogos, J.A., M. Karayiorgou, H. Aburatani, and F.C. Kafatos. 1990. Detection of single base mismatches of thymine and cytosine residues by potassium permanganate and hydroxylamine in the presence of tetralkylammonium salts. Nucleic Acids Res. 18:6807-6812.

321. Lambrinakos, A., K.E. Humphrey, J.J. Babon, T.P. Ellis, and R.G.H. Cotton. Reactivity of potassium permanganate and tetraethylammonium chloride with ismatched bases and a simple mutation detection protocol.(1999) Nucleic Acids Res. 27:1866-1874.

322. Rowley, G., S. Saad, F. Giannelli, and P.M. Green. 1995. Ultra rapid mutation detection by multiplex, solid-phase chemical cleavage. Genomics 30:574-582.

323. Roberts, E., V.J. Deeble, C.G. Woods, and G.R. Taylor. Potassium permanganate and tetraethylammonium chloride are a safe and effective substitute for osmium tetroxide in solid-phase fluorescent chemical cleavage of mismatch (1997) Nucleic Acids Res. 25:3377-3378.

324. Bui, C.T., A. Lambrinakos, J.J. Babon, and 18.Bui, C.T., A. Lambrinakos, J.J. Babon, and of DNA on solid support: application of DNA on solid support: application in mutation detection. BMC Chem. Biol. 3:1

325. Lambrinakos, A., M. Yakubovskaya, J.J. Babon, A.A. Neschastnova, Ya.V. Vishnevskaya, G.A. Belitsky, G. D'Cunha, et al. Novel TP53 gene mutations in tumors of Russian patients with breast cancer detected using a new solid phase chemical cleavage of mismatch method and identified by sequencing. (2004) Hum. Mutat. 23:186-192

326. Sadowski PD. Bacteriophage T7 endonuclease. I. Properties of the enzyme purified from T7 phage-infected *Escherichia coli* B.J Biol Chem. 1971 Jan 10;246 (1):209-16.

327. Dickie P, McFadden G, Morgan AR. The site-specific cleavage of synthetic Holliday junction analogs and related branched DNA structures by bacteriophage T7 endonuclease I. J Biol Chem. 1987 Oct 25;262(30):14826-36

328. Mizuuchi K, Kemper B, Hays J, Weisberg RA. T4 endonuclease VII cleaves holliday structures. Cell. 1982 Jun;29(2):357-65

329. Jensch F, Kemper B. Endonuclease VII resolves Y-junctions in branched DNA *in vitro*. EMBO J. 1986 Jan;5(1):181-9.

330. Kleff S, Kemper B. Initiation of heteroduplex-loop repair by T4-encoded endonuclease VII *in vitro*. EMBO J. 1988 May;7(5):1527-35.

331. Solaro PC, Birkenkamp K, Pfeiffer P, Kemper B. Endonuclease VII of phage T4 triggers mismatch correction *in vitro*. J Mol Biol. 1993 Apr 5;230(3):868-77.

332. Mashal RD, Koontz J, Sklar J. Detection of mutations by cleavage of DNA heteroduplexes with bacteriophage resolvases. Nat Genet. 1995 Feb;9(2):177-83.

333. Babon, J.J., M. McKenzie, and R.G. Cotton. 2003. The use of resolvases T4 endonuclease VII and T7 endonuclease I in mutation detection. Mol. Biotechnol. 23:73-81.

334. Lu AL, Hsu IC. Detection of single DNA-base mutations with mismatch repair enzymes. Genomics 1992;14:249-55.

335. Hsu IC, Yang QP, Kahng MW, Xu JF. Detection of DNA Point Mutations with DNA Mismatch Repair Enzymes. Carcinogenesis 1994;15:1657-62.

336. Lindahl, T. and Andersson, A. Rate of chain breakage at apurinic sites in double-stranded deoxyribonucleic acid (1972) *Biochemistry*, 11, 3618–3623.

337. Price, A. and Lindahl, T. Enzymatic release of 5'-terminal deoxyribose phosphate residues from damaged DNA in human cells (1991) *Biochemistry*, 30, 8631–8637.

338. Au KG, Cabrera M, Miller JH, Modrich P. *Escherichia coli* mutY gene product is required for specific A-G----C.G mismatch correction Proc Natl Acad Sci U S A. 1988 Dec;85(23):9163-6

339. Hennecke, F., Kolmar, H., Bründl, K., & Fritz, H.-J.The vsr gene product of *E. coli* K-12 is a strand- and sequence-specific DNA mismatch endonuclease (1991) Nature 353, 776-778

340. Wiebauer, K., & Jiricny, J *In vitro* correction of G-T mispairs to G-C pairs in nuclear extracts from human cells (1989) *Nature 339*, 234-236

341. Oleykowski, C.A., C.R. Bronson Mullins, A.K. Godwin, and A.T. Yeung. 1998. Mutation detection using a novel plant endonuclease. Nucleic Acids Res. 26:4597-4602.

342. Huang J, Kirk B, Favis R, Soussi T, Paty P, Cao W, Barany F. An endonuclease/ligase based mutation scanning method especially suited for analysis of neoplastic tissue. Oncogene. 2002 Mar 14;21(12):1909-21.

343. Tong J, Cao W, Barany F. Biochemical properties of a high fidelity DNA ligase from Thermus species AK16D. Nucleic Acids Res. 1999 Feb 1;27(3):788-94.

344. Friedberg, E.C., Walker, G.C. & Siede, W. (1995). DNA Repair and Mutagenesis, ASM Press Washington DC., USA.

345. R.B. Winter, O.G. Berg and P.H. von Hippel, Diffusion-driven mechanisms of protein translocation on nucleic acids. 3. The *Escherichia coli lac* repressor–operator interaction: kinetic measurements and conclusions. *Biochemistry* 20 (1981), pp. 6961–6977.

346. R.B. Winter and P.H. von Hippel, Diffusion-driven mechanisms of protein translocation on nucleic acids. 2. The *Escherichia coli* repressor–operator interaction: equilibrium measurements. *Biochemistry* 20 (1981), pp. 6948–6960.

347. O.G. Berg, R.B. Winter and P.H. von Hippel. Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory. *Biochemistry* 20 (1981), pp. 6929–6948.

348. Halford SE, Marko JF. How do site-specific DNA-binding proteins find their targets? Nucleic Acids Res. 2004 Jun 3;32 (10):3040-52. Print 2004

349. Gowers, D. M. & Halford, S. E. Protein motion from non-specific to specific DNA by three-dimensional routes aided by supercoiling (2003) *EMBO J* 22, 1410–1418.

350. Gowers, D. M., Wilson, G. G. & Halford, S. E. Measurement of the contributions of 1D and 3D pathways to the translocation of a protein along DNA (2005) *Proc. Natl. Acad. Sci. USA* 102, 15883–15888.

351. Blainey PC, van Oijen AM, Banerjee A, Verdine GL, Xie XS. A base-excision DNA-repair protein finds intrahelical lesion bases by fast sliding in contact with DNA Proc Natl Acad Sci U S A. 2006 Apr 11;103(15):5752-7. Epub 2006 Apr 3

352. T. Brown, O. Kennard, G. Kneale and D. Rabinovich, High-resolution structure of a DNA helix containing mismatched base pairs. *Nature* 315 (1985), pp. 604–606.

353. W.N. Hunter, T. Brown, N.N. Anand and O. Kennard, Structure of an adenine–cytosine base pair in DNA and its implications for mismatch repair. *Nature* 320 (1986), pp. 552–555.

354. T. Brown, W.N. Hunter, G. Kneale and O. Kennard, Molecular structure of the G:A base pair in DNA and its implications for the mechanism of transversion mutations. *Proc. Natl. Acad. Sci. USA* 83 (1986), pp. 2402–2406

355. G.A. Leonard, E.D. Booth and T. Brown, Structural and thermodynamic studies on the adenine guanine mismatch in B-DNA. *Nucleic Acids Res.* 18 (1990), pp. 5617–5623.

356. G.G. Prive, U. Heinemann, S. Chandrasegaran, L.S. Kan, M.L. Kopka and R.E. Dickerson, Helix geometry, hydration, and G:A Mismatch in a B-DNA decamer. *Science* 238 (1987), pp. 498–504.

357. J.V. Skelly, K.J. Edwards, T.C. Jenkins and S. Neidle, Crystal structure of an oligonucleotide duplex containing G:G base pairs: influence of mispairing on DNA backbone conformation. *Proc. Natl. Acad. Sci. USA* 90 (1993), pp. 804–808.

358. K. Kuwata, H. Liu, T. Schleich and T.L. James, Rotational correlation times of internuclear vectors in a DNA duplex with G–A mismatch determined in aqueous solution by complete relaxation matrix analysis of off-resonance ROESY (O-ROESY) spectra. *J. Magn. Res.* 128 (1997), pp. 70–81

359. V. Faibis, J.A.H. Cognet, Y. Boulard, L.C. Sowers and G.V. Fazakerley, Solution structure of two mismatches G:G and I:I in the K-ras gene context by nuclear magnetic resonance and molecular dynamics. *Biochemistry* 35 (1996), pp. 14452–14464.

360. Y. Boulard, J.A.H. Cognet, J. Gabarro-Arpa, M. LeBret, L.C. Sowers and G.V. Fazakerly, The pH dependent configurations of the C:A mispair in DNA. *Nucleic Acids Res.* 20 (1992), pp. 1933–1941.

361. V. Gervais, J.A.H. Cognet, M. LeBret, L.C. Sowers and G.V. Fazakerley, Solution structure of 2 mismatches A:A and T:T in the K-ras gene context by nuclear magnetic resonance and molecular dynamics. *Eur. J. Biochem.* 228 (1995), pp. 279–290

362. H.T. Allawi and J. SantaLucia, Jr., NMR solution structure of a DNA dodecamer containing single GT mismatches. *Nucleic Acids Res.* 26 (1998), pp. 4925–4934.

363. K.E. McAuleyhecht, G.A. Leonard, N.J. Gibson, J.B. Thomson, W.P. Watson, W.N. Hunter and T. Brown, Crystal structure of a DNA duplex containing 8-hydroxyguanine-adenine base-pairs. *Biochemistry* 30 (1994), pp. 10266–10270.

364. C. Bauer, T. Peleg-Shulman, D. Gibson and A.H.J. Wang, Monofunctional platinum amine complexes destabilize DNA significantly. *Eur. J. Biochem.* 256 (1998), pp. 253–260.

365. A. Gelasco and S.J. Lippard, NMR solution structure of a DNA dodecamer duplex containing a cis-diammineplatinum (II) d(GpG) intrastrand cross-link, the major adduct of the anticancer drug cisplatin. *Biochemistry* 37 (1998), pp. 9230–9239.

366. N.J. Gibson, J.A. Parkinson, T. Barlow, W.P. Watson and T. Brown, The stability of duplex DNA containing 3,$N^4$-etheno-2′-deoxycytidine (dC) — A UV melting and high-resolution 1H-NMR study. *J. Chem. Soc. Chem. Commun.* **19** (1994), pp. 2241–2242.

367. F.R. Setayesh, B.L. DeCorte, P. Horton, C.M. Harris, T.M. Harris and M.P. Stone, Styrene oxide adducts in an oligodeoxynucleotide containg the human N-ras codon 12: Minor groove structure of the R(12,1)- and S(12,1)-alpha-(N-2-guanyl) stereoisomers deteremined by H-1 nuclear magnetic resonance. *Chem. Res. Toxicol.* 11 (1998), pp. 766–777.

368. B. Mao, B.E. Hingerty, S. Broyde and D.J. Patel, Solution structure of the aminofluorene [AF]-intercalated conformer of the syn-[AF]-C-8-dG adduct opposite dC in a DNA duplex. *Biochemistry* 37 (1998), pp. 81–94.

369. Banerjee A,Yang W, Karplus M,Verdine GL. Structure of a repair enzyme interrogating undamaged DNA elucidates recognition of damaged DNA Nature. 2005 Mar 31;434(7033):612-8

370. P.C. Hanawalt, R.H. Haynes, Repair replication of DNA in bacteria: irrelevance of chemical nature of base defect, Biochem. Biophys.Res. Commun. 19 (1965) 462–467.

371. D. Gunz, M.T. Hess, H. Naegeli, Recognition of DNA adducts by human nucleotide excision repair, J. Biol. Chem. 271 (1996)

372. D.G. Vassylyev, T. Kashiwagi, Y. Mikami, M. Ariyoshi, S. Iwai, E. Ohtsuka and K. Morikawa, Atomic model of a pyrimidine dimer excision repair enzyme complexed with a DNA substrate: structural basis for damaged DNA recognition. *Cell* 83 (1995), pp. 773–782

373. P.J. O'Brien and T. Ellenberger, Dissecting the broad substrate specificity of human 3-methyladenine-DNA glycosylase, *J. Biol. Chem.* 279 (2004), pp. 9750–9757.

374. N.V. Bulychev, C.V. Varaprasad, G. Dorman, J.H. Miller, M. Eisenberg, A.P. Grollman and F. Johnson, Substrate specificity of *Escherichia coli* MutY protein. *Biochemistry* **35** (1996), pp. 13147–13156.

375. K.G. Berdal, R.F. Johansen and E. Seeberg, Release of normal bases from intact DNA by a native DNA repair enzyme. *EMBO J.* **17** (1998), pp. 363–367

376. Allawi HT, SantaLucia J Jr. Thermodynamics of internal C.T mismatches in DNA Nucleic Acids Res. 1998 Jun 1;26(11):2694-701.

377. Tikhomirova A, Beletskaya IV, Chalikian TV. Stability of DNA duplexes containing GG, CC, AA, and TT mismatches Biochemistry. 2006 Sep 5;45 (35):10563-71

378. Wang Y,Wang Y. Synthesis and thermodynamic studies of oligodeoxyribonucleotides containing tandem lesions of thymidine glycol and 8-oxo-2'-deoxyguanosine Chem Res Toxicol. 2006 Jun;19 (6):837-43

379. Gelfand CA, Plum GE, Grollman AP, Johnson F, Breslauer KJ. Thermodynamic consequences of an abasic lesion in duplex DNA are strongly dependent on base sequence Biochemistry. 1998 May 19;37 (20):7321-7

380. H.T. Allawi and J. SantaLucia, Jr., Nearest-neighbor thermodynamics of internal A·C mismatches in DNA: sequence dependence and pH effects. *Biochemistry* 37 (1998), pp. 9435–9444.

381. G.E. Plum and K.J. Breslauer, DNA lesions — a thermodynamic perspective. *Ann. NY Acad. Sci.* 726 (1994), pp. 45–56.

382. A.N. Lane and B. Peck, Conformational flexibility in DNA duplexes containing G:G mismatches. *Eur. J. Biochem.* 230 (1995), pp. 1073–1087.

383. J.Y. Kao, I. Goljer, T.A. Phan and P.H. Bolton, Characterization of the effects of a thymine glycol residue on the structure, dynamics, and stability of duplex DNA by NMR. *J. Biol. Chem.* 268 (1993), pp. 17787–17793.

384. Wood, R.D (1999) DNA damage recognition during nucleotide exision repair in mammalian cells Biochimie, 81, 39-44

385. Dip R, Camenisch U, Naegeli H. Mechanisms of DNA damage recognition and strand discrimination in human nucleotide excision repair DNA Repair (Amst). 2004 Nov 2;3 (11):1409-23

386. Sugasawa,K., Okamoto,T., Shimizu,Y., Masutani,C., Iwai,S. and Hanaoka,F. (2001) A multistep damage recognition mechanism for global genomic nucleotide excision repair. *Genes Dev.*, 15, 507–521.

387. Zou,Y., Luo,C. and Geacintov,N.E. (2001) Hierarchy of DNA damage recognition in *Escherichia coli* nucleotide excision repair. *Biochemistry*, 40, 2923–2931

388. Hess,M.T., Schwitter,U., Petretta,M., Giese,B. and Naegeli,H. (1997) Bipartite substrate discrimination by human nucleotide excision repair. *Proc. Natl Acad. Sci. USA.*, 94, 6664–6669

389. M.T. Hess, U. Schwitter, M. Petretta, B. Giese and H. Naegeli, Bipartite substrate discrimination by human nucleotide excision repair enzymes. *Proc. Natl. Acad. Sci. U.S.A.* 94 (1997), pp. 6664–6669

390. Yang W. Poor base stacking at DNA lesions may initiate recognition by many repair proteins DNA Repair (Amst). 2006 Jun 10;5(6):654-66. Epub 2006 Mar 29

391. G.E. Plum, A.P. Grollman, F. Johnson and K.J. Breslauer, Influence of the oxidatively damaged adduct 8-oxodeoxyguanosine on the conformation, energetics, and thermodynamic stability of a DNA duplex, *Biochemistry* 34 (1995), pp. 16148–16160.

392. C. Cao, Y.L. Jiang, J.T. Stivers and F. Song, Dynamic opening of DNA during the enzymatic search for a damaged base, *Nat. Struct. Mol. Biol.* 11 (2004), pp. 1230–1236.

393. D. Barsky, N. Foloppe, S. Ahmadia, D.M. Wilson III and A.D. MacKerell Jr., New insights into the structure of abasic DNA from molecular dynamics simulations, *Nucleic Acids Res.* 28 (2000), pp. 2613–2626.

394. M.A. Elizondo-Riojas and J. Kozelka, Unrestrained 5 ns molecular dynamics simulation of a cisplatin-DNA 1,2-GG adduct provides a rationale for the NMR features and reveals increased conformational flexibility at the platinum binding site, *J. Mol. Biol.* 314 (2001), pp. 1227–1243.

395. O. Kennard, Structural studies of DNA fragments: the G.T wobble base pair in A, B and Z DNA; the G.A base pair in B-DNA, *J. Biomol Struct. Dyn.* 3 (1985), pp. 205–226

396. Yakovchuk P, Protozanova E, Frank-Kamenetskii MD. Base-stacking and base-pairing contributions into thermal stability of the DNA double helix Nucleic Acids Res. 2006 Jan 31;34(2):564-74.

397. Krueger A, Protozanova E, Frank-Kamenetskii MD. Sequence-dependent base pair opening in DNA double helix Biophys J. 2006 May 1;90(9):3091-9.

398. Frank-Kamenetskii, M.D. How the double helix breathes (1987) *Nature*, 328, 17–18

399. G. Natrajan, M.H. Lamers, J.H. Enzlin, H.H. Winterwerp, A. Perrakis and T.K. Sixma, Structures of *Escherichia coli* DNA mismatch repair enzyme MutS in complex with different mismatches: a common recognition mode for diverse substrates, *Nucleic Acids Res.* 31 (2003), pp. 4814–4821

400. S.E. Tsutakawa, H. Jingami and K. Morikawa, Recognition of a TG mismatch: the crystal structure of very short patch repair endonuclease in complex with a DNA duplex, *Cell* 99 (1999), pp. 615–623

401. T.E. Barrett, R. Savva, G. Panayotou, T. Barlow, T. Brown, J. Jiricny and L.H. Pearl, Crystal structure of a G:T/U mismatch-specific DNA glycosylase: mismatch recognition by complementary-strand interactions, *Cell* 92 (1998), pp. 117–129.

402. Klimasauskas S, Kumar S, Roberts RJ, Cheng X. HhaI methyltransferase flips its target base out of the DNA helix. Cell. 1994 Jan 28;76(2):357–369.

403. D.P. Hornby and G.C. Ford, Protein-mediated base flipping. *Curr. Opin. Biotechnol.* 9 (1998), pp. 354–358

404. R. Roberts and X. Cheng, Base flipping. *Annu. Rev. Biochem.* 67 (1998), pp. 181–198

405. R.S. Lloyd and X. Cheng, Mechanistic link between DNA methyltransferases and DNA repair enzymes by base flipping. *Biopolymers* 44 (1997), pp. 139–151.

406. A.K. McCullough, M.L. Dodson, O.D. Scharer and R.S. Lloyd, The role of base flipping in damage recognition and catalysis by T4 endonuclease V. *J. Biol. Chem.* 272 (1997), pp. 27210–27217.

407. G. Slupphaug, C.D. Mol, B. Kavli, A.S. Arvai, H.E. Krokan and J. Tainer, A nucleotide flipping mechanism from the structure of human uracil-DNA glycosylase bound to DNA. *Nature* 384 (1996), pp. 87–92

408. J. Ramstein and R. Lavery, Energetic coupling between DNA bending and base pair opening. *Proc. Natl Acad. Sci. USA* 85 (1988), pp. 7231–7235

409. J. Bernet, K. Zakrzewska and R. Lavery, Modelling base pair opening: the role of helical twist. (1997) J. Mol. Struct. (THEOCHEM) 398, pp. 473–482.

410. G.L. Verdine, S.D. Bruner, How do DNA repair proteins locate damaged bases in the genome? (1997) Chem. Biol. 4.329–334.

411. Parikh SS, Mol CD, Tainer JA. Base excision repair enzyme family portrait: integrating the structure and chemistry of an entire DNA repair pathway. (1997) *Structure* 5:1543–50

412. Lau, A. Y., Wyatt, M. D., Glassner, B. J., Samson, L. D., and Ellenberger, T. Molecular basis for discriminating between normal and damaged bases by the human alkyladenine glycosylase, AAG (2000) *Proc. Natl. Acad. Sci. U. S. A.* 97**,** 13573-13578

413. Bruner, S. D., Norman, D. P., and Verdine, G. L. Structural basis for recognition and repair of the endogenous mutagen 8-oxoguanine in DNA (2000) *Nature* 403, 859-866

414. Mees, A., Klar, T., Gnau, P., Hennecke, U., Eker, A. P. M., Carell, T., and Essen, L. O. Crystal Structure of a Photolyase Bound to a CPD-Like DNA Lesion After in Situ Repair (2004) *Science* 306**,** 1789-1793

415. Malta E, Moolenaar GF, Goosen N. Base flipping in nucleotide excision repair. J Biol Chem. (2006) Jan 27;281 (4):2184-94. Epub 2005 Nov 9

416. Mees, A., Klar, T., Gnau, P., Hennecke, U., Eker, A. P. M., Carell, T., and Essen, L. O. Crystal Structure of a Photolyase Bound to a CPD-Like DNA Lesion After in Situ Repair (2004) *Science* 306**,** 1789-1793

417. Malta E, Moolenaar GF, Goosen N. Base flipping in nucleotide excision repair. J Biol Chem. (2006) Jan 27;281 (4):2184-94. Epub 2005 Nov 9

# CHAPTER TWO

## Switching Base Preferences of Mismatch Cleavage in Endonuclease V: An Improved Methods for Scanning Point Mutation

### I. Abstract

Endonuclease V recognizes a broad range of aberrations in DNA such as deaminated bases or mismatches. It nicks DNA at the second phosphodiester bond 3' to a deaminated base or a mismatch. Endonuclease V obtained from *Thermotoga maritima* preferentially cleaves purine mismatches in certain sequence context. Endonuclease V has been combined with a high fidelity DNA ligase to develop an enzymatic method for mutation scanning. A biochemical screening of site-directed mutants identified mutants in motifs III and IV that altered the base preferences in mismatch cleavage. Most profoundly, a single alanine substitution at Y80 position switched the enzyme to essentially a C-specific mismatch endonuclease, which recognized and cleaved A/C, C/A, T/C, C/T and even the previously refractory C/C mismatches. Y80A can also detect the G13D mutation in K-*ras* oncogene, an A/C mismatch embedded in a G/C rich sequence context that was previously inaccessible using the wild-type endonuclease V. This investigation offers insights on base recognition and active site organization. Protein engineering in endonuclease V may translate into better tools in mutation recognition and cancer mutation scanning.

## II. Introduction

Techniques to scan unknown single nucleotide polymorphisms (SNPs) or point mutations are an essential tool in the post-genomic era. Current mutation scanning methods include single-stranded conformational polymorphism (SSCP) and heteroduplex analysis (HA) (1,2), denaturing high performance liquid chromatography (DHPLC) (3), and chemical or enzymatic cleavage (4-7). Several enzymatic cleavage methods have been developed (7,8). T4 endonuclease VII and T7 endonuclease I, the two phage resolvases, have been used for mutation scanning with limited success due to high background generated by cleavage of non-mismatch sequences (9). Other enzymes such as MutY DNA glycosylase and thymine DNA glycosylase (TDG), and CEL1 nuclease have also been employed in mutation scanning (7,10).

Endonuclease V (endo V) is a DNA repair enzyme with unique enzymatic properties. Under physiological conditions, endo V cleaves deaminated bases at the second phosphodiester bond 3' to a lesion. By shifting reaction conditions to higher pH, metal cofactor to $Mn^{2+}$, using excess enzyme, and/or using solvents such as DMSO and betaine, this repertoire may be extended to include cleavage of most mismatched DNA base pairs (11-13). This enzymatic property has been explored for the development of mutation scanning techniques (8,14). We have devised a scheme that uses thermostable endonuclease V obtained from *Thermotoga maritima* (*Tma*) to cleave mismatches and a high-fidelity thermostable DNA ligase from *Thermus* species AK16D to seal nonspecific cleavage (8,15,16). Co-incubation of the two enzymes allows for endonucleolytic cleavage of mismatches with real-time resealing of matched nicks, allowing for detection

of low-abundance mutations in tumor tissue at a ratio of 1:50 mutant to wild-type DNA (8,15).

*Tma* endonuclease V preferentially cleaves purine bases in a mismatch in certain sequence context (13). The wild-type enzyme cleaves the C-containing mismatches the least and C/C mismatches are essentially resistant to cleavage (13). Even some A/C mismatches are refractory to cleavage when located in a G/C rich sequence context, as exemplified in the G13D mutation in K-*ras* (8). Identification of endo V variants that can cleave C-containing mismatches will broaden the applicability of the endo V/ligase mutation scanning technique. Although an endo V-DNA complex structure is not available, an extensive site-directed mutagenesis analysis has identified motifs and specific amino acid residues that influence base recognition and DNA-protein interactions (17). Taking advantage of a battery of over sixty endo V single-site mutants previously isolated, we screened for and identified endo V variants that possessed altered base preference in mismatch cleavage. Y80A in motif III converted endo V to essentially a C-specific mismatch cleavage variant that was capable of nicking refractory A/C mismatches in the K-*ras* gene.

## III. Materials and Methods

### A. Materials

Purified deoxyribooligonucleotides were ordered from Integrated DNA Technologies Inc. (Coralville, IA). Duplex deoxyoligonucleotide substrates were prepared as previously described (17). The wild-type and mutant *Tma* endonuclease V proteins and Tsp AK16D DNA ligase were purified as previously described (16-18).

### B. Endo V cleavage assays

The cleavage reaction mixtures (10 µl) containing 10 mM HEPES-KOH (pH 7.4), 1 mM dithiothreitol, 2% glycerol, 5 mM $MnCl_2$ unless otherwise specified, 10 nM oligonucleotide DNA substrate and 10 nM of *Tma* endo V protein unless otherwise specified were incubated at 65°C for 30 min. The reactions were terminated by addition of an equal volume of GeneScan Stop Buffer (80% formamide, 50 mM EDTA (pH 8.0), and 1% blue dextran). The reaction mixtures were then heated at 94°C for 3 min and cooled down on ice. Samples (3.5 µl) were loaded onto a 10% denaturing polyacrylamide gel containing 7 M urea. Electrophoresis was conducted at 1500 volts for 1.5 h using an ABI 377 sequencer (Applied Biosystems). Cleavage products and remaining substrates were quantified using the GeneScan analysis software version 3.0.

### C. PCR amplification of K-*ras* exon I

For detecting K-*ras* mutations, genomic DNA was extracted from cell lines as described (19). Cell lines HT 29 contains wild-type K-*ras* gene. SW480 contains G12V (G->T) mutation (pure G12V mutant). DLD-1 contains G13D (G->A) mutation. K-*ras*

exon I was amplified by PCR as described (8). To remove *Taq* DNA polymerase, 4 µl of 20 mg/ml proteinase K (QIAGEN) was added to the PCR mixtures (50 µl) and incubated at 70°C for 10 min. Proteinase K was inactivated by incubating at 80°C for 10 min. Amplicons containing wild-type sequence were added in approximately equal ratios when missing from the sample (i.e. pure mutant cell line DNA). The mixed PCR fragments, were heated at 94°C for 1 min to denature the DNA, and then cooled at 65°C for 15 min and at room temperature for 15 min to allow efficient formation of heteroduplex DNA.

To generate sticky ended PCR products, K-*ras* exon I was amplified as described with the exception that the PCR primers are as follows (8): Oligo 1, 5'-CCCCGCTGAGGATAGTGTATTAACCTTATGTGTGACATGTTC-3' (underlined: N.BbvC IA site); Oligo 2, 5'-Fam-CCCCCCTCAGCAAAATGGTCAGAGAAACCTTTATCTGTATC-3' (underlined: N.BbvC IB site, which is complementary to the N.BbvC IA site). After PCR, the top strand contained two N.BbvC IA sites and the bottom strand contained two N.BbvC IB sites (Fig. 2.4A). Post-PCR processing and formation of duplex DNA were carried out as described above. PCR products (6 µg) were then digested at 37°C overnight with 60 units of N.BbvC IA in NEB buffer 2 (New England Biolabs). The reaction mixtures were extracted with phenol/chloroform/isoamyl alcohol (25:24:1) to remove proteins and passed through microcon YM-50 spin column (Millipore) to remove the small DNA fragments generated by BbvC IA nicking.

## IV. Results and Discussion

### A. Examination of base preferences of mismatch cleavage in endonuclease V mutants.

Endonuclease V contains seven conserved motifs in which motifs III and IV play a major role in protein-DNA interactions (17). We screened a total of sixty-four mutants previously isolated for mismatch cleavage activity (Fig. 2.1) (17). The assays were performed in the presence of $Mn^{2+}$ instead of $Mg^{2+}$ since endo V enzymes show enhanced mismatch cleavage with $Mn^{2+}$ (12,13). As expected, a majority of mutants lost mismatch cleavage activity.

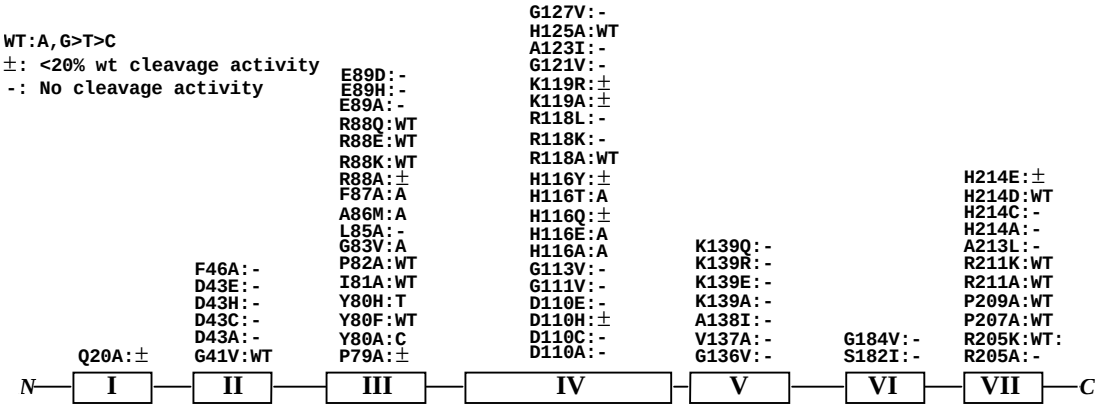



**Figure 2.1 Base preference of mismatch cleavage of *Tma* endonuclease V mutants.** Cleavage reactions were performed as described in Materials and Methods. Motifs are shown in Roman numbers. See (17) for sequence alignment.

Other mutants still maintained mismatch cleavage activity in a pattern similar to the wild-type enzyme, which included G41V in motif II; Y80F, I81A, P82A, R88K and R88Q in motif III, R118A and H125A in motif IV; R205K, P207A, P209A, R211A, R211K, and H214D in motif VII (Fig. 2.1). Yet, several mutants in motifs III and IV showed quite distinctively altered base preference in mismatch cleavage. An alanine

substitution at Y80 position essentially switches the base preference from purine mismatches to C-specific mismatches (Figs. 2.1-2.2). All five C-containing mismatches were cleaved by Y80A (Fig. 2.2, compare the band intensities in wild-type and Y80A). Most remarkably, even the refractory C/C mismatch in this sequence context was cleaved on both strands (Fig. 2.2, C/C lane in Y80A). On the other hand, cleavage of other mismatches was minimum or not detected. A histidine substitution at Y80 rendered the enzyme more active in cleaving T-containing mismatches, while reducing the cleavage of other mismatches (Fig. 2.2, Y80H). Apparently, A86M preferentially cleaved A-containing mismatches. All four A-containing strands G/A, C/A, A/G and A/C and both strands in A/A were cleaved by A86M. Other mutants such as G83V and F87A also showed preference for A bases (Fig. 2.1).

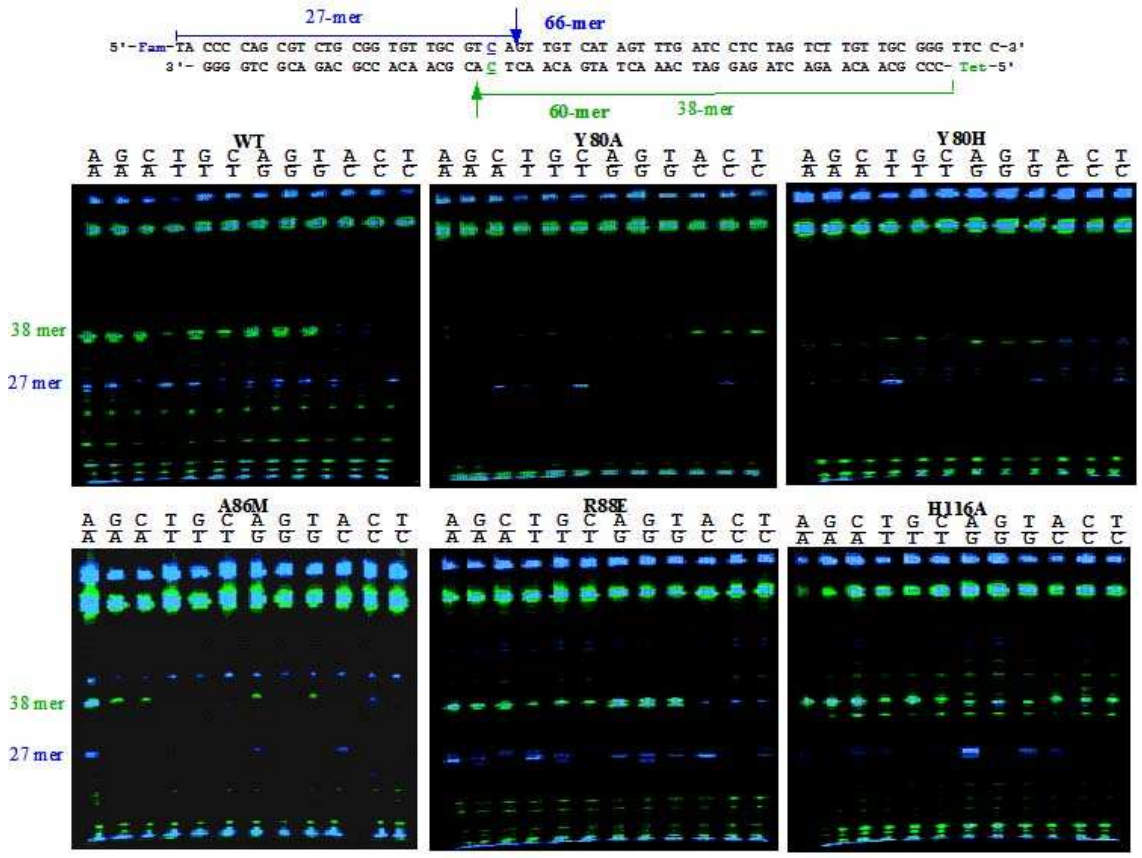



**Figure 2.2 Representative GeneScan gel pictures of mismatch cleavage**. Cleavage reactions were performed as described in Materials and Methods

The base preference in R88E remained similar to the wild-type enzyme, i.e., G and A bases were preferred. However, the cleavage site on the top strand (blue band) is more promiscuous. Cleavage at one nt closer or one nt further away from the mismatches was observed (Fig. 2.2, R88E). Similar cleavage site promiscuity occurred in R88Q (data not shown). A few H116 mutants such as H116A, H116E and H116T somewhat preferred the A base in a mismatch (Figs 2.1-2.2)

**B. Cleavage of A/C mismatches in synthetic K-*ras* substrates.**

Given the strong preference of Y80A for the C base in a mismatch, we tested its ability to cleave C-containing mismatches that were refractory for the wild-type enzyme. Previously, we developed an enzymatic mutation scanning method, which takes advantage of the mismatch cleavage of endo V and nicking joining activity of DNA ligase to seal nonspecific cleavage at matched bases (8). During the course of that study, we found G13D mutation in K-*ras* was completely refractory to endo V cleavage when using $Mg^{2+}$ as cofactor in the presence of both 5% (V/V) DMSO and 1.5 M betaine. G13D is a G to A transition that yields G/T and A/C mismatches. A closer look at the flanking sequence indicates that the mismatches are located in a G/C rich sequence context (TGGCG, the mutation site is underlined), which may make it difficult for endo V to cleave (11). To test the ability of Y80A to cleave this sequence, we synthesized an oligodeoxynucleotide substrate that was identical to the G13D sequence in K-*ras* (Fig. 2.3A, A/C). The overall design was consistent with the mismatch substrates used for initial activity screening (Fig. 2.2). When using $Mn^{2+}$ as the metal cofactor, the wild-type endo V exhibited non-specific fragmentation of both the top and the bottom strand as a result of non-specific cleavage, but did not yield correct length fragments from the mismatched base pair (Fig. 2.3B). Remarkably, Y80A generated a cleavage band from the bottom C-containing strand at about 38-39-mer position, indicating that the altered base preference has enabled the mutant to cleave the refractory sequence (Fig. 2.3B). To verify the specificity of the cleavage by Y80A, we synthesized a similar substrate but with the C base on the top strand, which would generate a 27-mer if cleaved (Fig. 2.3A, C/A). Again, the Y80A cleaved the C-containing strand in the C/A mismatch at the

anticipated position, while the wild-type enzyme generated lower molecular weight fragments (Fig. 2.3B). These results confirmed the C base preference of the Y80A mutant in the refractory sequence.

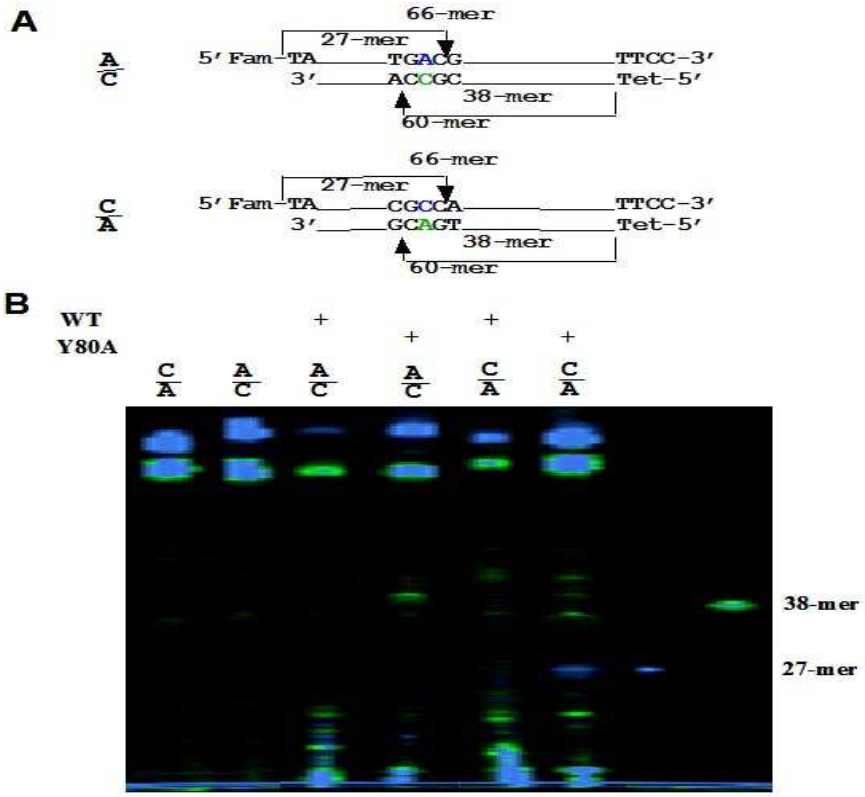



**Figure 2.3 Cleavage of A/C mismatch in synthetic K-*ras* G13D sequence by Y80A *Tma* endonuclease V mutant.** Cleavage reactions were performed as described in Materials and Methods with 2.5 mM $MnCl_2$. **A.** Schematic illustration of A/C cleavage. A/C heteroduplex was formed by annealing of 5'-FAM-TAACTTGTGGTAGTTGG AGCTGGTGACGTAGGCAAGAGTGCCTTGACGATACAGCTAATTCATTCC-3' and 3'-TGAACACCATCAACCTCGACCACCGCATCCGTTCTCACGGAACTGCGA TGTCGATTAAGT-TET-5'. C/A heteroduplex was formed by annealing of 5'-FAM-TATCGTCAAGGCACTCTTGCCTACGCCACCAGCTCCAACTACCACAAGTTTAT ATTCAGTCATTCC-3' and 3'-AGCAGTTCCGTGAGAACGGATGCAGTGGTCGAG GTTGATGGTGTTCAA ATATAAGTCAGT-TET-5'. **B.** Cleavage of A/C K-*ras* G13D mismatch by wild-type *Tma* endonuclease V and Y80A mutant.

**C. Cleavage of A/C mismatch in K-*ras* amplicons.**

To test the ability of the Y80A mutant to cleave PCR products, we amplified the exon 1 of the K-*ras* gene from both the wild-type, G12V, and G13D mutant cell lines. Heteroduplexes were generated by mixing the wild-type PCR amplicon with the mutant amplicons (Fig. 2.4A, left). The 286-bp long heteroduplexes containing T/C and G/A mismatches from G12V and A/C and G/T mismatches from G13D were treated with Y80A mutant endo V. Since Y80A acted as a C-specific mismatch endonuclease (Figs. 2.2-2.3), we scored the specific cleavage bands as resulting from cleaving C-containing mismatches. As expected, G12V was cleaved by Y80A on the C-containing strand to yield a 166-mer product (Fig. 2.4B, left). However, cleavage of A/C mismatch in the G13D was minimal (Fig. 2.4B, left).

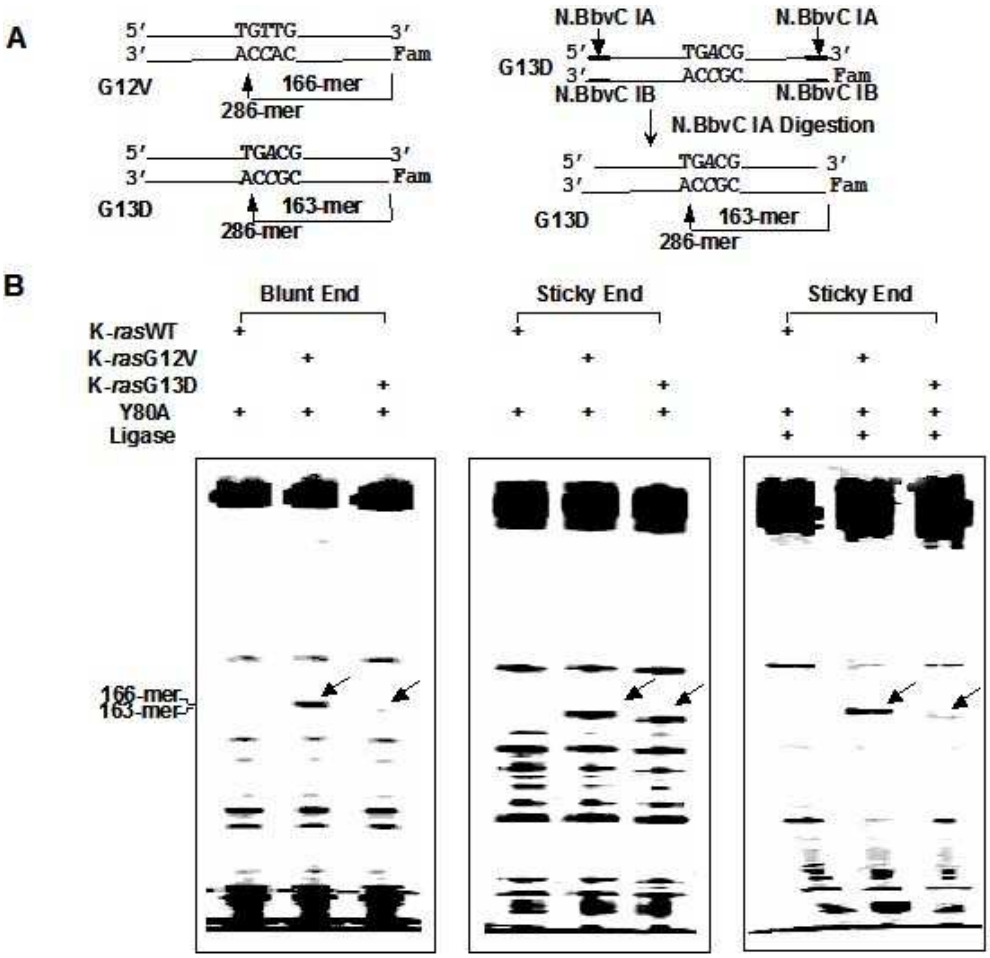



**Figure 2.4 Cleavage of A/C mismatch in K-*ras* G13D sequence amplified from colon cancer cell lines by *Tma* endonuclease V mutant Y80A. A.** Schematic illustration of blunt end and sticky end heteroduplex G12V and G13D PCR products. See Materials and Methods for details. **B.** Cleavage of G13D by Y80A *Tma* endonuclease V mutant. Cleavage reaction mixtures (10 µl) containing 10 mM HEPES-KOH (pH 7.4), 1 mM DTT, 2% glycerol, 2.5 mM $MnCl_2$, 100 ng of wild-type K-*ras* homoduplex or G12V heteroduplex or G13D heteroduplex and 100 nM *Tma* endo V mutant Y80A protein were incubated at 65°C for 30 min. For the reactions that were followed by ligation, the amount of K-*ras* homoduplex or heteroduplex was increased to 200 ng in the cleavage reactions. The cleavage reaction mixtures were filtered through an YM-10 microcon spin column and washed with TE buffer containing 10 mM Tris-HCl (pH 7.6) and 1 mM EDTA. To seal the nonspecific nicks, the washed cleavage reaction mixtures (in 6 µl TE) were supplemented with 1 µl of 10 x Taklig buffer (20 mM Tris-HCl (pH 7.6), 100 mM KCl, 10 mM dithiothreitol, 20 µg/ml bovine serum albumin), 1 µl of 100 mM $MgCl_2$, 1 µl of 10 mM $NAD^+$, and 1 µl of 20 nM Tsp AK16D ligase. The ligation mixtures were incubated at 65°C for 20 min.

Previously, we have observed a reduction in fluorescence signal in blunt end amplicons due to cleavage of the fluorescent label and the adjacent base by endo V, liberating the label from the amplicon (15). We suspected a similar cleavage event might have occurred in the blunt ended amplicons that have reduced the cleavage product signals (Fig. 2.4B left, bottom of gel). Given that the synthetic duplexes contained overhangs (Fig. 2.3), we thought the overhangs at the ends may reduce the loss of fluorescence signal by endo V. We, therefore, designed a method to convert the PCR amplicons to sticky ended duplexes (Fig. 2.4A, right). N.BbvC IA recognizes double-stranded 5'-GC↓TGAGG-3' sequence and nicks between the C and T. The recognition sequence was incorporated into the PCR primers for amplifying the exon I of K-*ras* gene (see Materials and Methods for details). The resulting PCR amplicons were then treated with N.BbvC IA to generate a two-base overhang at the 3' end and five-base overhang at the 5' end for the C-containing strand, respectively (Fig. 2.4A, right). Both the G12V and G13D heteroduplexes were cleaved by Y80A mutant endo V (Fig. 2.4B, middle). The nonspecific products were sealed by the high fidelity Tsp. AK16D ligase, thus reducing the background (Fig. 2.4B, right). Some of the mismatch cleavage products were also sealed by the DNA ligase (16), resulting in a reduction in the intensity of the specific band.

This work identified endo V variant enzymes with substantially altered base preferences in mismatch cleavage. Since all these variant enzymes contained changes in motifs III and IV, this underscores the important role these motifs play in base recognition (Fig. 2.1). Consistent with a previous study (17), Y80 and H116 appear to be important determinants of base recognition. Although an endo V-DNA co-crystal

structure is not available, secondary structure analysis indicates that both Y80 and H116 are located in loop regions (20). We speculate that motifs III and IV are components of recognition loops that are involved in specific base recognition.

Y80A is the most striking in that it essentially converts the enzyme to a C-specific mismatch endonuclease (Fig. 2.2). Consequently, the previous refractory C/C mismatch for the wild-type enzyme now becomes cleavable by the Y80A mutant. First, how does a single alanine substitution at Y80 position accomplish such a dramatic alteration in base preference? A simple model is illustrated in Fig. 2.5. In the wild-type enzyme, Y80 imposes an unfavorable interaction with a C base, in which the amino group at C4 position spatially clashes with the bulky tyrosine residue. This steric hindrance prevents the wild-type endo V from recognizing and cleaving C-containing strand in a mismatch. By substituting the phenol side-chain with a small methyl group, Y80A releases the steric tension and allows the C base to be accommodated in the recognition pocket (Fig. 2.5). A comparison with uracil DNA glycosylase (UDG) is illuminating. The N204 in the recognition site of human uracil DNA glycosylase forms hydrogen bonds with O4 and N3 of uracil via the amide side chain and the Y147 excludes a thymine base by steric complementarity (21). Interestingly, N204D confers cytosine DNA glycosylase activity to hUDG by forming hydrogen bonds with the C4-amino group and the N3-nitrogen via the carboxyl side chain, while Y147A switches the enzyme to thymine DNA glycosylase by preventing the steric clash with the C5-methyl group of the thymine base (22). It is possible that endo V and UDG adopted a similar strategy as part of base-specific recognition mechanism (22-26).

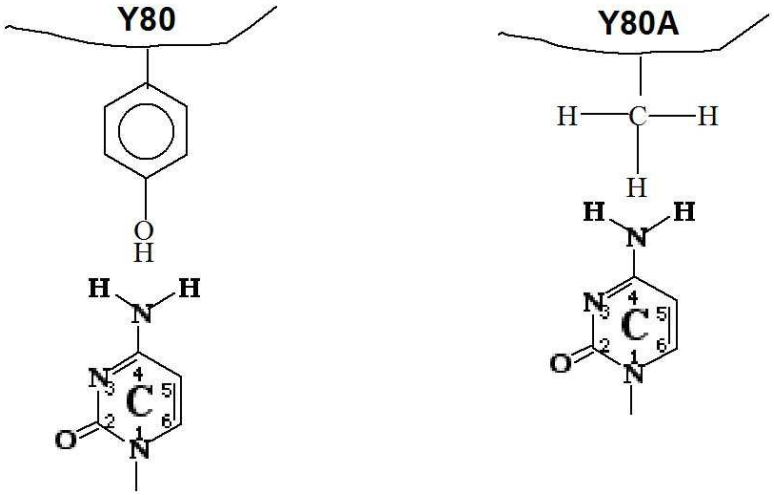



**Figure 2.5. A hypothetical model for alteration of base recognition by Y80A.** See text for details.

The surprising alteration in base preference of mismatch cleavage prompted us to investigate the potential implication in improving the endo V/ligase mutation scanning technique previously reported (8). The use of this technique in scanning K-*ras* mutations met with difficulty partly due to the inability of the wild-type endo V to cleave A/C mismatches in some G/C rich sequence context (8). Data presented here indicate that the Y80A is not only specific for C-containing mismatches, but also for those embedded in G/C rich environment (Fig. 2.3). Therefore, the C-specific mismatch cleavage ability may have enabled the Y80A to recognize and nick the C-strand previously not accessible by the wt endo V. Based on the model explained above, favorable interactions between Y80A and a C base may facilitate the base recognition process, which assists in guiding the complex to a catalytically competent path. Likewise, the previously inaccessible C/C mismatch now becomes a substrate for Y80A (Fig. 2.2). The difference in A/C mismatch cleavage efficiency between amplicons with blunt or overhang ends is due to loss of fluorescence signal by endo V cleavage. This problem was previously addressed by

synthesizing modified primers that are refractory to Endo V cleavage (15). Introducing a nicking site into a PCR primer provides a simple alternative method to maintain mismatch cleavage signal. This study demonstrates how malleable endonuclease V is, allowing for alteration of base preference in mismatch cleavage by single amino acid changes. Some of these mutants offer the potential for developing base-specific Endo V/DNA ligase mutation scanning assays.

## V. Acknowledgments

This project was supported in the Cao lab in part by CSREES/USDA (SC-1700274, technical contribution No. 5189), the NIH (GM 067744), the Concern Foundation and DOD-Army Research Office (W911NF-05-1-0335). Work in the Barany lab was supported by the NCI (P01 CA65930-08). We thank Dr. Pu-Chun Ke for critically reading the manuscript and members of Cao and Barany labs for stimulating discussions.

## VI. References


1. Highsmith, W.E., Jr., Jin, Q., Nataraj, A.J., O'Connor, J.M., Burland, V.D., Baubonis, W.R., Curtis, F.P., Kusukawa, N. and Garner, M.M. (1999) Use of a DNA toolbox for the characterization of mutation scanning methods. I: construction of the toolbox and evaluation of heteroduplex analysis. *Electrophoresis*, 20**,** 1186-1194.

2. Highsmith, W.E., Jr., Nataraj, A.J., Jin, Q., O'Connor, J.M., El-Nabi, S.H., Kusukawa, N. and Garner, M.M. (1999) Use of DNA toolbox for the characterization of mutation scanning methods. II: evaluation of single-strand conformation polymorphism analysis. *Electrophoresis*, 20**,** 1195-1203.

3. Hecker, K.H., Taylor, P.D. and Gjerde, D.T. (1999) Mutation detection by denaturing DNA chromatography using fluorescently labeled polymerase chain reaction products. *Anal Biochem*, 272**,** 156-164.

4. Bui, C.T., Lambrinakos, A., Babon, J.J. and Cotton, R.G. (2003) Chemical cleavage reactions of DNA on solid support: application in mutation detection. *BMC Chem Biol*, 3**,** 1.

5. Babon, J.J., McKenzie, M. and Cotton, R.G. (2003) The use of resolvases T4 endonuclease VII and T7 endonuclease I in mutation detection. *Mol Biotechnol*, 23**,** 73-81.

6. Glare, E.M., Jenkins, A.J. and Cotton, R.G. (2003) Plastic contaminant masquerades as DNA in mutation detection by denaturing HPLC. *Biotechniques*, 34**,** 59-60.

7. Yeung, A.T., Hattangadi, D., Blakesley, L. and Nicolas, E. (2005) Enzymatic mutation detection technologies. *Biotechniques*, 38**,** 749-758.

8. Huang, J., Kirk, B., Favis, R., Soussi, T., Paty, P., Cao, W. and Barany, F. (2002) An endonuclease/ligase based mutation scanning method especially suited for analysis of neoplastic tissue. *Oncogene*, 21**,** 1909-1921.

9. Mashal, R.D., Koontz, J. and Sklar, J. (1995) Detection of mutations by cleavage of DNA heteroduplexes with bacteriophage resolvases. *Nat Genet*, 9**,** 177-183.

10. Zhang, Y., Kaur, M., Price, B.D., Tetradis, S. and Makrigiorgos, G.M. (2002) An amplification and ligation-based method to scan for unknown mutations in DNA. *Hum. Mutat.*, 20**,** 139-147.

11. Yao, M. and Kow, Y.W. (1994) Strand-specific cleavage of mismatch-containing DNA by deoxyinosine 3'-endonuclease from *Escherichia coli*. *J Biol Chem*, 269**,** 31390-31396.

12. Feng, H., Klutz, A.M. and Cao, W. (2005) Active Site Plasticity of Endonuclease V from Salmonella typhimurium. *Biochemistry*, 44**,** 675-683.

13. Huang, J., Lu, J., Barany, F. and Cao, W. (2001) Multiple Cleavage Activities of Endonuclease V from *Thermotoga maritima*: Recognition and Strand Nicking Mechanism. *Biochemistry*, 40**,** 8738-8748.

14. Bazar, L., Collier, G., Vanek, P., Siles, B., Kow, Y., Doetsch, P., Cunningham, R. and Chirikjian, J. (1999) Mutation identification DNA analysis system (MIDAS) for detection of known mutations. *Electrophoresis*, 20**,** p1141-1148.

15. Pincas, H., Pingle, M.R., Huang, J., Lao, K., Paty, P.B., Friedman, A.M. and Barany, F. (2004) High sensitivity EndoV mutation scanning through real-time ligase proofreading. *Nucleic Acids Res*, 32**,** e148.

16. Tong, J., Cao, W. and Barany, F. (1999) Biochemical properties of a high fidelity DNA ligase from Thermus species AK16D. *Nucleic Acids Res.*, 27**,** 788-794.

17. Feng, H., Dong, L., Klutz, A.M., Aghaebrahim, N. and Cao, W. (2005) Defining Amino Acid Residues Involved in DNA-Protein Interactions and Revelation of 3'-Exonuclease Activity in Endonuclease V. *Biochemistry*, 44**,** 11486-11495.

18. Huang, J., Lu, J., Barany, F. and Cao, W. (2002) Mutational analysis of endonuclease V from *Thermotoga maritima*. *Biochemistry*, 41**,** 8342-8350.

19. Khanna, M., Park, P., Zirvi, M., Cao, W., Picon, A., Day, J., Paty, P. and Barany, F. (1999) Multiplex PCR/LDR for detection of K-ras mutations in primary colon tumors. *Oncogene*, 18**,** 27-38.

20. Rand, T.A., Ginalski, K., Grishin, N.V. and Wang, X. (2004) Biochemical identification of Argonaute 2 as the sole protein required for RNA-induced silencing complex activity. *Proc Natl Acad Sci U S A*, 101**,** 14385-14389.

21. Mol, C.D., Arvai, A.S., Slupphaug, G., Kavli, B., Alseth, I., Krokan, H.E. and Tainer, J.A. (1995) Crystal structure and mutational analysis of human uracil-DNA glycosylase: structural basis for specificity and catalysis. *Cell*, 80**,** 869-878.

22. Kavli, B., Slupphaug, G., Mol, C.D., Arvai, A.S., Peterson, S.B., Tainer, J.A. and Krokan, H.E. (1996) Excision of cytosine and thymine from DNA by mutants of human uracil-DNA glycosylase. *EMBO J*, 15**,** 3442-3447.

23. Handa, P., Acharya, N. and Varshney, U. (2002) Effects of mutations at tyrosine 66 and asparagine 123 in the active site pocket of *Escherichia coli* uracil DNA glycosylase on uracil excision from synthetic DNA oligomers: evidence for the occurrence of long-range interactions between the enzyme and substrate. *Nucleic Acids Res*, 30**,** 3086-3095.

24. Savva, R., McAuley-Hecht, K., Brown, T. and Pearl, L. (1995) The structural basis of specific base-excision repair by uracil-DNA glycosylase. *Nature*, 373**,** 487-493.

25. Kwon, K., Jiang, Y.L. and Stivers, J.T. (2003) Rational engineering of a DNA glycosylase specific for an unnatural cytosine:pyrene base pair. *Chem Biol*, 10**,** 351-359.

26. Hitchcock, T.M., Gao, H. and Cao, W. (2004) Cleavage of deoxyoxanosine-containing oligodeoxyribonucleotides by bacterial endonuclease V. *Nucleic Acids Res*, 32**,** 4071-4080.

# CHAPTER THREE

## Role of Tyrosine 80 in Base Recognition in Endonuclease V from *Thermotoga maritima*

## I. Abstract

Endonuclease V recognizes and cleaves deaminated DNA and mismatch base pairs. Tyrosine 80 located in motif III is involved in protein-DNA interactions. This study investigates the role of Y80 in base recognition by substituting it with all nineteen natural amino acids. The resulting mutants were characterized biochemically using deaminated base-containing and mismatch-containing DNA. Substitutions with aromatic amino acid residues retained partial binding affinity as demonstrated by gel mobility shift analysis with double-stranded inosine-containing DNA. Uridine-containing double-stranded DNA as an adenosine/uridine base pair was only cleaved by the wild type endo V in the presence of $MgCl_2$. Uridine-containing double-stranded DNA as a guanosine/uridine base pair was cleaved by the wild type endo V and to a much less extent by Y80F in the presence of $MgCl_2$, indicating an essential role of the aromatic ring and the hydroxyl group in base recognition. For mismatched base pairs, Y80H preferred cleavage of thymidine-strand in a thymidine-containing mismatch; Y80P and Y80S, along with previously identified Y80A, enhanced the cleavage of the cytosine-strand in a cytosine-containing mismatch, Y80R tended to cleave thymidine- and cytosine-strand in a mismatch. The ability to alter cleavage preferences in a mismatch further underscores the

role of Y80 in base recognition. Base recognition by Y80 and its relationship to amino acid substitutions were discussed.

## II. Introduction

Endonuclease V (endo V) is a DNA repair enzyme that recognizes deaminated bases and mismatch base pairs and in general nicks the second phosphodiester bond 3' to the aberrant site (1-5). Its broad substrate specificity raises important questions as to how endo V recognizes deaminated bases and mispairs and the role of specific amino acid residues in base recognition. Based on the biochemical analysis of enzymatic activities towards different deaminated bases, we previously proposed that endo V may use a three-element base recognition mechanism to distinguish a deaminated base from a regular base (6). This model suggests that recognition element A of endo V may interact with the 6-keto in deaminated purine bases or 4-keto in uracil, while element C recognizes the $N^7$ position in purine bases (6). Element B may be snug between the 2-keto and $N^3$ in uracil.

To define the DNA-protein interactions in endo V, we previously conducted a systematic site-directed mutagenesis and biochemical analysis using the thermostable endo V from *Thermotoga maritima* (*Tma*) as a model system (7). This study reveals that some conserved residues such as Y80, G83 and L85 in motif III and H116, R118 and G121 in motif IV make significant contribution to base recognition. In particular, Y80 in motif III and H116 in motif IV are interesting as the hydrogen bond donor capacity of these residues appear to be important for DNA-protein interactions (7). It was speculated that Y80 and H116 might act as elements A and C in base recognition.

The notion that motifs III and IV contain important base recognition elements is further supported by biochemical analysis of mismatch cleavage activities in endo V mutants (8). In a biochemical screening using $Mn^{2+}$ as the metal cofactor, we found that several mutants in these two motifs that altered base preference in mismatch cleavage.

While the wild type endo V prefers A and G in a mismatch, Y80H showed some degree of preference for a T base (8). Y80A, on the other hand, became a C-specific mismatch endonuclease. Several mutations at H116 converted the preference to an A base. Consistent with the above-mentioned three-element model and the implication of Y80 as recognition element A, we proposed that Y80A might prevent the steric clash with a C base due to its smaller side chain, thus altering the base preference in mismatch cleavage.

The interesting biochemical properties exhibited by some Y80 mutants prompted us to survey the effects of all possible amino acid changes at position 80 in *Tma* endo V. Here, we report a biochemical analysis of nineteen Y80 mutants using deaminated base- and mismatch-containing DNA substrates. This comprehensive amino acid substitution at a single site underscores the importance of aromatic ring and hydrogen bond donor capacity in base recognition, reveals additional mutants with altered base preference in mismatch cleavage, and offers new insight on the role of Y80 in base recognition.

## III. Materials and Methods

### A. Reagents, media, and strains

All routine chemical reagents were purchased from Sigma Chemicals (St. Louis, MO), Fisher Scientific (Suwanee, GA) or VWR (Suwanee, GA). Restriction enzymes and T4 DNA ligase were purchased from New England Biolabs (Beverly, MA). DNA sequencing kits were purchased from Applied Biosystems (Foster City, CA). BSA and dNTPs were purchased from Promega (Madison, WI). *Taq* DNA polymerase was purchased from Eppendorf (Hamburg, Germany). Deoxyoligonucleotides were ordered from Integrated DNA Technologies Inc. (Coralville, IA). Duplex deoxyoligonucleotide substrates were prepared as previously described (7). LB medium was prepared according to standard recipes. GeneScan Stop Buffer consisted of 80% formamide (Amresco, Solon, OH), 50 mM EDTA (pH 8.0), and 1% blue dextran (Sigma Chemicals). TB Buffer (1X) consisted of 89 mM Tris base and 89 mM boric acid. TE buffer consisted of 10 mM Tris-HCl, pH 8.0 and 1 mM EDTA. *Escherichia coli* host strain AK53 (*mrr*B$^-$, MM294) was from our laboratory collection.

### B. Site-directed mutagenesis of *Thermotoga maritima* endo V

An overlapping extension PCR procedure was used for site-directed mutagenesis (9). PCR products digested with a pair of *Nde*I and *Bam*HI were ligated to cloning vector pEV5 treated with the same pair of restriction endonucleases. The ligated vectors were transformed into *E. coli* strain AK53 (*mrr*B$^-$, MM294). Plasmids containing inserts were reisolated and sequenced on an ABI sequencer using dye-dideoxy terminator chemistry to identify mutated sequence and ensure that the constructs were free of PCR error. An

overnight *E. coli* AK53 LB culture containing the desire site-directed mutation was diluted 100-fold into MOPS medium supplemented with 50 µg/ml ampicillin. The *E. coli* cells (1 liter) were grown at 37°C while shaking at 250 rpm overnight. The cells were collected by centrifuging at 4000 rpm at 4°C and washed once with pre-cooled PBS buffer and stored at -20°C. Protein purification and quantitation were carried out essentially as previously described (4, 10).

**C. Endonuclease V cleavage assays**

The cleavage reaction mixtures (10 µl) containing 10 mM HEPES-KOH (pH 7.4), 1 mM dithiothreitol, 2% glycerol, 5 mM $MgCl_2$ or 5 mM $MnCl_2$, 10 nM oligonucleotide DNA substrate and 10 nM of *Tma* endo V protein unless otherwise specified were incubated at 65°C for 30 min. The reactions were terminated by addition of an equal volume of GeneScan Stop Buffer (80% formamide, 50 mM EDTA (pH 8.0), and 1% blue dextran). The reaction mixtures were then heated at 94°C for 3 min and cooled down on ice. Samples (3.5 µl) were loaded onto a 10% denaturing polyacrylamide gel containing 7 M urea. Electrophoresis was conducted at 1500 volts for 1.5 h using an ABI 377 sequencer (Applied Biosystems). Cleavage products and remaining substrates were quantified using the GeneScan analysis software version 3.0.

**E. Gel mobility shift assays**

The binding reaction mixtures (20 µl) contained 100 nM fluorescently labeled oligonucleotide DNA substrates, 5 mM $CaCl_2$, 20% glycerol, 10 mM HEPES-KOH (pH 7.4), 1 mM DTT, and 100 nM of *Tma* endo V protein. The binding reactions were carried

out at 65°C for 30 min. Samples were electrophoresed on a 6% native polyacrylamide gel in 1 x TB Buffer supplemented with 5 mM $CaCl_2$. The bound and free DNA species were analyzed using a Typhoon 9400 Imager (Amersham Biosciences) with the following settings: PMT at 600 Volts, excitation at 495 nm, emission at 535 nm.

## IV. Results

### A. Cleavage of inosine-, xanthosine-, oxanosine-containing DNA

Endonuclease V contains seven conserved motifs in which motifs III and IV play a major role in protein-DNA interactions (7). Y80 located in motif III is implicated in base recognition (8). To better define the role of Y80 in protein-DNA interactions and base recognition, we systematically substituted Y80 with all nineteen amino acids. Since endo V recognizes and cleaves all three deaminated purine bases, hypoxanthine, xanthine and oxanine, we tested the cleavage activity of these mutants against inosine (I)-, xanthosine (X)-, oxanosine (O)-containing oligonucleotide substrates first. The deaminated purine bases were placed in the bottom strand at position 37 (Fig.3.1A). Endo V in general cleaves at the second phosphodiester bond from the lesion at the 3' side (1-4). Therefore, cleavage by Y80 and its mutants would generate a 38-mer product (Fig.3.1). Under that assay condition that the enzyme:substrate ratio of 1:1, all the mutants achieved a close-to-complete cleavage of the T/I substrate (Fig.3.1B), suggesting that inosine in DNA is still recognized by Y80 mutants. A similar cleavage pattern was observed with C/X substrate, i.e., all the Y80 mutants as well as the wild type enzyme, achieved a close-to-complete cleavage of the T/X substrate (data not shown). On the other hand, C/O substrate was only partially cleaved by the wt enzyme. None of the nineteen Y80 mutants exhibited detectable cleavage of the C/O substrate (data not shown), suggesting that the any amino acid substitution at Y80 disables its oxanosine endonuclease activity. These results are consistent with a previous large-scale mutagenesis analysis, which indicated that oxanosine cleavage is vulnerable to amino acid substitutions at the positions that affect base recognition (7).

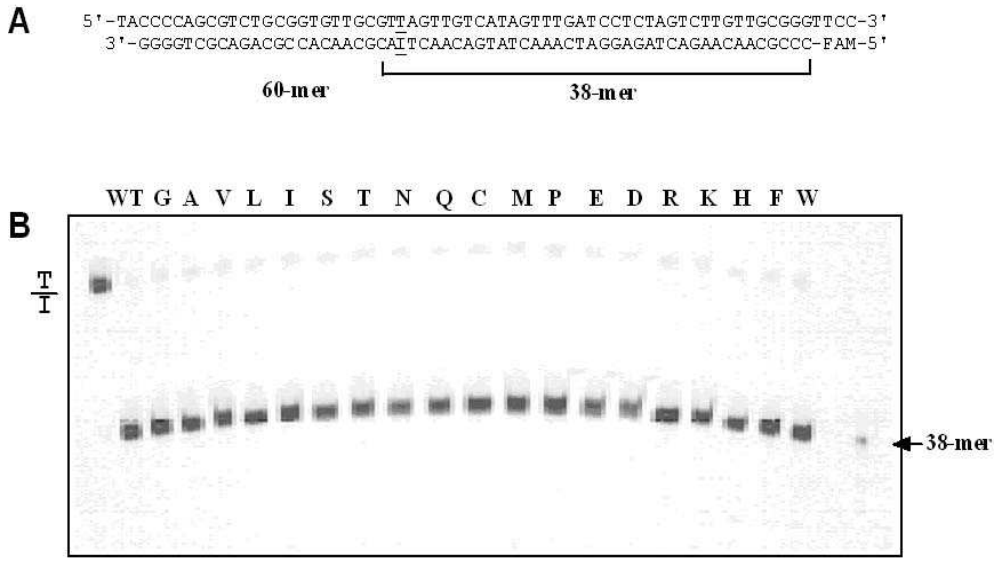



**Figure 3.1 Cleavage of T/I substrate by Y80 *Tma* endonuclease V mutants.** Cleavage reactions were performed as described in Materials and Methods with 5 mM $MgCl_2$. **A.** Lesion-containing deoxyribooligonucleotide substrate. I: deoxyinosine. The 5' end of the bottom strand was labeled with FAM fluorophore. **B.** Cleavage activity of wt and Y80 *Tma* endo V mutants on inosine-containing substrate (T/I). The single letter amino acid code is shown above the gel

**B. Binding of inosine-containing DNA and kinetic analysis of T/I cleavage**

One of the unique enzymatic features endo V possesses is that the enzyme retains tight binding to inosine- and xanthosine-containing DNA after cleavage, resulting in single-turnover of these substrates (4, 5, 7). To compare the binding affinity of the Y80 mutants with the wt enzyme, we performed gel mobility shift assay using the T/I substrate (Fig. 3.2A). Among the three aromatic ring-containing mutants (Y80F, Y80H and Y80W), Y80F retained the highest degree of binding affinities to the T/I substrate (Fig. 3.2B). The binding affinity of Y80R to the T/I substrate was about one fourth of level demonstrated by the wt enzyme (Fig. 3.2B). Other mutants either showed no detectable retarded band by gel mobility shift or very low level binding to the T/I substrate (Fig. 3.2B). To evaluate the binding affinities to the cleavage product, we

performed gel mobility shift assay using the nicked T/I substrate (Fig. 3.2C). The four mutants (Y80F, Y80H, Y80W and Y80R) that retained significant binding to the T/I substrate also retained binding to the nicked T/I substrate (Fig. 3.2D). Y80F again exhibited highest degree of binding affinity among all nineteen mutants, while the other mutants showed low or no detectable binding to the nicked T/I substrate (Fig. 3.2D). In repeated experiments, Y80R consistently showed some smearing in the retarded band, indicating that the Y80R-T/I complex or the Y80R-nicked T/I complex is less stable that the complexes formed by the wt enzyme and some other mutants such as Y80F (Fig. 3.2A, 3.2C).

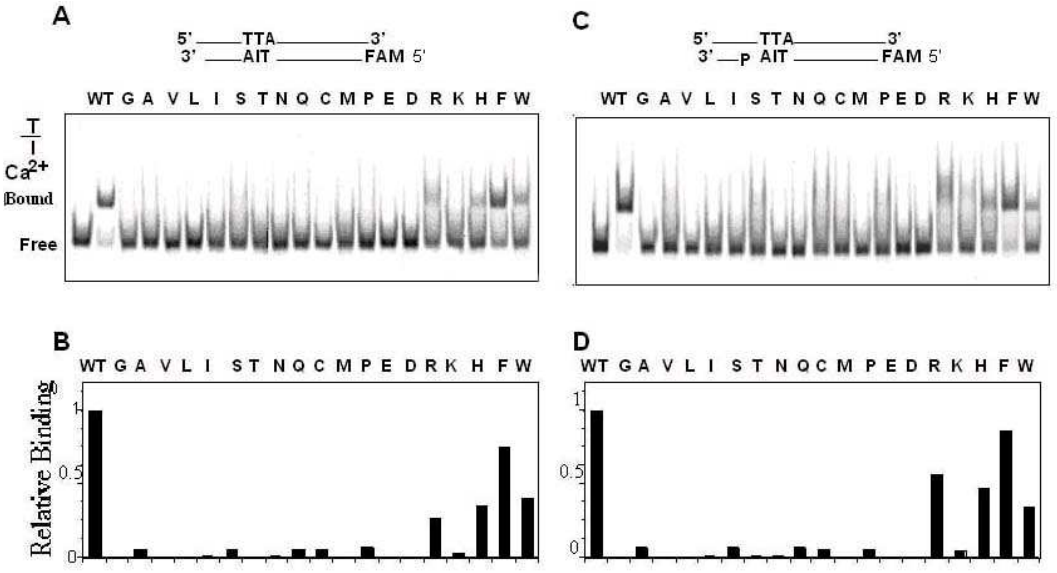



**Figure 3.2 Binding analysis of Y80 *Tma* endo V mutants on double-stranded inosine-containing DNA.** Gel mobility shift assays were performed as described in Materials and Methods. **A.** Gel mobility shift analysis of binding of Y80 *Tma* endo V mutants to double-stranded inosine-containing substrate (T/I) with 5 mM $CaCl_2$. **B.** Quantitative analysis of binding affinity of Y80 *Tma* endo V mutants to T/I substrate. The binding affinity of the wt *Tma* endo V was taken as 1.0. **C.** Gel mobility shift analysis of binding of Y80 *Tma* endo V mutants to nicked double-stranded inosine-containing product (nicked T/I) with 5 mM $CaCl_2$. **D.** Quantitative analysis of binding affinity of Y80 *Tma* endo V mutants to nicked T/I product. The binding affinity of the wt *Tma* endo V was taken as 1.0.

Previously, we have observed that tight binding may limit the turnover of the enzyme to inosine-containing DNA substrates (7, 10). To test the kinetic behavior of the Y80 mutants, we conducted a time-course analysis at the enzyme: substrate ratio of 1:10. As expected, the wt enzyme behaved as a single-turnover enzyme (Fig. 3.3). Among the nineteen mutants examined, Y80F showed limited turnover of the T/I substrate, in keeping with the retention of tight binding to the nicked T/I (Fig. 3.2D). Other mutants achieved close-to-complete cleavage, suggesting that the reduced binding affinity caused the mutants to dissociate from the nicked T/I (Fig. 3.3 and data not shown).

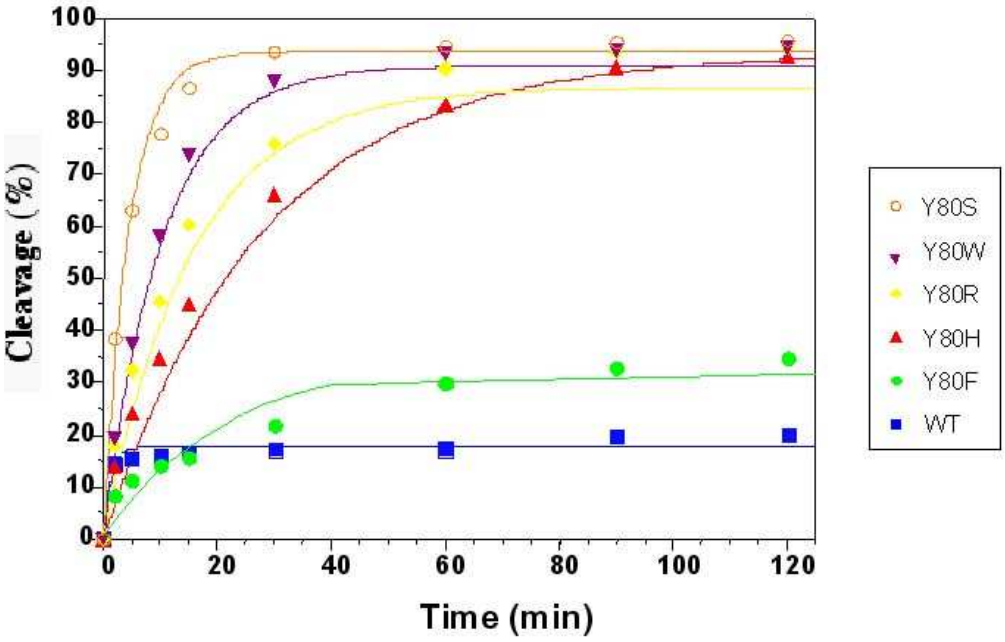



**Figure 3.3 Representative time course analysis of T/I cleavage by wt and Y80 *Tma* endo V mutants.** Cleavage reactions were performed as described in Materials and Methods with 1 nM *Tma* endo V protein (E:S = 1:10) and 5 mM $MgCl_2$. Reactions were stopped on ice at indicated time points and followed by adding equal volume of GeneScan Stop Buffer. Time course of (■) wt *Tma* endo V; (●) Y80F; (▲) Y80H; (♦) Y80R; (▼) Y80W; (○) Y80S.

**C. Cleavage of uridine-containing DNA**

Deamination of cytosine yields uracil in DNA, which is recognized and cleaved by endo V (4, 5, 11, 12). However, a uridine in a mismatched base pair seems to be more recognizable than an A/U base pair, which maintains Watson-Crick base pair symmetry (4, 5). To investigate the effects of amino acid substitutions may have on uridine endonuclease activity, we tested cleavage of both A/U and G/U substrates. In the presence of $Mg^{2+}$, the wt enzyme was the only one that cleaved A/U to a limited extent (Fig. 3.4A). Both the wt endo V and the Y80F cleaved the G/U substrate, albeit the latter was about 15% of the wt activity (Fig. 3.4B). Since $Mn^{2+}$ often relaxes the specificity of nucleases, we then tested uridine endonuclease activity in the presence of $Mn^{2+}$. For the A/U substrates, nine mutants exhibited detectable cleavage activity (Fig. 3.4C). The general trend was that substitutions by aromatic rings (Y80H, Y80F, Y80W) retained higher activity followed by two substitutions with small side chains (Y80C and Y80A). Other substitutions (Y80I, Y80S, Y80R and Y80K) showed lowest activity (Fig. 3.4C). For the G/U mismatched base pair, Y80G, Y80V, Y80L, Y80T, Y80M, Y80E and Y80D did not show detectable activity (Fig. 3.4D). The cleavage activity of the rest of Y80 mutants followed the order of Y80A = Y80F > Y80A, Y80S > Y80C, Y80H > Y80W, Y80P, Y80R > Y80K > Y80I, Y80N, Y80Q, Y80V, Y80L (Fig. 3.4D).

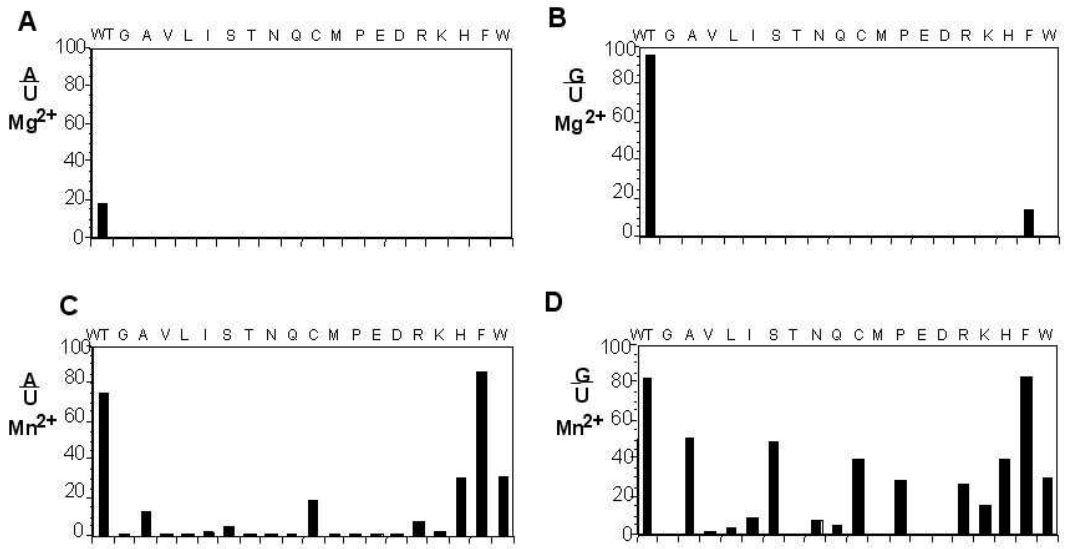



**Figure 3.4 Cleavage of uridine-containing DNA by Y80 *Tma* endonuclease V mutant.** Cleavage reactions were performed as described in Materials and Methods with inosine-containing substrate substituted by uridine-containing substrate. **A.** A/U cleavage by endo V in the presence of 5 mM $MgCl_2$. **B.** G/U cleavage by endo V in the presence of 5 mM $MgCl_2$. **C.** A/U cleavage by endo V in the presence of 5 mM $MnCl_2$. **D.** G/U cleavage by endo V in the presence of 5 mM $MnCl_2$.

## C. Cleavage of mismatched base pairs

The analysis of uridine cleavage illustrates that endo V recognize uridine more readily in a mismatched base pair than a Watson-Crick base pair. In fact, in addition to recognition and cleavage of all four deaminated bases, endo V is also capable of cleaving mismatched base pairs, especially with $Mn^{2+}$ as the metal cofactor (3). Previously, we reported that a single alanine substitution at Y80 position switched the enzyme to essentially a C-specific mismatch endonuclease, which recognized and cleaved A/C, C/A, T/C, C/T and even the previously refractory C/C mismatches (8). Therefore, it is of interest to evaluate the cleavage of mismatched base pairs by all Y80 mutants. In addition to Y80A that prefers C-containing mismatches and Y80H that prefers T-containing mismatches (8), several mutants showed interesting cleavage patterns for the mismatched base pairs in the presence of $Mn^{2+}$ (Fig. 3.5). The majority of mutants

showed low or no detectable mismatch cleavage (Fig. 3.5). However, similar to Y80A, Y80P and Y80S showed preference to C-containing mismatches as demonstrated by cleavage of C-containing strands in C/A, C/T, A/C, T/C and weak cleavage of C/C (Fig. 3.6). Y80R somewhat reduced cleavage of purine-containing mismatches while enhanced cleavage of pyrimidine-containing mismatches (Fig. 3.6). These data underscore the role of Y80 in mismatch recognition and cleavage.

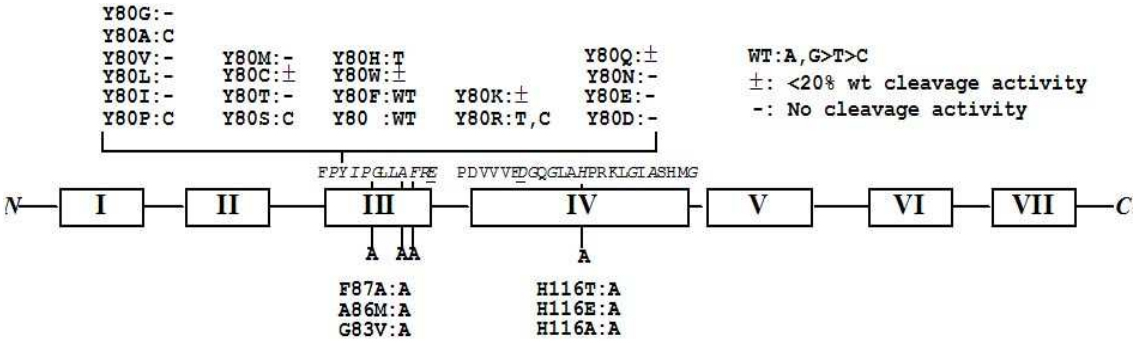



**Figure 3.5 Base preference of mismatch cleavage of Y80 *Tma* endonuclease V mutants.** Cleavage reactions were performed as described in Materials and Methods with 5 mM $MnCl_2$. Motifs are shown in Roman numbers. The sequence of motif III is FPYIPGLLAFRE (The catalytic residue E89 is underlined). See (7) for sequence alignment.

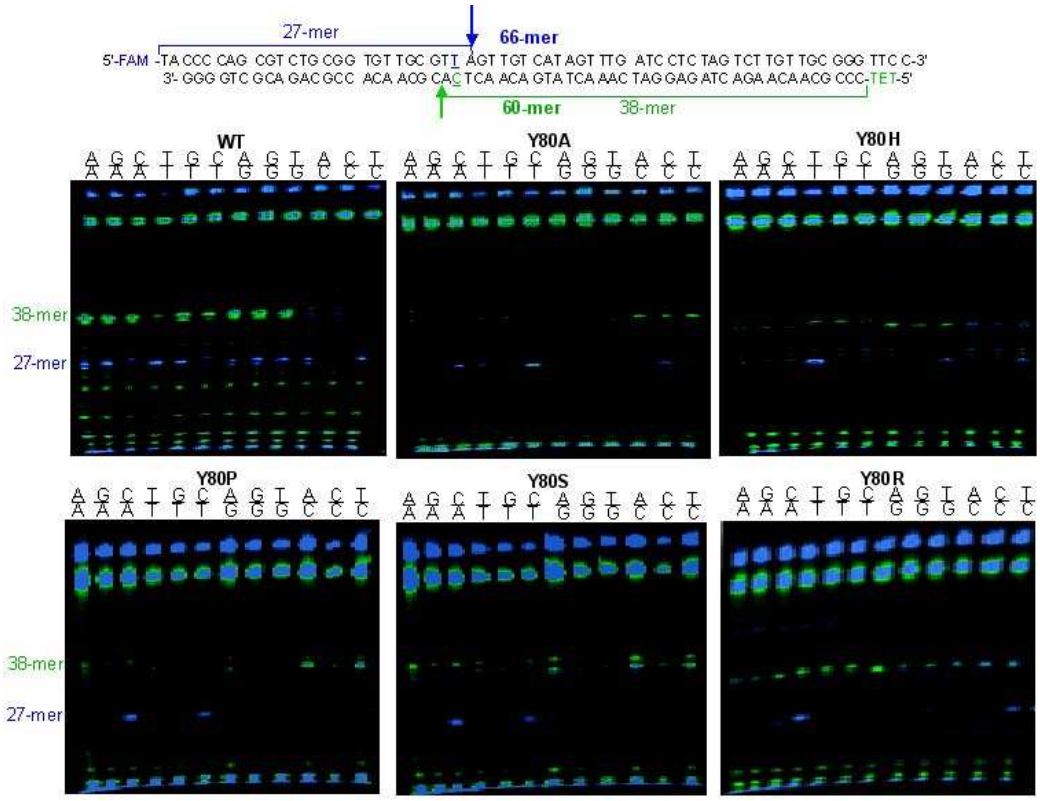



**Figure 3.6 Representative GeneScan gel pictures of mismatch cleavage**. Cleavage reactions were performed as described in Materials and Methods with 100 nM endo V protein and 5 mM $MnCl_2$.

## V. Discussion

Y80 in endonuclease V is an invariant residue located in a highly conserved region we defined as motif III [Fig. 3.5 and (7)]. Several amino acids in this motif affect DNA-protein interactions (7). E89 in this motif is part of the active site that may involve coordination of a catalytic metal ion (12). Previously, we proposed a three-element base recognition mechanism for endo V, which suggests that recognition elements A and C may interact with the 6-keto (or 4-keto in uracil) and $N^7$ position of a deaminated purine base while element B may be snug between the 2-keto and $N^3$ in uracil (6) (Fig. 3.7). A biochemical analysis implicates that Y80 may potentially play a role as element A in base recognition (7). This notion is further strengthened by the observation of altered base preference in mismatch cleavage (8). Prompted by insight from previous studies, this work intended to conduct a systematic analysis by substituting Y80 with all nineteen possible amino acids to gain a deeper understanding of the role of Y80 in base recognition.

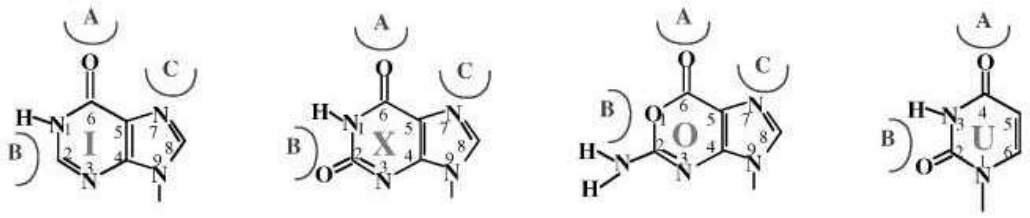



**Figure 3.7 A hypothetic model of deamianted base recognition mechanism by endonuclease V.** See text for detail. I: Inosine X: Xanthosine O: Oxanosine U: Uridine. Adapted from (6).

**A. Substitutions with aromatic residues at Y80 position**

One of the major findings revealed from this study is that aromatic amino acid residues play an important role at this position. The Y80F mutant retains much of the binding affinity to both the T/I substrate and the nicked T/I product (Fig. 3.2), indicating that the aromatic benzene ring of the tyrosine residue in the wt enzyme obviously is involved in Protein-DNA interactions. The requirement for the tyrosine residue is specific and can only be substituted partially by other aromatic residues as seen in Y80H and Y80W. From the kinetic analysis of the T/I cleavage, it is clear that the wt enzyme and the Y80F mutant are the only ones that maintain single-turnover character (Fig. 3.3).

This conclusion is also well illustrated by cleavage of uridine-containing DNA (Fig. 3.4). The wt enzyme is the only one that shows cleavage of A/U, while the wt enzyme and the Y80F mutant are the only ones that are active on the G/U in the presence of $Mg^{2+}$ (Fig. 3.4A-3.4B). The discrepancy in A/U and G/U cleavage is likely due to mismatch nature of the G/U base pair. For all uridine-containing substrates, endo V enzymes prefers mismatched uridine-containing base pairs (4, 5). It appears that the ability of endo V to recognize a Watson-Crick base paired A/U is highly dependent on the tyrosine residue at this position (Fig. 3.4A). How a tyrosine residue play such a role is still unknown. However, among the other DNA repair enzymes, when F114 in MutM (a glycosylase that recognizes 8-oxoguanine) was inserted into duplex DNA, causing loss of helix stacking and localized destabilization of target base pair (14). Whether Y80 in endo V plays a similar role in base recognition or not remains to be seen.

The three aromatic side chains show different cleavage patterns of mismatched base pairs (Fig. 3.5), while Y80F maintains wt level of cleavage, Y80W is not quite as

active on mismatches. Y80H is more distinct as it somewhat prefers T-containing mismatches (Fig. 3.6). The structural basis of such an alteration is not clear. Yet, we speculate that the imidizole ring may facilitate the recognition of a thymine base in a mismatch by formation of a hydrogen bond with the C4-keto group.

**B. Substitutions with small side-chain residues at Y80 position**

The systematic site-directed mutagenesis allows a complete view of amino acid substitutions. Another finding emerged from this study is that substitutions with small side-chain amino acids help maintain mismatch cleavage or alter base preference in mismatch cleavage. The first scenario is exemplified in the case of G/U cleavage in the presence of $Mn^{2+}$ (Fig. 3.4D). Y80A, Y80S and Y80C retain greater than 50% of cleavage activity as compared with the wt enzyme. The second scenario is observed in mismatch cleavage by Y80A and Y80S (Fig. 3.6). Both mutants prefer C-containing mismatches rather than the purine-containing mismatches as seen in the wt enzyme. It is not clear why substitutions with small side-chain residues retain greater uridine endonuclease activity in the G/U base pair. However, the alteration of base preferences in mismatch cleavage may be explained by steric factors. In the wt enzyme, cytosine in a mismatch may spatially clash with the bulky tyrosine residue. Substitutions with Ala and Ser may prevent the steric hindrance, thus accommodate a C base in a mismatch. How to explain the narrowing of mismatch cleavage specificity to almost exclusively a C-only endonuclease? Both T and G bases may be excluded because they lose a favorable interaction between the hydroxyl group in Y80 and the keto group at the C4 or C6 position (6, 7). How the A base is excluded is not obvious from the model. Yet, we

speculate that with a small side-chain substitution at Y80, the recognition pocket may become more compact that makes the A base (and possibly the G base as well) difficult to fit into it. Consequently, the A base is also excluded. A caveat is that this model would not explain all the small side-chain substitutions. The lack of mismatch cleavage by Y80G is a case in point. Amino acid substitutions may cause effects beyond steric factors.

**C. Substitutions with Arg and Pro at Y80 position**

Y80R and Y80P exhibit some unique effects on endo V. First, Y80R still retains some binding affinity to the T/I substrate and the nicked T/I product (Fig. 3.2). Second, both Y80R and Y80P cleave the uridine strand in G/U in the presence of $Mn^{2+}$ (Fig. 3.4D). Third, both mutants show altered base preference in mismatch cleavage (Fig. 3.6). While Y80R prefers T and C, Y80P predominantly cleaves the C-containing strand in a mismatch. These observations can not be simply explained by steric effects, perhaps, Y80R and Y80P cause localized conformational changes so that only pyrimidine base is accommodated in the active site during mismatch cleavage.

In summary, this work probes the effects of all nineteen amino acid substitutions on endonuclease V. Data presented here further underscore the role of Y80 in recognition of deaminated bases and mismatched base pairs. The aromatic ring and the hydroxyl group in Y80 play important role in base recognition. Although Y80A, Y80S, Y80P and Y80R cause a similar alteration in base preference in mismatch cleavage, but the underlying mechanism may not be the same. These observations, thus, highlight the dynamic nature of the active site in endonuclease V.

## VI. Acknowledgments

This project was supported in the Cao lab in part by CSREES/USDA (SC-1700274), the NIH (GM 067744), the Concern Foundation and DOD-Army Research Office (W911NF-05-1-0335). We thank members of Cao lab for stimulating discussions.

## VII. References


1. Yao, M., Hatahet, Z., Melamede, R. J., and Kow, Y. W. (1994) Purification and characterization of a novel deoxyinosine-specific enzyme, deoxyinosine 3' endonuclease, from *Escherichia coli, J Biol Chem 269*, 16260-8.

2. Yao, M., Hatahet, Z., Melamede, R. J., and Kow, Y. W. (1994) Deoxyinosine 3' endonuclease, a novel deoxyinosine-specific endonuclease from *Escherichia coli*, *Ann N Y Acad Sci 726*, 315-6.

3. Yao, M., and Kow, Y. W. (1994) Strand-specific cleavage of mismatch-containing DNA by deoxyinosine 3'-endonuclease from *Escherichia coli, J Biol Chem 269*, 31390-6.

4. Huang, J., Lu, J., Barany, F., and Cao, W. (2001) Multiple Cleavage Activities of Endonuclease V from *Thermotoga maritima*: Recognition and Strand Nicking Mechanism, *Biochemistry 40*, 8738-8748.

5. Feng, H., Klutz, A. M., and Cao, W. (2005) Active Site Plasticity of Endonuclease V from Salmonella typhimurium, *Biochemistry 44*, 675-83.

6. Hitchcock, T. M., Gao, H., and Cao, W. (2004) Cleavage of deoxyoxanosine-containing oligodeoxyribonucleotides by bacterial endonuclease V, *Nucleic Acids Res 32*, 4071-80.

7. Feng, H., Dong, L., Klutz, A. M., Aghaebrahim, N., and Cao, W. (2005) Defining Amino Acid Residues Involved in DNA-Protein Interactions and Revelation of 3'-Exonuclease Activity in Endonuclease V, *Biochemistry 44*, 11486-95.

8. Gao H, Huang J, Barany F, Cao W. Switching base preferences of mismatch cleavage in endonuclease V: an improved method for scanning point mutations. Nucleic Acids Res. 2007;35(1):e2. Epub 2006 Nov 27.

9. Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K., and Pease, L. R. (1989) Site-directed mutagenesis by overlap extension using the polymease chain reaction, *Gene 77*, 51-59.

10. Huang, J., Lu, J., Barany, F., and Cao, W. (2002) Mutational analysis of endonuclease V from *Thermotoga maritima*, *Biochemistry 41*, 8342-50.

11. Yao, M., and Kow, Y. W. (1997) Further characterization of *Escherichia coli* endonuclease V, *J Biol Chem 272*, 30774-30779.

12. Gates, F. T., 3rd, and Linn, S. (1977) Endonuclease V of *Escherichia coli, J Biol Chem 252*, 1647-53.

13. Feng, H., Dong, L., and Cao, W. (2006) Catalytic mechanism of endonuclease v: a catalytic and regulatory two-metal model, *Biochemistry 45*, 10251-9.

14. Banerjee, A., Santos, W. L., and Verdine, G. L. (2006) Structure of a DNA glycosylase searching for lesions, *Science 311*, 1153-7.

# CHAPTER FOUR

## A Cell-Free System from *Escherichia coli* for Study of Endonuclease V Initiated DNA Repair


## I. Abstract

Deamination of deoxyadenosine can occur spontaneously, generating highly mutagenic deoxyinosine which can cause A:T to G:C transition mutation. *Escherichia coli* endonuclease V (*E. coli* Endo V) recognizes deoxyinosine-containing DNA and cleaves the DNA strand at the second phosphodiester bond 3' to the damage, leaving a nick with 3'-hydroxyl and 5'-phosphoryl groups. Endo V remains bound to the nicked lesion-containing DNA, exhibiting similar affinity to both the substrate and the product. Little is known about the subsequent repair processes after the nicking step initiated by endo V. In order to understand the repair pathway initiated by endo V, we established a cell-free system from *E. coli* to monitor the repair *in vitro*. Plasmid containing an inosine lesion was incubated with cell-free extracts made from wild type and *nfi* mutants *E. coli* strains in the presence of $\alpha^{32}$P-dCTP and non-radiolabed dNTPs to characterize the *E. coli* endo V mediated DNA repair pathway. The cell-free reactions showed significant repair synthesis using wild type *E. coli* cell-free extracts while little repair synthesis was monitored using *nfi* mutant cell-free extracts. The preliminary results obtained from the cell-free assays indicated that the repair patch of endo V mediated DNA repair pathways may consist of a long patch (longer than 2.3 kb, estimated) and a short patch (less than 330 bp, estimated).

## II. Introduction

The DNA molecules that carry the vital genetic material of cells are constantly subject to spontaneous decay under physiological conditions. Hydrolytic deamination occurs at non-negligible rates (1-2). For example, hydrolytic deamination of cytosine has been estimated to introduce between 100 and 500 uracil residues in the form of U:G mismatches per cell per day (3). Nitric oxide (NO), nitrous acid and high temperature can further enhance this process (4-5). It is speculated that chronic inflammation is associated with cancer. Tissues with chronic inflammation have high concentration of NO secreted by activated macrophage which can induce nitrosative damage to DNA in cells, thereby increasing cancer risk (6-8).

Uracil, hypoxanthine, and xanthine are the deamination products of cytosine, adenine and guanine, respectively. Deamination of adenine to hypoxanthine causes A:T to G:C transition mutation. Endonuclease V, an evolutionarily conserved DNA repair enzyme, found in bacterial, archaeal and eukaryotic cells, recognizes deaminated DNA. This enzyme incises the DNA at the second phosphodiester bond 3' to a deoxyinosine site (9). The enzyme has almost equal affinity for the substrate and the product (10). It stays bound with the product after the cleavage. Genetic studies of *E. coli* endo V suggest that *E. coli* endo V is involved in the repair of hypoxanthine, xanthine and abasic sites in DNA (11-12). It was showed that *E. coli* endo V is the primary DNA repair enzyme deals with deoxyinosine lesion *in vivo* while AlkA is not involved in the repair of deoxyinosine *in vivo* (11-12). It was suggested that endo V may use a novel alternative excision repair pathway, in which an exonuclease is required to remove the lesion (13-15). The gap is then filled in by polymerase and sealed by ligase to finish the repair. Both

*in vivo* and *in vitro* studies of DNA repair mechanism have used DNA substrates which contain specific DNA lesions at defined sites. For *in vitro* studies using purified proteins, short linear synthetic substrates are useful. For *in vivo* studies or *in vitro* studies using cell-free extracts, closed circular duplex substrates are more desirable to avoid substrate degradation and other effects that short linear substrates may encounter (16).

Studies of the repair patch size associated with gap-filling DNA synthesis using cell-free extracts and closed circular plasmids have been conducted by many researchers (16-18). In order to understand the repair pathway initiated by endo V, we established a cell-free system to monitor the repair *in vitro*. We constructed a pUC19-derived plasmid which contained two nicking sites (N.BstNB I and N.Bbvc IA) at the same strand with 32 flanking bases. The plasmid was nicked by the two nicking enzymes and a fragment of 32 bases was removed, producing a gapped plasmid. 32-nt oligonucleotides containing a single deoxyinosine was annealed and ligated to the gapped plasmid. Closed circular inosine-containing plasmid was purified by the CsCl purification method. The inosine-containing substrate was confirmed by converting closed circular form of the plasmid to nicked form by *E. coli* endo V. We incubated inosine-containing plasmid with different endonuclease V (*nfi*) mutants (Table 4.1) [BW1034 (*nfi*$^-$,*ung*$^-$) (ung is the gene for uracil DNA glycosylase), BW1162 (*nfi*$^-$,*nfo*$^-$) (nfo is the gene for endonuclease V), BW1163 (*nfi*$^-$,*nfo*$^-$,*xth*$^-$) (xth is the gene for exonuclease III), BW1179 (*nfi*$^-$,*alkA*$^-$) (alkA is the gene for AlkA)] cell-free extracts and AB1157 (*nfi*$^+$) cell-free extract together with $\alpha^{32}$P-dCTP and non-radiolabed dNTPs. The results showed that significant repair synthesis occurred using wt *E. coli* cell-free extracts, while using *nfi*$^-$ *E. coli* cell-free extracts only

a little repair synthesis was monitored. This confirmed that endo V is the major enzyme responsible for inosine repair.

To determine the repair patches of endo V initiated repair, we used Hpa II and EcoR I to digest the plasmids recovered from the cell-free assays. Following the restriction enzyme digestion and non-denaturing gel electrophoresis, the extent of incorporation of the radioactively labeled dCTP in the plasmid was determined by autoradiography. The results from the cell-free assays indicated that the repair patches of endo V-initiated DNA repair pathways may consist of a long (longer than 2.3 kb, estimated) and a short patch (less than 330 bp, estimated)

## III. Materials and Methods

### A. Materials

All routine chemical reagents were purchased from Fisher Scientific (Suwanee, GA), Sigma Chemicals (St. Louis, MO), VWR (Suwanee, GA). Restriction enzymes, nicking enzymes, T4 DNA ligase, T4 polynucleotide kinase and BSA were purchased from New England Biolabs (Beverly, MA). dNTPs were purchased from Promega (Madison, WI). $\alpha^{32}$P-dCTP and $\gamma^{32}$P-ATP were purchased from MP Biomedicals (Irvine, CA). Purified deoxyribooligonucleotides were ordered from Integrated DNA Technologies Inc. (Coralville, IA). YM-3 and YM-100 spin columns were purchased from Milipore Corporation (Bedford, MA). LB medium was prepared according to standard recipe. 1 X TBE buffer contains 89 mM Tris-Boric acid and 2 mM EDTA. 10 X agarose gel loading buffer contains 0.02% bromophenol blue, 25% sucrose, 0.01 M EDTA, 1% SDS, 0.25% xylene cyanol. 1 X T4 DNA ligase buffer contains 50 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, 10 mM dithiothreitol, 1 mM ATP, 25 ug/ml BSA. 1 X T4 polynucleotide kinase Buffer contains 70 mM Tris-HCl (pH 7.6), 10 mM $MgCl_2$, 5 mM dithiothreitol. Protein purification and quantitation were carried out as previously described (19).

### B. Preparation of *E. coli* cell-free extracts

*E. coli* wt strain AB1157 was obtained from the *E. coli* Genetic Stock Center (MCDB Department, Yale University). BW1034 (*nfi* ¯, *ung* ¯), BW1162 (*nfi* ¯, *nfo* ¯), BW1163 (*nfi* ¯, *nfi* ¯, *xth* ¯), BW1179 (*nfi* ¯, a*lkA* ¯) were kind gifts from B. Weiss (Emory University) (Table 4.1). Cell-free extracts were prepared according to (17). *E. coli* cells

were grown to an $A_{590}$=1.0-1.2 in 1-Liter cultures of LB media [with or without antibiotics, (wt: without antibiotic), (BW1034: 34ug/ml chloramphenicol and 50ug/ml thymine), (BW1162: 34ug/ml chloramphenicol and 10ug/ml kanamycin), (BW1163: 34ug/ml chloramphenicol and 10ug/ml kanamycin), (BW1179: 34ug/ml chloramphenicol)] supplemented with 0.1% glucose. Cells were collected by centrifugation at 4°C and suspended in 5 ml of 0.05 M Tris-HCl (pH 7.6) with 10% sucrose, and frozen in liquid nitrogen and stored at -70°C. The thawed cells were supplemented with 1.2 mM DTT, 0.15 M KCl, 0.23 mg/ml lysozyme, followed by heat shock at 37°C until a final suspension temperature of 20°C. After cell debris was removed by centrifugation at 13,000 rpm for 20 min at 4°C, proteins were precipitated from the supernatant by addition of 0.5 g of powdered ammonium sulfate per ml of the extract, and the precipitate was recovered by centrifugation at 13,000 rpm for 20 min at 4°C. The pellet was resuspended in 0.3 ml of a buffer contains 0.025 M HEPES (pH7.6)/0.1 mM EDTA/ 2 mM DTT/100 mM KCl, and dialysized against the same buffer until the conductivity reaches the equivalent of 0.2 M KCl as measured by a hand held conductivity meter (model PM6304  Control Company). The resulting solution was aliquoted and frozen by liquid nitrogen and stored at -70°C. Protein concentration was estimated by Bradford method.

| Strain | Description[a] | Source | Note |
|---|---|---|---|
| AB1157 | *hisG4 argE3* | *E. coli* genetic stock center | WT |
| BW1034 | BW1138 *nfi[b]-1::cat ung[c]-153::kan* | B. Weiss | P1(BW1160) × BW1138 |
| BW1162 | AB1157 *nfi-1::cat nfo[d]-1::kan* | B. Weiss | P1(BW1160) × RPC500 |
| BW1163 | AB1157 *nfi-1::cat Δ(xth[e]-pncA)90 nfo-1::kan* | B. Weiss | P1(BW1160) × RPC501 |
| BW1179 | KD1092 *(trpA58) nfi-1::cat alkA[f]1 his+* | B. Weiss | P1(BW1185) × BW1178 |

a. All the strains used are derivatives of *E. coli* K-12

b. nfi is the gene for endonuclease V

c. ung is the gene for uracil DNA glycosylase

d. nfo is the gene for endonuclease IV

e. xth is the gene for exonuclease III

f. alkA is the gene for 3-Methyladenine-DNA Glycosylase II (AlkA)

**Table 4.1 Bacterial strains used.**

**C. Construction of a pUC19 derivative plasmid encoding two different nicking sites**

pUC19HE (2 ug) which is a pUC19 derivative containing one N.BstNB1 nicking site (a gift from John B. Hays, Oregon State University) and pUC19 were digested with 20 U Afl III and Aat II at 37°C for 3 hours. The 1.8 kb fragment of pUC19HE and 0.87 kb fragment of pUC19 recovered from 1% agarose gel were ligated together to generate pUC19E. A synthesized double-stranded oligo (5'-CGGGGTACCGCTGAGGAGATC TGGATCCACATGTGGG-3', complementary strand 5'-CCCACATGTGGATCCAGAT CTCCTCAGCGGTACCCCG-3') and pUC19E were digested with Kpn I and Afl III and ligated together to generate a plasmid which has two nicking sites N.BstNB1 and N.BbvcIA in the same strand with 32 flanking bases. The standard procedures of

molecular cloning were carried out as described in (20). The new plasmid was designated as pUC19EN3.

**D. Preparation and purification of a single inosine-containing plasmid**

pUC19EN3 (1.5 mg) was purified using Qiagen Plasmid Mega Kit (Qiagen Science, Maryland), starting with 1 L of bacteria culture. 60 ug of pUC19EN3 was completely digested by 200 U of N.Bbvc IA at 37°C overnight as determined by 1% agrose gel. The nicked plasmid was purified by phenol/chloroform/isoamyl alcohol extraction followed by ethanol precipitation. The purified nicked plasmid was then digested by 200 U of N.BstNB I at 37°C overnight (Fig 4.1). Previous experiments showed that 200 U of N.BstNB I was sufficient to completely digest 60 ug of pUC19EN3 at 37°C (Data not shown). The doubled nicked plasmid was purified by phenol extraction followed by ethanol precipitation. The recovered plasmid was dissolved in TE buffer (pH 7.4) to a final concentration of 100 ng/ul. It was heated at 85°C for 5 minutes and put on ice quickly.

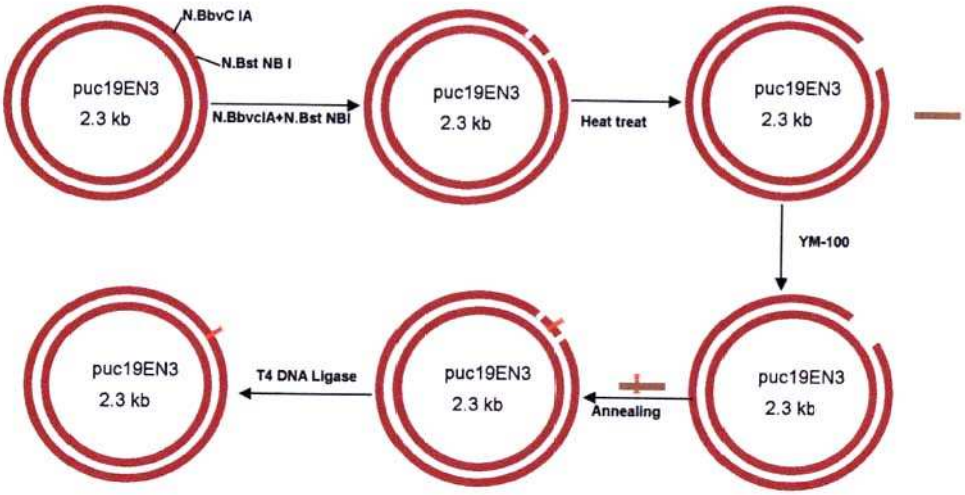



**Figure 4.1 A scheme for preparing inosine-containing circular substrates.** pUC19EN3 was digested by nicking enzymes N.BstNBI and N.BbvcIA respectively, the double nicked plasmid was denatured by heat treatment at 85°C for 5 minutes and the small fragment was removed by YM-100 spin column. The resulting gapped pUC19EN3 was annealed with an inosine-containing oligo and ligated to yield the closed circular pUC19EN3-I.C which contains a single inosine lesion.

To purify the resulting gapped plasmid, the small fragment was removed by passing through an YM-100 spin column and washed with 500 ul TE (pH 7.4) five times. Gapped plasmid (200 ng) was ligated overnight at 4°C using T4 DNA ligase. Agarose gel analysis (1%) showed that there was no significant ligation, indicating the gapped plasmid was relatively pure. The gapped plasmid was annealed with 20 fold excess of oligo containing a deoxyinosine (5'-p-TGAGGAGATCTGIATCCACATGTGAGTCCG AT-3') and then ligated overnight at 4°C using T4 DNA ligase. The ligated products were then subject to CsCl-Ethidium Bromide Gradient Purification and the deoxyinosine-containing plasmid was collected and purified according to (20). Specifically, closed circular plasmid was withdrawn by using a disposable syringe; Ethidium bromide was removed by extracting the DNA solution with water-saturated n-butanol until all the pink color disappeared from both the aqueous phase and the organic phase. CsCl was removed

from the DNA solution by passing through a YM-100 microcon spin column (Milipore Corporation, MA) and washed five times with 500 ul TE (pH 7.4). The product was aliquoted, quickly frozen by liquid nitrogen and stored at -70°C. DNA samples were run on 1% agarose gels and quantified by Quantity One Software of Gel Doc system (Bio-Rad, CA)

**E. *E. coli* endonuclease V cleavage assays**

The reaction mixtures (10 ul) containing 10 mM HEPES-KOH (pH 7.4), 1 mM DTT, 2% glycerol, 5 mM $MgCl_2$, 200 ng deoxyinosine-containing closed circular substrate or control plasmid, and 20 nM *E. coli* endo V protein were incubated at 37°C for 30 minutes. The products were resolved by electrophoresis in 1% agarose gel in 1 X TBE buffer (Tris-Boric acid 89 mM, EDTA 2 mM).

**F. *In Vitro* cell-free assays**

Reactions mixture (50 ul) contained 10 mM HEPES (pH 7.4), 2 mM DTT, 0.5 mM EDTA, 2 mM ATP, 20 uM of dNTP, 10 uci of $\alpha^{32}$P-dCTP, 0.5 mM NAD, 50 ug/ml BSA, 5 mM $MgCl_2$, 1 mM glutathione, 0.2 mM spermidine, 300 ng inosine-containing plasmid or control plasmid and 4 mg/ml of cell-free extract proteins. The mixture was incubated at 37°C for 1 hour and terminated by adding 50 ul of 40 mM EDTA and 2 ul of 10 mg/ml protease K. The mixture was incubated for an additional 30 minute at 37°C. The reaction product was purified by phenol/chloroform/isoamyl alcohol extraction followed by ethanol precipitation and dissolved in 10 ul TE (pH 7.4). Restriction digestions were carried out in 20 ul mixtures containing purified substrates, different

restriction enzymes and buffers at 37°C for 5 hours. 2 ul of 10 X agarose gel loading buffer (0.02% bromophenol blue, 25% sucrose, 0.01 M EDTA, 1% SDS, 0.25% xylene cyanol) was added to the mixture and the products were separated by electrophoresis in 1% agarose gel in 1 X TBE buffer (Tris-Boric acid 89 mM, EDTA 2 mM). For restriction digestion with Hpa II and EcoR I, 10 ul of 50% glycerol was added to the mixture and the products were resolved by electrophoresis in 6% native polyacrylamide gel in 1 X TBE buffer. The gels were then analyzed by autoradiography and quantified by ImageQuant software using Typhoon Imager (Filter, 390BP; Laser, Red 633 nM).

**G. Preparation of radiolabeled inosine-containing plasmid**

The reaction mixture (50ul) containing 200 nM single-stranded oligo [5'-IATCCACAGTGAGTCCGAT-3') (10 pmol), 120 uCi $\gamma^{32}$P-ATP (26.4 pmol), 1 X T4 polynucleotide kinase Buffer (70 mM Tris-HCl (pH 7.6), 10 mM $MgCl_2$, 5 mM dithiothreitol] and 30 U of T4 polynucleotide kinase was incubated at 37°C for 4 hours. The mixture was then heat treated at 60°C for 20 minutes to denature the T4 polynucleotide kinase. The mixture was passed through a YM-3 spin column and washed extensively with TE buffer (pH 7.6) till there was little radioactivity in the flow-through. The purified labeled oligo (5 pmol) and another single-stranded oligo (5'-phos-TGAGGAGATCTG-3') (10 pmol) were annealed with 10 ug gapped pUC19EN3 in the presence of 1 X T4 DNA ligase buffer [50 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, 10 mM dithiothreitol, 1 mM ATP, 25 ug/ml BSA] at 68°C for 2 minutes and gradually lowered temperature to 25°C in 30 minutes. T4 DNA ligase (24 U) was added to the mixture and

incubated at 4°C overnight. The ligated closed circular plasmid was run on 1% agarose gel and recovered by glass wool spin column method.

**H. Determination of the released inosine-containing fragments in cell-free assay**

The reaction mixture contained 100 ng labeled plasmid, 10 mM HEPES (pH 7.4), 2 mM DTT, 0.5 mM EDTA, 2 mM ATP, 20 uM of dNTP, 0.5 mM NAD, 50 ug/ml BSA, 5 mM $MgCl_2$, 1 mM glutathione, 0.2 mM spermidine, and 4 mg/ml of *E. coli* AB1157 cell-free extract proteins. The mixture was incubated at 37°C for 1 hour and passed through a YM-100 spin column (cutoff: single-stranded DNA 300 bps, double-stranded DNA 125 bps) and washed three times with 25 ul TE buffer (pH 7.4). The flow-through was then passed through a YM-3 spin column (cutoff: single-stranded DNA 10 bps, double-stranded DNA 10 bps) and washed three times with 25 ul TE buffer (pH 7.4). The flow-through from the YM-3 spin column, which contains DNA fragments smaller than 10 bps, was collected. The remaining radioactivity, which contains DNA fragments smaller than 300 bps and bigger than 10 bps, on YM-3 was also collected. The radioactivities were determined by scintillation counter (Model LS 6500 Beckman Coulter).

## IV. Results

### A. Substrate quality and yield

pUC19EN3 (60 ug) can generate about 9.6 ug of inosine-containing plasmid. The inosine-containing substrate and control plasmid pUC19EN3 were treated with *E. coli* endo V, and over 95% of the closed circular form of inosine-containing substrate was converted to the nicked form (Figure 4.2, Lane 5) while less than 10% of the control plasmid was converted to the nicked form (Figure 4.2, Lane 3). It is reported that endo V has weak nicking activity on plasmid DNA (21).

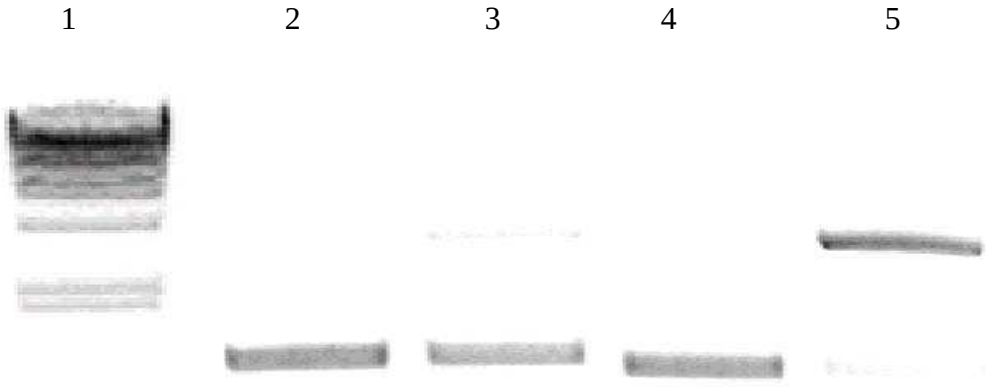



**Figure 4.2 Digestion of inosine-containing circular substrate by *E. coli* endonuclease V.** Inosine-containing substrate (200 ng) was digested with 20 nM *E. coli* endoV in a 10ul reaction buffer for 30 minutes at 37°C. Lane 1, λDNA Hind III digestion. Lane 2, pUC19EN3 Lane 3, pUC19EN3 digested by *E. coli* endoV Lane 4, pUC19EN3-I.C Lane 5, pUC19EN3-I.C digested by *E. coli* endo V

### B. DNA repair synthesis in inosine-containing plasmid after incubation with *E. coli* cell-free extracts

pUC19EN3-I.C (Figure 4.3A) which was identical except for the presence of an incorrect I.C pair was used to determine whether endo V is responsible for the repair

synthesis induced by inosine lesion. Cell-free assays were carried out using wt *E. coli* cell-free extracts in the presence of $\alpha^{32}$P-dCTP. After incubation, $\alpha^{32}$P-dCMP containing products were isolated and digested with restriction endonuclease Afl III and AlwN I. The DNA fragments were resolved by electrophoresis in 1% agarose gel, followed by Phosphor Image analysis. Figure 4.3B showed that control plasmid pUC19EN3 did not show significant repair synthesis (Fig 4.3B, Lane 1), while pUC19EN3-I.C, which contains a single inosine lesion, had significant amount of radioactive incorporation (Fig 4.3B, Lane 2), indicating that the repair synthesis is inosine dependent. Lanes 3-5 showed that *E. coli* cell extracts from *nfi* mutants did not induce repair synthesis, suggesting that endo V is responsible for initiating the inosine repair.

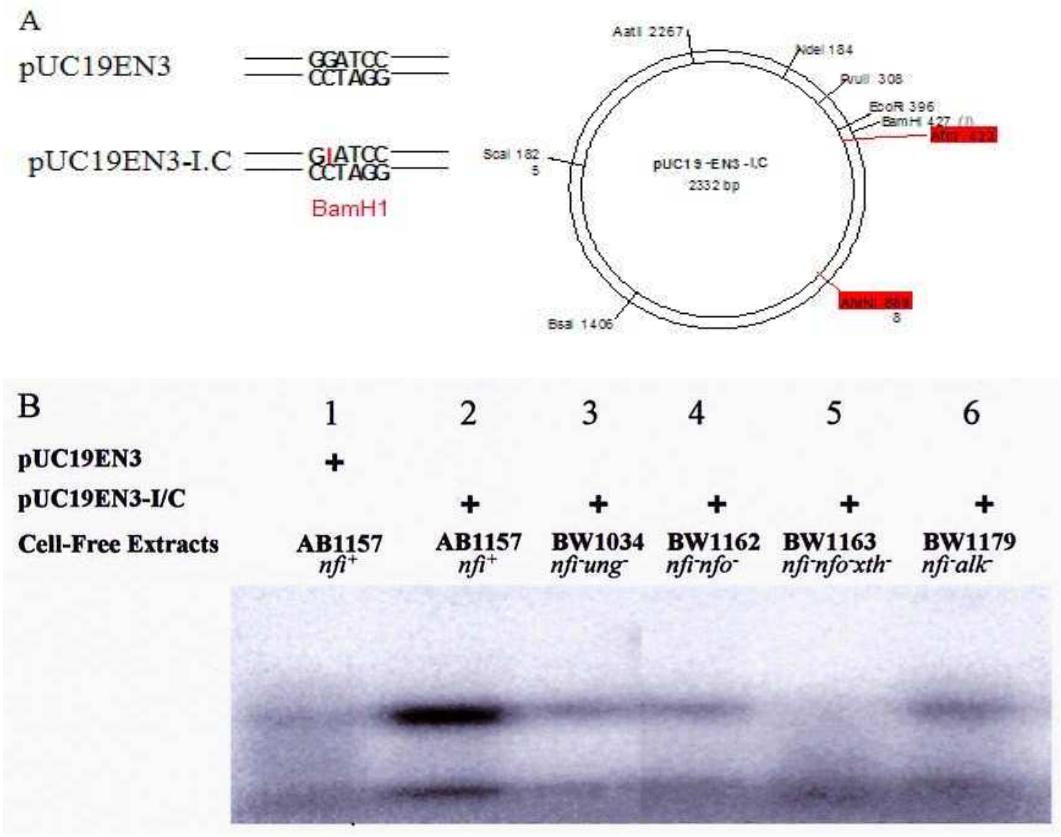



**Figure 4.3 Inosine-lesion mediated DNA repair synthesis in *E. coli* cell-free extracts.** **A.** pUC19EN3-I.C restriction map. **B.** pUC19EN3 and pUC19EN3-I.C were incubated with *E. coli* cell-free extracts at 37°C for 1 hour and terminated by adding 50 ul of 40 mM EDTA and 2 ul of 10 mg/ml protease K. The mixture was incubated for additional 30 minute at 37°C. The reaction products were purified by phenol/chloroform/isoamyl alcohol extraction followed by ethanol precipitation and dissolved in 10 ul TE (pH 7.4) and subject to AlwN 1 and Afl III digestion and resolved by 1% agarose gel. The gel was then analyzed by autoradiography. Lane 1, *E. coli* AB1157 (wt) cell-free extract incubate with pUC19EN3 (homoduplex, G.C pair). Lane 2, *E. coli* AB1157 (wt) cell-free extract incubate with pUC19EN3-I.C (heteroduplex I.C pair). Lane 3, *E. coli* BW1034 (*nfi*$^-$ *ung*$^-$) cell-free extract incubated with pUC19EN3-I.C. Lane 4, *E. coli* BW1162 (*nfi*$^-$ *nfo*$^-$) cell-free extract incubated with pUC19EN3-I.C. Lane 5, *E. coli* BW1163 (*nfi*$^-$ *nfo*$^-$*xth*$^-$) cell-free extract incubate with pUC19EN3-I.C. Lane 6, *E. coli* BW1179 (nfi$^-$ alkA$^-$) cell-free extract incubate with pUC19EN3-I.C.

There are twelve Hpa II recognition sites and one EcoR I recognition site located in the pUC19EN3-I.C, but only 8 fragments generated by Hpa II and EcoR I digestion are larger than 100 bp (Figure 4.4A-B). Among these fragments, the 301-bp fragment generated by EcoR I and Hpa II digestion is located 31 nt 5' to the inosine lesion. In

order to define the distribution of DNA repair synthesis associated with inosine-mediated DNA repair involving the inosine-containing plasmid, cell-free assays were carried out using *E. coli* cell-free extracts in the presence of $\alpha^{32}$P dCTP. After incubation, $\alpha^{32}$P-dCMP containing products were isolated and digested with restriction endonuclease Hpa II and EcoR I. The DNA fragments were separated by a 6% native polyacrylamide gel electrophoresis (Figure 4.4). Following the electrophoresis, Phosphor Image analysis of the $\alpha^{32}$P labeled DNA fragments showed that $\alpha^{32}$P-dCMP was incorporated into the whole plasmid, while the 301 bp fragment had the strongest radioactivity incorporation. The results suggested that multiple pathways may be involved in the repair of the inosine-containing substrates in the cell-free assays. Control reactions were conducted using a pUC19EN3 plasmid (homoduplex) containing a G:C base pair at the target site to determine the extent of $\alpha^{32}$P-dCMP incorporation at the same assay condition. The results showed that there was no significant $\alpha^{32}$P-dCMP incorporation in the plasmid, suggesting that the repair synthesis of inosine-containing plasmid is lesion mediated.

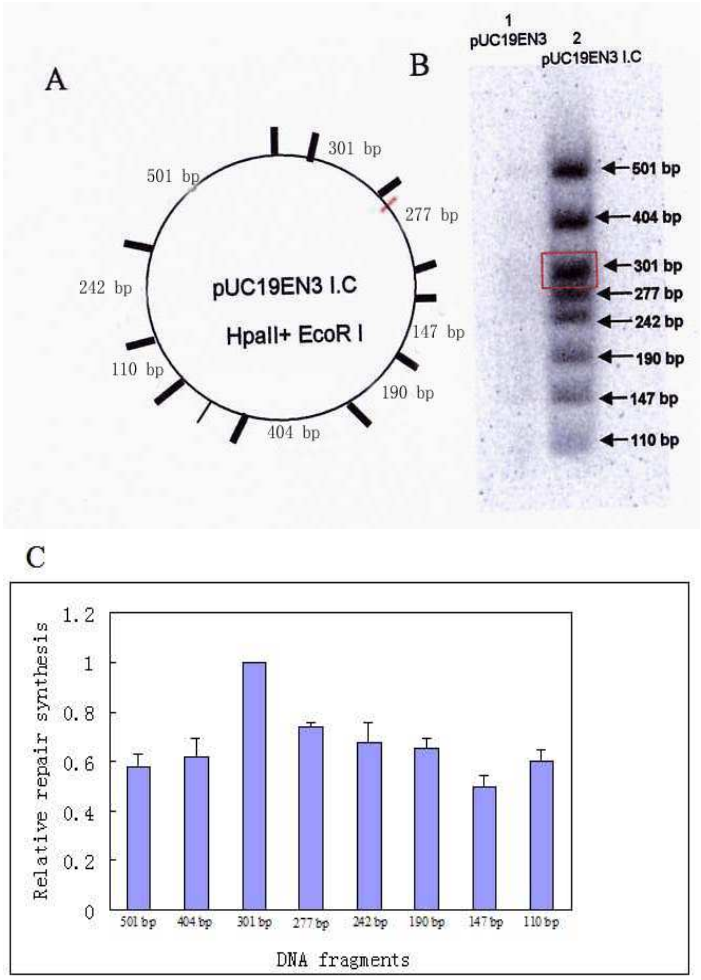



**Figure 4.4 Characterization of inosine mediated repair synthesis in *E. coli* AB1157 (wt) cell-free extract.** pUC19EN3-I.C and pUC19EN3 were incubated with *E. coli* AB1157 cell-free extract at 37°C for 1 hour. The reaction products were purified and subject to Hpa II and EcoR I digestion and resolved in 6% native polyacrylamide gel. The gel was then analyzed by autoradiography and quantified by ImageQuant software using Typhoon Imager. **A.** pUCEN3-I.C restriction map. **B**. Autoradiography of the gel Lane 1, *E. coli* AB1157 cell-free extract incubate with pUC19EN3 Lane 2, *E. coli* AB1157 cell-free extracts incubated with pUC19EN3-I.C. **C**. Relative intensity of radioactivity incorporation. The incorporation of $\alpha^{32}$P-dCMP in each DNA fragment was normalized by dividing the amount of $\alpha$-$^{32}$P radioactivity detected by the number of cytosine residues in each corresponding DNA fragment. The 301 bp fragment which has the highest normalized value was designated as 1, and the $\alpha^{32}$P-dCMP incorporation in other DNA fragments was calculated relative to that value. The data are the average from three reactions and the standard error is calculated.

**C. Analysis of the released inosine-containing fragments in *E. coli* cell-free extracts**

A synthesized single-stranded oligo was phosphorylated and labeled with $\gamma^{32}$P-ATP at the inosine position. The labeled oligo and another single-stranded oligo were annealed with gapped pUC19EN3 and ligated using T4 DNA ligase to form a closed circular plasmid. The closed circular plasmid was incubated with wt *E. coli* AB1157 cell-free extracts and subsequently passed through YM-100 spin column and YM-3 column (Figure 5.5). The radioactivity remained on the YM-3 spin column was 55000 cpm and the radioactivity in the flow-through was 293000 cpm determined by scintillation counter. The results showed that the sizes of most of the released fragments are less than 10 bps.

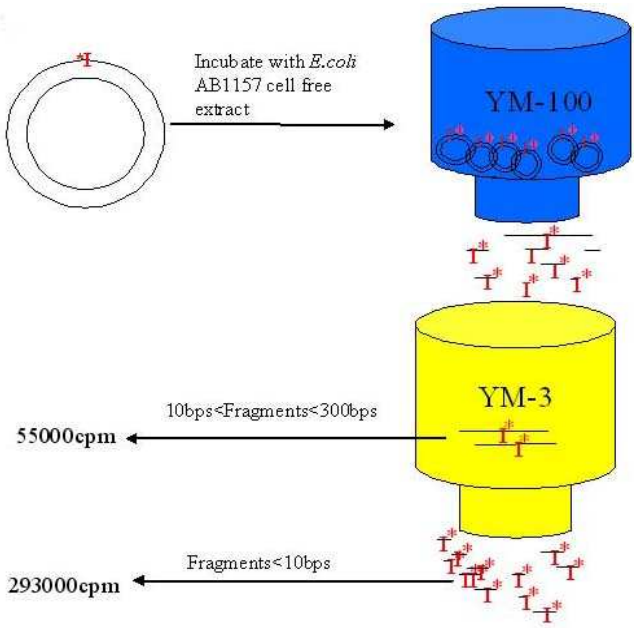



**Figure 4.5 Analysis of the released inosine-containing fragments by *E. coli* cell-free extract.** Inosine-containing pUC19EN3 were labeled with $\gamma^{32}$P-ATP at the inosine position, and incubated with *E. coli* AB1157 wt cell-free extract at 37°C for 1 hour. The mixture was passed through YM-100 spin column; the flow-through was then passed through a YM-3 spin column. The radioactivity in the flow-through and the radioactivity remained on the YM-3 column were counted using scintillation counter.

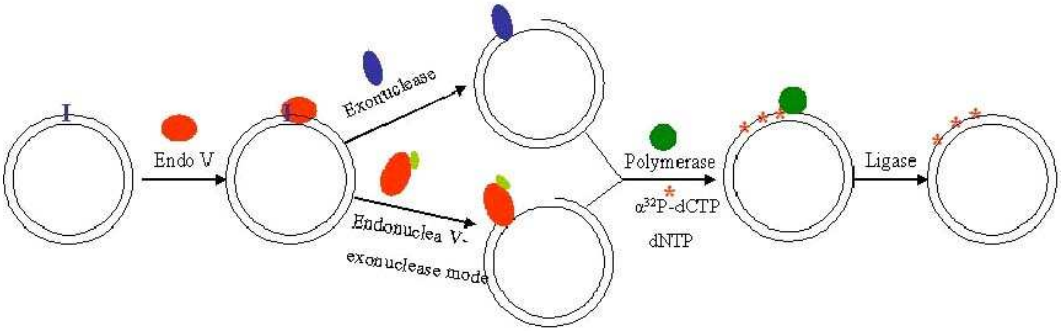
## V. Discussion

### A. Endonuclease V-initiated DNA repair

The pathway of endonuclease V-mediated DNA repair is still unknown. In order to unravel the problem, we first constructed a closed circular substrate which contained a single inosine lesion. We then developed an *in vitro* cell-free system which uses *E. coli* cell-free extracts, inosine-containing closed circular substrates, dNTP, and $\alpha^{32}$P-dCTP to analyze endo V-mediated DNA repair synthesis. Our preliminary results showed that endo V-mediated DNA repair may involve multiple pathways. The preliminary results from the cell-free assays suggested that the repair patches of *E. coli* initiated DNA repair pathways may consist of a long patch (longer than 2.3 kb, estimated) and a short patch pathway (less than 330 bps, estimated). After endo V nicks the inosine-containing substrate, there may be two scenarios (Figure 4.6). It is possible that endo V binds with the nicked inosine-containing substrate and recruits some exonuclease (Figure 4.6A). It is also possible that endo V recruits partner protein (s) and switches endonuclease V to an exonuclease mode. Recently, endonuclease V was shown to have a 3' exonuclease activity in certain conditions (22). Cao and colleagues proposed an alternative model that after endo V binds with the nicked inosine-containing substrate and recruits some protein(s), endo V changes to an exonuclease mode and removes the lesion (22). The 3' exonuclease activity removes the inosine lesion together with a portion of the plasmid. A polymerase subsequently fills in the gap and the ligase seals the nick. This will result in the incorporation of the radioactivity into the gap that is 5' to the inosine lesion site, a short patch repair. The model has been proposed by several groups (13-15). Analysis of released fragments by *E. coli* cell-free extracts showed that most of the released

fragments were less than 10 bps, suggesting that exonuclease activity is involved in the pathway. Another scenario is that after the polymerase fills in the gap, it starts strand displacement reaction. After the synthesis of a new DNA strand, an endonuclease nicks the strand and releases the displaced old strand, resulting in a long patch of the incorporation of radioactivity (Figure 4.6 B).

A




B

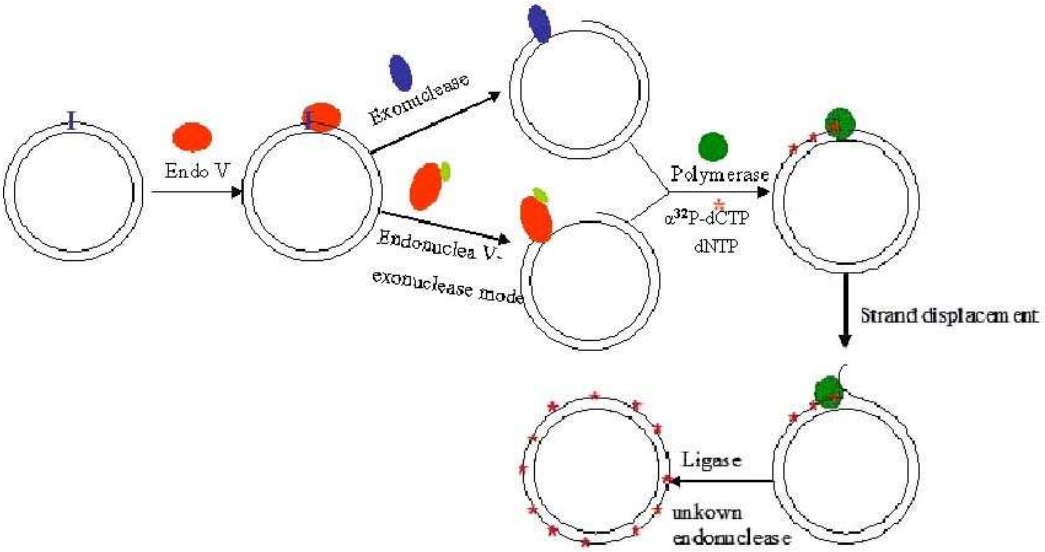



**Figure 4.6 Hypothetical Model for Endonuclease V initiated DNA repair pathway in a single lesion-containing plasmid (from preliminary studies).** A, in this pathway, endonuclease V first nicks at the second phosphodiester bond 3' to the inosine lesion and stay bound with the substrate. An exonuclease displaces endonuclease V and removes the inosine lesion together with a portion of the substrates; alternatively, a partner protein switches endonuclease V to an exonuclease mode and removes the inosine lesion, generating a gap in the plasmid. A DNA polymerase fills in the gap and *E. coli* ligase seals the nick. B, in this pathway, after the exonuclease activity removes the inosine lesion and generates a gap, a polymerase fills in the gap and starts strand displacement. After the synthesis of a new DNA strand, an endonuclease nicks the strand and releases the displaced old strand. A DNA polymerase fills in the gap and *E. coli* ligase seals the gap.

**B. Recent research progresses on endonuclease V**

An alternative way to study the endo V mediated pathway is to find the key components of this pathway. We have used gel shift assays to test different possible enzymes but failed to identify one that interacts with endo V (data not shown). Genetic screening of the strains that are sensitive to nitrosative stress may also provide us important information. It is conceivable that multiple sites will be deaminated when the cells encounter significant nitrosative stress. If the two sites are very close to each other and located in different strands, double strand breaks may occur. In this case, recombinational repair is a pathway to repair the lesions. RdgB is an NTPase with a 100-fold preference for ITP, XTP and dITP than the canonical nucleotides (23-24). RdgB mutants accumulate Endo V-recognized modification and show elevated double strand breaks (25). It is speculated that RdgB may play a role in the 5' incision around the lesion site after endo V nicks the lesion containing DNA (25). More experiments need to be done to understand this pathway.

## VI. Acknowledgements

We thank Bernard Weiss (Emory University School of Medicine, Atlanta, GA) for *E. coli* strains and John B.Hays (Oregon State University) for pUC19HE. We also thank Huixian Wang (Oregon State University) for his advice on preparing inosine-containing closed circular substrates. We thank members of Cao lab for stimulating discussions. This project is supported in part by the NIH, the Concern Foundation and SC Experimental Station.

## VII. References


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

2. Shen,J.C., Rideout,W.M. 3rd and Jones,P.A. (1994) The rate of hydrolytic deamination of 5-methylcytosine in double-stranded DNA. *Nucleic Acids Res.*, 22, 972–976

3. Krokan HE,Otterlei M, Nilsen H, Kavli B, Skorpen F, Andersen S, Skjelbred C, Akbari M, Aas PA, Slupphaug G. Properties and functions of human uracil-DNA glycosylase form the *Ung* gene. *Prog. Nucleic Acids Res.* 2001;68:366–386.

4. Shapiro, R., and Yamaguchi, H. Nucleic acid reactivity and conformation. I. Deamination of cytosine by nitrous acid (1972) *Biochim. Biophys. Acta* 281, 501-506

5. Lindahl, T., and Nyberg, B. Heat-induced deamination of cytosine residues in deoxyribonucleic acid. (1974) *Biochemistry* 13, 3405-3410

6. Ohshima,H. and Bartsch,H. (1994) Chronic infections and inflammatory processes as cancer risk factors: possible role of nitric oxide in carcinogenesis. *Mutat. Res.*, 305, 253–264

7. Shacter,E. and Weitzman,S.A. (2002) Chronic inflammation and cancer. *Oncology*, 16, 217–226

8. Hussain SP, Hofseth LJ, Harris CC. 2003. Radical causes of cancer. Nat Rev Cancer 3:276 –285.

9. M. Yao and Y.W. Kow, Further characterization of *Escherichia coli* endonuclease V. Mechanism of recognition for deoxyinosine, deoxyuridine, and base mismatches in DNA. *J. Biol. Chem.* 272 (1997), pp. 30774–30779.

10. M. Yao and Y.W. Kow, Interaction of deoxyinosine 3′-endonuclease from *Escherichia coli* with DNA containing deoxyinosine. *J. Biol. Chem.* 270 (1995), pp. 28609–28616

11. Guo, G. and B. Weiss, *Endonuclease V (nfi) Mutant of Escherichia coli K-12*. Journal of Bacteriology, 1998. 180(1): p. 46-51.

12. Karen A. Schouten and B. Weiss, Endonuclease V protects *Escherichia coli* against specific mutations caused by nitrous acid. Mutation Research 435 (1999) 245-254

13. He B, Qing H, Kow YW. Deoxyxanthosine in DNA is repaired by *Escherichia coli* endonuclease V. Mutat Res. 2000 Mar 20;459 (2):109-14.

14. Burgis, N. E., Brucker, J. J., and Cunningham, R. P. (2003) Repair system for noncanonical purines in *Escherichia coli*, *J. Bacteriol. 185*, 3101-10.

15. Moe, A., Ringvoll, J., Nordstrand, L. M., Eide, L., Bjoras, M., Seeberg, E., Rognes, T., and Klungland, A. (2003) Incision at hypoxanthine residues in DNA by a mammalian homologue of the *Escherichia coli* antimutator enzyme endonuclease V, *Nucleic Acids Res. 31*, 3893-900

16. Wang H, Hays JB. Signaling from DNA mispairs to mismatch-repair excision sites despite intervening blockades. EMBO J. 2004 May 19;23(10):2126-33. Epub 2004 Apr 22.

17. Lu AL, Clark S, Modrich P. Methyl-directed repair of DNA base-pair mismatches *in vitro*. Proc Natl Acad Sci U S A. 1983 Aug;80(15):4639-43.

18. Sanderson RJ, Mosbaugh DW. Fidelity and mutational specificity of uracil-initiated base excision DNA repair synthesis in human glioblastoma cell extracts. J Biol Chem. 1998 Sep 18;273(38):24822-31.

19. Feng H, Klutz AM, Cao W. Active site plasticity of endonuclease V from Salmonella typhimurium.Biochemistry. 2005 Jan 18;44(2):675-83.

20. Joseph Sambrook, David W. Russell Molecular Cloning: A Laboratory Manual 2000

21. Huang J, Lu J, Barany F, Cao W. Multiple cleavage activities of endonuclease V from *Thermotoga maritima*: recognition and strand nicking mechanism. Biochemistry. 2001 Jul 31;40(30):8738-48

22. Feng H, Dong L, Klutz AM, Aghaebrahim N, Cao W. Defining amino acid residues involved in DNA-protein interactions and revelation of 3'-exonuclease activity in endonucleaseV. Biochemistry. 2005 Aug 30;44(34):11486-95.

23. Chung JH, Back JH, Park YI, Han YS. Biochemical characterization of a novel hypoxanthine/xanthine dNTP pyrophosphatase from Methanococcus jannaschii. Nucleic Acids Res. 2001 Jul 15;29(14):3099-107.

24. Chung JH, Park HY, Lee JH, Jang Y. Identification of the dITP- and XTP-hydrolyzing protein from *Escherichia coli*. J Biochem Mol Biol. 2002 Jul 31;35(4):403-8.

25. Bradshaw JS, Kuzminov A. RdgB acts to avoid chromosome fragmentation in *Escherichia coli*. Mol Microbiol. 2003 Jun;48(6):1711-25.

# CHAPTER FIVE

## Research Significance and Concluding Remarks

DNA damages may arise as a consequence of exposure to various DNA damage agents. In order to maintain the genetic integrity of chromosomes, it is essential that DNA damages are repaired efficiently. Fortunately, the genome is under continuous surveillance by DNA repair mechanisms. Defects in DNA repair cause hypersensitivity to DNA-damaging agents, accumulation of mutations in the genome and finally to the development of cancer and various metabolic disorders (1). So, it is important to understand different DNA repair pathways in order to prevent or cure these diseases. Endonuclease V is a ubiquitous enzyme found in different species from bacteria to mammals. It recognizes deaminated lesions including deoxyinosine (2), deoxyuridine (3), deoxyxanthosine (4), and deoxyoxanosine (5). It was suggested that endo V prevents mutations from nitrosative deamination *in vivo* (6).

Endo V also has mismatch cleavage activity (7). The detection of the mutations in DNA is important in the diagnosis of both hereditary diseases and acquired diseases. Barany and colleagues in Cornell University developed an endo V/ligase mutation scanning method which has very high sensitivity and is suitable for low-frequency known or unknown mutation detection (8). However some A/C mismatches embedded in G/C rich sequences are refractory to endo V cleavage, such as the G13D mutation in K-*ras* (8). Despite intensive studies by researchers, the DNA repair pathway of endonuclease V

remains unknown. Several models have been put forward (9-12). Most interestingly, in the presence of $Mn^{2+}$, endo V displays a 3' exonuclease activity. A model was proposed by Cao and colleagues that endo V may play a dual role in the repair pathway (12). After endo V cleaves the lesion-containing DNA and forms a complex with the product, downstream proteins may be recruited to the repair complex, and change the conformation of endo V. This may trigger the endo V's exonuclease activity and remove the lesion and generate a gap for polymerase and ligase to finish the repair.

My projects were aimed to broaden our knowledge for the biochemical properties of endo V and its application in mutation scanning. We are also interested in understanding endo V initiated DNA repair pathway. We have screened over 60 endo V single-site mutants previously isolated and identified endo V mutants with altered base preference in mismatch cleavage. Y80A, a mutant in motif III, exhibits strong preference for C in mismatches. Taking advantage of this property, Y80A was successfully used to detect G13D mutation in K-*ras* (13), an A/C mismatch embedded in a G/C rich sequence context which was previously inaccessible using the wild-type endo V. We also devised a method to covert blunt end PCR products to sticky end products thus reduced the fluorescence signal loss during mismatch cleavage assay using endo V. A hypothetic model was proposed to explain the dramatic change in base preference (12). In the wild-type enzyme, tyrosine 80 (Y80) imposes an unfavorable interaction with a C base, in which the amino group at C4 position spatially clashes with the bulky tyrosine residue. This steric effect prevents the wild-type endo V from recognizing and cleaving C-containing strand in a mismatch. By replacing the phenol side-chain with a small methyl group, Y80A releases the steric tension and allows the C base to be accommodated in the

recognition pocket. This investigation provides us insights on base recognition and active site organization. Protein engineering in endonuclease V may translate into better tools in mutation recognition and cancer mutation scanning.

In view of the interesting properties possessed by several Y80 mutants, we further mutated and characterized the other sixteen Y80 mutants (14). The nineteen mutants were characterized biochemically using mismatch-containing DNA and deaminated base-containing. Substitutions with aromatic amino acid residues retained partial binding affinity as revealed by gel mobility shift analysis with double-stranded inosine-containing DNA. An adenosine/uridine base pair in a uridine-containing double-stranded DNA was only cleaved by the wild type endo V in the presence of $MgCl_2$. A guanosine/uridine base pair in a uridine-containing double-stranded DNA was cleaved by the wild type endo V and to a much less extent by Y80F in the presence of $MgCl_2$, indicating an essential role of the aromatic ring and the hydroxyl group in base recognition. For mismatched base pairs, Y80H preferred to cleave thymidine-strand in a thymidine-containing mismatch; Y80P and Y80S, along with previously identified Y80A, tended to cleave the cytosine-strand in a cytosine-containing mismatch, Y80R preferred to cleave thymidine- and cytosine-strand in a mismatch. The ability to alter cleavage preferences in a mismatch further underscores the role of Y80 in base recognition. The studies showed that aromatic ring and hydrogen bonding are very important in base recognition for endo V and offers new insight on the role of Y80 in base recognition.

Endonuclease V initiated DNA repair pathway remains unknown. No other component of this pathway is identified. In order to understand the DNA repair pathway initiated by endo V, we developed an *in vitro* cell-free system which uses *E. coli* cell-free

extracts, inosine-containing closed circular substrates, dNTP, and $\alpha^{32}$P-dCTP to analyze endo V mediated DNA repair synthesis. The preliminary results from cell-free assays indicated that the repair patch of endo V initiated DNA repair pathways may consist of a long patch (longer than 2.3 kb, estimated) and a short patch (less than 330 bps, estimated). More experiments need to be done to understand the pathway mediated by endo V.

**References**


1. Bohr VA. DNA damage and its processing, relation to human disease. J Inherit Metab Dis. 2002 May;25(3):215-22. Review.

2. M. Yao, Z. Hatahet, R.J. Melamede and Y.W. Kow , Purification and characterization of a novel deoxyinosine-specific enzyme, deoxyinosine 3′ endonuclease, from *Escherichia coli*. *J. Biol. Chem.* 269 (1994), pp. 16260–16268.

3. M. Yao and Y.W. Kow, Further characterization of *Escherichia coli* endonuclease V. Mechanism of recognition for deoxyinosine, deoxyuridine, and base mismatches in DNA. *J. Biol. Chem.* 272 (1997), pp. 30774–30779.

4. He B, Qing H, Kow YW. Deoxyxanthosine in DNA is repaired by *Escherichia coli* endonuclease V. Mutat Res. 2000 Mar 20;459 (2):109-14.

5. Hitchcock, T.M., Gao, H., Cao, W. (2004) Cleavage of deoxyoxanosine-containing oligodeoxyribonucleotides by bacterial endonuclease V *Nucleic Acids Res.*, 32, 4071–4080

6. Weiss, B., Endonuclease V *of Escherichia coli* prevents mutations from nitrosative deamination during nitrate/nitrite respiration. Mutat Res, 2001. 461(4): p. 301-9.

7. M. Yao and Y.W. Kow, Strand-specific cleavage of mismatch-containing DNA by deoxyinosine 3′-endonuclease from *Escherichia coli*. J. Biol. Chem. 269 (1994), pp. 31390–31396.

8. Huang J, Kirk B, Favis R, Soussi T, Paty P, Cao W, Barany F. An endonuclease/ligase based mutation scanning method especially suited for analysis of neoplastic tissue. Oncogene. 2002 Mar 14;21(12):1909-21.

9. Kow YW. Repair of deaminated bases in DNA. Free Radic Biol Med. 2002 Oct 1;33(7):886-93

10. Burgis NE, Brucker JJ, Cunningham RP. Repair system for noncanonical purines in *Escherichia coli*. J Bacteriol. 2003 May;185(10):3101-10.

11. Moe A, Ringvoll J, Nordstrand LM, Eide L, Bjoras M, Seeberg E, Rognes T, Klungland A. Incision at hypoxanthine residues in DNA by a mammalian homologue of the *Escherichia coli* antimutator enzyme endonuclease V. Nucleic Acids Res. 2003 Jul 15;31(14):3893-900

12. Feng H, Dong L, Klutz AM, Aghaebrahim N, Cao W. Defining amino acid residues involved in DNA-protein interactions and revelation of 3'-exonuclease activity in endonuclease V. Biochemistry. 2005 Aug 30;44(34):11486-95.

13. Gao H, Huang J, Barany F, Cao W. Switching base preferences of mismatch cleavage in endonuclease V: an improved method for scanning point mutations. Nucleic Acids Res. 2007;35(1):e2. Epub 2006 Nov 27.

14. Gao H and Cao W. Role of Tyrosine 80 in Base Recognition in Endonuclease V from *Thermotoga maritime*. Manuscript in preparation